The Bell System Technical Journal

Vol. XVIII *January, 1939* *No. 1*

Electrostatic Electron-Optics

By **FRANK GRAY**

Certain types of electrostatic fields may be used as lenses to focus electron beams. The theory of these lenses is developed for electric fields that are symmetrical about a central axis. The introduction of two velocity functions exactly reduces the partial differential equations of electron motion to a series of ordinary differential equations. The first equation describes the action of a lens for electron paths near the axis; the remaining equations determine the higher order aberration terms. Sections on the following subjects are included: the general equations of electron-optics, thin lenses, thick lenses, aberration, the reduction of aberration, apertured plates, and concentric tubes. A list of symbols and lens equations is also included at the end of the article.

IN certain types of modern vacuum tubes, a beam of electrons is brought to a focus by an electrostatic field whose action on the beam is analogous to that of an optical lens on a beam of light. An electrostatic field which acts in this manner is called an electron lens. Such lenses are rapidly finding applications in amplifier tubes, television and oscillograph tubes, electron microscopes, and various types of experimental apparatus. As the extent of their application widens, the theory of these lenses naturally assumes a corresponding importance.

The first articles on the new science of electron-optics were published by Bush[1] in 1926–1927, and the next important step in its development was taken by Davisson and Calbick[2] and by Brüche and Johannson[3] working independently in 1931–1932. The following years marked an increased interest in the subject, with comprehensive articles by various authors, and its literature expanded rapidly. An

---

[1] H. Bush, *Ann. d. Physik, 81*, 974, 1926 and *Arch. f. Elecktrotech., 18*, 583, 1927.

[2] C. J. Davisson and C. J. Calbick, *Phys. Rev., 38*, 585, 1931 and *Phys. Rev., 42*, 580, 1932.

[3] E. Brüche and N. Johannson, *Ann. d. Physik, 15*, 145, 1932.

excellent review of this literature and the history of electron-optics are given in a symposium[4] of papers published in 1936, and the various practical applications of electron lenses are well described in the books on that subject.[5]

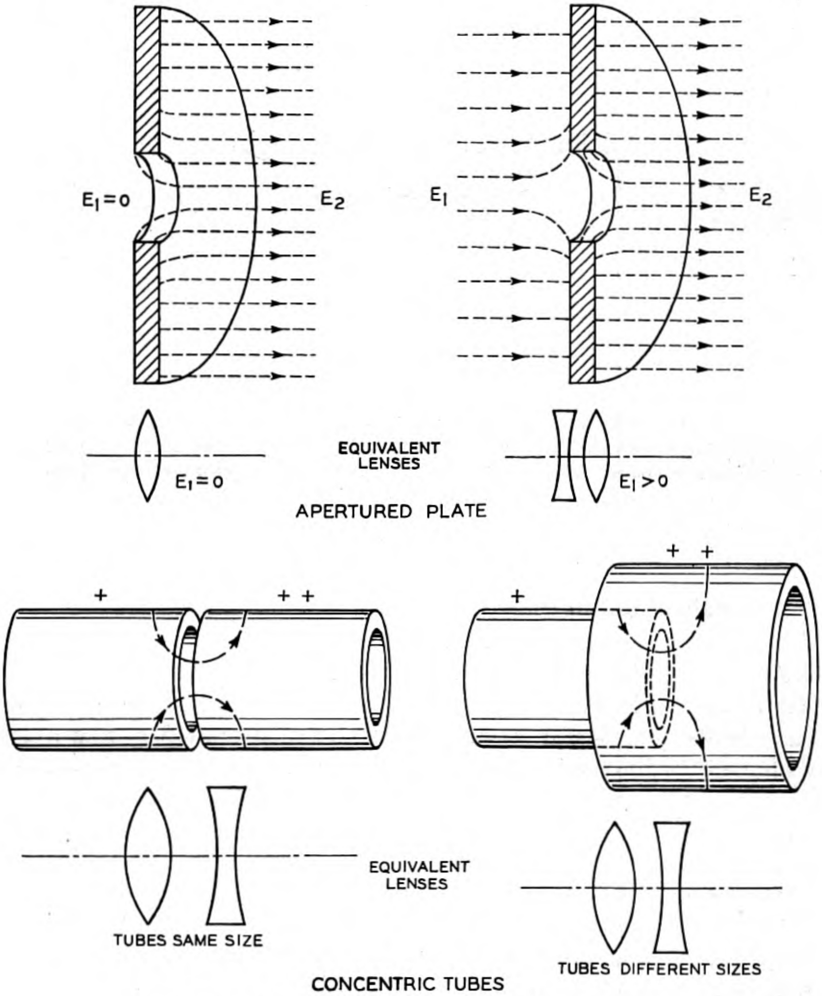


Fig. 1—Lines of force in typical electron lenses.

---

[4] *Zeit. f. techn. Physik*, *17*, 584–645, 1936.

[5] E. Brüche and O. Scherzer, *Geometrische Electronenoptik* (Springer, 1934). J. T. MacGregor-Morris and J. A. Henley, "Cathode Ray Oscillography" (Monographs on Electrical Engineering), 1936. Maloff and Epstein, "Electron Optics in Television" (McGraw-Hill, 1938).

The theory of electron-optics is thus well established and any further attempts at the subject must lead to substantially the same results. There is, however, a need for a precise development of the theory in a simpler manner. With this need in mind, the present article approaches the subject in a manner that appeals to the reader who is more familiar with electrical theory than he is with the concepts of geometrical optics, and this approach leads clearly to the various approximations that are needed in the development of the theory. With the aid of two velocity functions, the partial differential equations of electron motion are briefly and exactly reduced to a series of ordinary differential equations; the theory is then developed in terms of their approximate solutions.

Attention is confined to systems in which the electric fields are symmetrical about a central axis. In such systems any field having a radial component of electric intensity changes the radial velocity of an electron passing through it, and thus behaves—to some extent at least—as an electron lens. A uniform field parallel to the axis and field-free space are the only regions in which there is no lens action. Typical electron lenses are shown in the figures on the second page. As illustrated by these examples, a practical electron lens is characterized by a short region in which there is an abrupt change in the electric intensity parallel to the axis. Lines of force are continuous, and the field parallel to the axis can change only by lines of force coming into it, or going out from it, in a radial direction. In the region of the abrupt change, there are consequently strong radial fields which can deflect an electron in a radial direction. The region changes the focus of an electron beam passing through it, and its action is analogous to that of an optical lens.

SECTION I—THE GENERAL EQUATIONS

In the present paper it is assumed that the initial electron source has perfect symmetry of form about the central axis, and that the electrons have no appreciable velocities of emission from the source. An electron thus has no angular velocity about the axis, and its motion may be described in terms of a coordinate z taken along the axis and a radial coordinate r measured from the axis.

If an electron's velocity vector is projected at any point along its path, it intersects the axis at some point p, as illustrated in Fig. 2, and the electron may be regarded as instantaneously moving either away from, or else toward, this point of intersection. The distance d along the axis from the electron to the point of intersection is called the instantaneous focal distance.[6] Defined in this manner, the focal distance conforms with the optics of light; it is positive when an electron is moving toward a focal point; and is negative when the electron is moving away from such a point. From the geometry of the figure, it is seen that

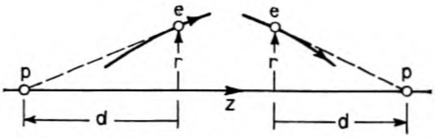


Fig. 2—Focal distance.

$$d = -\frac{r\dot{z}}{\dot{r}}, \tag{1}$$

where $\dot{r}$ and $\dot{z}$ are the instantaneous components of electron velocity.

The focal distance of an electron varies continuously as the electron moves along. The simplest variation occurs in field-free space, where the electron travels in a straight line and the focal point remains stationary; but even then the focal distance varies as the electron moves; for the focal distance is measured from the moving electron to the stationary focal point, as illustrated in Fig. 3. In an electron lens, the focal point of an electron also shifts continuously as the electron moves through the lens and the focal distance varies in a complicated manner.

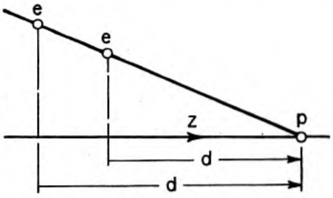


Fig. 3—Focal distances, field-free space.

The values of d at the two sides [7] of an electron lens, for any electron path, are called conjugate focal distances of the lens, and are usually designated as d_1 and d_2. The theory of electron-optics is largely concerned with the derivation of an equation relating these conjugate focal distances.

---

[6] The term is here used in a broad sense to include the distance to any intersection point on the z-axis, even though the latter is not the point of convergence of an electron beam.

[7] The value of d as an electron enters the non-uniform field of the lens, and the value of d as it leaves the non-uniform field.

Before passing on to such a derivation, it is well to introduce another quantity, which is analogous to focal distance and very useful in making approximations. Suppose that, from any point along its path, an electron were to continue on with its instantaneous velocity in a straight line. Its velocity along the axis would continue to have the instantaneous value $\dot{z}$, and the electron would travel over the distance d and arrive at the focal point in a period of time T given by

$$T = d/\dot{z} \tag{2}$$

or from equation 1

$$T = -r/\dot{r}. \tag{3}$$

This period of time is analogous to focal distance, and we therefore call it focal time. The values of T at the two sides of an electron lens, for any electron path, are in a corresponding manner called conjugate focal times of the lens.

To obtain an equation relating the conjugate focal distances of a lens, we must consider the path of an electron through the lens. The path is determined by the initial velocity and coordinates of the electron as it enters the lens and by its acceleration in the electric field of the lens. By defining electrical units in the proper manner the ratio e/m is eliminated from the equations of acceleration and they assume the simple form

$$\ddot{r} = \frac{\partial \Phi}{\partial r}, \tag{4}$$

$$\ddot{z} = \frac{\partial \Phi}{\partial z}, \tag{5}$$

where Φ is the potential at points in space.[8]

The first solution of these equations gives the well known energy relation

$$\dot{r}^2 + \dot{z}^2 = 2\Phi, \tag{6}$$

where the electron source is taken as zero potential.

With the exception of special cases, the equations are not further soluble in the usual sense, and one resorts to solution in series.

As they stand, the two equations for acceleration are inconvenient; they involve partial derivatives of potential with respect to space and ordinary derivatives of velocity with respect to time, and the latter cannot be transformed to partial derivatives with respect to space, for the simple reason that the velocity of an electron does not exist

---

[8] The final equations of electron-optics involve the potentials only in the form of ratios which are independent of the electrical units.

at points off its path. The equations may, however, be reduced to a more convenient form by the introduction of two velocity functions[9] defined as follows.

Let u and w be any two functions of r and z that satisfy the equations

$$\frac{\partial u}{\partial z} - \frac{\partial w}{\partial r} = 0, \tag{7}$$

$$u^2 + w^2 = 2\Phi. \tag{8}$$

Consider now an imaginary point moving with velocity components

$$\dot{r} = u, \qquad \dot{z} = w. \tag{9}$$

The derivative of $\dot{r}$ with respect to time is

$$\ddot{r} = \frac{\partial u}{\partial r}\dot{r} + \frac{\partial u}{\partial z}\dot{z} = \frac{\partial u}{\partial r}u + \frac{\partial u}{\partial z}w \tag{10}$$

and from equations 7 and 8

$$\ddot{r} = \frac{\partial u}{\partial r}u + \frac{\partial w}{\partial r}w = \frac{\partial \Phi}{\partial r} \tag{11}$$

the component $\dot{r}$ thus satisfies differential equation (4) for electron motion. In a similar manner it may be shown that the velocity component $\dot{z}$ satisfies equation (5). The motion of the imaginary point is thus the same as the motion of an electron, and the velocity functions u and w are therefore the velocity components of electron motion.

The velocity functions are solutions of equations 7 and 8, one of which is a simple algebraic equation and the other a partial differential equation with respect to space alone. The inconvenient time derivatives have been eliminated in these new equations for electron velocity.

The existence of a velocity function is not confined to a single electron path; it exists over the electric field in general. Any pair of particular solutions for u and w thus corresponds to an infinite number of possible electron paths. In the converse manner, there are an infinite number of particular solutions for any electric field, and there is a pair of particular solutions corresponding to any given electron path through the field.[10]

---

[9] These functions are the components of the generalized vector function described in Appendix 4.

[10] The existence of such solutions is proved by the existence of the series solutions, which are derived in the following pages.

Solutions for the velocity functions are obtained by expressing them as power series in r.

$$u = Ar + Br^3 + Cr^5 + \cdots, \tag{12}$$

$$w = a + br^2 + cr^4 + \cdots, \tag{13}$$

where the coefficients are functions of z alone. The above powers of r are the ones required in a system symmetrical about the z-axis. In such a system $\dot{r}$ reverses in sign with r and the u-series is odd; $\dot{z}$ does not reverse sign with r and the w-series is even. Aside from such reasoning, the choice of the two series is justified provided they lead to solutions of the differential equation in a form suitable for the purposes of electron-optics.

The potential Φ obeys the equation

$$\Delta\Phi = \frac{\partial^2\Phi}{\partial z^2} + \frac{\partial^2\Phi}{\partial r^2} + \frac{1}{r}\frac{\partial\Phi}{\partial r} = 0 \tag{14}$$

and it may likewise be expressed as a power series in r. This well known series is

$$\Phi = v - \frac{v''}{2^2}r^2 + \frac{v''''}{(2\cdot 4)^2}r^4 \cdots, \tag{15}$$

where v is the potential on the axis of the system, and the primes indicate differentiation with respect to z.

On substituting the three series in equations 7 and 8 and equating the coefficients of the various powers of r in each equation we obtain a series of ordinary differential equations for the coefficients of the u-series.

$$\sqrt{2v}A' + A^2 = -\frac{v''}{2}, \tag{16}$$

$$\sqrt{2v}B' + 4AB = \frac{v''''}{16} - \frac{(A')^2}{2}, \tag{17}$$

$$\sqrt{2v}C' + 6AC = -\frac{v''''''}{384} - 3B^2 - 3/4A'B', \tag{18}$$

$$\cdots\cdots\cdots\cdots\cdots$$

and the coefficients of the w-series are

$$\begin{aligned} a &= \sqrt{2v}, \\ b &= A'/2, \\ c &= B'/4. \\ &\cdots \end{aligned} \tag{19}$$

The solution of the partial differential equations for electron velocity is thus reduced to the solution of a series of ordinary differential equations, which in themselves contain no approximations.

From equation 1, the inverse focal distance is now obtained by dividing u by r and w, which gives

$$\frac{1}{d} = -\frac{A + Br^2 + Cr^4 + \cdots}{\sqrt{2v} + \frac{A'}{2}r^2 + \frac{B'}{4}r^4 + \cdots}. \tag{20}$$

This is the general equation for focal distance as it is affected by aberration. In using this equation, we are at liberty to set the higher coefficients equal to zero at the incident side of the lens. This determines the initial value of A in terms of the first conjugate focal distance. The second conjugate focal distance is then determined by solving for the coefficients at the exit side of the lens. Due to the presence of the terms in r, this second focal distance varies slightly with the radial distance at which an electron passes through the lens, and the focus is therefore diffused along the axis. This diffusion of the focus is called aberration.

The coefficient A is of particular importance in the theory of electron-optics. For paraxial rays, that is, rays near the axis, the higher terms in the two series are negligibly small compared to their first terms, and for such rays

$$\frac{1}{d} = -\frac{A}{\sqrt{2v}}. \tag{21}$$

With the exception of aberration, the single coefficient A thus determines the complete performance of a lens, and the principal constants of a lens are determined by its differential equation alone. In lenses where the rays are confined to a region near the axis with proper diaphragms, the aberration terms are small and the coefficient A describes the performance of a lens sufficiently well.

The next section is devoted to the derivation of the principal lens equations from this coefficient. The aberration terms are considered only in the last section of the paper.

SECTION II—RAYS NEAR THE AXIS

For rays near the axis the optical characteristics of an electric field are determined by the differential equation for A alone,

$$\sqrt{2v}A' + A^2 = -\frac{v''}{2}. \tag{16}$$

For such rays the higher terms in r may be neglected in the general equations and we obtain the following useful relations

$$A = \dot{r}/r = -\frac{\sqrt{2v}}{d}, \tag{22}$$

$$\dot{z} = \sqrt{2v}, \tag{23}$$

$$T = d/\sqrt{2v} = -\frac{1}{A}, \tag{24}$$

$$dt = dz/\dot{z} = dz/\sqrt{2v}. \tag{25}$$

A Uniform Electric Field

A uniform electric field parallel to the axis is not usually regarded as an electron lens,[11] but it does shift the focal point of a beam of electrons passing through it. In a uniform field, v'' is zero and the differential equation for A may be written in the form

$$\frac{dA}{A^2} = -\frac{dz}{\sqrt{2v}}. \tag{26}$$

An integration of this equation from any point z_1 to any other point z_2, in the uniform field, gives

$$\frac{1}{A_2} - \frac{1}{A_1} = \frac{2(z_2 - z_1)}{\sqrt{2v_2} + \sqrt{2v_1}}, \tag{27}$$

where A_1 and A_2 are the values of A at z_1 and z_2. On substituting $-\sqrt{2v}/d$ for A in this equation, it may be transformed to

$$\left(1 + \sqrt{\frac{v_2}{v_1}}\right) d_1 - \left(1 + \sqrt{\frac{v_1}{v_2}}\right) d_2 = 2(z_2 - z_1), \tag{28}$$

which is the equation relating the conjugate focal distances at any two planes—located at z_1 and z_2—in the uniform field.

The shift in the focal point of an electron beam as it passes through a uniform field is illustrated in Fig. 4.

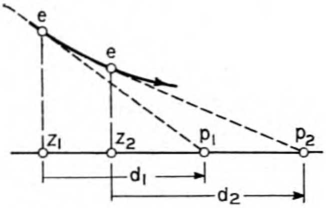


Fig. 4—Focal distances in a uniform field.

---

[11] Electron rays parallel to the axis are not bent by the field, and it does not magnify an electron image.

Thin Lenses

Approximate solutions of the differential equation 16 for A are obtained more clearly by first changing the space variables to time variables. This is done by using relations 24 and 25, which transform the equation to

$$\frac{1}{T^2}\left(\frac{dT}{dt} + 1\right) = -\frac{v''}{2} \tag{29}$$

or

$$\frac{1}{T^2}\frac{d}{dt}(T + t) = -\frac{v''}{2}\cdot \tag{30}$$

The new equation tells how the focal time T varies with time as an electron moves along.[12]

A thin lens is defined as a region of non-uniform field extending over such a short distance along the axis that an electron traverses it in a period of time small compared to the focal times involved: the thickness of the lens is small compared to the conjugate focal distances. By taking the origin of time t at the middle of an electron's period of transit through a lens, t in the lens is not greater than half the period of transit, and t may therefore be neglected in comparison to T in a thin lens. With this approximation in equation 31, it reduces to

$$\frac{d}{dt}\left(\frac{1}{T}\right) = \frac{v''}{2}\cdot \tag{31}$$

In integrating this equation through a lens we choose two points z_1 and z_2 at the approximate boundaries of the non-uniform field, that is, the points where v'' substantially drops to zero as illustrated in Fig. 5. Then, remembering that dt is $dz/\sqrt{2v}$, an integration from z_1 to z_2 gives

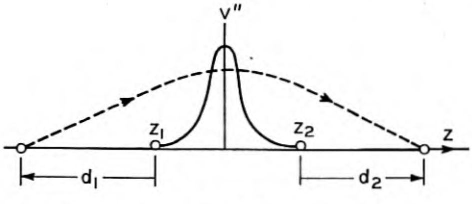


Fig. 5—Conjugate focal distances, thin lens.

$$\frac{1}{T_2} - \frac{1}{T_1} = \frac{1}{F}, \tag{32}$$

---

[12] The period of time that a train requires to reach its terminal point also varies with time as the train moves along.

where the inverse focal term is

$$\frac{1}{F}=\int_{z_1}^{z_2}\frac{v''}{2\sqrt{2v}}dz \tag{33}$$

or on integration by parts

$$\frac{1}{F}=\frac{1}{2}\left[\left(\frac{v'}{\sqrt{2v}}\right)_2-\left(\frac{v'}{\sqrt{2v}}\right)_1\right]+\frac{1}{2}\int_{z}^{z_2}(v')^2(2v)^{-3/2}dz. \tag{34}$$

The substitution in equation 32 of the values for T_1 and T_2 as given by equation 24 now gives the lens equation

$$\frac{\sqrt{2v_2}}{d_2}-\frac{\sqrt{2v_1}}{d_1}=\frac{1}{F}. \tag{35}$$

This equation is analogous to the equation for a thin optical lens

$$\frac{\mu_2}{d_2}-\frac{\mu_1}{d_1}=\frac{1}{F}, \tag{36}$$

bounded on its two sides by media with different refractive indices μ_1 and μ_2, the $\sqrt{2v}$ corresponding to refractive index.

Electron rays parallel to the axis do not come to a focus at a distance F from an electron lens; in other words, F is not a principal focal distance. There are, in general, two principal focal points on opposite sides of an electron lens. Their principal focal distances f_1 and f_2 are found by setting first d_1, and then d_2, equal to infinity in equation 35. This gives

$$f_1=-\sqrt{2v_1}F, \qquad f_2=\sqrt{2v_2}F \tag{37}$$

as the two principal focal distances. It may be shown from equation 33 that these principal focal distances really involve the voltages only in the form of the ratio v_2/v_1. By substituting them in the lens equation 35, it may be written in the convenient form

$$\frac{f_2}{d_2}+\frac{f_1}{d_1}=1, \tag{38}$$

which likewise involves the voltages only in the form of a ratio.

There are two types of electron lenses that deserve special consideration. The first is a small aperture in a thin plate separating two uniform fields of different intensities—as a special case one of the fields may be zero. An example of such a lens is illustrated in Fig. 6.

In this type of lens, the non-uniform field at the aperture covers a distance along the axis about equal to its diameter.[13] If the diameter is small compared to v/v', there is little change in potential throughout the lens and the $\sqrt{2v}$ may be considered as a constant in the integration 33 for the inverse focal term. With this approximation,

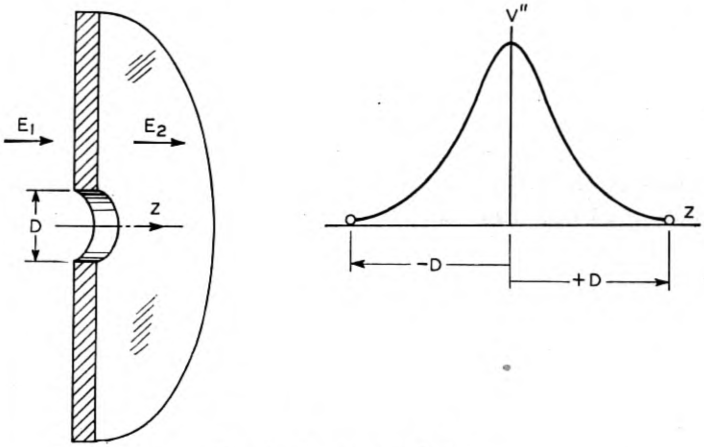


Fig. 6—An apertured plate.

$$\frac{1}{F} = \frac{1}{2}\left[\frac{v_2' - v_1'}{\sqrt{2v}}\right], \tag{39}$$

when v is the potential of the plate and v_1' and v_2' are the electric intensities of the two uniform fields. The lens equation 35 is then

$$\frac{\sqrt{2v}}{d_2} - \frac{\sqrt{2v}}{d_1} = \frac{1}{2}\left[\frac{v_2' - v_1'}{\sqrt{2v}}\right], \tag{40}$$

which may be written in the simpler form

$$\frac{1}{d_2} - \frac{1}{d_1} = \frac{v_2' - v_1'}{4v}. \tag{41}$$

In this type of lens, the electrical refractive index $\sqrt{2v}$ is the same on both sides of the lens and the two principal focal distances are equal, just as they are for a thin optical lens when it is bounded by air on both sides.[14]

---

[13] "Two Problems in Potential Theory," T. C. Fry, *Bell Telephone System Monograph B-671.*

[14] A complete electron-optical system usually involves a combination of lenses. The calculations for a combination are illustrated by the example in Appendix 1.

The apertured plate between two uniform fields is the only lens that permits such a simple calculation of focal distances. In all other lens structures the potential varies appreciably throughout the lens and the integration for the focal term is complicated. The actual numerical calculations have been carried out for only a few of these cases.

The second type of lens deserving special consideration is a lens bounded on both sides by field-free space. For such boundaries the first term in the last member of equation 34 vanishes, and $1/F$ is determined by the integral term alone. This integral in inherently positive, and a lens bounded on both sides by field-free space is thus always a convergent lens. The two concentric tubes of Fig. 7 give a lens of this type, the electric field in each tube dropping to zero at a short distance from its end. It is true that there is a divergent field of the same intensity as the convergent field; but an electron is at a higher potential in the divergent field and traveling faster, so it receives a smaller radial deflection in that field and the lens is convergent. It is interesting to note that the lens is still positive even when the potentials on the electrodes are reversed; in other words, a lens of this type is positive irrespective of the direction of the electric field.[15]

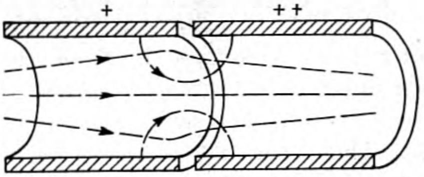


Fig. 7—Concentric tubes—lines of force and electron paths.

An Approximation for Certain Thick Lenses

In certain electron lenses there is a short region of strong lens action accompanied by more extended regions of weaker action; the large values of the derivative v'' are confined to a short distance along the axis, but the derivative does have appreciable values over a more extended region. A lens of this type can be treated in the following approximate manner, provided that there is but one maximum of $|v''|$ in the lens.

For this purpose, the differential equation 31 is rewritten in the form

$$d\left(\frac{1}{T+t}\right) = \frac{v''}{2}\left(1+\frac{t}{T}\right)^{-2} dt, \qquad dt = dz/\sqrt{2v} \qquad (42)$$

---

[15] The principal focal distances of concentric tubes are calculated in Appendix 2.

and the lens equation is derived by integrating it from a point z_1 to a point z_2, where the two points are taken at the substantial boundaries of the non-uniform field. In carrying out this integration, the origin of time is taken at the instant that the electron is at the maximum of $|v''|$, as illustrated in Fig. 8, and for convenience the origin of z is also taken at that point. With this choice of the origin, the term t/T in the second member of the equation is small compared to unity in the region where v'' is large and not very important in the regions where v'' is small. This term may therefore be neglected in lenses when the time of transit is not too great a fraction of the focal times involved. The integration of the equation then gives

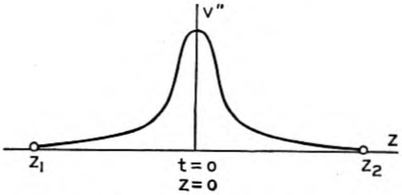


Fig. 8—Coordinates for a thick lens.

$$\frac{1}{T_2 + t_2} - \frac{1}{T_1 + t_1} = \frac{1}{F} \tag{43}$$

when the inverse focal term is again

$$\frac{1}{F} = \int_{z_1}^{z_2} \frac{v''}{2\sqrt{2v}}\,dz \tag{44}$$

and

$$t_2 = \int_0^{z_2} \frac{dz}{\sqrt{2v}}, \qquad t_1 = \int_0^{z_1} \frac{dz}{\sqrt{2v}}. \tag{45}$$

A transformation to space variables by means of equations 24 and 25 gives the lens equation in a form analogous to that for a thick optical lens,

$$\frac{\sqrt{2v_2}}{d_2 - \alpha_2} - \frac{\sqrt{2v_1}}{d_1 - \alpha_1} = \frac{1}{F}, \tag{46}$$

where

$$\begin{aligned} \alpha_2 &= -\sqrt{2v_2}t_2 = -\int_0^{z_2} \sqrt{\frac{v_2}{v}}\,dz, \\ \alpha_1 &= -\sqrt{2v_1}t_1 = -\int_0^{z_1} \sqrt{\frac{v_1}{v}}\,dz. \end{aligned} \tag{47}$$

A plane located at a distance α_1 from the point z_1 is the approximate first principal plane of the lens; and a plane located at a distance α_2 from the point z_2 is the approximate second principal plane of the lens. In the lens equation, $d_1 - \alpha_1$ and $d_2 - \alpha_2$ are the conjugate focal distances measured from the principal planes. If the focal distances measured in this manner are designated as D_1 and D_2 respectively, the lens equation assumes the simpler form

$$\frac{\sqrt{2v_2}}{D_2} - \frac{\sqrt{2v_1}}{D_1} = \frac{1}{F}. \tag{48}$$

An electron lens frequently has both a positive and a negative maximum of v'', and the preceding approximation cannot be applied to the lens as a whole. There is, however, necessarily a point between the two maxima where v'' is zero and by taking this as a division point, the lens can be separated into two components. The approximation can then be separately applied to each component, and the whole lens treated as a combination of two lenses.

The General Theory of Thick Lenses

The equation for the coefficient A,

$$\frac{dA}{dz} + \frac{A^2}{\sqrt{2v}} = -\frac{v''}{2\sqrt{2v}},$$

is a Racciti equation and, with the exception of special cases, it has no exact solution in the usual sense. Particular solutions can be obtained only by integration in series. It is, however, possible to express the general solution of a Racciti equation in terms of any two particular solutions, and this property enables us to develop the general theory of a thick lens in terms of its principal focal distances.

Fig. 9—Paths corresponding to X and Y.

In considering a thick lens, two points z_1 and z_2 are again taken at the substantial boundaries of the non-uniform field constituting the lens. The differential equation for A necessarily has a particular solution equal to zero at z_1. This solution is designated as Y, and it corresponds to an electron ray entering the lens parallel to the axis. At z_2 this solution is equal to $-\sqrt{2v_2}/f_2$, where f_2 is the second principal focal distance measured from z_2. The path of such a ray is illustrated in Fig. 9. This particular solution obeys the same differential equation as A. By subtracting the differential equation of A from that of Y and making a slight transformation, we obtain

$$\frac{d}{dz}\log(A - Y) = -\frac{A + Y}{\sqrt{2v}}, \tag{49}$$

and it should be noted that

$$A/\sqrt{2v} = \frac{\dot{r}}{r\sqrt{2v}} = \frac{d}{dz}\log r. \tag{50}$$

An integration from z_1 to z_2 and a transformation to focal distances then gives the relation

$$\frac{(f_2 - d_2)d_1}{f_2 d_2} = k_2\sqrt{\frac{v_1}{v_2}\frac{r_1}{r_2}}, \tag{51}$$

where k_2 is a constant of the lens, given by

$$1/k_2 = \exp.\int_{z_1}^{z_2}\frac{Ydz}{\sqrt{2v}}. \tag{52}$$

By proceeding in the same manner with a particular solution X for a ray leaving the lens parallel to the axis, we obtain a second relation

$$\frac{(f_1 - d_1)d_2}{f_1 d_1} = k_1\sqrt{\frac{v_2}{v_1}\frac{r_2}{r_1}}, \tag{53}$$

where

$$k_1 = \exp.\int_{z_1}^{z_2}\frac{Xdz}{\sqrt{2v}}. \tag{54}$$

The differential equations of X and Y may also be subtracted and integrated, and this gives a third relation

$$f_1/f_2 = -\frac{k_2}{k_1}\sqrt{\frac{v_1}{v_2}}. \tag{55}$$

A multiplication of the first two relations 51 and 53 gives

$$(f_2 - d_2)(f_1 - d_1) = k_1 k_2 f_1 f_2, \tag{56}$$

which is one form of the equation relating the conjugate focal distances of a lens. This equation may be converted into a more useful form by the following considerations.

A combination of the three preceding relations gives

$$\frac{r_2}{d_2}(d_2 - \alpha_2) = \frac{r_1}{d_1}(d_1 - \alpha_1), \tag{57}$$

where

$$\alpha_1 = f_1(1 - k_1), \qquad \alpha_2 = f_2(1 - k_2). \tag{58}$$

To interpret this equation, we erect two imaginary planes as shown in Fig. 10. The first plane is located at a distance α_1 from z_1. If the path of the incident ray is projected it intersects this plane at some radial distance R_1. The second plane is erected at a distance α_2 from z_2. The path of the exit ray intersects it at a radial distance R_2. The equation says—from simple geometry—that the two radial distances R_1 and R_2 are equal. The path of an electron through the lens is therefore the same as if the electron proceeded in a straight line to the first plane, passed parallel to the axis to the second plane, and then proceeded again in a straight line to the second conjugate focal point. These two planes are called the first and second principal planes of the lens. The action of a thick lens is the same as if the space between the principal planes were non-existent, leaving them in coincidence, and a thin lens were located at the plane of coincidence.

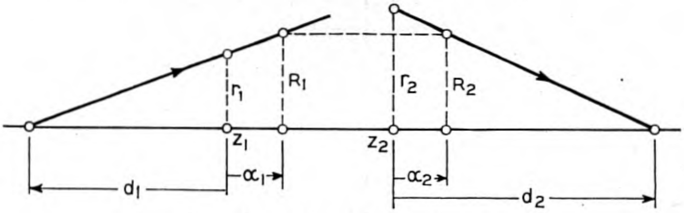


Fig. 10—The principal planes.

The principal planes of a lens may lie either inside or outside of the lens. In most convergent lenses, α_1 is positive and α_2 negative, and the two planes both lie inside the lens.

The first conjugate focal distance measured from the first principal plane is designated as D_1, and the second conjugate focal distance measured from the second principal plane is designated as D_2. When they are measured in this manner, the two conjugate distances are

$$D_1 = d_1 - \alpha_1, \qquad D_2 = d_2 - \alpha_2. \tag{59}$$

The two principal focal distances measured from the principal planes are, in a similar manner, designated as F_1 and F_2; then, from equation 58,

$$\begin{aligned} F_1 &= f_1 - \alpha_1 = k_1 f_1, \\ F_2 &= f_2 - \alpha_2 = k_2 f_2, \end{aligned} \tag{60}$$

and, from equation 55,

$$F_1/F_2 = -\sqrt{\frac{2v_1}{2v_2}}. \tag{61}$$

Substitution of the new focal distances in the lens equation 56 now gives

$$(F_2 - D_2)(F_1 - D_1) = F_1 F_2 \tag{62}$$

or

$$\frac{F_2}{D_2} + \frac{F_1}{D_1} = 1. \tag{63}$$

This is the general equation relating the conjugate focal distances in any lens. With the aid of equation 61, it may be written in the more familiar form

$$\frac{\sqrt{2v_2}}{D_2} - \frac{\sqrt{2v_1}}{D_1} = \frac{1}{F}, \tag{64}$$

where

$$\frac{1}{F} = \frac{\sqrt{2v_2}}{F_2} = -\frac{\sqrt{2v_1}}{F_1}. \tag{65}$$

The Principal Points of a Lens

The points locating the two principal planes on the axis of a lens and its two principal focal points are called the cardinal points of the lens. The preceding theory of a thick lens shows that its performance is completely determined by the locations of these four points.[16] The theory does not furnish a general method for calculating their locations, but it does show that they can be determined from a knowledge of two so-called principal rays. The first is a ray leaving the lens parallel to the axis. If its entrance and exit paths are projected, they intersect as shown in Fig. 11, and the intersection locates the first principal plane. The projected incident ray also intersects the axis, and this intersection locates the first principal focal point. The second principal plane and the second principal focal point may be located in a similar manner from the entrance and exit paths of a ray entering the lens parallel to the axis.

[16] We are here speaking only of rays near the axis.

The required paths of the two rays must in general be determined either by a series or step-by-step integration of the differential equation for A, or else by actual measurements on the physical lens.[17]

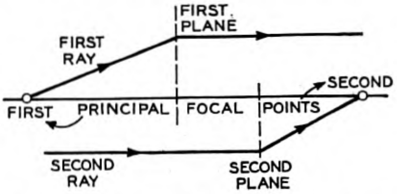


Fig. 11—The cardinal points of a lens.

Magnification

Electron object and electron image are defined as their optical analogies. The electron object may be an actual source of electrons, or the real image of such a source, or it may be a virtual image from which the electrons are apparently coming as they enter a lens.

The magnification by an electron lens may be treated in the following manner. Let S_1 be the size of an electron object located at a distance D_1 from the first principal plane of a lens. Two electron rays from the edge of the object are considered—as shown in Fig. 12. The ray entering the lens parallel to the axis may be regarded as passing on to the second principal plane and then bending sharply to pass through the second principal focal point; the ray through the first principal focal point may be considered as passing on to the first principal plane and then proceeding parallel to the axis. The intersection of the two rays locates the electron image and determines its size S_2. The magnification M is defined as S_2/S_1, and it follows from simple geometry that

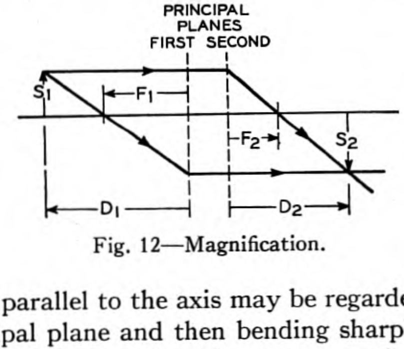


Fig. 12—Magnification.

$$M = \frac{D_2 - F_2}{F_2} = \frac{F_1}{D_1 - F_1}. \tag{65}$$

---

[17] Other step-by-step methods can be used when a map of the equipotential surface is available.

A more convenient expression for magnification is now obtained by combining the two preceding expressions to give

$$M = \frac{F_1 D_2}{F_2 D_1} \tag{66}$$

and from equation 61

$$M = -\sqrt{\frac{2v_1}{2v_2}}\frac{D_2}{D_1}. \tag{67}$$

The magnification is not in general equal to the ratio of the image distance to the object distance, as it is for an optical lens in air. It is only equal to that ratio when the voltage is the same on both sides of the lens.

SECTION III—ABERRATION IN A LENS

Returning to the first section, we see that the general expression for focal distance is

$$\frac{1}{d} = -\frac{A + Br^2 + Cr^4 \cdots}{\sqrt{2v} + \frac{A'}{2}r^2 + \frac{B'}{4}r^4 \cdots}. \tag{20}$$

The exact focal distance of an electron thus depends on its radial coordinate r, and a ray passing through a lens at a distance from the axis does not come to the same focus as a ray near the axis. A precise, general theory for rays at a distance from the axis could—in theory at least—be derived by solving the differential equations for as many of the higher coefficients as desired and substituting them in the above equation. Such a general solution would, however, be very difficult indeed, and one is content—as he usually is in optics—to treat the performance of a lens in a much more restricted manner.

The equation for focal distance can be simplified to some extent by noting that its denominator is the velocity component $\dot{z}$. With the aid of the energy equation 6, this component can be written in the form

$$\dot{z} = \sqrt{2v}\left[1 + \left(\frac{r}{d}\right)^2\right]^{-1/2}. \tag{68}$$

In most lenses r is small compared to d, and the last factor in the above equation may be approximately set equal to unity. This approximation is accurate to one per cent even for a lens with an angular aperture corresponding to F3.5—an F2 lens is a very fast camera lens. With this approximation the inverse focal distance is

$$\frac{1}{d} = -\frac{A + Br^2 + Cr^4 + \cdots}{\sqrt{2v}}. \tag{69}$$

The presence of the terms in r causes a diffusion of the focus in a lens, and a clearer picture of this diffusion is obtained by expressing it as lateral aberration. So we now proceed to derive an expression for this aberration, and the meaning of the term becomes apparent from the derivation. For this purpose we consider electrons entering a lens as if they all came from a point source at a distance d_1 from the first side of the lens. We are at liberty to set the higher coefficients equal to zero at that side of the lens, and this gives

$$A_1 = -\frac{\sqrt{2v_1}}{d_1}, \tag{70}$$

$$B_1 = C_1 = \cdots = 0.$$

At the exit side of the lens, the focal distance is

$$\frac{\sqrt{2v_2}}{d_2} = -(A_2 + B_2r^2 + C_2r^4 + \cdots), \tag{71}$$

where the coefficients are solutions of their differential equations subject to the initial conditions 70. The focal distance d_0 for rays near the axis is given by

$$\frac{\sqrt{2v_2}}{d_0} = -A_2. \tag{72}$$

The difference between the focal distance d_2 of a ray leaving the lens at a distance r from the axis and the focal distance d_0 of a ray near the axis is

$$d_2 - d_0 = \frac{d_0^2}{\sqrt{2v_2}}(B_2r^2 + C_2r^4 + \cdots). \tag{73}$$

This difference is called the longitudinal aberration of the lens. It is the distance that the focal point is diffused along the axis, when the lens is limited by an exit diaphragm or radius r.

If a screen is placed at a distance d_0 from the lens, rays near the axis will come to a point focus on the screen; but rays leaving the lens at a distance r from the axis will strike the screen along a circular line. The radius s of this circle is called the lateral aberration of the lens. It follows rather simply, from the value of the longitudinal aberration, that the lateral aberration is

$$s = \frac{d_0}{\sqrt{2v_2}}(B_2r^3 + C_2r^5 \cdots). \tag{74}$$

This is the radius of the diffuse image of a point source, when the lens is limited by a diaphragm of radius r.

The differential equations 17, 18 ··· for the aberration coefficients are linear and subject to solution in the usual manner when A and v are known functions of z. The solutions for the higher coefficients would of course be quite involved. The higher terms are, however, small compared to the second term, which causes most of the aberration, and the approximate distortion is given by the second term alone. This term is called the second order aberration term.

The Reduction of Aberration

The coefficient of any aberration term vanishes when conditions are arranged so that the last member of its differential equation is zero, for the coefficient may be arbitrarily set equal to zero at the first side of the lens, and the solution of its linear equation is then zero throughout the lens.

The important second order aberration term can thus be made to vanish by arranging conditions so that

$$\frac{V''''}{16} - \frac{(A')^2}{2} = 0. \tag{75}$$

In a lens that is not too thick compared to the focal distances involved, we have seen that the term A^2 may be neglected in the differential equation for A, and

$$A' = -\frac{v''}{2\sqrt{2v}}. \tag{76}$$

The substitution of this value for A' in the above equation gives

$$v'''' - \frac{(v'')^2}{v} = 0 \tag{77}$$

as the differential equation for electric fields that are approximately free from second order aberration, when the focal distances are reasonably large compared to the length of the field along the axis.

The general solution of this equation is a series solution, but several particular solutions have been obtained in terms of known functions. The potentials corresponding to these particular solutions are given by the following equations:

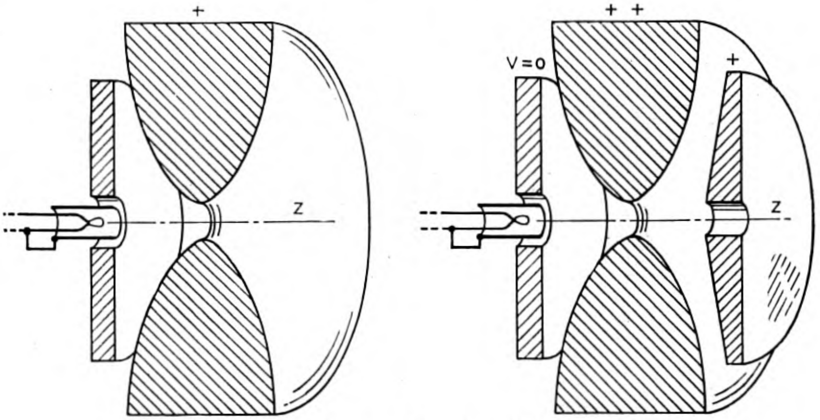
$$\Phi = ae^{\pm\omega z}J_0(\omega r), \tag{78}$$

$$\Phi = (a \sin \omega z + b \cos \omega z)J_0(i\omega r), \tag{79}$$

$$\Phi = (a \sinh \omega z + b \cosh \omega z)J_0(\omega r), \tag{80}$$

$$\Phi = 3az^{3/2}\left[\frac{1}{3} - \frac{1}{4}\left(\frac{r}{2z}\right)^2 + \frac{3}{64}\left(\frac{r}{2z}\right)^4 \cdots\right]. \tag{81}$$

Any one of these electric fields can be produced by shaping and positioning electrodes to correspond with two of its equipotential surfaces. These fields are, however, in general not well adapted to production with practical electrode structures. The one exception is the field defined by equation 79, and electrodes for producing it in a practical form are shown in Fig. 13. They are suitable for giving an electron stream its initial acceleration. The electric field constitutes a divergent lens, as do practically all initial accelerating fields.



Figs. 13, 14—Lenses with reduced aberration.

As expressed by equation 79, this field is followed by a symmetrically reversed field, and for some purposes it may be desirable to include the reversed field. This is done by locating a low potential electrode along its corresponding equipotential surface as shown in Fig. 14. A small aperture may be cut in this electrode for the passage of electrons. The aperture then acts as a lens to bring the beam to a focus, but this lens has its own aberration, and the whole system is then only partially free from first order aberration.

Appendix I—Calculations for a Complete System

The electrode arrangement of Fig. 15 is chosen for giving a simple example of the calculations for a complete optical system. The final focal distance is found by calculating the focal distances at the points m, n, o, p in succession. Electrons leave the cathode and travel parallel to the axis in the uniform field between the first and second plates, so their focal distance is $-\infty$ when they arrive at the point m. The electrons then pass through the aperture in the second plate, and their focal distance at n is calculated from the lens equation 41, which gives

$$\frac{1}{d_n} + \frac{1}{\infty} = \frac{1}{4v_1}\left[\frac{v_1 - v_2}{l} - \frac{v_1}{l}\right] \tag{1}$$

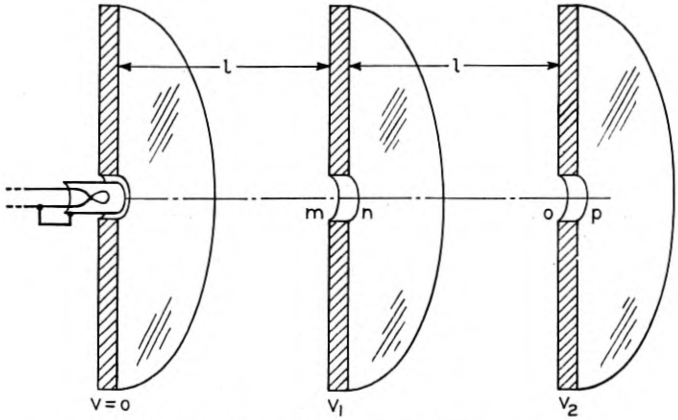


Fig. 15—Example of a complete system.

and the focal distance at n is

$$d_n = \frac{4l}{\beta^2 - 2}, \tag{2}$$

where β is $\sqrt{v_2/v_1}$. The beam then passes through the uniform field between the second and third plates, and the focal distance at 0 is calculated from equation 28 for a uniform field, which becomes

$$(1 + \beta)d_n - (1 + 1/\beta)d_0 = 2l \tag{3}$$

and gives

$$d_0 = 2\beta l \frac{4 + 2\beta - \beta^2}{(1 + \beta)(\beta^2 - 2)}. \tag{4}$$

The beam then passes through the aperture in the third plate into field-free space, and the lens equation for this aperture is

$$\frac{1}{d_p} - \frac{1}{d_0} = \frac{1}{4v_2}\left[0 - \frac{v_2 - v_1}{l}\right]. \tag{5}$$

Substitution for d_0 now gives

$$\frac{1}{d_p} = \frac{1 + \beta}{2\beta l}\left[\frac{\beta^2 - 2}{4 + 2\beta - \beta^2} + \frac{1 - \beta}{2\beta}\right], \tag{6}$$

which is the reciprocal of the final focal distance measured from the last plate.

In complete lens systems, where the symbolic calculations are complicated, it is frequently simpler to introduce specific numerical values and carry the successive steps of the calculation through in a numerical manner. By doing this for a few suitably chosen numerical values one can obtain the particular information that is desired.

APPENDIX II—CONCENTRIC TUBES

Two concentric tubes at different potentials form an electron lens that is well adapted to practical tube construction. When the two tubes are of the same diameter, the approximate constants of the lens may be determined as follows.[18]

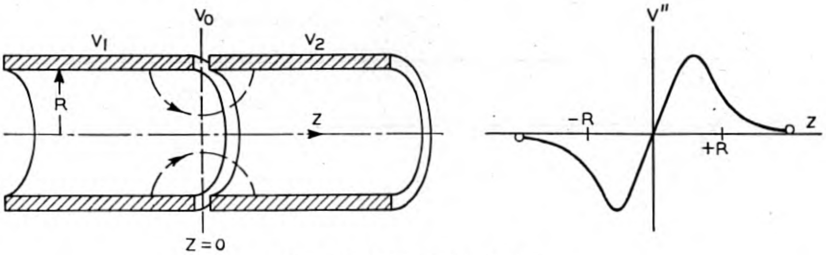


Fig. 16—Concentric tubes.

In this type of lens, the electric intensity is symmetrical with respect to an imaginary plane drawn between the two tubes—as illustrated in Fig. 16—and the plane is therefore an equipotential surface. Its potential v_0 is the mean potential of the two tubes. This plane is regarded as a division plane separating the lens into two component electric fields.

We first consider the component to the right of the plane. The solution for the potential inside of the tube may be obtained in the form of a Bessel Function series, and it follows from this series that the potential on the axis is

$$v = v_2 - (v_2 - v_0)\sum_{\mu} \frac{2}{\mu J_1(\mu)} \exp.\left(-\frac{\mu z}{R}\right), \tag{1}$$

where R is the radius of the tubes, and μ takes on discrete values equal to the successive roots of

$$J_0(\mu) = 0. \tag{2}$$

We find that an approximation to the exponential series is given by

$$\sum_{\mu} \frac{2}{\mu J_1(\mu)} \exp.\left(-\frac{\mu z}{R}\right) = 1 - \tanh \omega z, \tag{3}$$

---

[18] We assume that the separation between their ends is negligibly small compared to their diameter.

where ω is equal to $1{,}32/R$. The closeness of this approximation is shown in Fig. 17. Its introduction gives

$$v = v_0 + (v_2 - v_0) \tanh \omega z. \tag{4}$$

A similar approximation is found for the potential on the axis to the left of the division plane,

$$v = v_0 - (v_1 - v_0) \tanh \omega z, \tag{5}$$

and it turns out that the potential on the axis of both tubes can be expressed by the single equation

$$v = \tfrac{1}{2}[(v_2 + v_1) + (v_2 - v_1) \tanh \omega z]. \tag{6}$$

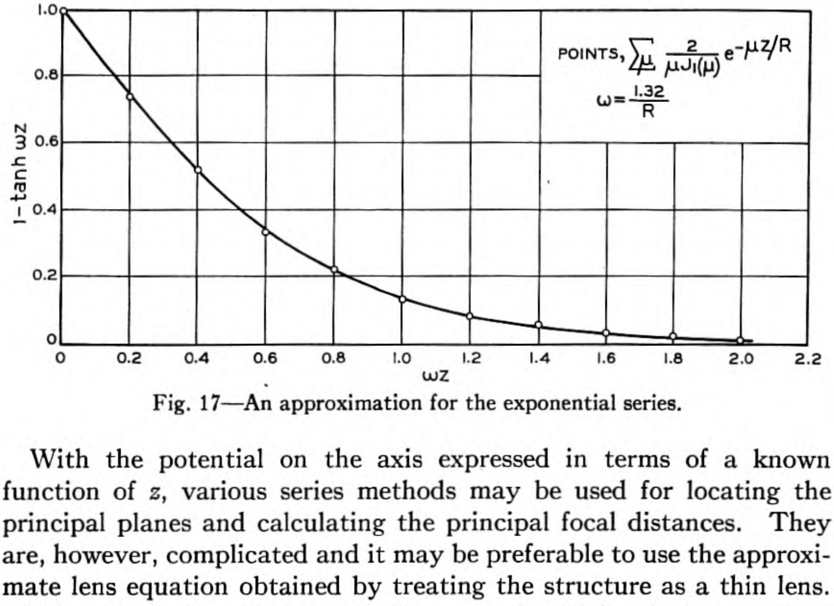


Fig. 17—An approximation for the exponential series.

With the potential on the axis expressed in terms of a known function of z, various series methods may be used for locating the principal planes and calculating the principal focal distances. They are, however, complicated and it may be preferable to use the approximate lens equation obtained by treating the structure as a thin lens.

When treated in this manner, the expression 33 for the inverse focal term can be exactly integrated, and the lens equation is

$$\frac{\sqrt{2v_2}}{d_2} - \frac{\sqrt{2v_1}}{d_1} = \frac{\omega\sqrt{2}}{3(\sqrt{v_2} + \sqrt{v_1})}(\sqrt{v_2} - \sqrt{v_1})^2, \qquad \omega = 1{,}32/R. \tag{7}$$

Division by either $\sqrt{2v_2}$ or $\sqrt{2v_1}$—as desired—reduces this equation to one that involves the voltages only in the form of a ratio. The error in a focal distance d calculated from this equation is of the order of R, when the focal distance is measured from the division plane of the tubes.

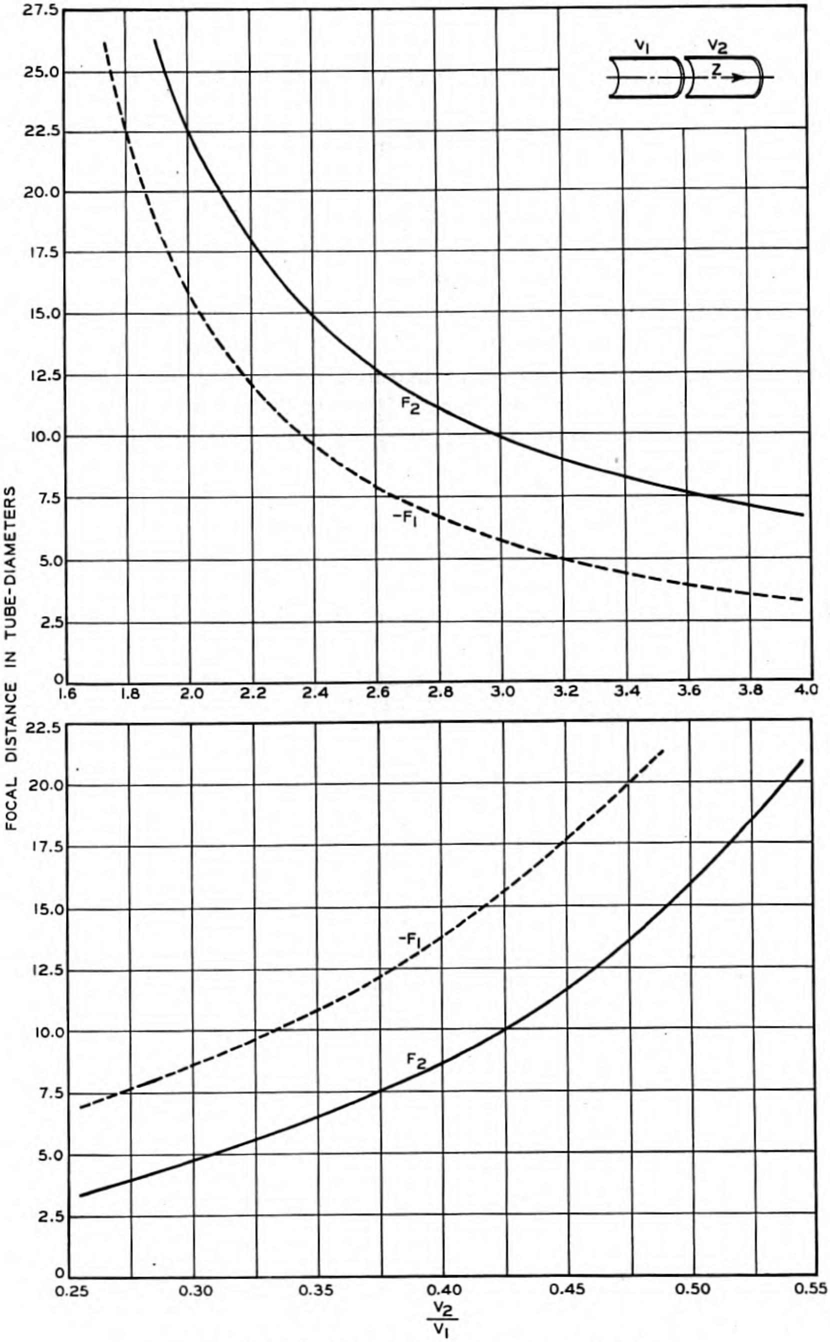


Fig. 18—Principal focal distances—concentric tubes.

The principal focal distances for various voltage ratios are given in terms of the tube diameter by the curves of Fig. 18. For rough calculations, these plotted values may be used in the lens equation

$$\frac{f_2}{d_2} + \frac{f_1}{d_1} = 1, \tag{8}$$

where focal distances are again measured from the division plane of the tubes.

The electric field of the concentric tubes has two maxima of $|v''|$ located symmetrically with respect to the division plane, as illustrated in Fig. 16. Each maximum is located at a distance $.5R$ from the plane. The electron lens may therefore be treated in a somewhat more exact manner by considering it as two thin lenses located at these points. The inverse focal term of the equivalent lens to the left of the plane is

$$\frac{1}{F} = \frac{\omega\sqrt{2}}{v_0 - v_1}\left[v_0(\sqrt{v_0} - \sqrt{v_1}) - \frac{v_0^{3/2} - v_1^{3/2}}{3} \right], \tag{9}$$

and the inverse focal term of the equivalent lens to the right of the plane is

$$\frac{1}{F} = \frac{\omega\sqrt{2}}{v_2 - v_0}\left[v_0(\sqrt{v_2} - \sqrt{v_0}) - \frac{v_2^{3/2} - v_0^{3/2}}{3} \right]. \tag{10}$$

The final focal distance in any particular case is found by carrying out the calculations for the two lenses in succession, with their separation taken equal to R.

APPENDIX III

A Plane Electrode at the End of a Tube.—In addition to their above application, the last two equations may be used for other purposes. In electron devices, one frequently puts a plane electrode at the end of another, tubular electrode.[19] The approximate lens action of the electric field between the plate and tube is then described by one or the other of these equations. Equation 9 applies when the plane follows the tube in the direction of electron motion; and equation 10 applies when the plane precedes the tube.

In structures of this type, the plate is usually pierced with an aperture for the passage of electrons. When the aperture is small compared to the tube diameter, the lens system can be treated in the following manner.

---

[19] We assume the separation between the plate and the end of a tube to be negligible compared to the tube diameter.

We first consider the case of the plane preceding the tube. The electric intensity at the plate is found by differentiating equation 4 with respect to z and then setting z equal to zero. A substitution of this intensity in equation 41 of the text gives the lens equation of the aperture. In addition to this lens there is an equivalent thin lens located inside the tube at a distance $.5R$ from the plate, and having the inverse focal distance of equation 10. The system is considered as a combination of the two lenses and the calculations are carried through in the usual manner. When the plane follows the tube, the constants of the two lenses are determined from equations 5 and 9, and the combination is treated in a similar manner.

APPENDIX IV—THE VELOCITY FUNCTION

The auxiliary functions u and w are a special case of the components of a generalized vector function that is useful in developing series solutions for electron motion. The equations of this function are equivalent to the Hamilton-Jacobi equation; they are briefly outlined in the present system of units as follows.

In a field that may comprise both an electric intensity E and a magnetic intensity H, let v be any vector function of x, y, z that satisfies the equations

$$\mathrm{curl}\ v = H/c, \tag{1}$$

$$1/2|v|^2 = \phi + W, \tag{2}$$

where W is a constant equal to the energy of electron emission from the source. Then v is a possible vector velocity for electron motion in the field.

If the magnetic intensity is zero, the vector function v has a potential ψ, which may be any solution of the equation

$$1/2|\mathrm{grad}\ \psi|^2 = \phi + W \tag{3}$$

and grad ψ is then a possible vector velocity for electron motion in the field.

The validity of these equations is established by transforming them to the usual equations for electron acceleration.

A LIST OF THE MORE IMPORTANT SYMBOLS AND EQUATIONS

In the present theory of electron-optics, *all distances* along the axis are measured in the direction of motion, as they are in the optics of light.

r, z —cylindrical coordinates
t —time

Φ —potential at point in space, the electron source taken as zero potential
v —potential on the axis
v' —derivative of v with respect to z
v —is also used for the voltage of electrodes
d —focal distance in general
T —focal time in general
A —the important coefficient for rays near the axis, a function of z alone
u, w —velocity functions corresponding to $\dot{r}$ and $\dot{z}$
d_1, d_2—conjugate focal distances measured from the two sides of a lens
f_1, f_2 —principal focal distances measured from the two sides of a lens

As an approximation in thin lenses, the focal distances are measured either from the mid-point of the lens, or from the point where $|v''|$ is a maximum, provided that there is but one maximum in the lens.

α_1, α_2 —location of the principal planes with respect to the sides of a lens
D_1, D_2—conjugate focal distances measured from the principal planes
F_1, F_2—principal focal distances measured from the principal planes
F —the focal term of a lens, not a focal distance

Equations for Rays Near the Axis

$$\dot{z} = \sqrt{2v},$$

$$A = -\frac{\sqrt{2v}}{d} = -\frac{1}{T} \doteq \dot{r}/r,$$

$$\sqrt{2v}A' + A^2 = -\frac{v''}{2},$$

$$\frac{1}{T^2}\frac{d}{dt}(T + t) = -\frac{v''}{2}.$$

The important equations for a thin lens are:

$$\frac{\sqrt{2v_1}}{d_2} - \frac{\sqrt{2v_1}}{d_1} = \frac{1}{F},$$

$$\frac{f_2}{f_1} = -\sqrt{\frac{2v_2}{2v_1}},$$

$$\frac{f_2}{d_2} + \frac{f_1}{d_1} = 1,$$

$$\frac{1}{F} = \int_{z_1}^{z_2} \frac{v''}{2\sqrt{2v}}\,dz.$$

The following equations hold for any lens:

$$\frac{\sqrt{2v_2}}{d_2 - \alpha_2} - \frac{\sqrt{2v_1}}{d_1 - \alpha_1} = \frac{1}{F},$$

$$\frac{\sqrt{2v_2}}{D_2} - \frac{\sqrt{2v_1}}{D_1} = \frac{1}{F},$$

$$\frac{F_2}{F_1} = -\sqrt{\frac{2v_2}{2v_1}},$$

$$\frac{F_2}{D_2} + \frac{F_1}{D_1} = 1,$$

$$M = \frac{F_1 D_2}{F_2 D_1} = -\sqrt{\frac{2v_1}{2v_2}}\frac{D_2}{D_1},$$

where M is magnification.

Equivalent Modulator Circuits

By E. PETERSON and L. W. HUSSEY

Equivalent modulator circuits are developed in the form of linear resistance networks. They are equivalent in the sense that the current magnitude in any mesh of the network is equal to the current amplitude of a corresponding frequency component in the modulator. The elements of the network are determined by the properties of the modulator, while the terminating resistances are those physically existent in the connected circuit.

With this correspondence demonstrated, the operating features of the modulator may be deduced from the known properties of linear networks. Among the properties considered are the transfer efficiency from signal to sideband, and the input resistance to signal as affected by the sideband load resistance.

Equivalent networks are worked out for a number of interesting cases, involving different impedances to unwanted modulation products, together with different non-linear characteristics. The equivalents come out comparatively simple in form under the restrictions noted and followed in the text, which make the carrier large compared to the signal, and the circuit elements purely resistive.

CONSIDERED from the circuit standpoint, a number of modulator performance features are important in any application. Among these features might be mentioned the efficiency of power transfer from signal input to sideband output, and associated with it the question of how the signal input energy is distributed among the different frequency components and dissipated in the modulator itself. Then, too, we need to know how the impedance of the modulator to any component depends upon the modulator structure and upon the connected impedances to other products.

In attempting to get answers to these questions by mathematical analysis, we encounter lengthy and cumbersome expressions in general which do not lend themselves to ready physical interpretation. The physical interpretation of these equations may be facilitated by introducing equivalent circuits of familiar form. One form commonly used in the past replaces the non-linear system by a circuit including a series of generators and linear impedances.

This may be illustrated by reference to a simple non-linear circuit, in which carrier and signal generators are connected through an external resistance to a two-terminal non-linear element such as a diode, or a copper oxide rectifier. The effects of non-linearity show up in the change of modulator resistance with changes in applied potentials and in the appearance of new frequency components. These effects may be reproduced quantitatively if we replace the non-linear element by its equivalent consisting of a linear internal resistance together with a series of internal generators—indicated at the right of the dashed line of Fig. 1-*A*. It is easy to see from this circuit[1] what the amplitude of any current component should be; for the general component

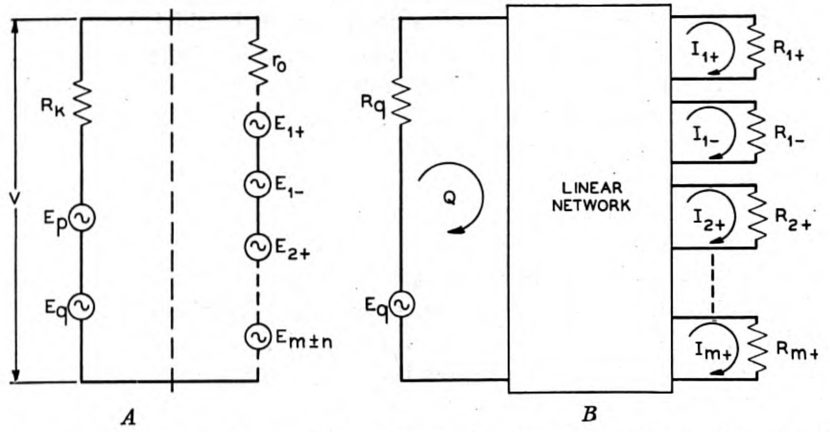


Fig. 1—(*A*) Equivalent modulator circuit in which the modulator is replaced by a fixed internal resistance together with a series of generators. (*B*) Equivalent modulator circuit replacing the modulator by a network of linear elements which serves to couple the signal circuit into the paths followed by the modulation products. In this circuit the mesh currents represent the amplitudes of the various frequency components.

$$I_{m\pm n} = E_{m\pm n}/(r_0 + R_{m\pm n}).$$

Despite the apparent simplicity of this relation, a difficulty arises as soon as we attempt to state the internal generator e.m.f.'s explicitly, since they are found to be tied up with the impressed potentials, the

---

[1] Here we denote the carrier frequency by $p/2\Pi$ and the signal frequency by $q/2\Pi$; the corresponding generator potentials are E_p and E_q respectively, and the external resistance is R_k where k indicates the frequency at which the resistance is effective.

The new frequencies are usually made up of sums and differences of integral multiples of carrier and signal frequencies. In general, they may be represented by $(mp \pm nq)/2\Pi$, where m and n are integers or zero. It is advantageous to adopt an abbreviated notation for the voltage component, say, of any frequency by which the general component is indicated as E_{m+n}. Further when n is unity it is omitted from the subscript, so that the generator e.m.f. of frequency $(mp \pm q)/2\Pi$ is indicated as E_{m+}. One of the restrictions mentioned further on results in limiting n to unity.

modulator characteristics, and the external circuit impedances. For this reason the equivalent circuit of Fig. 1-*A* reveals only part of the story, and in general the relation between the amplitudes of impressed and generated components remains somewhat obscured.

In a number of cases of practical interest it is possible to represent the connection between the amplitudes of various frequency components by means of a different type of equivalent circuit. Figure 1-*B* is an illustration of this type, in which the paths of the current components are shown individually. The connection between the various circuits is effected by means of a linear network which contains no internal generators. In this equivalent network the magnitude of any mesh current is equal to the amplitude of a corresponding frequency component in the modulator circuit. The purpose of this paper is to demonstrate the validity of this representation, and to show in detail what the linear network looks like when applied to various types of modulating elements, in a variety of interesting cases.

In order to develop such equivalent networks in simple and useful form, the following restrictions are imposed. The system includes only one non-linear element.[2] The terminating impedances are purely resistive, although they may be functions of frequency. The signal amplitude must be much smaller than that of the carrier. Finally, the slope of the modulator current-voltage characteristic never becomes negative. Under these conditions a number of modulating systems can be treated, including variable resistance modulators with a variety of current-voltage characteristics, and the variable resistance microphone.

The section following deals with the modulator as a resistance (or conductance) varying at carrier frequency. Succeeding sections consider the behavior of such variable elements under different circuit conditions.

I. CARRIER CONTROLLED RESISTANCE

In setting up equations from which the equivalent networks are obtained, the restriction on signal amplitude permits us to assume the modulator to be a resistance or conductance varying at carrier frequency. This commonly used assumption may be arrived at with the aid of Fig. 2, which shows a typical non-linear current-voltage

---

[2] Other cases are to be found in a paper by R. S. Caruthers on "Copper Oxide Modulators in Carrier Telephone Systems," presented at the A.I.E.E. Winter Convention, January 1939.

Modulators including a plurality of elements can frequently be replaced by an equivalent structure with a single modulating element. This is true of the rectifier type of modulator. In the double-balanced or ring type, however, under certain conditions the equivalent circuit involves merely an ideal transformer connecting signal and sideband circuits.

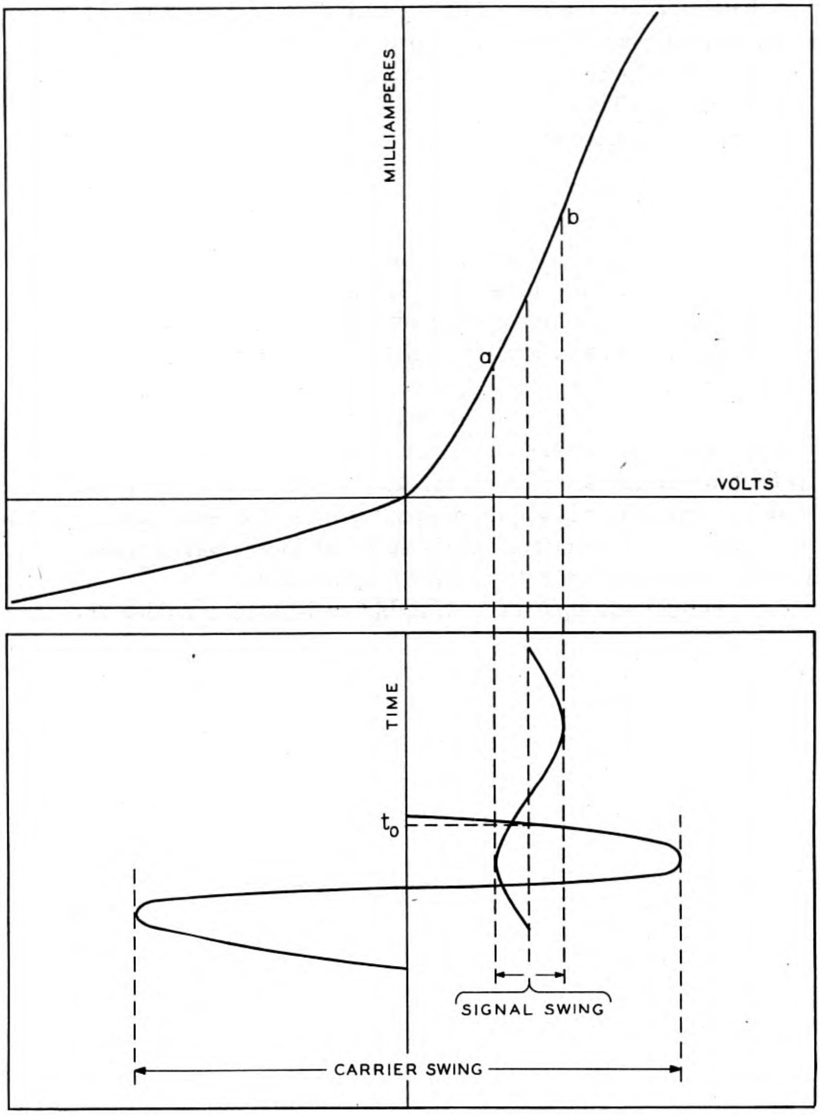


Fig. 2—Non-linear current-voltage relation representative of certain types of modulators. The impressed carrier voltage is large in comparison to the signal, the variation of which is represented in the neighborhood of the potential reached by the carrier wave at time t_0.

curve. The variation with time of a large carrier voltage and a small signal voltage are also indicated. Now if we consider the carrier to provide a varying bias, then at any typical point t_0 we can consider the signal voltage to sweep over a small segment (a, b) of the modulator characteristic. The resistance is practically constant over this segment and of magnitude

$$R = \frac{dv}{di}, \tag{1}$$

the derivative being evaluated at the carrier voltage under consideration. In general this resistance varies from point to point of the carrier cycle. Thus the carrier enters the signal-sideband relation only through the variation of a resistance facing the signal and modulation products.

If the current-voltage characteristic is a smooth curve the resistance varies smoothly over the carrier cycle. If the characteristic is made up of two straight lines the resistance switches between two constant values. This latter is approximated by most rectifiers, such as diode and copper-oxide rectifiers with suitably large carrier amplitudes; it is called by analogy a commutator modulator.

As a simple example of a variable resistance consider the characteristic

$$i = av + bv^2, \tag{2}$$

from which

$$1/R = \frac{di}{dv} = a + 2bv. \tag{3}$$

If the impressed carrier potential, v, is $P \cos pt$, the conductance is

$$G = \frac{1}{R} = a + 2bP \cos pt. \tag{4}$$

More generally, the resistance or conductance for a given characteristic can be expressed as a series

$$R = r_0 + \sum_1^\infty 2r_n \cos npt, \tag{5a}$$

or

$$G = g_0 + \sum_1^\infty 2g_n \cos npt. \tag{5b}$$

Here the coefficients depend only on the modulator characteristic and the carrier amplitude. In special cases some of the coefficients vanish.

Thus an expansion for the linear rectifier includes only those coefficients for which n is odd, whereas one for a modulator exhibiting odd symmetry in its current-voltage relation, such as thyrite, includes only coefficients for which n is even.

The choice between (5a) and (5b) in any given case is usually a matter of convenience.[3] This will be made clear by the forthcoming examples. In every case use will be made of Ohm's law in one of the two forms

$$v = Ri$$

or

$$i = Gv.$$

For simplicity, we select the relation which leads to the smallest number of terms in the expansion. Thus if, from the form of the terminating impedance, we know that i involves only a small number of significant frequency components, whereas the voltage involves a large number, (5a) will be used. If the potential, v, across the modulating element is known to be the simpler, (5b) will be used.

In the practical application of modulators to carrier systems the impedance characteristics of the connected selective circuits for taking out the desired sideband energy provide, to a good approximation, just such simplification. Thus a filter is substantially resistive in its pass band and the suppression regions may be designed to have either a very high or a very low impedance. If very high, no currents flow in these frequency regions and (5a) applies; if very low no potentials appear across it in these frequency regions so that (5b) applies.

II. Single Sideband—High Impedance Outside Band

We will first consider a single sideband modulator involving any variable resistance which can be expressed in the form (5a). The terminating resistance is R_q to signal and R_{1+} to the upper second order sideband. Because of the high terminating impedance which we assume to all other products, all current components other than signal (Q) and sideband (I_{1+}) are negligibly small.

The total current flowing in the circuit is then

$$i = Q \cos qt + I_{1+} \cos (p + q)t. \tag{6}$$

The potential across the non-linear element ($v = Ri$) is obtained from (5a) and (6) as

$$v = \left[r_0 + \sum_1^\infty 2r_n \cos npt \right] [Q \cos qt + I_{1+} \cos (p + q)t]. \tag{7}$$

---

[3] Except for those cases in which the occurrence of an infinity in any one of these two quantities prohibits its use.

After multiplying, and separating out the different frequency components, each frequency component of v is equated to the corresponding terminating generator e.m.f. minus the potential drop across the external impedance. Carrying out this process for signal and sideband, respectively,

$$\begin{aligned} E_q &= (R_q + r_0)Q + r_1 I_{1+}, \\ 0 &= r_1 Q + (R_{1+} + r_0) I_{1+}. \end{aligned} \tag{8}$$

If Q and I_{1+} are considered as mesh current amplitudes in a simple linear circuit, it is evident that r_1 represents a mutual resistance, and that Fig. 3 represents an equivalent network. In this system the signal source is connected to the sideband load by a simple T network.[4] It will be found in subsequent cases similarly, that the connection between signal and sideband circuits may be effected by a network comparatively simple in form. Hence we can make our deductions concerning the performance of modulator circuits by reference to the well known properties of such equivalent networks. In the present case, for example, we can draw the following conclusions.

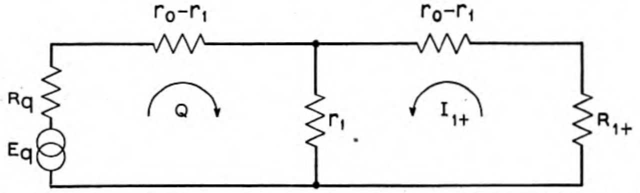


Fig. 3—Equivalent modulator network connecting signal and sideband when other modulation current components are suppressed by high circuit impedances.

1. The modulator loss becomes negligibly small if the series arm resistance, $r_0 - r_1$, is very small and the shunt arm resistance, r_1, is relatively large.
2. Considering the modulator network as fixed, maximum power is transferred when the signal and sideband resistances match the characteristic resistance of the network, so that

$$R_q = R_{1+} = \sqrt{r_0^2 - r_1^2}. \tag{9}$$

3. Under matched impedance conditions the power efficiency is

$$\eta = \left[\frac{r_1}{r_0 + \sqrt{r_0^2 - r_1^2}}\right]^2. \tag{10}$$

---

[4] While the results come out most simply in terms of a T network, the various possible transformations (for example to a π or to a lattice network) are of course equally valid.

The term *power efficiency*—as used here—means the ratio of the power delivered to the load resistance (R_{1+}) to that introduced at the input side of the network by the signal source. The corresponding current ratio of sideband to signal is the square root of η. If the sideband resistance is shorted the current ratio rises to its maximum. The ratio of voltage at the network output to that at the network input when the load resistance is made infinite coincides with this value (r_1/r_0).

If we consider the various possible kinds of resistance variation with time under the restrictions noted, it appears that the greatest attainable value of the power ratio is unity. It is evident from the equivalent circuit that this limit corresponds to no loss from signal to sideband. The closest approach to this no-loss condition is obtained in a modulating element presenting a resistance which, over a carrier cycle, varies between widely different resistances, taking on one extreme value for a small fraction of the cycle and remaining near the other extreme for the remainder of the cycle. Under these conditions the series arm of the equivalent net tends to zero, the shunt arm to infinity. There are practical limitations to the extent to which these conditions can be approached in practical modulators. For example the best attainable values of the two resistance extremes usually depend upon the modulator characteristic, and upon the carrier amplitude employed which may be limited by heat dissipation or by voltage breakdown. Further limitation is imposed by parasitic capacitances, which effectively limit the maximum attainable modulator resistance.

The commutator modulator may be used to illustrate the results of analysis above. Figure 4 shows the variation of the resistance of such a modulator over a carrier cycle. B and A are the two values of the resistance ($B > A$) and β is the fraction of the carrier cycle

Fig. 4—Variation of resistance with time in a commutator modulator, in which the resistance is switched from A to B, remaining at the higher value B for the fraction β of the carrier cycle.

over which the resistance is B. The coefficients of the resistance expansion are readily shown to be

$$r_0 = \beta B + (1 - \beta)A, \tag{11}$$

$$r_k = (B - A)\frac{\sin k\beta\pi}{k\pi}. \tag{12}$$

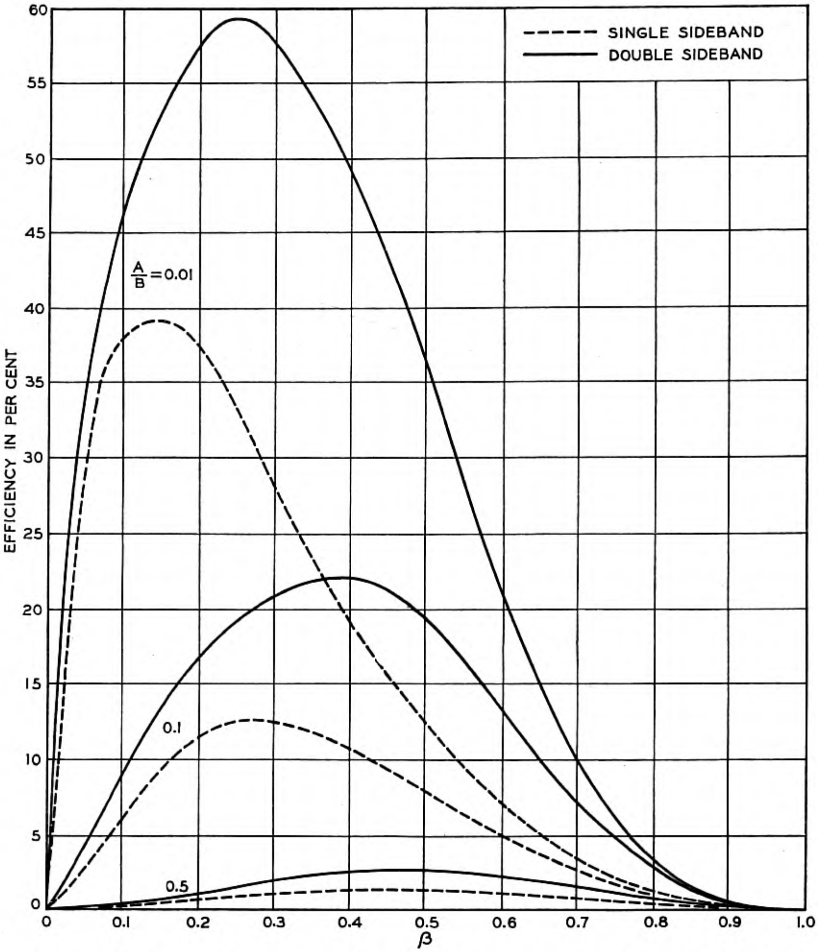


Fig. 5—Efficiency of a commutator modulator shown as a function of the pulse fraction β with the ratio of the two commutator resistances as parameter. Here the resistance terminations to signal and to sideband are optimum and the circuit impedance is made high to unwanted components. Full line applies to double sideband output, dashed line to single sideband output.

The efficiency (with optimum termination) depends only on β and the ratio A/B. The dashed curves on Fig. 5 show the variation of efficiency with these parameters. It is evident from the equivalent circuit that best efficiency is obtained with $r_0 - r_1$ small and r_1 large. In this case r_0 increases linearly with β, and r_1 varies sinusoidally with β—having its maximum at $\beta = 0.5$. The immediate conclusion is that the optimum conditions are obtained when β is less than 0.5. In fact the efficiency approaches the limiting value of 100 per cent when β is very small and B is much greater than A, as may be seen from (11) and (12).

III. DOUBLE SIDEBAND—HIGH IMPEDANCE OUTSIDE BAND

In double sideband operation both upper and lower sideband currents flow, but all other modulation products are suppressed as in the previous case. Here the equations for signal and upper and lower sideband respectively are

$$\begin{aligned} (r_0 + R_q)Q + r_1 I_{1+} + r_1 I_{1-} &= E_q, \\ r_1 Q + (r_0 + R_{1+})I_{1+} + r_2 I_{1-} &= 0, \\ r_1 Q + r_2 I_{1+} + (r_0 + R_{1-})I_{1-} &= 0. \end{aligned} \tag{13}$$

Comparing (13) with the equations for a three-mesh circuit, we obtain the equivalent network of Fig. 6. It is obvious from the symmetry of this network that the two sidebands are equal when $R_{1+} = R_{1-}$. Conditions for optimum efficiency may be put in form permitting convenient comparison with the single sideband case when we assume equal resistances to both sidebands.

Efficiency curves of a commutator modulator are shown on Fig. 5 for both single and double sideband cases. They differ primarily in that the utilization of two sidebands gives greater efficiency, except in limiting cases. The outstanding difference is that the unsymmetric network has optimum signal and sideband resistances which are not equal except at three values of β equal to 0, 1/2 and 1. Modulators are often operated with β approximately 1/2, so that in this case the results here check with the common experience that the two terminating resistances should be equal. It may be remarked that only in highly efficient modulators would unequal terminations make an appreciable difference in the efficiency of power transfer.

A comparison of Figs. 3 and 6 gives some light on the difference in efficiency of the single and double sideband cases. The comparison is made when $R_{1+} = R_{1-}$. From the symmetry of the circuit of Fig. 6, I_{1+} then equals I_{1-} and the mutual resistance $(-r_2)$ may be eliminated, leaving a simple T network connecting the input and the load. This is, with two exceptions, the T network of Fig. 3 with all elements doubled in magnitude. The input series arm is decreased ($r_0 - 2r_1$ instead of $2r_0 - 2r_1$) and the output series arm is increased by an element $2r_2$. Since $2r_2$ is generally much smaller than r_0 there is a net gain in efficiency.

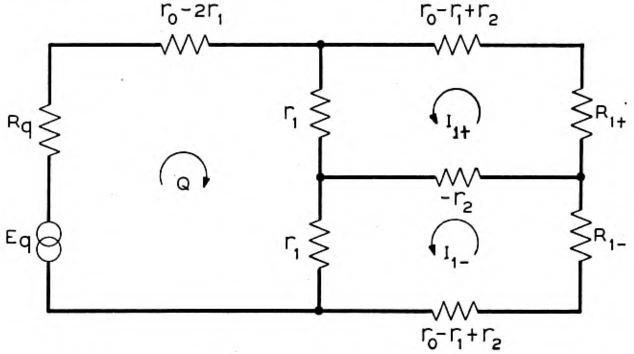


Fig. 6—Equivalent modulator circuit showing the connection between signal and two sideband circuits. All other modulation current components are suppressed by means of high circuit impedances.

IV. Low Impedance Outside Band

The foregoing systems involved high impedances, suppressing the flow of current at all but two or three frequencies. Circuits are as readily obtained in the case of low impedance to all but two or three of the modulation products. The physical systems this approximates are the same as those previously discussed except that the terminating filters must present a low impedance to frequencies outside the band.

All the external potential drops across the modulating element are taken as negligible except components at the signal and one or two sideband frequencies. The analysis, corresponding to that of the previous sections, uses Ohm's law in the form $i = Gv$. Thus this analysis, and the equivalent circuits, involve the expansion of a conductance instead of a resistance. The equations corresponding to (8) and (13) are equations in V_q, V_{1+} and V_{1-}. In the single sideband case a resulting equivalent circuit is that of Fig. 7. This is a simple symmetric π network. From its well known characteristics the optimum terminating conductance and maximum efficiency are immediately available:

$$G_{1+} = \frac{1}{R_{1+}} = \sqrt{g_0^2 - g_1^2}, \tag{14}$$

$$\eta = \left[\frac{g_1}{g_0 + \sqrt{g_0^2 - g_1^2}} \right]^2. \tag{15}$$

These expressions are identical in form with corresponding ones obtained for the high-impedance single-sideband case, in which conductance components replace resistance components. This confirms what has been observed in several special cases: that there is no theoretical advantage of either impedance extreme over the other in the general case.[5] Which one to use in any particular case depends upon the special characteristics of the modulator or upon the practicability of obtaining required impedance conditions.

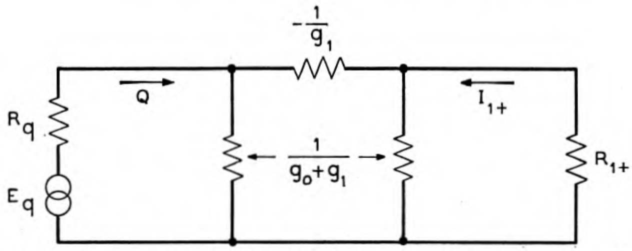


Fig. 7—Equivalent modulator circuit showing connection between signal and a single sideband. All other modulation voltage components are suppressed by means of low circuit impedances. The use of a Π network and the specification of element values as conductances are both matters of convenience.

The equivalent network for the corresponding double sideband system is shown in Fig. 8. Similarly to the previous double sideband case, the symmetry shows that when $R_{1+} = R_{1-}$ the sideband current amplitudes are equal and there is no potential across the coupling resistance $-(1/g_2)$. Thus it may be shorted, reducing the circuit to a simple unsymmetric π. The matching resistances and maximum efficiency may be obtained as before.

The results again are identical with the high-impedance case, if resistances are replaced by conductances. The comment made on the single-sideband case still holds—that there is no general theoretical advantage of either a high- or low-impedance system over the other as far as maximum possible efficiency is concerned.

The curves of Fig. 5 are evidently immediately applicable to the low-impedance circuits provided all resistances are replaced by conductances.

There are, of course, practical advantages of the high- or the low-impedance circuit in particular cases. For example, it is commonly easier to make the terminating impedance to unwanted frequency

---

[5] The form of the equations in corresponding high- and low-impedance cases suggests that the impedance and efficiency relations for one case could be deduced from those of the other through the principle of duality. See Guillemin, "Communication Networks," Vol. 2.

components very small rather than very large. The impedance matching may be simpler in one case than in the other since, for the same efficiency, the matching resistances are quite different in the two cases.

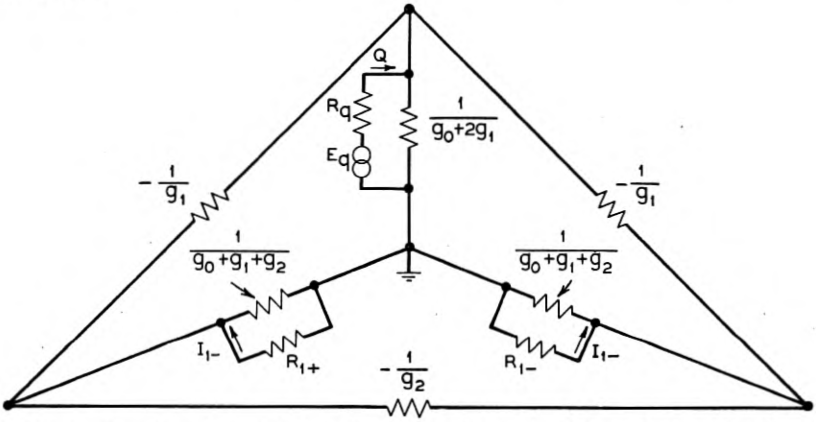


Fig. 8—Equivalent modulator circuit showing connection between signal and two sidebands. All other modulation voltage components are suppressed by low circuit impedances. When one of the sideband paths is shorted, the network reduces to that of Fig. 7.

V. FINITE RESISTANCE TO DISTORTION PRODUCTS

One more extension may be made, without excessively complicating the equivalent circuit. This is an extension to the case of a constant resistance R to all unwanted products, which yields information unobtainable from the limiting cases previously treated of $R = \infty$ and $R = 0$.

This problem can be handled in a simple way by the artifice of incorporating R within the modulator proper. In that case the external impedances to signal and sideband must be reduced by R to keep the total circuit resistance at its correct value, while the external resistance to any other modulation product then becomes zero. This brings the situation down to that considered and solved in the section immediately preceding.

By this manipulation the equivalent circuit is obtained as that of Fig. 9 in the single sideband case. The primes indicate coefficients in the expansion of the modified characteristic. These coefficients are immediately available in the case of the commutator, since the sole change there is an increase in both values of the variable resistance by the amount R. Comparing the efficiency of transformation with that obtained with extreme values of R as in the cases preceding, it appears that the use of extreme values of R results in the maximum obtainable power.

The impedance conditions specified above permit the flow of any generated products, so that an infinite network would be required in general to bring out the relation of any one product to the others. Discussion was limited, however, to signal and upper sideband; the amplitudes of other components did not appear explicitly.

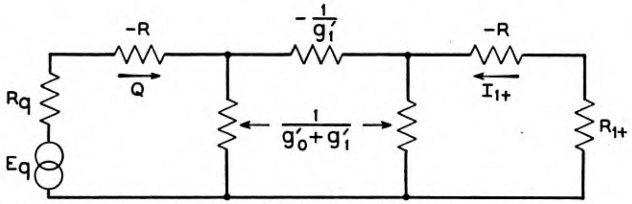


Fig. 9—Equivalent modulator circuit showing connection between signal and single-sideband circuits when a constant resistance R is effective to all other modulation products. The primed coefficients (g_k') apply to the conductance coefficients when the series resistance R is included in the original modulator.

The section following deals with a different case involving an infinite network in which individual meshes are treated explicitly.

VI. EQUIVALENT NETWORK FOR THE IDEALIZED RESISTANCE MICROPHONE

Other variable resistance systems may be put into equivalent form. For example, from the electrical side, an idealized variable resistance microphone actuated by a sinusoidal acoustic wave can be represented by the resistance

$$r = R_0 + R \cos qt.$$

This is exactly the form of the variable resistance already discussed, with $r_0 = R_0$, $2r_1 = R$, and $r_n = 0$ for $n > 1$.

If one were interested in the modulation products with other frequency components impressed electrically, the systems would be of the same type as those of the previous section. In the case of the microphone the d-c. voltage impressed leads to current components which are d-c. and harmonics of the signal q. In this case, the equations are

$$\left.\begin{aligned} V_0 &= E_0 - R_bI_0 &&= R_0I_0 + RI_1/2, \\ V_q &= - R_qI_1 &&= R_0I_1 + RI_0 + RI_2/2, \\ V_{2q} &= - R_{2q}I_2 &&= R_0I_2 + R(I_1 + I_3)/2, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ V_{nq} &= - R_{nq}I_n &&= R_0I_n + R(I_{n-1} + I_{n+1})/2, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned}\right\} \qquad (16)$$

in which R_b represents the external d-c. resistance, and R_{nq} represents the external resistance to the nth harmonic.

Figure 10 shows the equivalent circuit for this system in the form of an infinite ladder structure. From this circuit relative magnitudes of the various frequency components are readily perceived; evidently the successive harmonic current components decrease progressively in magnitude. A large value of R_{nq} makes the nth harmonic and all successive harmonics small. A case in which quantitative information is obtainable in simple form is that in which the resistances to all harmonics are equal ($R_{nq} = R_t$, $n > 1$). In this case the equivalent network beyond terminals (1) and (2) is a simple recurrent structure and the resistance is obtainable by the customary methods of handling infinite recurrent networks. The admittance looking in at the terminals (1, 2) is the iterative admittance of the latter and is given by [6]

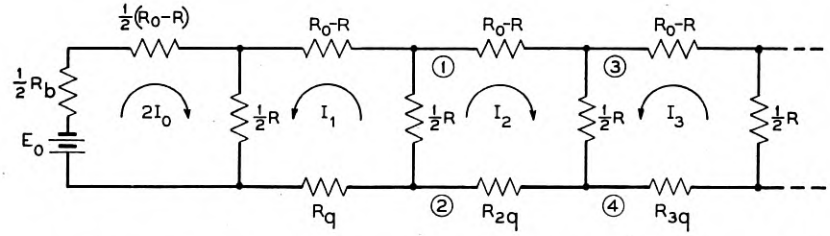


Fig. 10—Equivalent circuit of idealized variable resistance microphone. Here mesh currents represent amplitudes of the various frequency components. R_0 represents the fixed and R the variable internal microphone resistance, while R_{nq} represents the external circuit resistance to the nth harmonic of the signal.

$$\frac{1}{R_i} = \frac{2}{R} + \frac{1}{R_0 - R + R_t + R_i} \cdot \tag{17}$$

Solving for R_i gives

$$R_i = \frac{R}{1 + \sqrt{1 + \dfrac{2R}{R_0 - R + R_t}}} \cdot$$

As far as d-c. and fundamental current components are concerned, the network beyond (1), (2) may be replaced by R_i and the system reduces to a two-mesh circuit from which Q, I_0 and power relations are obtainable very simply. If harmonic current amplitudes are

---

[6] In an infinite recurrent structure, the resistance R_i must be the same looking in at successive point pairs (1, 2) (3, 4) etc. This is stated in (17).

desired, they may be obtained from equations (16), using the known values of I_0 and I_1.

The two-mesh circuit for I_0 and Q can immediately be put in the form of an unsymmetric T terminated at one end by the battery and its internal impedance and at the other end by R_q. The optimum terminating resistances and corresponding efficiencies are obtainable as in previous cases but it is evident from the network, without further computation, that losses are minimized (with suitable termination) if $R_0 - R$ and R_i are small compared with R. R_i is decreased by decreasing $R_{n\,q}$ $(n > 1)$. These conditions mean that the best electrical efficiency is obtained when the resistance variation is large and the unwanted signal harmonics are short circuited.

VII. Extensions and Summary

Equivalent networks can be obtained in some cases when the restrictions on the relative amplitudes of signal and carrier are removed. It is evident from Fig. 2 that the value of the variable resistance at any instant then depends not only on the carrier amplitude but also on the signal amplitude. Thus the equivalent networks are no longer made up of constant resistances, but depend upon the magnitudes of both signal and sideband components. Further, new components appear involving multiples of the signal frequency. The equivalent for this case lacks the simplicity of those discussed here, a simplicity which appears when one of the two input components is much greater than the other.

The reason for the restriction to pure resistances becomes evident when one attempts to generalize the results. The current components will then have phase angles differing from zero in general. Consideration of lower sidebands then shows that the phase angles must have their signs reversed in certain circumstances, which leads to obvious complexities. Again in purely resistive circuits it is possible to determine the instantaneous current-voltage relation and hence to specify the resistance variation as a function of time. In a reactive circuit, however, additional difficulty arises in that the relations are much more complex and in general impossible to specify in simple terms.

To summarize, the presentation has been limited to the simplest circuits used for modulation by means of a variable resistance. In each example, the inter-relations between modulation product amplitudes, terminating resistances, and types of modulator characteristics are shown in terms of familiar linear resistance networks. From these, qualitative information concerning the properties of the system is more readily obtained than from the equations and, in some cases, the solutions for effective impedances and current and voltage amplitudes are obtainable without further recourse to the equations.

ACKNOWLEDGMENT

The writers are indebted to several of their associates in the Bell Telephone Laboratories for the use of unpublished material in this paper. In particular, acknowledgment is due to Mr. R. S. Caruthers, Mr. J. M. Manley, Dr. G. R. Stibitz, and Mr. R. O. Wise, who originally obtained some of the impedance and efficiency relations.

An Improved Three-Channel Carrier Telephone System

By J. T. O'LEARY, E. C. BLESSING and J. W. BEYER

This paper describes an improved three-channel carrier telephone system for use on open-wire lines. It employs recent advances in the telephone art to bring about many economies and circuit simplifications as compared with previous models of the three-channel system. A new type of automatic regulating equipment is included.

INTRODUCTION

THERE are now in service in the Bell System approximately 750,000 miles of telephone circuit which are furnished by carrier systems. Of this total, almost 90 per cent is provided by some 600 Type C systems, ranging from about 75 miles to over 2000 miles in length. Basically designed to add three carrier channels to the normal voice channel on open-wire lines, the Type C system has also been used in special cases to provide additional circuits over deep sea cables of moderate lengths.

The system was first described in this Journal in the July 1928 issue.[1] Improved designs and the application of new circuit elements have recently permitted a very extensive revision of the terminal and repeater equipment which results not only in striking reductions in size and cost as compared with the older equipment, but also gives a considerable improvement in transmission performance. A new type of automatic regulating equipment has been provided for both the terminal and repeater.

The improved system employs heater type pentode tubes, copper-oxide modulators and demodulators and makes use of the negative feedback type of amplifier at both terminal and repeater points. The terminal band filters are newly designed to give improved transmission frequency characteristics on all channels. Each channel is arranged to terminate on a four-wire basis in the same manner as the Type K system for cables.[2]

An outstanding feature of the modified design is the large saving in space in comparison with the previous equipment. As shown on Fig. 1, the complete terminal with its regulating equipment occupies a single bay, whereas the older system without regulating equipment required two and one-half bays. The repeater space savings, while not so large, are nevertheless substantial. The number of vacuum

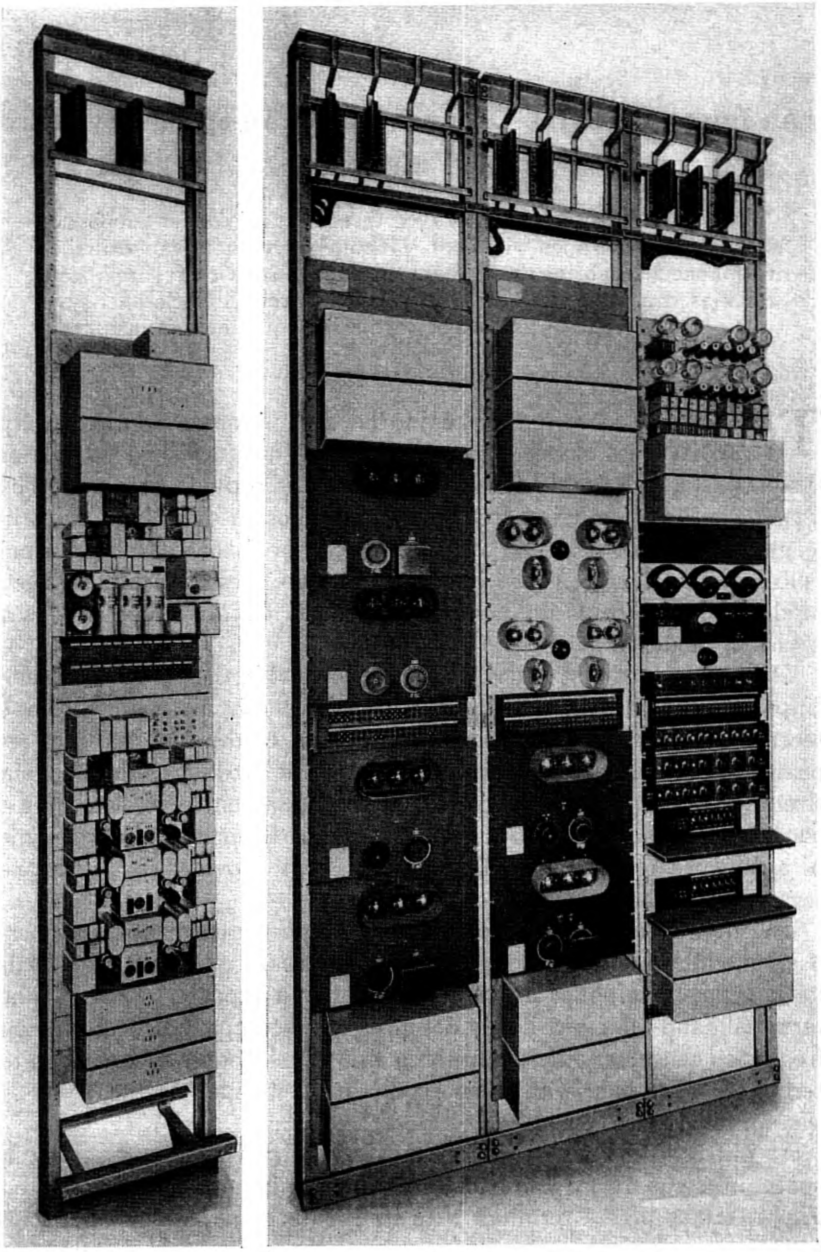


Fig. 1—New and old type C carrier telephone terminals.

tubes required in the system has been reduced, which results in a material saving in power.

Certain features of the improved equipment, notably the automatic regulation, can also be used on the older types of systems and the design objectives were set up with this in view.

Frequency Allocation

The frequency range employed by the system and the allocation of channel bands within that range are shown in Fig. 2. The allocations used in the older systems are also shown for comparison.

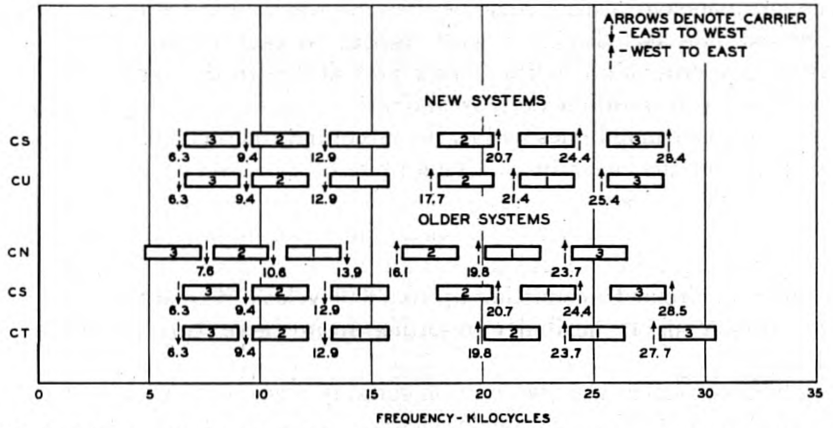


Fig. 2—Frequency allocation of the new systems relative to that of the older systems.

The original selection of the frequency range for the Type C system was the result of many different factors. Foremost among these was the desire to keep the frequencies low in order to minimize line crosstalk and attenuation and changes in the latter due to weather and temperature. On the other hand was the greater filter cost that results from crowding the channels close together. Different frequency bands were used for transmitting in opposite directions in order to avoid the problem of near-end crosstalk and to give the advantages of four-wire transmission. Although consideration was given to the general desirability of increasing the band of frequencies effectively transmitted by the individual channels, the requirement for coordinated operation with older systems already in use precluded any material increase in the frequency space allocated to each channel of the new system. Nevertheless, as will be seen from Fig. 4, the channels show very little distortion within the transmitted band and represent a material improvement over the older systems.

Because the line crosstalk tends to be greater at the higher frequencies, past experience has indicated the advantage of having available two systems between which the crosstalk in the higher frequency group will be unintelligible. Two allocations are provided for this purpose, designated CS and CU. The channel bands are identical in the lower frequency group (East to West) and in the upper frequency group (West to East) differ only in that the carrier frequencies are at opposite ends of the bands. In this group crosstalk between similar bands will have the speech frequencies inverted and will therefore be unintelligible.

This arrangement does not give as high a crosstalk advantage as the arrangement used previously where the bands were not only inverted but also displaced with respect to each other. However, better line crosstalk conditions now prevail due to the application of improved transposition designs and line configurations to the more recently constructed lines and to the use of new methods of mitigating crosstalk on the older lines. This permits the simplification of the frequency allocation, as a result of which one system may be readily converted into the other with fairly simple equipment changes. It will also be possible to use the voice frequency circuit on all pairs as a program circuit transmitting up to 5000 cycles. The advantages of the greater plant flexibility resulting from these two factors are obvious.

The new system may be used on suitably transposed lines with the Type D[3] and Type H[4] single channel systems, whose frequency bands are such that no serious near-end crosstalk problem will arise.

OVERALL SYSTEM

A block diagram of a complete system, consisting of the two terminals and a single intermediate repeater, is shown on Fig. 3. In practice there might be as many as ten or more such repeaters. The two terminals differ from each other only in the frequencies for which their respective transmitting and receiving circuits are designed. The west terminal transmits the high-frequency group of Fig. 2 and receives the lower frequency group while the east terminal does the reverse. The repeater is provided with means for separating the frequencies in the two directions of transmission, amplifying the current to the desired level, and passing them on to the next line section.

A typical overall frequency characteristic for one of the circuits derived from the new system is shown on Fig. 4. This characteristic illustrates the relative freedom from distortion in the transmitted frequency range.

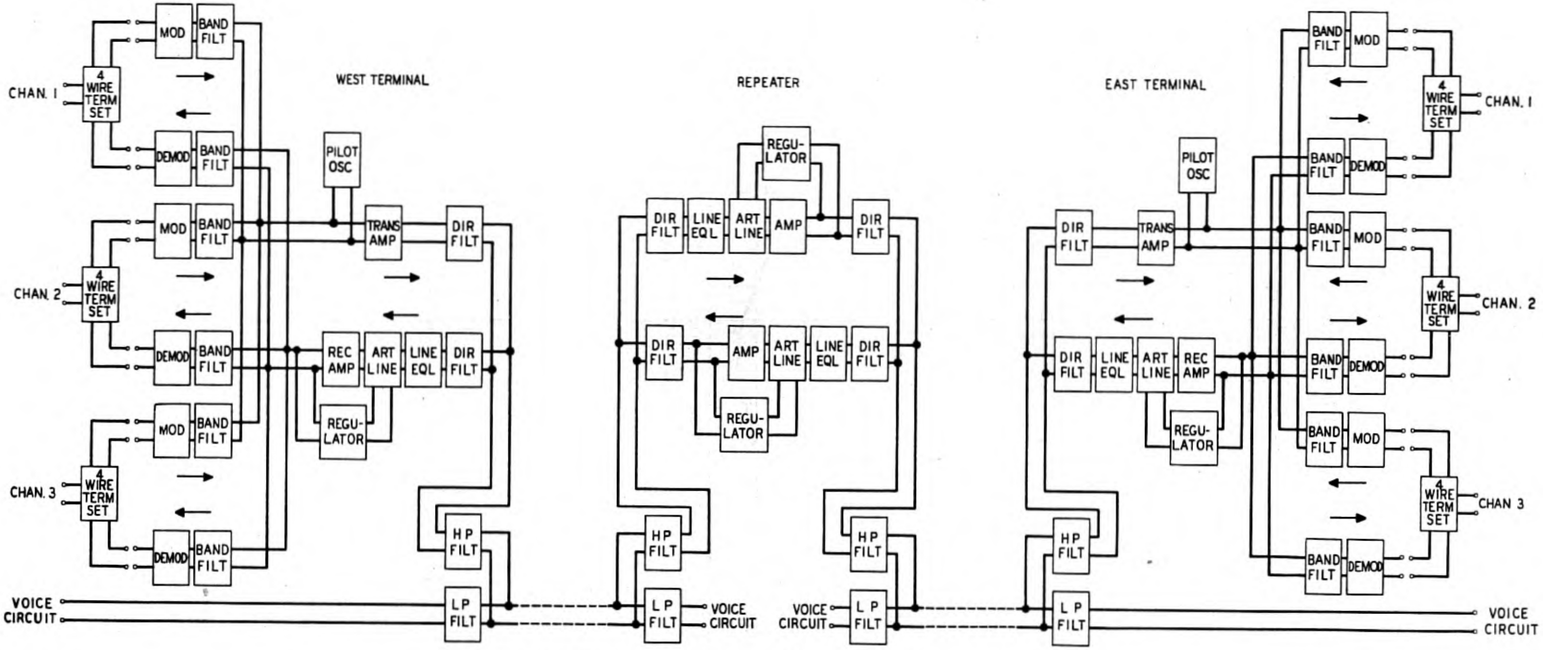


Fig. 3—Schematic of an overall system with one repeater.

The carrier channels are separated from the voice frequency circuit on the same pair of wires by means of a high-pass and low-pass filter combination as shown on Fig. 3. Several different filter sets are available for this purpose differing from each other in the frequency band which is desired in the voice circuit. Where the voice circuit is an ordinary message circuit the filter will have a cutoff around 3 kc. Where it may be used for program transmission a filter set having a cutoff above 5 kc is provided. For still wider program bands there is a filter set cutting off above 8 kc. The use of this latter filter would, of course, require the sacrifice of the lowest carrier channel since it would be overlapped by the program band.

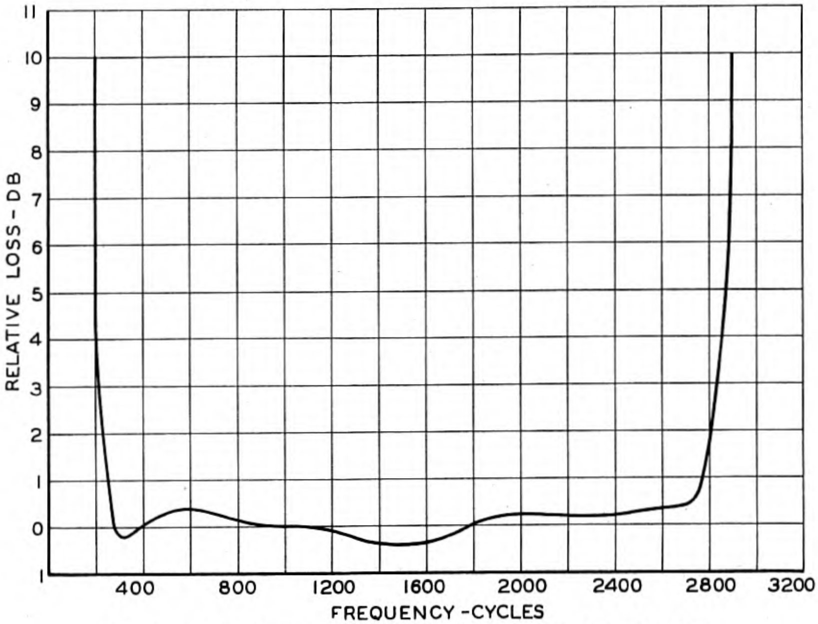


Fig. 4—Typical overall transmission-frequency characteristic.

An important feature of the system is the method of stabilizing the overall transmission. Ahead of the terminal transmitting amplifier in each direction of transmission there is connected to the circuit an oscillator which generates a pilot current. This pilot current has a frequency adjacent to the band of the middle channel. The oscillator is designed to have a relatively high degree of stability with respect to output and frequency. At the output of each repeater and at the receiving terminal the pilot frequency is selected by a high-impedance bridging filter, which has little effect on the through transmission, and is then used to actuate a regulating mechanism. Changes in the line transmission at this frequency are indicative of the changes at all frequencies and the regulator functions to maintain a nearly constant output level in all three-channel bands.

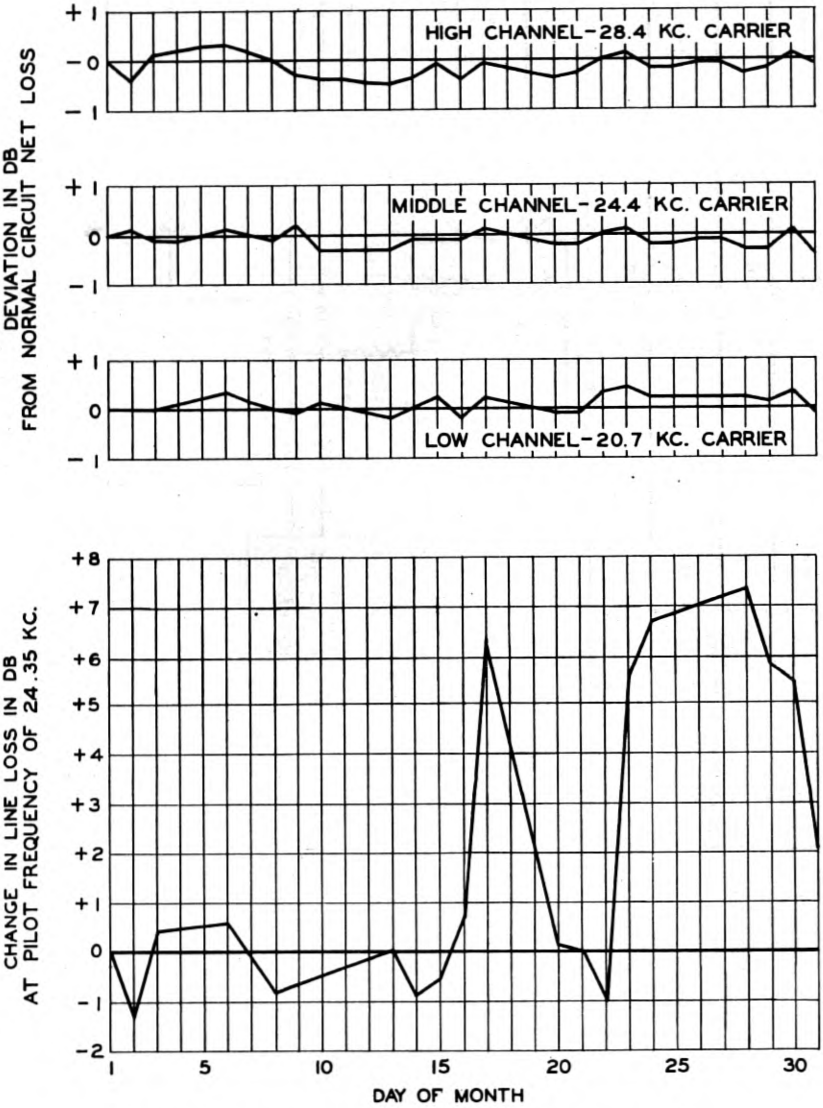


Fig. 5—Chart showing regulation for a period of one month on system with one repeater.

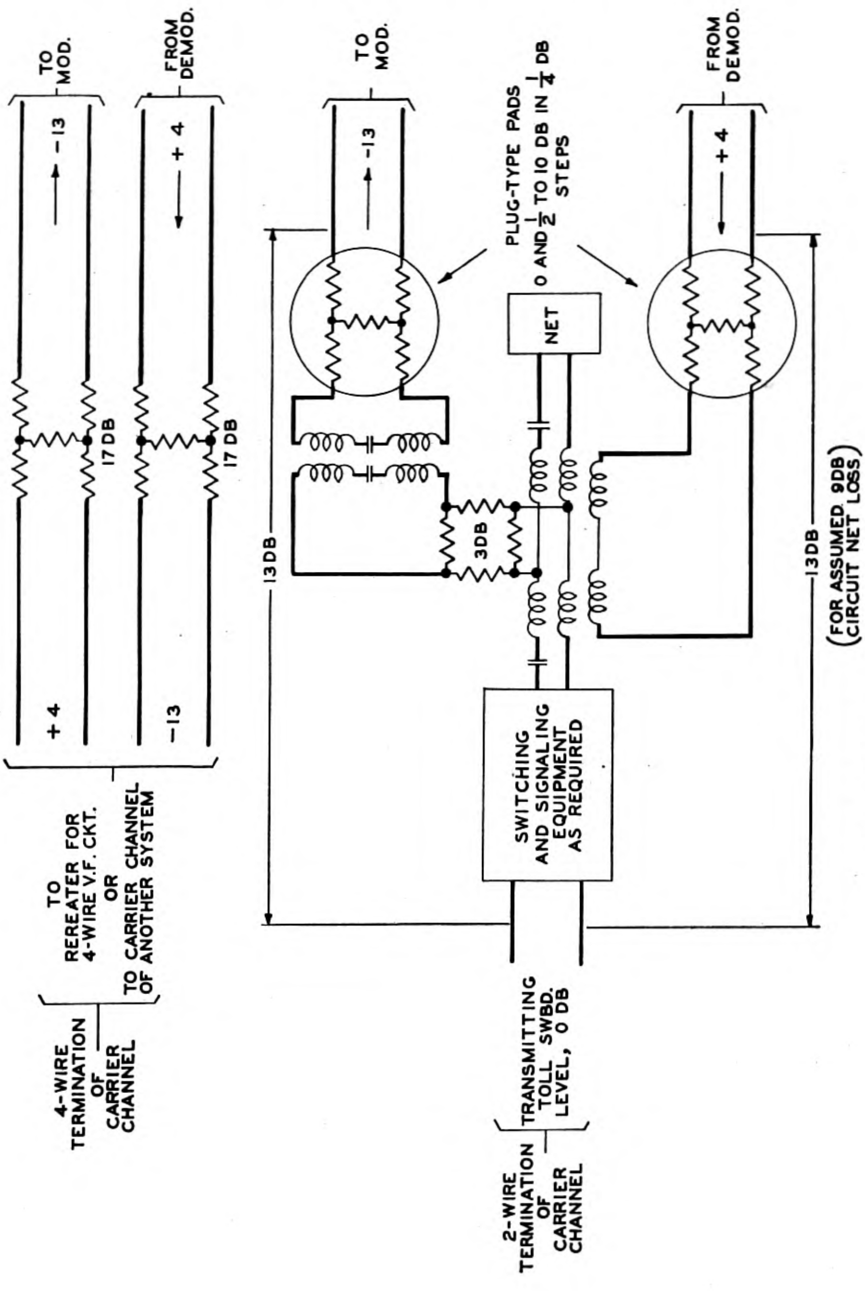


Fig. 6—Voice terminating arrangements.

The pilot current is also used to indicate through an audible or visual alarm any trouble which results in large sudden changes in transmission such as would be occasioned by an open or short circuit on the line itself.

The ability of the regulating mechanism to stabilize the transmission over the system is shown on Fig. 5 which shows the deviations recorded in daily measurements on all three channels of a 250-mile system over a period of one month. The actual changes in line loss at the pilot frequency are also shown for comparison. During this period various conditions of temperature, rain and fog were experienced.

With the transmitting level that has been provided and for ordinary line conditions it is found practicable to employ repeater spacings of from 125 miles to over 250 miles. The exact distance in any particular case depends upon many factors, such as: wire size, length of toll entrance or intermediate cables, location of existing offices and the susceptibility of the line to sleet or frost. Where this latter condition is prevalent conservative spacings are desirable.

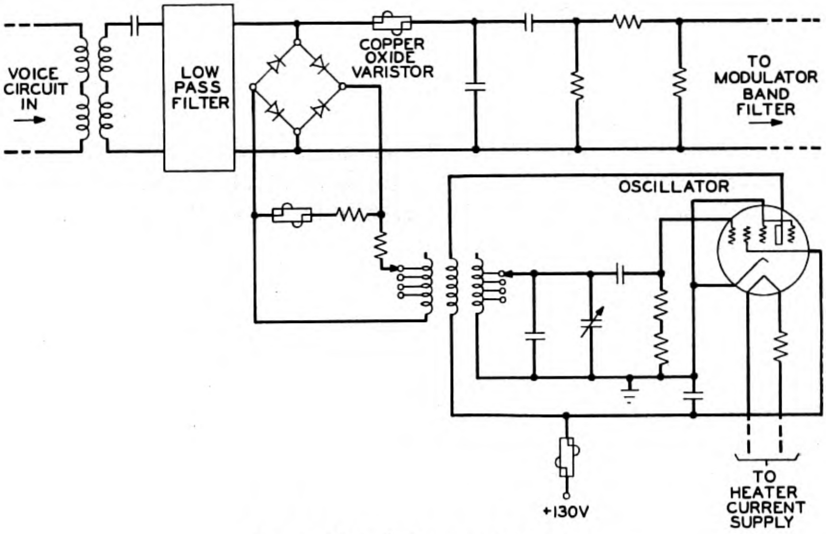


Fig. 7—Schematic of modulator.

TERMINALS

The general theory of operation of the terminal may be understood from the block diagram shown in Fig. 3. On the voice-frequency side each channel terminates as a four-wire circuit. The input to the carrier system from the voice circuit is designed to operate at a level

13 db below the transmitting toll switchboard which is the common reference point. The output from the system is at a level 4 db above that point. Equipment for coupling the system to both two-wire and four-wire circuits has been designed. The circuits employed in each case may be seen in Fig. 6.

The modulator circuit, shown in Fig. 7, uses copper-oxide varistors[5] for converting the voice frequencies to the higher line frequencies. The high degree of balance obtained in the copper-oxide varistors has the important advantage of making carrier leak a practically negligible factor. This is of particular importance in the case of that channel to which the pilot current is adjacent in the frequency spectrum. The modulator circuit is also designed to limit the peaks of very loud talkers which would otherwise overload the common amplifiers. The effect of this limitation on the quality of the speech transmitted is not noticeable.

The oscillator which supplies the carrier to the modulator is designed to be stable in both output and frequency. When it is once adjusted with the oscillator at the distant end, departures from synchronism will be relatively small. Part of this stability is due to a new circuit design employing coil and condenser elements having opposite temperature coefficients so that changes in one will be compensated for by changes in the other.

The band filters use coils wound on magnetic core material, having improved modulation characteristics, instead of the solenoidal air core coils previously used. This results in a considerable reduction in the space which they occupy.

The transmitting and receiving filters associated with each channel are identical as to band width. They are further characterized by a more abrupt increase in discrimination immediately below and above the pass-band frequencies than was realized in the channel filters for the previous Type C systems and also by less distortion across the pass band. Most of this distortion is in the form of higher loss in the vicinity of the band limiting frequencies. It was deliberately included in the design of the filters for the purpose of masking delay distortion effects on overall transmission quality which might otherwise become noticeable when four or five type C carrier telephone systems are connected in tandem.

The uniformity and symmetry of the various filters are shown by Fig. 8 which gives the characteristics of those in the upper frequency group. This symmetry is required in this group in order to make the CS allocation convertible into the CU by moving the carrier from one end of the band to the other.

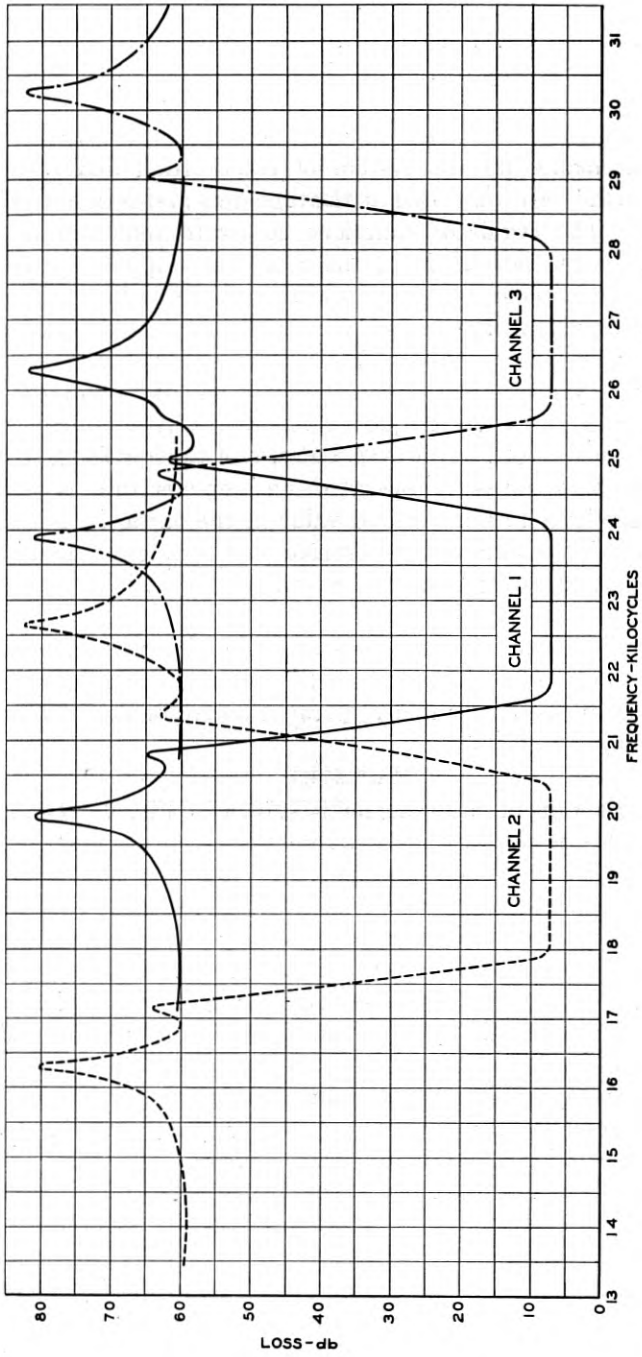


Fig. 8—Typical channel band filter characteristics.

The transmitting amplifier is the same as that used on the receiving side and in both directions of transmission in the repeater. This amplifier is capable of operating at a level 18 db above the transmitting toll switchboard.

On the receiving side, following the directional filters, is the equipment which makes up the system of equalization and regulation. This is identical with that used at the repeaters and is described more fully later. The regulator functions so as to maintain a nearly constant level at the output of the receiving amplifier. The band filters differ from those on the transmitting side only in the frequency bands which are transmitted and are the same as those used at the distant transmitting terminal. The demodulator circuit is of the same general type as the modulator circuit and the oscillator which supplies it with a carrier frequency is practically identical to that used by the modulator. However, because of the low levels at which these copper-oxide units are operated, an amplifier tube is necessary to restore the level to the required value at the output. The gain of this amplifier is continuously adjustable over a range of about 10 db so that precise adjustments of the overall circuit net loss can be made on each channel individually.

On very short non-repeatered systems the transmission variations may not be great enough to require the automatic regulating equipment. In such cases a manually operated potentiometer will be used for controlling the gain.

Repeater

A block diagram of the repeater is shown in Fig. 9. Directional filters on each side separate the two directions of transmission. As in the case of the receiving terminal, the equalizing and regulating equipment maintains all channels at the proper level at the amplifier output. The high cut-off filter shown on the circuit in the west-to-east direction limits the transmission to frequencies below 30 kc. This may be desirable when a system employing still higher frequencies is used on the same pair as the Type C system or on other pairs on the same line.

The repeater provides a maximum gain of about 49 db at the highest frequency in the upper group of the new system and about 43 db at the similar point in the lower group. The exact gains at different points in the frequency range are adjusted by the regulator so as to compensate for the attenuation of the line section preceding the repeater.

A complete repeater with its regulating equipment is mounted on a single equipment bay. A photograph of this bay is shown in Fig. 10.

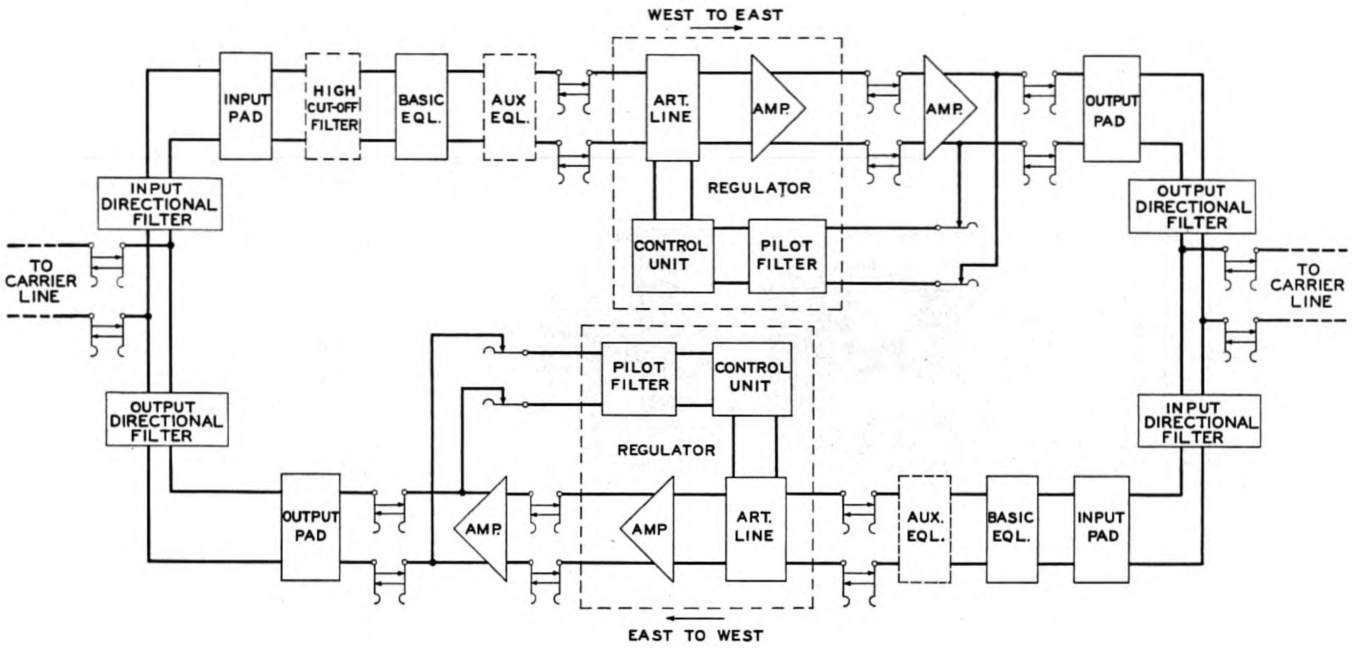


Fig. 9—Schematic of carrier repeater.

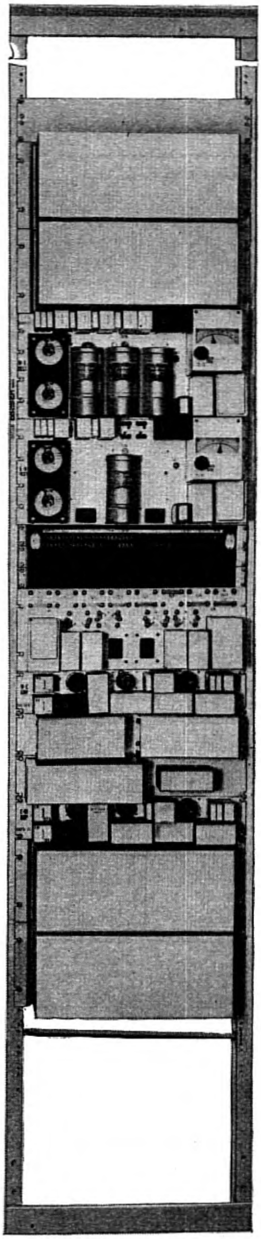


Fig. 10—New repeater for Type C systems.

Regulation

At the transmitting terminal each channel is adjusted to the same output level and in the operation of the system it is desirable to restore this equality of levels at each repeater point and at the receiving terminal. In each direction of transmission the line losses at the upper end of the frequency range will be higher than at the lower

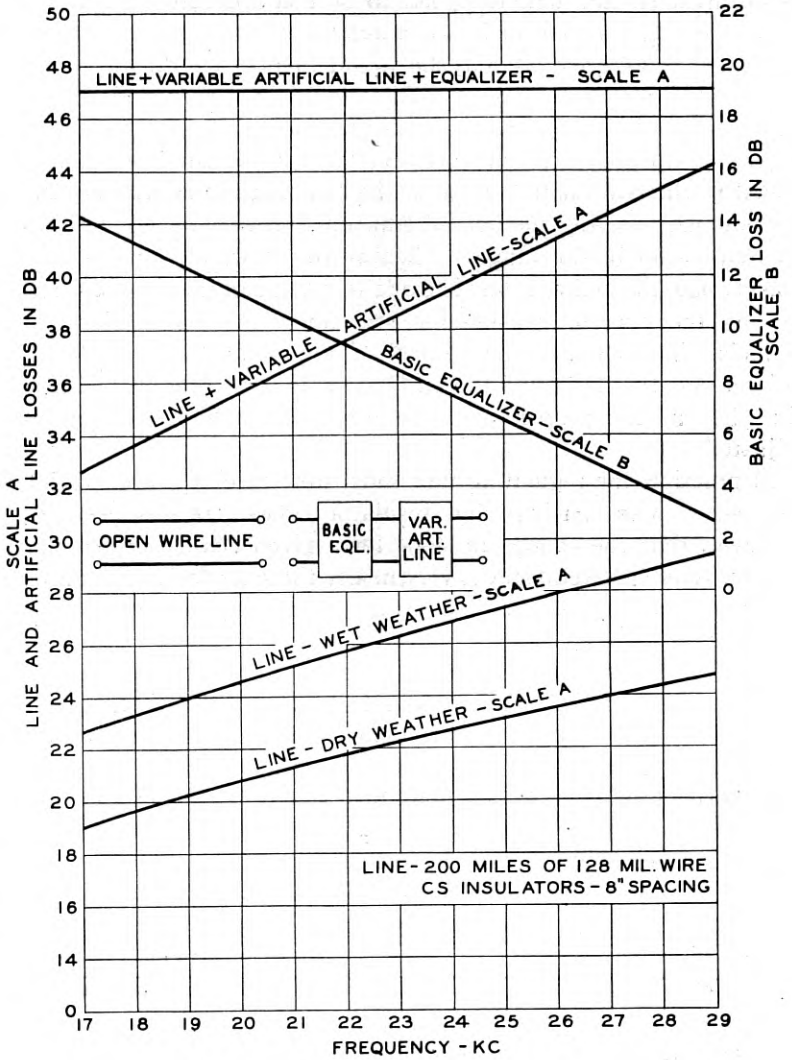


Fig. 11—Theory of regulator.

end and, furthermore, these losses change by varying amounts with temperature and weather conditions. It is the function of the regulating equipment to provide the needed adjustment and equalization of levels over the wide range of line conditions.

The theory of the operation of the regulating system is shown in a simplified manner in Fig. 11. The line circuit is connected at the end of each repeater section to a line equalizer, which is in turn connected to an artificial line unit designed to be continuously variable. The slope of the line equalizer loss characteristic is the reverse of the line slope and is as great as may be found in practice on any ordinary length of line section except under conditions of severe ice or frost. The artificial line slopes in the same direction as the line circuit itself. In lining up a system the artificial line is adjusted so that when added to the real line the slope of the combination neutralizes that of the equalizer leaving the overall transmission very nearly uniform for all frequencies in the range. Then as the loss and slope of the real line change, the artificial line is made to change in the reverse direction leaving the overall transmission still uniform. The action of this artificial line is under the control of the pilot current, referred to before, and the design of the equipment is such that in maintaining the pilot at the proper level the other channels are also properly adjusted.

In practice, the regulating unit must take care of a wide variety of wire sizes, wire spacings and insulator types. It has been found, however, that the change in slope for a given change in attenuation at some reference frequency is very nearly the same for all combinations of the above during ordinary weather conditions. As a result a single unit can be made to serve all cases.

Where conservative repeater spacings are employed there is a large amount of regulating range available to take care of sleet or frost formations on the wires. For the particular use illustrated in Fig. 11 the total range is about twice that required for ordinary wet weather. For shorter sections the available range would be still greater.

Some details of the regulating equipment are shown on Fig. 12. The pilot frequency is selected from the other frequencies on the line by a narrow band filter bridged across the output of the amplifier. This filter has a high impedance so that the bridging loss is small and does not interfere with regulation at succeeding stations.

A copper-oxide varistor is used to convert the selected pilot frequency into direct current which in turn actuates two Weston Sensitrol relays connected in series. The relays are equipped with meter scales on which a needle attached to the armature serves as a pointer. One of these relays controls the action of the regulator, the other functions

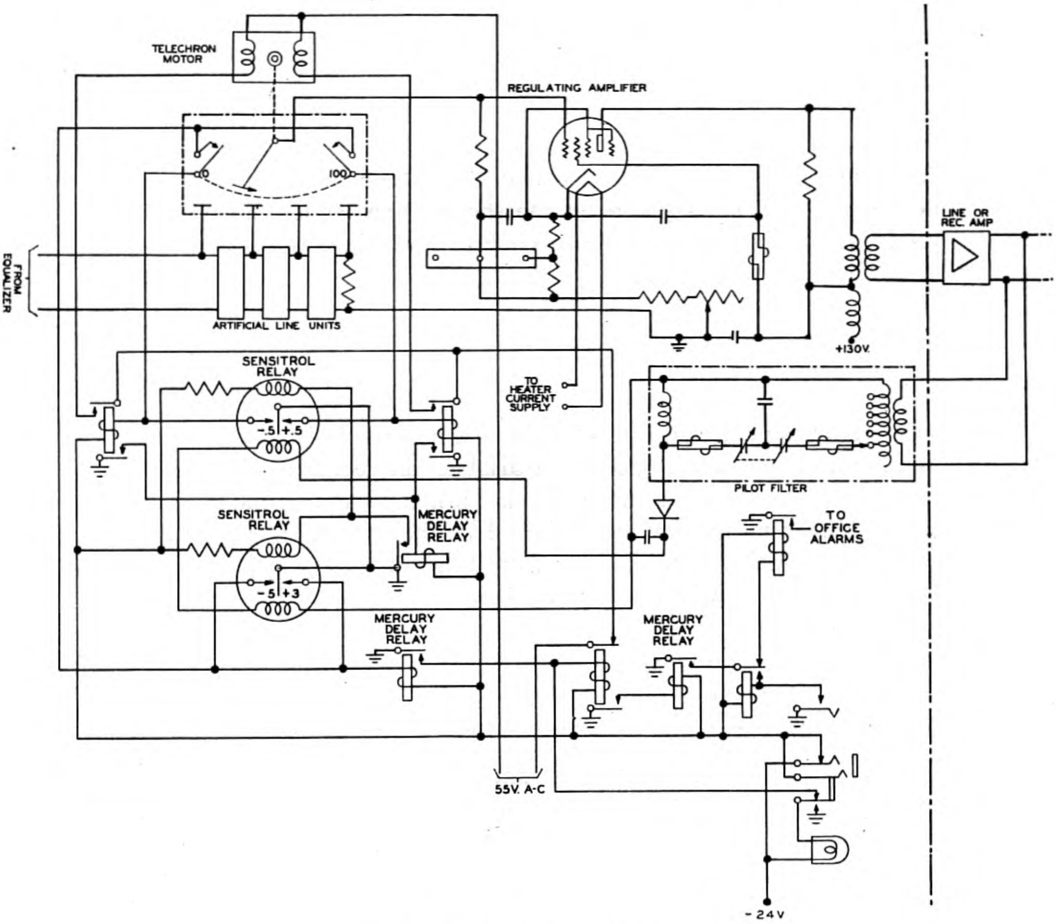


Fig. 12—Regulator circuit.

in an alarm circuit described later. With normal current flowing, corresponding to the proper pilot level at the output of the amplifier, the pointers of these relays are on midscale with a reading of 0 db. When the pilot level changes by more than .5 db in either direction, contacts close on the relay controlling the regulator and cause power to be supplied to a Telechron motor which controls the continuously variable artificial line mentioned above. This will be driven intermittently at a rate of about 1 db per minute until the output level has been restored to within ± .5 db.

As will be seen from the sketch the variable line consists of three sections connected in tandem. The ends of these sections are connected to the four stators of a variable air condenser. The rotor of the condenser meshes with any stator or with parts of two adjacent stators. The condenser, therefore, serves as a potential divider across the regulating network sections, the loss introduced depending upon the position of the rotor with respect to the stators. Basically these artificial line sections consist of units having the same loss characteristic. The first section, however, may be supplemented by additional units which will be required on the shorter line sections in order to build out the line slope as shown on Fig. 11. Since this section will be the last one to be cut out by the regulator, the less accurate part of the regulating range is thereby reserved for the periods of very high line loss, such as during ice or frost formation which occurs only infrequently.

The second sensitrol relay has contacts which close only on much larger changes in the pilot level such as would result from a failure of the line itself. When it operates it disables the regulating circuit and through a slow operating mercury relay brings in an audible or visual alarm indicating to the attendant that the system is in trouble. The slow operating relay introduces a delay in the operation so that short interruptions will not operate the alarm system.

The principal function of the regulating amplifier shown on Fig. 12 is to provide a high-impedance termination for the regulating network and condenser combination. There is, however, a small amount of gain available which may be useful in some cases.

NEW LINE ANPLIFIER

The amplifier which is used in the repeaters and in the transmitting and receiving branches of each terminal is one of the outstanding developments of the system. It was designed to have satisfactory transmission characteristics over both upper and lower frequency groups. It employs the principle of negative feedback [6] to achieve a high degree of stability, freedom from modulation and stabilized input and output impedances.

The advances made in the design of this amplifier can be seen by the comparison in the following table with the push-pull amplifier which was used in the older systems. In some cases the latter was supplemented by an auxiliary amplifier where higher gains were needed.

	24-Volt Power—Watts	130-Volt Power—Watts	Panel Space—Inches	Gain db	No. of Tubes
Push-Pull Amplifier	52.8	15.1	12¼	32	6
Push-Pull Amplifier Plus Auxiliary Amplifier	76.1	17.2	17½	48	8
New Feedback Amplifier	16.4	6.3	3½	50	2

There should be added to the comparison the fact that the new amplifier does not require selection of tubes to obtain satisfactory modulation results. It is also more stable with variations of power voltages and changes of vacuum tubes. The large space saving will be evident from Fig. 13, which pictures the new amplifier with the old push-pull amplifier and its auxiliary.

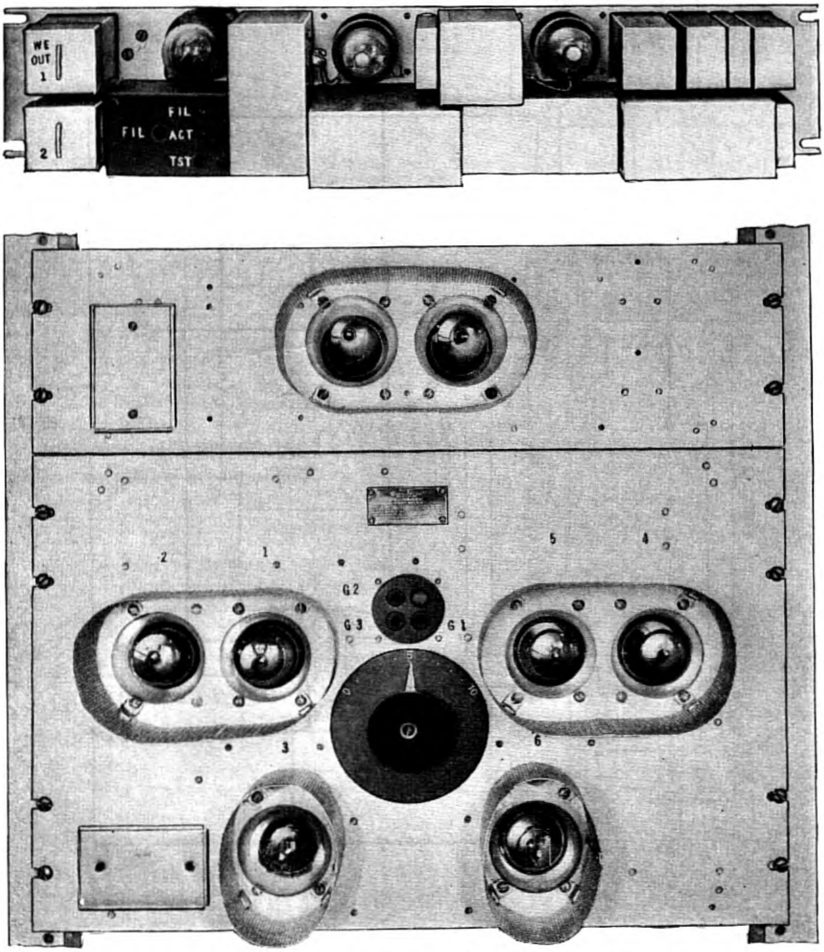


Fig. 13—New amplifier compared with the older type which it replaces.

The circuit of the amplifier is shown in Fig. 14. It is a two-stage circuit using pentode tubes. As will be seen from the circuit the feedback connection is obtained through the use of input and output transformers which are essentially hybrid coils. These coils are

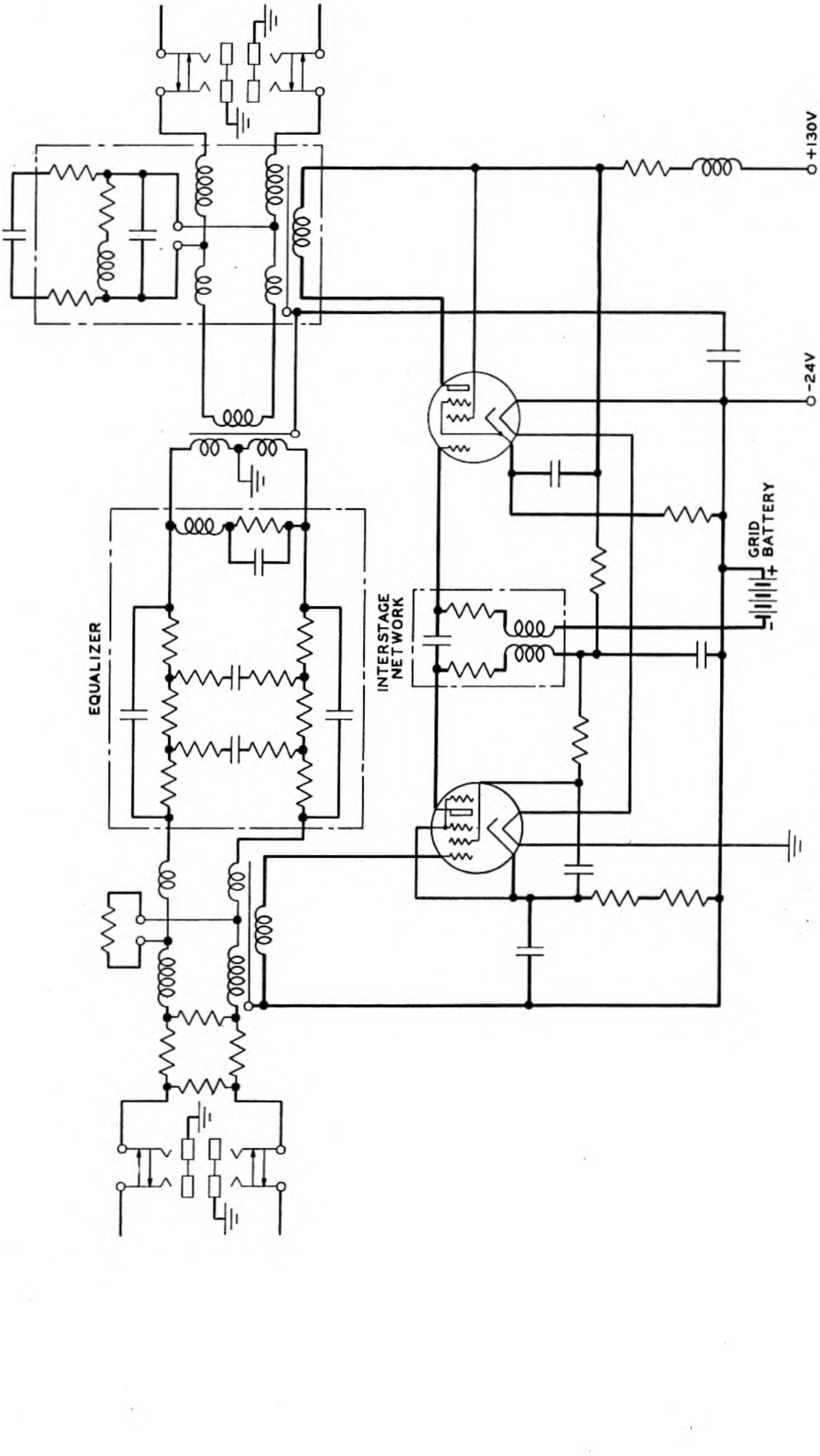


Fig. 14—Schematic of amplifier used in terminal and repeater.

designed to have unequal ratios, minimizing the loss to through transmission at the expense of greater loss in the feedback circuit. Including the transformers in the feedback path makes them also beneficiaries of the feedback with a resultant reduction in impedance irregularities, transmission distortion and modulation products.

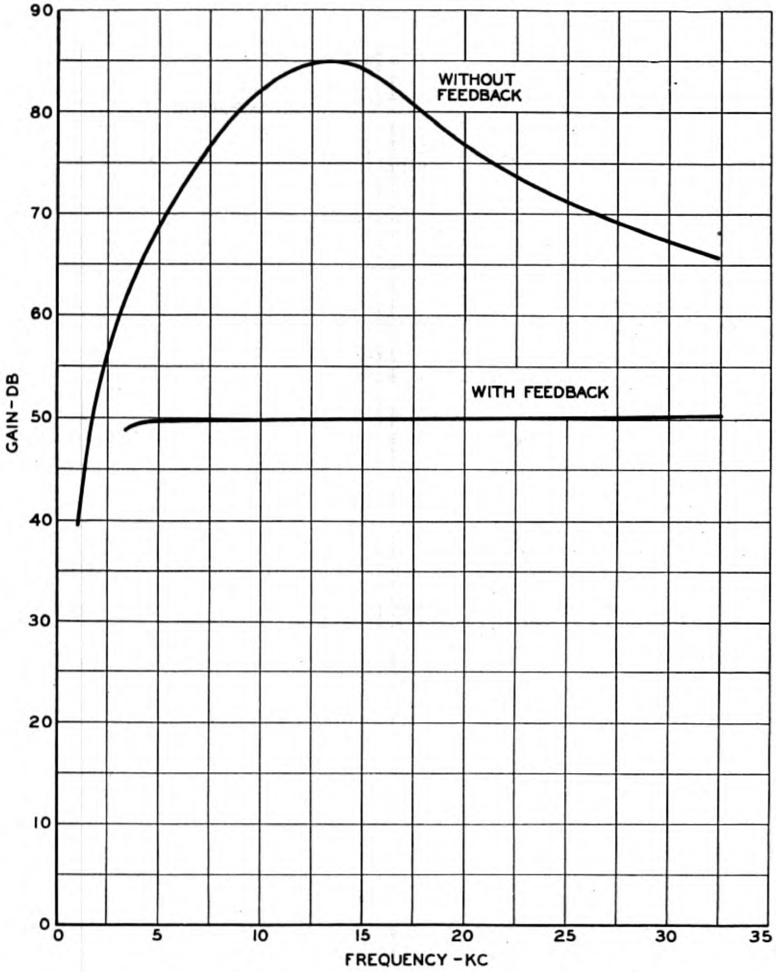


Fig. 15—Gain frequency characteristic of amplifier.

The amplifier has a gain of 50 db with provision for increasing it to 52 db when used as a terminal transmitting amplifier. Gain-frequency characteristics are shown in Fig. 15 for the 50-db gain condition with and without the feedback connection. The effect of feedback on transmission distortion is evident in this figure.

The second and third harmonic products in the amplifier are illustrated in Fig. 16 which shows their relation to the fundamental for different values of fundamental output. This harmonic production is a good measure of the modulation performance. In this respect the new amplifier is as good as or better than the push-pull amplifier which was used in the older system when a periodic selection of tubes was necessary to insure a satisfactory reduction of second harmonics.

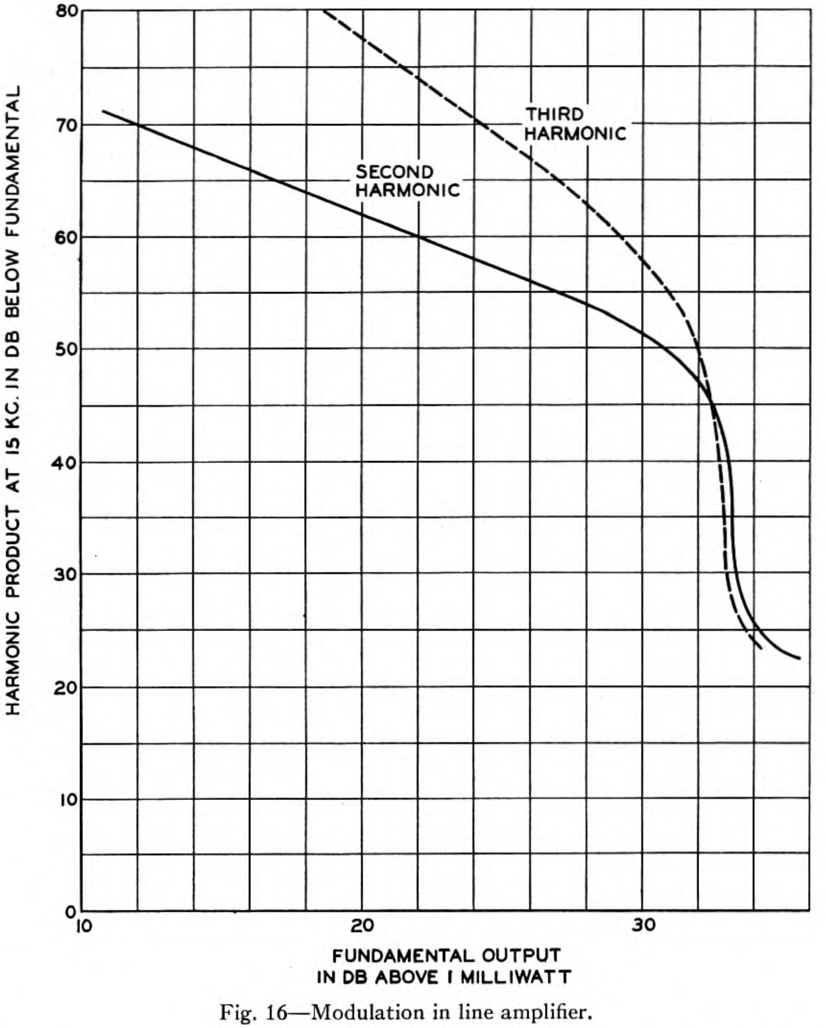


Fig. 16—Modulation in line amplifier.

A gain-load characteristic of the amplifier, with respect to a single-

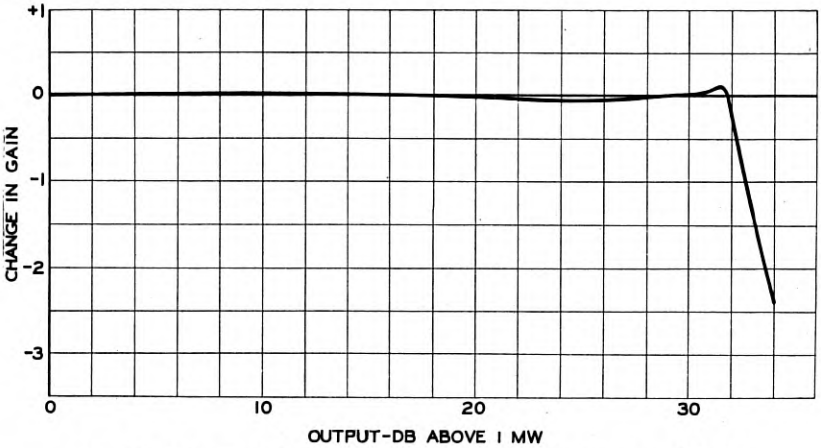


Fig. 17—Gain-load characteristic of amplifier at frequency of 15 kc.

frequency output, is shown in Fig. 17. Translated into terms of three-channel operation this means a permissible level of 18 db above the transmitting toll switchboard without noticeable distortion or interference between channels due to modulation.

POWER SUPPLY

The new system is designed to operate at both terminals and repeaters on standard telephone office battery supply of 24 and 130 volts. A dry-cell battery is required to supply grid bias to the output tube of each amplifier. A small amount of 110-volt a-c power is also required to drive the Telechron motor in the regulator circuit. The total steady power consumption in a terminal is 88 watts and in a repeater is 56 watts. These figures compare with 220 watts and 164 watts, respectively, in the older terminal and repeater.

Where regulated 24-volt battery is not available, tubes having a slightly greater current consumption are used in combination with ballast lamps. In this case the power used will be somewhat greater.

Provision has been made for a separate a-c power conversion unit to be used with the system in offices where the usual d-c voltage is not available. This should prove of great value where the provision of a battery reserve is not warranted.

SIGNALING

Standard voice-frequency signaling equipment can be used with the new terminals. There is also available a new type of ringer circuit in which a single tube functions both as the source of power for an outgoing signal and as a detector for an incoming signal. This circuit employs heater type tubes and will operate from the a-c power conversion unit mentioned above. Since it also obviates the need for a 1000-cycle generator as a source of signaling current it is particularly well adapted to the small office type of installation.

Equipment Features

As mentioned before the new terminal is much more compact than those previously used. Formerly 2½ bays of standard size were required for the terminal proper and an additional bay for the automatic regulating equipment. Now a complete terminal including the regulating equipment can be mounted in one such bay with some space left for miscellaneous equipment.

The same degree of compactness has been applied to the new repeater. A bay of standard size was formerly required for the repeater proper with another bay for the automatic regulating equipment when provided. Both are now provided in one bay and, as in the terminal, some space is available for mounting other equipment.

The assembly of the equipment panels of the carrier terminal and repeater generally follows conventional practices, the repeating coils, condensers, vacuum tubes, etc., being mounted on the front of steel panels with the electrical terminals projecting through and the wiring placed on the rear. The filters are in sealed cans with soldering terminals brought out on the rear for wiring connections.

In view of the wide field of use anticipated for the new system, somewhat more than the usual flexibility of assembly and arrangement of parts has been provided. In small terminal offices, that is, offices having one or two systems, the four-wire terminating sets and associated patching jacks and the carrier line equipment and associated jacks may all be in one bay, using for this purpose the miscellaneous bay space referred to above. Similarly, the line filter equipment may be mounted on the repeater bay.

In the larger terminal offices, in order to facilitate operation and maintenance, the four-wire voice-frequency patching jacks can be located in a central patching bay with similar jacks from other carrier channels. Testing and monitoring equipment provided at such a point will, therefore, be common to many circuits. In the same manner the carrier frequency patching jacks of a large number of terminals or repeaters in an office can be grouped at a central point.

Line Considerations

Since the new system occupies practically the same frequency range as its predecessor, it can be applied to open-wire routes in very much the same manner. Wire sizes of 104 mil, 128 mil and 165 mil are commonly used in the Bell System plant, the particular size often being governed by mechanical rather than by transmission considerations. These lines are carried into the terminal and repeater offices through cables which are usually loaded to maintain smooth impedance relations and reduce the transmission losses.

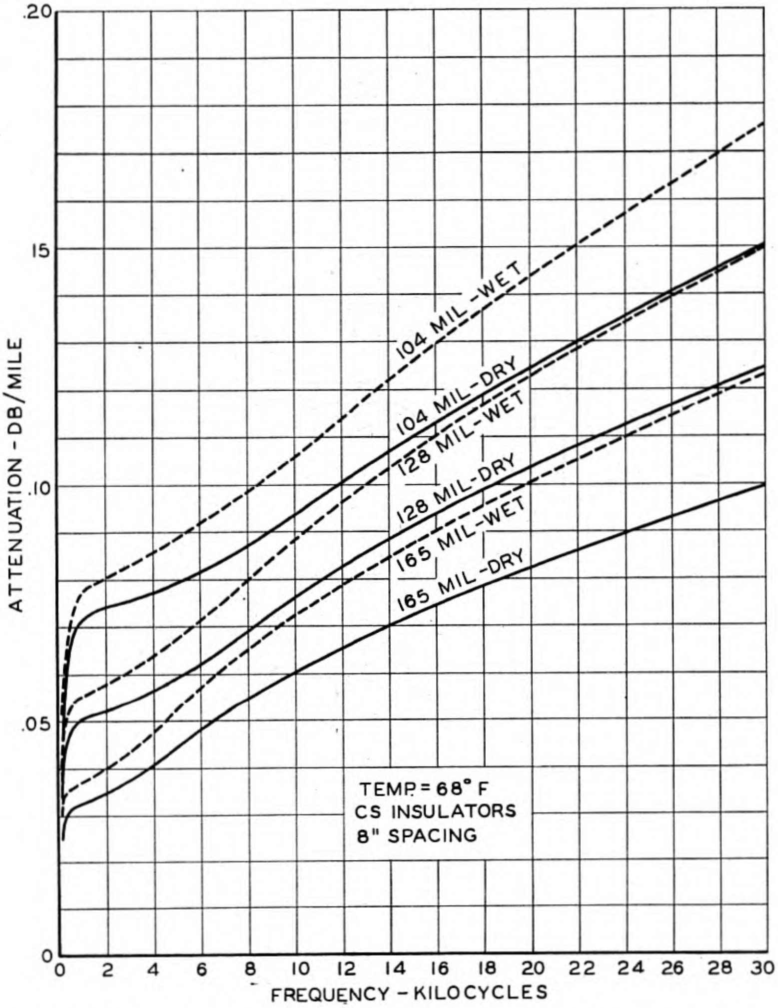


Fig. 18—Attenuation characteristics of 104, 128 and 165 mil open-wire lines.

The control of crosstalk[7] is one of the major problems in the application of carrier to open-wire lines. On short lines where only a

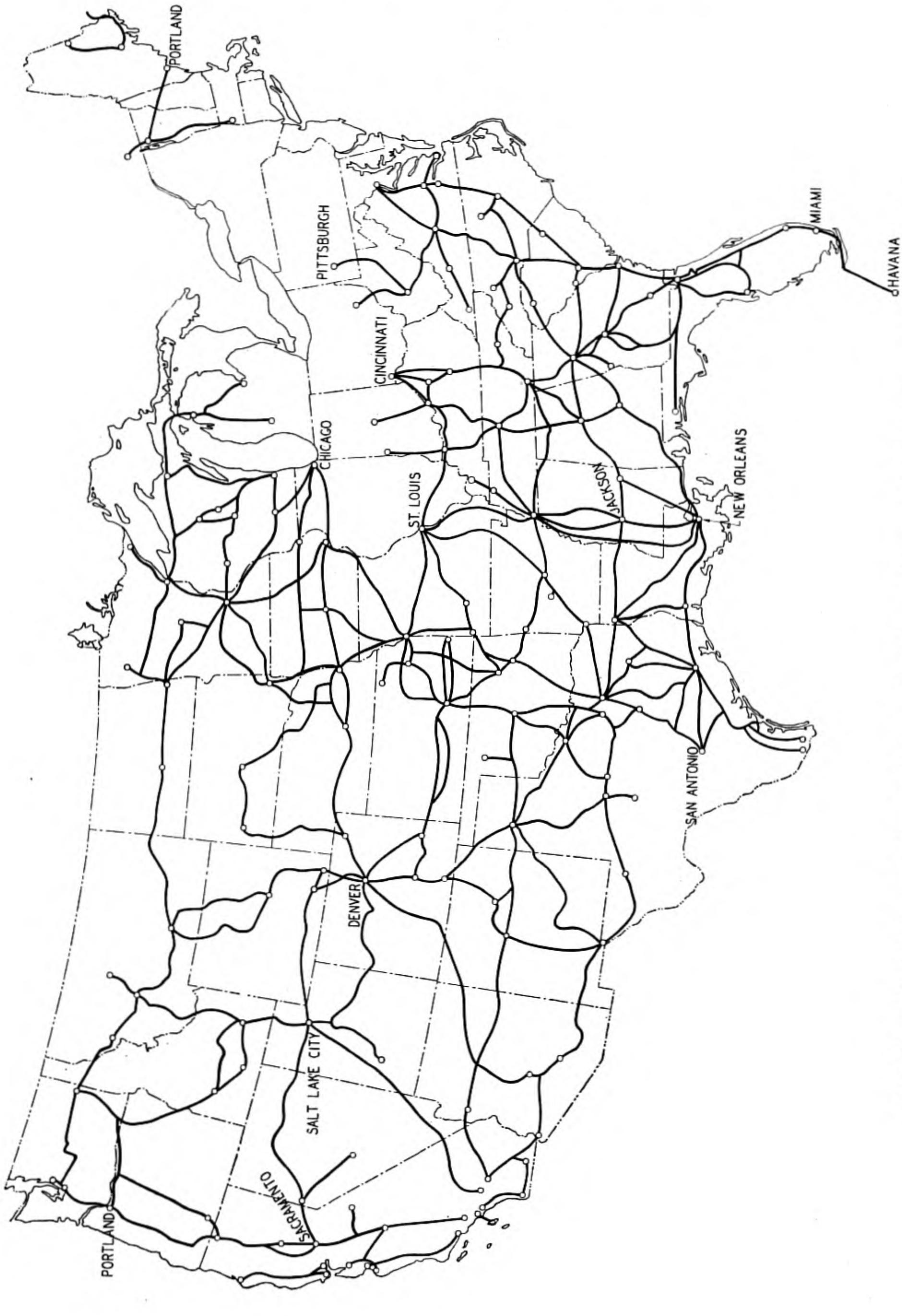


Fig. 19—**Bell Telephone System routes on which Type C carrier telephone systems are operated.**

few systems are involved, the engineering solution may be quite different from that on a long line on which it is desired to operate many systems. Several different plans for transposing wires have been developed which can be applied with various pole line configurations so as to meet the necessary requirements in any practical case. The more recently constructed carrier lines have, in general, employed a spacing of 8 inches between the wires of a pair, with from 16 to 26 inches between horizontally adjacent pairs. Besides contributing to the crosstalk reduction, the closer spacing is also less susceptible to interference from outside sources, which is an advantage from a noise standpoint.

Attenuation characteristics[8] for typical eight-inch spaced pairs using the CS type of insulator are given in Fig. 18. Normal or dry weather characteristics are shown, together with the losses that are assumed for ordinary wet weather. Temperature changes also result in sizable transmission changes. It is also important to note that the losses when the wires are coated with sleet or frost may go far beyond those indicated for the wet weather condition.

Conclusion

The widespread use that has been made of the Type C system up to this time is pictured in Fig. 19, which shows the routes in the Bell System over which systems are now operating. The new design with its lower costs, improved performance and greater flexibility should find increased application not only on these routes but on shorter lines on which the system has hitherto not been economical.

References

1. "Carrier Systems on Long Distance Telephone Lines," H. A. Affel, C. S. Demarest and C. W. Green, *Bell Sys. Tech. Jour.*, volume 7, July 1928, pages 564–629; *A.I.E.E. Transactions*, volume 47, October 1928, pages 1360–1386.
2. "Cable Carrier Telephone Terminals," R. W. Chesnut, L. M. Ilgenfritz and A. Kenner, *Bell Sys. Tech. Jour.*, volume 17, Jan. 1938, pages 106–124; *Electrical Engineering* (*A.I.E.E. Transactions*), volume 57, May 1938, pages 237–244.
3. "Carrier Telephone System for Short Toll Circuits," H. S. Black, M. L. Almquist and L. M. Ilgenfritz, *Electrical Engineering* (*A.I.E.E. Transactions*), volume 48, January 1929, pages 117–140.
4. "A New Single-Channel Carrier Telephone System," H. J. Fisher, M. L. Almquist and R. H. Mills, *Bell Sys. Tech. Jour.*, volume 17, January 1938, pages 162–183; *Electrical Engineering* (*A.I.E.E. Transactions*), volume 57, January 1938, pages 25–33.
5. "Copper-Oxide Modulators," R. S. Caruthers, Presented at Winter Convention of A.I.E.E., January 1939.
6. "Stabilized Feedback Amplifiers," H. S. Black, *Bell Sys. Tech. Jour.*, volume 13, January 1934, pages 1–18; *Electrical Engineering*, volume 53, January 1934, pages 114–120.
7. "Open-Wire Crosstalk," A. G. Chapman, *Bell Sys. Tech. Jour.*, volume 13, January 1934, pages 19–58; April 1934, pages 195–238.
8. "The Transmission Characteristics of Open-Wire Telephone Lines," E. I. Green, *Bell Sys. Tech. Jour.*, volume 9, October 1930, pages 730–759; *Electrical Engineering* (*A.I.E.E. Transactions*), volume 49, October 1930, pages 1524–1535.

Crossbar Dial Telephone Switching System *

By F. J. SCUDDER and J. N. REYNOLDS

This paper describes the crossbar dial telephone switching system recently adopted by the Bell System for large cities where the panel system has been used for nearly twenty years. Central offices of the crossbar type can be introduced in panel areas without changes in existing offices and without changes in existing dial telephone instruments. Crossbar offices and panel offices in the same building will operate on a common power plant and utilize other equipment in common, such as "A" and "B" operator switchboards and outgoing trunks.

The precious metal contact crossbar switches are used for all switching purposes as contrasted with the base metal contact panel switches. The switches operate with relay-like movements under control of common control or marker circuits consisting primarily of multi-contact, *U* and *Y* type relays. The control and marker circuits, which are connected to the switch frames by means of multi-contact relays, perform their operations in a fraction of a second. The switches, the *U* and *Y* type relays and the multi-contact relays are equipped with twin contacts of precious metal. Senders similar to those of the panel system are employed.

The system will be used for new offices in larger cities as manufacturing and plant conditions permit.

INTRODUCTION

IT is the purpose of this paper to describe briefly the crossbar dial telephone central office switching system which has recently been developed by the Bell System for use in large cities. Sixteen years ago, in February 1923, a paper was read before the Institute, by Messrs. E. B. Craft, L. F. Morehouse and H. P. Charlesworth of the Bell System which outlined the history and the problems involved in telephone central office switching and described the panel dial central office system which had just been developed and was being introduced in the large cities. The first central office of this type was placed in service in December 1921, and since that time 456 panel dial offices serving nearly four and one-half million subscriber stations have been installed in 26 different cities throughout the country. During these years many improvements have been made in the panel system to make it more serviceable to the telephone public and to meet the new

---

* Presented at Winter Convention of A.I.E.E., New York, N. Y., January 23–27, 1939.

problems which have arisen, but in addition the engineers of the Bell System have continued their search to find new and better means for meeting telephone switching demands. This work has resulted in the adoption of the crossbar type central office switching equipment. Two offices of this type were placed in successful operation during 1938 and others are in process of manufacture and installation.

It will be appreciated that for large metropolitan areas, the development and economic introduction of a central office switching system which differs materially from the existing systems is a rather large undertaking. The system must fit into the existing plant as a whole without material change. Generally any important changes affecting the subscribers' use of the telephone or the methods used by switchboard operators should be avoided. Existing numbering plans should not be affected, existing classes of service should be continued, and the addition of others made feasible in case they should be required.

All of these and many other factors have been taken into account and all requirements have been met by the crossbar system which offers important improvements in telephone switching, both in operation and maintenance. Its introduction does not make any of the existing equipments obsolete in the sense that these equipments will be less serviceable nor will it cause their replacement. Central offices of the crossbar type can be installed in the same building with existing penel central offices without loss in operating economies in either type of office. Certain equipment, such as the existing and additional outgoing trunks to other central offices, manually operated switchboard positions, operating room and maintenance desks, power plant and alarms, can be used in common by the two types of offices in the same building.

GENERAL

Before describing the crossbar system it is desirable first to give a brief outline of the principal functions of a dial central office equipment. Such an office is capable of serving 10,000 subscriber line numbers, and is provided with a sufficient number of connecting switches, trunks and associated circuits so that under usual peak loads of traffic, calls will be completed promptly.

The central office circuits, in response to the lifting of the receiver by the calling subscriber, connect the subscriber line to the switching equipment. This equipment then extends the calling line, "link by link," through several switching stages to the called line as determined by the called office code and line number dialed by the calling subscriber. When the connection has been established to the called line, the subscriber bell is rung and, when the subscriber answers, the talking connection is completed. During the conversational period the connection is held under control of the calling telephone, and when the telephone receivers are replaced, the central office equipment and the telephones are released for use on other calls. The equipment, of course, transmits the busy tone signal to the calling subscriber if the called line is found busy, and automatically routes a call for a discontinued or an unassigned line to an operator who informs the subscriber of the status of such a line.

Operators and associated switchboards are provided in the dial system to handle certain classes of calls and to render assistance to subscribers when required. Calls to these operators are established in response to the dialing of operators' codes in a manner similar to the establishment of calls to other subscribers.

Operators are usually provided to complete calls terminating in a dial office which are originated by subscribers served by manual offices.

Prior to the introduction of the crossbar system, the Bell System employed two general types of dial central offices. These are the well known step-by-step and panel systems.

The step-by-step system has been used generally in the smaller cities which are frequently served by a single central office or by a relatively small number of offices and where the trunking problems are consequently less complicated. The switches of the step-by-step system are controlled directly by the impulses from the subscriber dials and, necessarily in conformity with the dial, the system operates on a decimal basis. The selectors of this system are first moved under control of the dial to any one of ten vertical positions, corresponding to the numeral of the digit dialed, and in the case of trunk hunting switches is then automatically rotated over a row of ten trunk terminals to find an idle trunk during the interdigital time of the dial.

The step-by-step switch thus has access to ten different groups of trunk terminals with ten terminals each. The location of the trunk groups on the switches is governed by the digits dialed and consequently the relocation of a group necessitates directory changes. These limitations in trunk access and flexibility are not material handicaps in the smaller cities throughout the country where the system is giving excellent service.

The panel system meets the complex service requirements of the larger cities with their large volume of traffic and multiplicity of central offices. In these cities the number of trunk groups is large and the number of trunks in the groups varies widely. Further, the number of groups and their sizes are frequently changed by the introduction of new offices and changes in the character or extent of existing central office areas.

In the panel system, senders are provided which record and store the dial pulses as they are dialed and then independently control the operation of the switching units. The large panel type switches provide access to large groups of trunks and to a large number of groups, and at the same time permit considerable variation in the sizes of the groups. The necessary flexibility in the size and location of the trunk groups is obtained by flexibly wired routing equipment provided in decoder circuits which are associated with the senders. These facilities permit trunk group locations on the switches as dictated by traffic regardless of the office codes listed in the directories and dialed by the subscribers. The panel system also readily provides for the routing of calls through intermediate or tandem offices where the traffic between offices can be more economically handled in this manner.

The crossbar system also makes use of the sender and decoder method of operation and provides a still greater flexibility in the trunking arrangements than is obtained by the panel system.

THE CROSSBAR SYSTEM

The two outstanding features of the crossbar system are the "crossbar switch" which is used for all major switching operations, and the "marker" system of control which is used in the establishment of all connections throughout the crossbar office.

The crossbar system is essentially a relay system employing simple forms of relays and relay type structures for all switching operations. The apparatus consists almost wholly of crossbar switches, multi-contact relays and the usual small relays similar to those generally employed in all telephone systems. The switching circuits are wired to the contacting springs of the switches, and the connections through the switches are made by pressing contacts together by means of simple electromagnetic structures instead of the moving brushes and associated fixed bank terminals of other systems.

The use of relay type apparatus with its small, pressure type contact surfaces economically permits the use of twin or double contacts with thin layers of precious metal for all contact points. Obviously, double precious metal contacts make for reliable operation, especially with the low speech and signaling currents inherent to a telephone system.

The short mechanical movements and the inherently small operating time intervals of the "relay-like" crossbar switch permit the use of common circuits or "markers" to control the operation of the switches. This has permitted the use of large assemblies of switches and associated relays on unit frames which can be wired and completely tested for operation in the factory before the units are shipped.

In the design of the switching frames and associated control circuits, one of the objectives realized has been the standardization of a relatively small number of different types of equipment units, thereby simplifying manufacture and merchandizing. This also simplifies the engineering of the equipment by the Telephone Companies in the preparation of their specifications to meet the particular traffic requirements of the various central offices.

The marker system used for controlling the switching operations has many advantages, the more important of which will be disclosed later in the general description of the operation of the equipment. It might be mentioned here, however, that the marker is an equipment unit consisting almost entirely of relays, which completes its functional operations in the establishment of a call in a fraction of a second. This short operating time permits a few markers to handle the entire traffic in the largest office. The markers are connected momentarily by means of multi-contact relays to the various switching units of the office to control the establishment of the calls through the crossbar switches.

An outstanding advantage of the marker system of control is the "second trial" feature, by means of which two or more attempts can be made to establish a call over alternate switches and trunks when the normally used paths are all busy. The markers are arranged to detect short-circuited, crossed, grounded and open-circuit conditions at all vital points, and before releasing from a connection they make circuit checks to insure that the connection has been properly established. When trouble conditions are detected, they make a second attempt to complete the connection, after sounding an alarm and recording the location and nature of the trouble encountered. The marker system facilitates the introduction of new service features and changes in operation, which may be found desirable from time to time, due to the fact that the principal controlling features of the entire system are vested in a small number of markers.

APPARATUS

Crossbar Switch

The crossbar switch from which the system derives its name is the basic switching unit of the system. Figure 1 shows the front view of a 200-point crossbar switch.

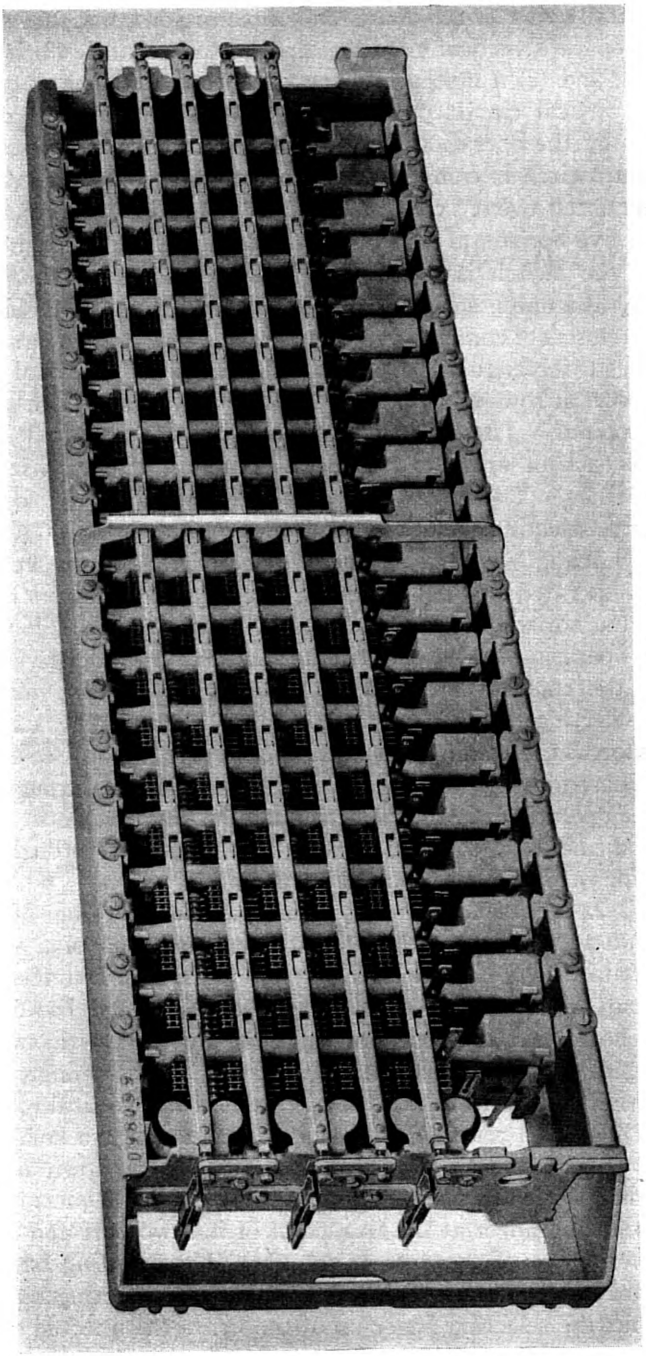


Fig. 1—Crossbar switch—front view.

Fundamentally this switch consists of three major functional parts: (*a*) twenty separate vertical circuit paths, (*b*) ten separate horizontal circuit paths, and (*c*) a mechanical means for connecting any one of the twenty vertical circuit paths to any one of the ten horizontal circuit paths by the operation of electromagnets. From a structural viewpoint the switch is comprised of a rectangular welded frame on which are mounted twenty vertical units and the selecting mechanism consisting of five horizontal bars operated by ten selecting magnets.

Primarily the switch is a multiple relay structure with twenty vertical relay-like units, each unit having an operating or "holding" magnet and ten sets of contacts in a vertical row. The switch arrangement provides a rectangular field of contacts in twenty vertical rows and ten horizontal rows or a total of 200 sets of contacts, one set at each "crosspoint." These crosspoint contacts are operated independently of each other by a coordinate operation of the horizontal and vertical bars. The horizontal bars are controlled by the ten horizontal or "selecting" magnets and the vertical bars by twenty vertical or "holding" magnets. Any set of contacts in any vertical row may be operated by first operating the selecting magnet corresponding to the horizontal row in which the set of contacts is located, and then by operating the holding magnet associated with the vertical row. Since the contacts are held operated by the holding magnet alone, the selecting magnet is operated but momentarily and is released as soon as the holding magnet is operated. After the selecting magnet is released, other connections may be established through the switch by the operation of other selecting and holding magnets. It is thus apparent that ten connections can be established through the switch, one for each of the horizontal paths.

From Fig. 2 the rather simple mechanical interlocking of the horizontal and vertical bars which causes the operation of a set of crosspoint contacts will be understood. The ten sets of contacts in a vertical row are associated with the vertical or "holding" bar of the row. Each horizontal or "selecting" bar is provided with twenty selecting fingers which are made of flexible wire. These fingers are mounted at right angles to the bar, one at each of the vertical rows of contacts. Thus when a selecting bar is rotated through a small arc by its magnet, the selecting fingers will move up or down into a position so that when a holding bar is operated by its magnet, it will engage the selecting finger at the crosspoint of the two bars and cause the corresponding set of contacts to operate. The selecting bar and the fingers not used will then be released when the selecting magnet is released, but the selecting finger used to operate the selected set of crosspoint contacts will remain latched and the contacts held closed by the holding bar until the holding magnet is released at the end of the connection. The selecting fingers are each provided with a damping spring to reduce vibration on the operation and release of the fingers.

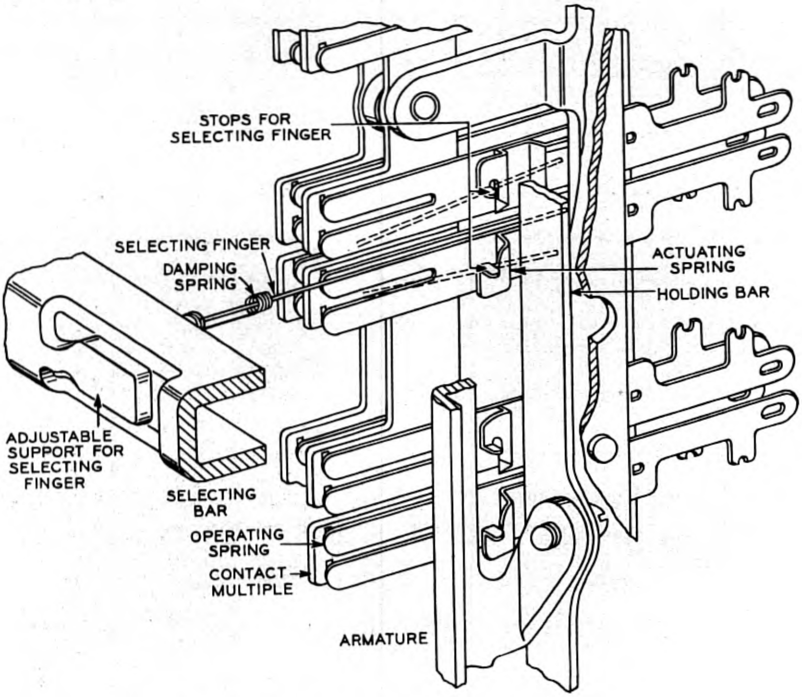


Fig. 2—Crossbar switch selecting mechanism.

It will be noted that the selection operation is performed by five horizontal bars although there are ten horizontal rows of contacts. This is accomplished by operating the bars in either of two directions. As shown in Fig. 1, two magnets are associated with each bar, one whose armature is on top of the bar, the operation of which causes the selecting fingers to move in a downward direction, and the other whose armature is below the bar causing the fingers to move upward. The selecting bars are restored to the normal or mid-position by the centering springs located on the end of the switch adjacent to the magnets.

Figure 3 shows the vertical unit of the crossbar switch with its ten sets of normally open "make" type contact springs, the holding magnet at the bottom, and the long vertical armature to which is attached

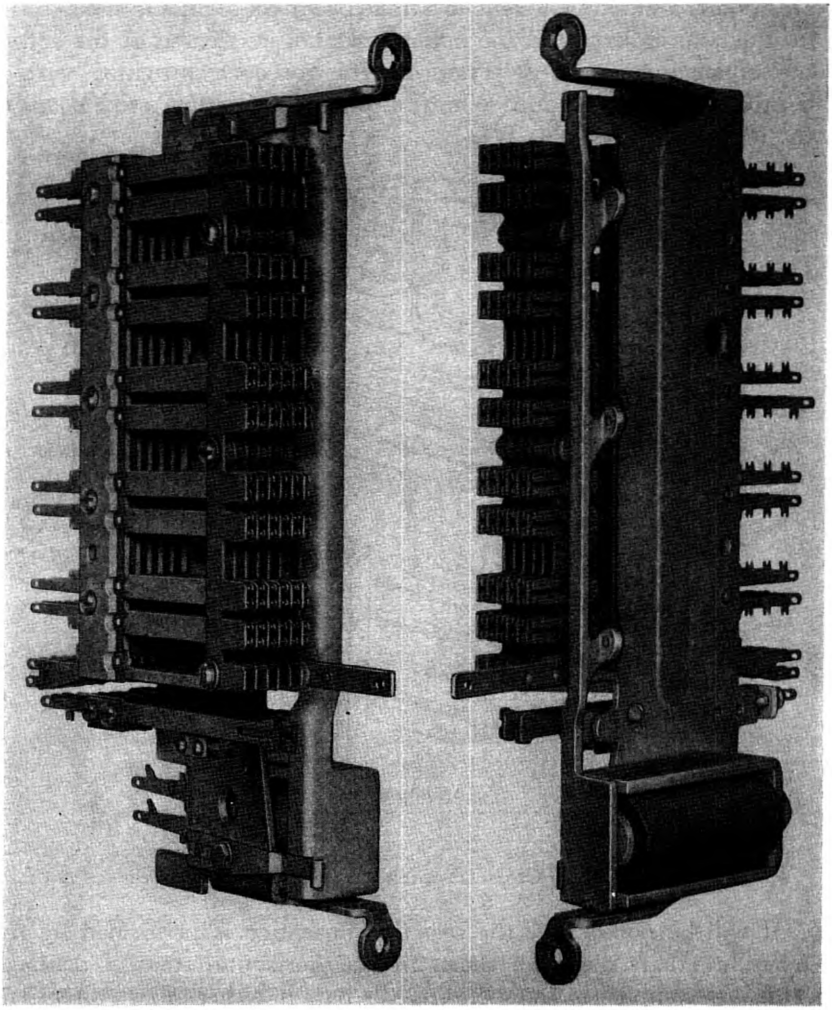


Fig. 3—Crossbar switch vertical unit.

the vertical holding bar. The vertical unit shown has six pairs of contacts at each of its ten crosspoints. Other vertical units are provided with ten sets of 3, 4 and 5 pairs of contacts per set. One spring of each pair as shown is a fixed spring consisting of a projection of an insulated vertical metal strip, made in the shape of a comb. This strip extends from the top to the bottom set of contacts of a vertical row. Wiring lugs are provided at the lower end of these vertical strips facing the rear to which are wired the lines or trunks of the vertical circuit paths. At the lower end of these strips and facing the front is another projection used by the maintenance force for testing purposes. The mate or movable spring of each pair is individually insulated from all other springs. These springs extend to the rear of the switch for wiring purposes and may be strapped horizontally to the corresponding springs of adjoining vertical units to extend the horizontal circuit path through the switch.

The contacting ends of the thin movable contacting springs are bifurcated to provide two flexible contacts in parallel. The contacting surfaces on these springs as well as the mating fixed springs are provided with a thin layer of palladium. The use of the double precious metal contacts is an important feature of the crossbar system in providing more reliable contacting surfaces. Experience has shown that the chance of simultaneous failures of both contacts of a pair is extremely small. The actual contacting surfaces of each pair of springs consist of small bars of contact metal located at right angles to each other. These bars are composed of a ribbon of nickel capped with a thin layer of palladium. This crossbar arrangement of contacts provides a rather large area over which the two springs can make contact with each other, and thereby permits considerable tolerance in the manufacture and adjustment of the contact spring assemblies.

The switch may be equipped with "off normal" contact spring assemblies. When these are furnished they are associated with each selecting or holding magnet and are operated like relay contacts when the associated magnet operates, regardless of which crosspoint contact is closed. They are used to perform circuit functions as required in the various uses of the switch.

In the design of the switch special attention was given to the problem of wiring and cabling. Figure 4 shows the wiring terminals on the rear of the switch. These terminals are arranged for individual wiring and also have staggered, notched projections so that the terminals can be readily strapped together horizontally with bare wire as shown. This is an important feature of the switch since it permits a multiple of terminals to be easily soldered together and reduces the wire congestion on the switch.

The 200-point crossbar switch is 9¼ inches in height and 30½ inches in length. In addition a 100-point switch 20½ inches in length is provided. This switch is similar to the 200-point switch but is equipped with 10 vertical units.

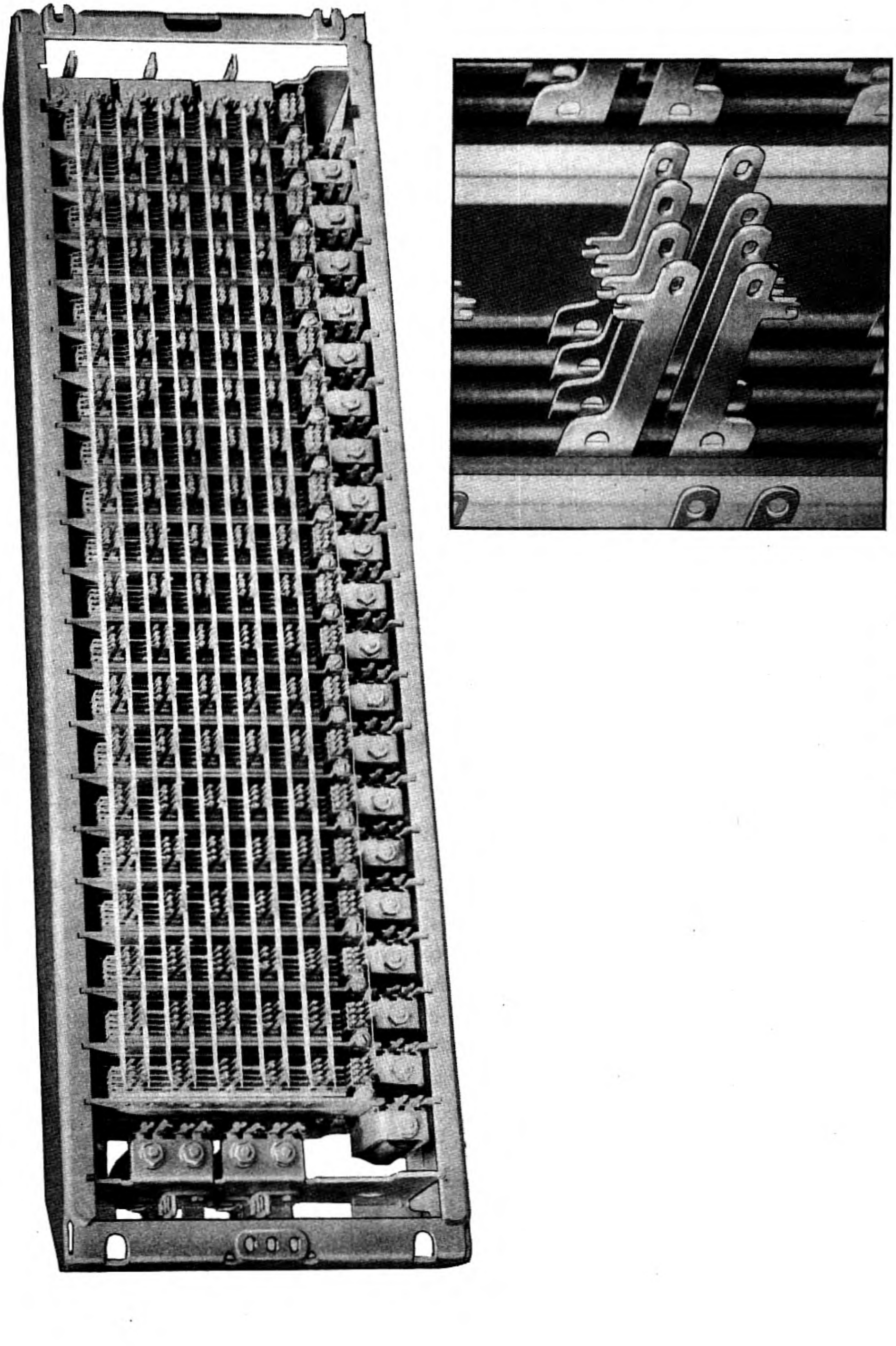


Fig. 4—Crossbar switch—rear view.

Multi-Contact Relay

The multi-contact relay used in the crossbar system is shown in Fig. 5. It resembles in design the vertical unit of a crossbar switch.

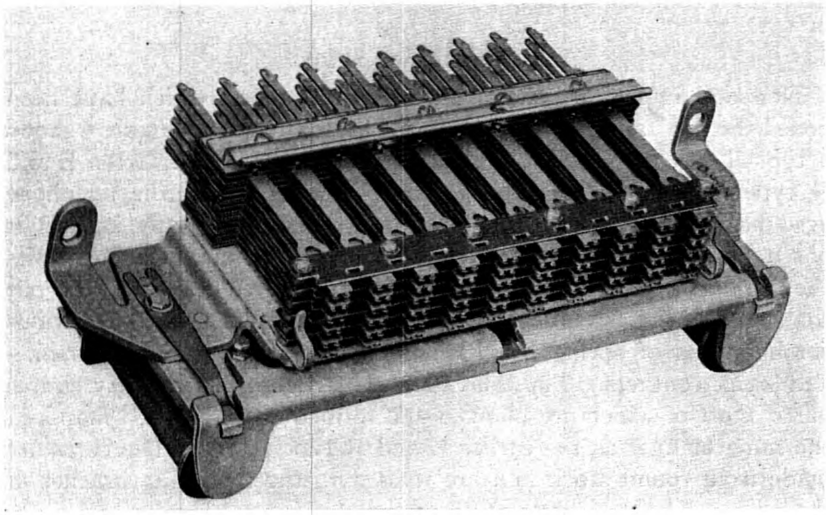


Fig. 5—Multi-contact relay.

The relay is provided in four sizes in respect to the number of contacts, namely, in 30, 40, 50 and 60 sets of individually insulated contacts, all of which are of the normally open type which are closed when the magnets of the relay are operated. Each relay is provided with two separate magnets, armatures and associated groups of springs, and both magnets are energized in parallel in order to close all of the contacts. By operating the two magnets independently the structures can be used as two separate relays, each equipped with 15, 20, 25 or 30 sets of contacts. The relay occupies a mounting space approximately 2″ x 11″ and is provided with a cover.

All contact springs are equipped with twin contacting surfaces similar to the contacts used on the crossbar switch except that they are composed of solid bars of precious metal due to the heavy duty requirements. To facilitate wiring, these relays are manufactured with two types of wiring terminals. In one type the movable springs are of graduated lengths and are provided with notched lugs for bare wire strapping to permit the multipling of springs horizontally to corresponding springs on other relays mounted adjacent. In the second type the strapping lugs are omitted and all springs are of the same length and are provided with soldering eyelets for individual or non-multiple wiring.

The multi-contact relay finds its chief use in the common connector circuits where a large number of leads must be connected simultaneously to a common circuit.

U and Y Type Relays

New and improved general purpose small relays which have been coded the "U" and "Y" type are used in this system. Figure 6 shows a typical "U" type relay. Although somewhat similar to the E and R type relays which have been in a common use in the telephone systems for many years, it differs from them principally in that it has a heavier and more efficient magnetic structure which permits the use of a greater number of contact springs. These relays permit the use of spring assemblies up to a maximum of 24 springs in various combinations of springs, including transfer contacts, simple make-and-break contacts. The relays are constructed of relatively simple parts, most of which are blanked and formed in the desired shapes in the same manner as the earlier E and R type relays. The cores are made from round stock and are welded to the mounting bracket of the relay. The structures of all of these relays are similar and differ principally in their spring assemblies and windings.

In order to insure more reliable contact closures, the relays are equipped with twin contacts. Various types of contact metal and sizes of contacts are provided, depending upon the characteristics of the circuit controlled by the contacts.

Improved methods of clamping the springs in their assemblies, together with the design of the springs, provide stability and minimize manufacturing and maintenance adjusting effort.

Contacts practically free from chatter on both the operation and release of the relay have been obtained by the use of relatively heavy stationary springs, short thin movable springs, and a pivoted arrangement of the armature suspension. By reference to Fig. 6 it will be seen that the rear ends of the armature are pivoted by two pins which project through holes in the hinge bracket mounted on the rear spring assembly. In the earlier flat type relays of the E and R type, the armature was suspended at the rear by means of a reed type armature hinge.

The Y type relays make use of the same manufacturing tools and processes as the U type. Copper or aluminum sleeves are provided over the cores beneath the windings to secure the slow-release characteristics required on these relays. The relay armature is embossed so that when the relay is operated satisfactory contact is made between the metal surfaces of the magnetic circuit which insures uniform time characteristics.

In both the U and Y type relays the cylindrical cores permit the use of form wound coils which are wound on special machines and slipped over the cores when completed. In the manufacture of these coils a removable mandrel is used. It is covered with a layer of sheet cellulose acetate and accommodates several coils. These coils are then automatically wound on the mandrel from different spools of insulated wire. Separations are left between adjacent coils so that when the winding operation has been completed the individual coils can be separated. A very thin sheet of cellulose acetate is automatically interleaved between successive layers of wire to hold the wire in place and to provide insulation between layers. This general method of winding coils also is used for the magnets of the crossbar switches and multi-contact relays.

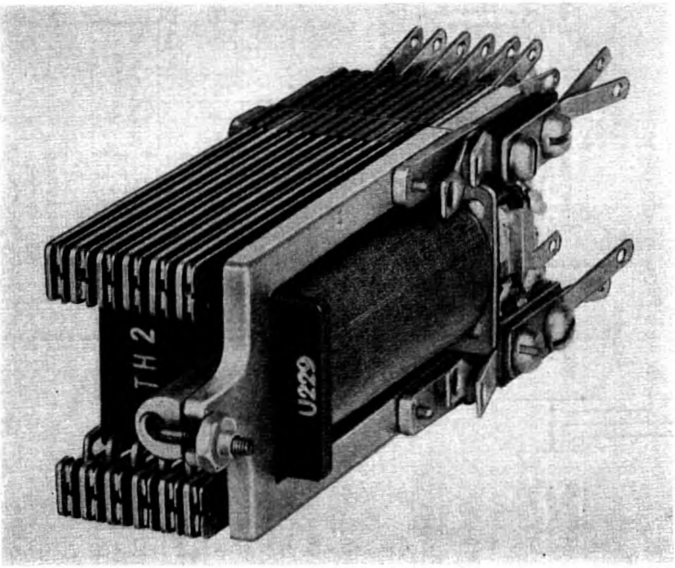


Fig. 6—"U" type relay.

Functions of the Equipment Units

The general operation of the system as a whole may be more easily understood by first describing the principal equipment units in the system and their functions before proceeding with a description of the

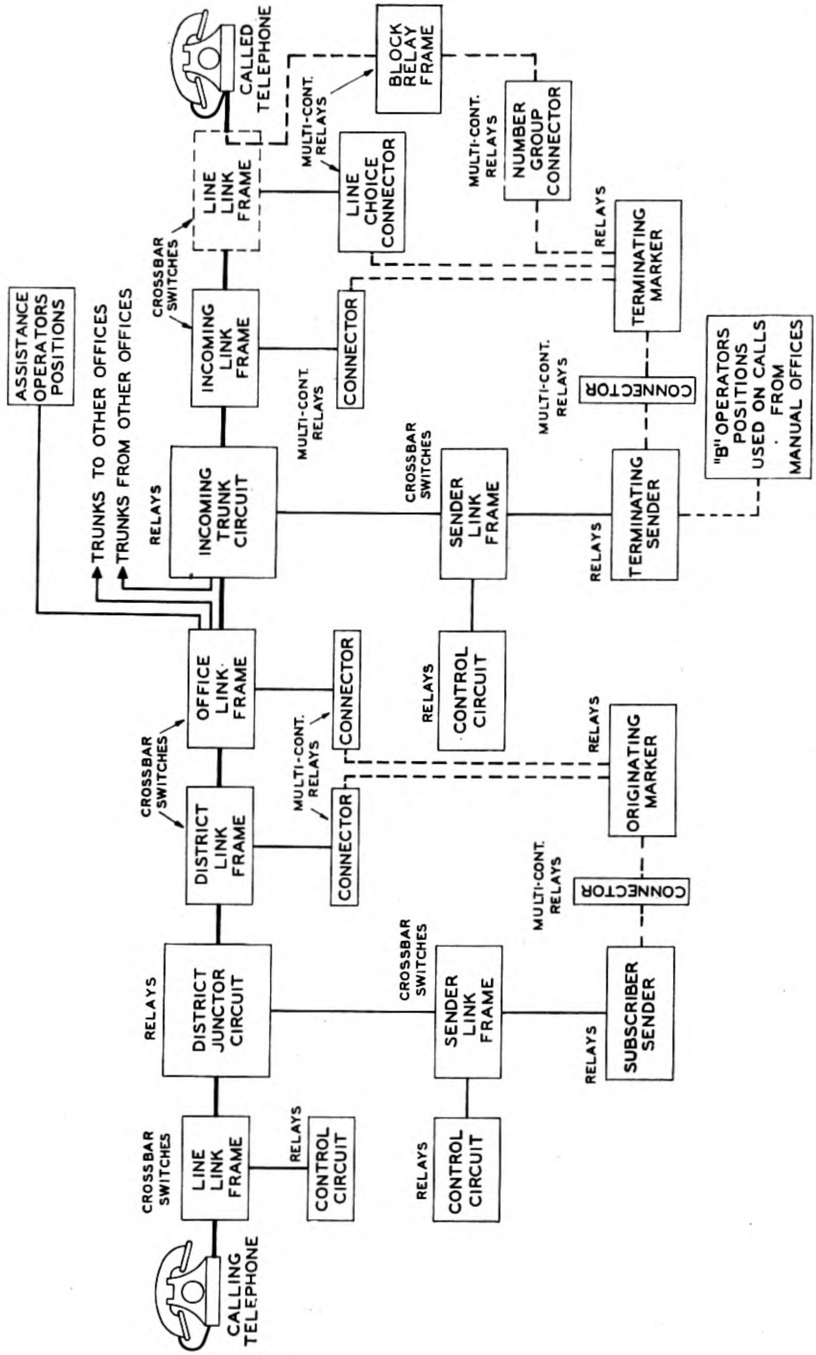


Fig. 7—Functional arrangement of equipment units.

operation of the circuits. A simplified block diagram of the principal equipment units of the system is shown in Fig. 7. It will be noted that in general there are three types of equipment units:

1. The transmission battery supply and supervisory circuits consisting of the "district junctors" and the "incoming trunks."
2. The crossbar switch frames.
3. The common "control" circuits, the "senders" and the "markers."

The "district junctor" and the "incoming trunk" circuits are composed principally of small relays. The district junctors furnish the talking battery for the calling subscribers and supervise the originating end of connections. The incoming trunks control the ringing of the called subscriber bells, furnish talking battery for the called subscribers, and supervise the terminating end of connections.

The switch frames, which consist almost entirely of crossbar switches, provide the means for switching between the subscriber lines, the district junctors and the incoming trunks. Switch frames also are used for switching the district junctors and the incoming trunks to the senders.

The "senders" consist principally of small relays and their functions are similar to those of the operators at a manual switchboard. The "subscriber senders" register the called numbers from the subscriber dials and transmit the necessary information to the "markers," to the "terminating senders" and to the manual operator positions in manual offices for completing connections to the called lines. The subscriber senders also control the operation of the selectors in distant panel offices. The "terminating senders" in the terminating end of the crossbar office receive the numerical digits of the called numbers from the subscriber senders of any dial office and transmit the required information to the "terminating markers" for setting up the connections to the called lines.

The "markers" are the most important control circuits in the system. They are composed of both small and multi-contact relays. There are two types, one for originating traffic and one for terminating traffic. The operating time of the markers is short, considerably less than one second, and consequently only three or four markers of each type are required in the average office.

The "originating markers" determine the proper trunk routes to the called office. They have access to all outgoing trunk circuits and all the crossbar switch frames that are used for establishing the connections to the called office trunks. They test the trunk group to find an idle trunk to the called office, and also test and find an idle channel through the switch frames, and finally operate the proper selecting and holding magnets of the crossbar switches to establish the connections from the subscriber line to the trunk circuit.

The "terminating markers" perform similar functions in the terminating end of the office to set up the connection from the incoming trunk circuit to the called subscriber line. They have access to all of the subscriber lines terminating in the office, and to all crossbar switch frames used for connecting to subscriber lines. They test the called line to determine whether it is idle, and also test for and find an idle channel through the switch frames and finally operate the proper magnets of the crossbar switches and establish the connection to the called subscriber line.

In addition there are common "control" circuits associated with the "line link" and the "sender link" frames for controlling the operation of the switches on these frames. There are also the common "connector" circuits, consisting mainly of multi-contact relays, which are used for connecting the markers to the senders, to the switch frames and to the test terminals of the called subscriber lines.

It should be noted that the line link frames, although shown separately, are used for both originating and terminating traffic.

After the talking connection has been established between two subscribers, all of the common control units, such as the senders, markers, connectors, line link control circuit, and the sender link frames and their associated control circuits, will have been released, and the talking connection will be maintained in this condition by the holding magnets of the crossbar switches used on the line link, district, office and incoming link switch frames. These switch magnets are held operated under control of the supervisory relays in the district junctor and the incoming trunk circuits and are released when the subscribers replace the receivers.

TRUNKING ARRANGEMENTS

The fundamental method of using the crossbar switch for setting up connections is illustrated in Fig. 8. This figure shows a 200-point crossbar switch with twenty vertical units each wired to a subscriber line and ten trunks strapped horizontally across the switch. With such an arrangement, any one of the twenty lines may be connected to any one of the ten trunks. The number of lines which can be connected to the same ten trunks may be increased to forty by adding a second 200-point crossbar switch with twenty different lines connected to its verticals and by wiring the horizontal contact multiple of this second switch to the horizontal multiple of the switch shown

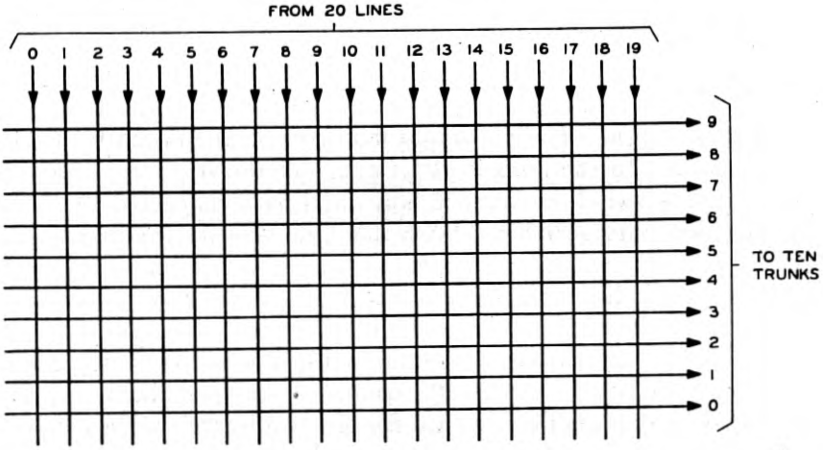


Fig. 8—Simple trunking arrangement with a single 200-point crossbar switch.

in Fig. 8. By adding other switches in this manner, any number of lines may be given access to the ten horizontal trunks.

To obtain greater trunking access, two groups of switches known as "primary" and "secondary" are used. Figure 9 illustrates this primary and secondary switch arrangement as used in the "line link" switch frames and in various forms throughout the crossbar office.

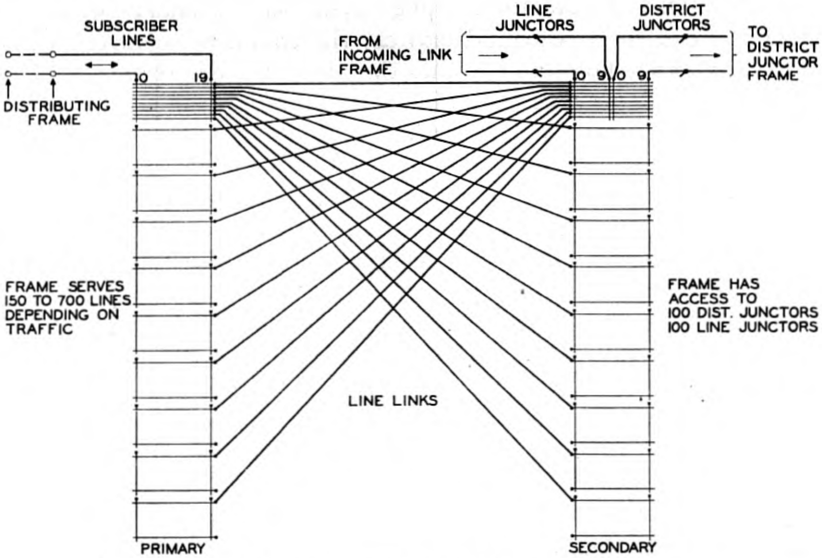


Fig. 9—Primary-secondary trunking arrangement.

The switches are arranged in two vertical files of ten primary switches and ten secondary switches. There are twenty subscriber lines connected to the verticals of each of the ten primary switches and twenty trunk circuits are connected to the twenty verticals on each secondary switch. The horizontal multiples on the primary switches are connected to the horizontal terminals of the secondary switches, each primary switch having one horizontal path connected to each of the ten secondary switches. With this arrangement, the twenty lines of any primary switch have access to all 200 trunks connected to the secondary switches. Since all of the primary switches are wired in this manner, that is, with their ten horizontal paths distributed over the ten secondary switches, then all of the 200 lines on the primary switches have access to the 200 trunks on the secondary switches. It is evident that another vertical file of ten primary switches may be added with twenty subscriber lines connected to the verticals of each switch, and with the horizontal paths strapped and connected to the horizontal paths of the primary switches shown. This would give 400 lines access to the 200 trunks on the secondary switches. In actual practice on a line link frame, several files of primary switches may be connected together in this manner depending upon the traffic volume of the subscriber lines. The circuit paths connecting the horizontal rows of terminals of the primary switches to the horizontal rows of terminals of the secondary switches are called "line links."

To establish a path from a line circuit on a primary switch to a trunk circuit on a secondary switch, the common "control" circuit serving this line link frame locates the subscriber line to be served and then simultaneously selects an idle "line link" on the primary switch on which the subscriber line appears and a group of trunks wired to a secondary switch in which there are one or more idle trunks. Thus the selection of the line link is made contingent upon the availability of trunks, and by means of this together with the primary-secondary distribution of the links a very efficient usage of the links and trunks is obtained.

In the "line link" frame shown in Fig. 9, it will be seen that the trunks on the verticals of the secondary switches are split into groups of 100 trunks each, one group being connected to the "district junctors" and used for originating traffic and the other group of 100 trunks being connected to "line junctors" and used for terminating traffic.

It will be noticed that there is but one crossbar switch appearance of a subscriber line in the office. This is on a vertical unit of a primary crossbar switch where both the originating and terminating calls are

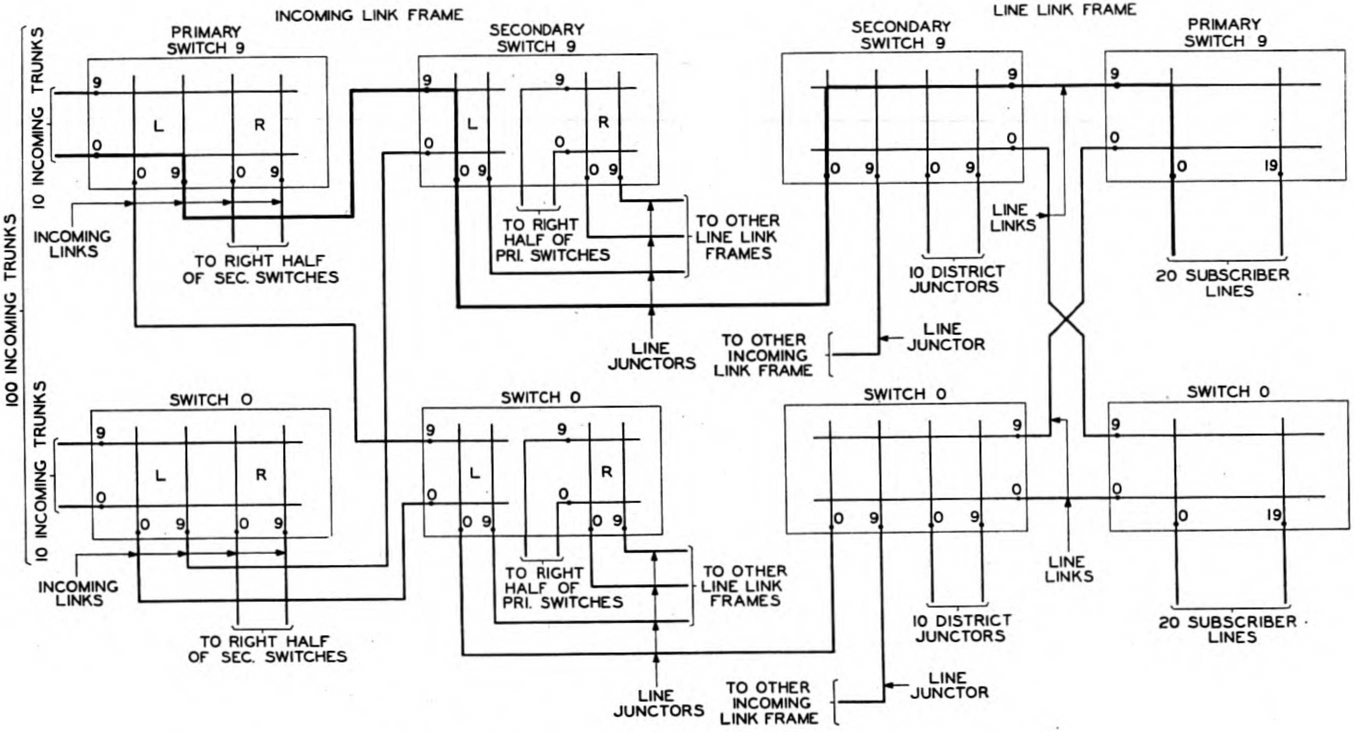


Fig. 10—Double primary-secondary trunking arrangement.

completed by means of the same line link circuits. Thus all originating traffic from any of the twenty lines on any primary switch flows through the associated ten line links to the 100 district junctors and all terminating traffic to these twenty lines flows through the same ten line links from the 100 line junctors.

This single "primary and secondary" trunking arrangement also is used at other points in the system, such as in the originating and terminating sender link switch frames, where the circuits reached are non-directional, that is, where any one of the selectable circuits wired to the frame can be used for setting up a connection.

For the switch frames where the circuits reached are directional, that is, where a particular called line or a particular group of trunks must be used in order to complete a connection, the problem of trunking becomes more complex and it is necessary to provide a trunking arrangement using two "primary and secondary" switch frames arranged in tandem.

Figure 10 shows a typical arrangement of this kind which is necessary to secure the required trunking flexibility and efficiency. This figure shows an "incoming link" frame to which incoming trunks are connected and a "line link" frame to which subscriber lines are connected as described above. These two frames are used in tandem for establishing the terminating connections between the incoming trunks and the called subscriber lines. As is indicated, 100 incoming trunks are connected to the 100 horizontal paths of the ten incoming link frame primary switches, there being ten incoming trunks connected to each of the primary switches. A total of 150 to 700 subscriber lines may appear on the verticals of the primary switches of the line link frame; however, only 200 lines or twenty on the verticals of each of the ten primary switches are shown in the figure.

In order to connect a particular incoming trunk to a particular called line, an idle channel is selected through these two switch frames, consisting of an "incoming link" on the incoming link frame, a "line junctor" between the two frames and a "line link" on the line link frame, and all are connected in series as shown in the figure. It will be noted that the incoming trunks on each of the primary switches have access to twenty incoming links appearing on the twenty verticals of the switch. These twenty incoming links are distributed over the ten secondary switches of the frame, two links being connected to each switch, one to each half switch. It will be observed that in order to provide for the distribution of the twenty incoming links over the ten secondary switches, the horizontal paths of the secondary switches are separated between the tenth and eleventh verticals, thus taking advantage of the flexibility of the crossbar switch by providing twenty horizontal paths instead of ten on each switch. The incoming links, on each half of these secondary switches, have access to "line junctors" appearing on the verticals of these switches. These line junctors are in turn distributed over the secondary switches of all the line link frames in the office. There will be at least one line junctor as shown, from each secondary switch on an incoming link frame to a secondary switch on every line link frame in the office, or a minimum of ten line junctor paths between any incoming link frame and any line link frame. The number of the line junctors between these frames will vary depending upon the number of frames required in an office. The line junctors on the verticals of each of the line link frame secondary switches in turn have access to ten line links on the horizontal paths. These ten line links are, as described above, distributed over the primary switches of the line link frame, one to each primary switch. These line links then have access to the called subscriber lines which appear on the verticals of the primary switches. With this arrangement of switches and the three groups of interconnecting link paths, any incoming trunk can be connected to any called line on the line link frame shown, or by means of other groups of line junctors, to a called line on any other line link frame in the office.

Terminating markers are employed for selecting the paths through these switches to connect an incoming trunk to a called subscriber line. The marker, as will be explained later, records information which permits it to connect to the test wire and holding magnet of a called line and to the test wires and switch magnets of the groups of incoming links, line junctors and line links through which the incoming trunk may be connected to the called line. The marker simultaneously tests these three groups of paths and "marks" an incoming link, a line junctor and a line link which are idle and are accessible to one another, and then operates the switch magnets to connect these three paths and the incoming trunk and the called line together. The paths are selected in an ordered arrangement, so that the lowest numbered incoming links, line junctors and line links are preferred and are used as long as they are available. This increases the efficiency of the paths as compared with a random selection, since it reduces the chance that one or two of them although idle cannot be used because the third one is busy.

A double primary and secondary trunk arrangement similar to the one shown in Fig. 10 is employed for connecting district junctors to outgoing trunks in the originating end of the office.

Brief Description of Circuit Operation

The operation of the system will be described by tracing the progress of a call through the system. The establishment of a call from one crossbar subscriber to another crossbar subscriber may be divided into four stages: two in the originating end of a connection and two in the terminating end.

1. The calling subscriber is connected to a sender for the purpose of registering the called number which is dialed.

2. The subscriber sender is connected to an originating marker and the marker selects the switch frames for establishing the connection to an outgoing trunk.

3. The outgoing trunk circuit is connected to a sender in the terminating end to register the called number.

4. The terminating sender is connected to a terminating marker and the marker selects the switch frames for establishing the connection to the called subscriber line.

The first stage in the progress of a call is illustrated in Fig. 11. It will be seen that the line of a calling subscriber terminates on a vertical unit of a primary crossbar switch located on a line link switch frame. When the subscriber receiver is lifted from the telephone preparatory to dialing, a line relay is operated, as in other systems, and the circuits proceed with the establishment of the connection to an idle subscriber sender which will register the called number when it is dialed.

The circuit functions on this stage of the call are as follows:

1. The subscriber line is located by the "line link control" circuit which is common to the line link frame, by a coordinate method of testing. That is, the control circuit determines the primary crossbar switch in which the line is located and the particular vertical unit in the switch on which the line is terminated. This operation is similar to the line finder operation in other dial systems, except that the operation is accomplished by relay operations instead of by a mechanically traveling brush.

2. The line link control circuit then simultaneously selects an idle line link between the primary switch in which the line appears, and a secondary switch on which a group of district junctors appears which has at least one idle district junctor in the group and which has access to idle senders and an idle sender link.

3. This will bring into operation the common "sender link control" circuit of the sender link switch frame to which the selected group of district junctors is connected. This control circuit will select an idle district junctor in this group which appears on a primary switch on

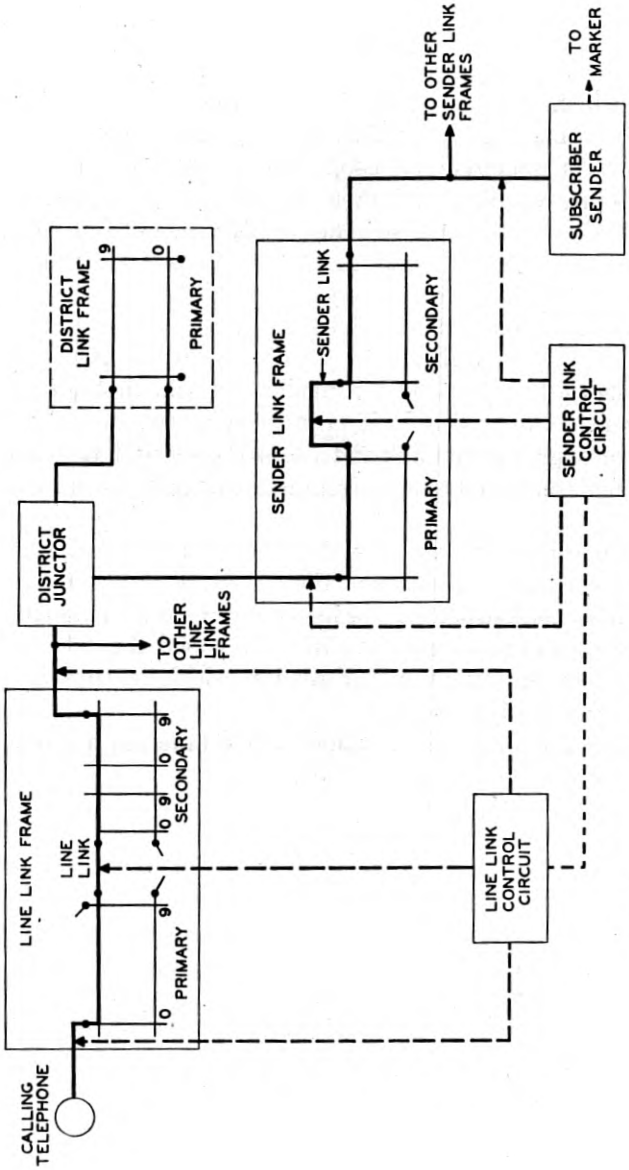


Fig. 11—Calling line connected to district junctor and subscriber sender.

the sender link frame. There are ten sender links serving the selected district junctor. These ten sender links and the ten sender groups to which they have access on the secondary switches are then tested simultaneously to find an idle sender link with access to a group of senders in which there are one or more idle senders. When this choice has been made an idle sender in the group is then selected.

4. The two control circuits in cooperation with each other first operate the selecting magnets and then the holding magnets associated with the paths selected on the switches of both the line link and sender link frames, and thereby establish the connection from the calling subscriber to an idle subscriber sender.

This connection may be traced by referring to Fig. 11, from the calling line on the vertical unit on a primary switch of the line link frame, through a line link and a secondary switch, through a district junctor circuit, to a vertical unit on a primary switch of the sender link frame, through a sender link and a secondary switch to a subscriber sender which is connected to a horizontal circuit path on the secondary switch.

5. The two control circuits are then released and made available for use on other calls. The connections through the switches to the sender are held established by means of the holding magnets which are held operated over a signal control lead, called the "sleeve" lead, under control of the relays in the sender, which in turn are under control of the subscriber telephone.

Upon completion of these operations which take but a fraction of a second the subscriber sender transmits the dial tone to the calling subscriber as an indication to dial the number. When the subscriber dials, electrical impulses are transmitted to the sender, which receives and registers them. When the sender has registered the office code, which in New York City for example is contained in the first three digits dialed, the sender will connect itself to an idle originating marker by means of multi-contact relays of a marker connector circuit.

Before proceeding further it is desirable to mention several other functions of the two common control circuits used for setting up this part of the connection.

1. The control circuits signal to the sender the class of the calling line, that is, for example, whether the line is a coin line or a non-coin line.

2. The sender link control circuit signals to the sender the number of the district link switch frame on which the selected district junctor appears, since this identification will be used later in the establishment of the connection.

3. The sender link control circuit tests the circuit paths chosen from the line circuit to the sender before disconnecting from the connection, in order to insure the proper establishment of the connection. In case of a failure the control circuits will make repeated trials to establish the connection over different paths and give an alarm to the maintenance force.

4. Emergency control circuits are provided for use in case the regular control circuits are removed from service for maintenance reasons.

The next stage in the progress of the call is illustrated in Fig. 12. In this stage of the call the principal control unit is the originating marker. Its major function is to control the switches in the establishment of the connection to an idle outgoing trunk circuit to the called office, which may terminate in a distant office or in the same office as the calling subscriber.

When the subscriber sender connects to the originating marker through the connector circuit, the sender transfers the called office code indication and the district link frame identification to the marker circuit. The called office code indication causes the operation of a "route" relay in the marker corresponding to the particular office called.

There are a number of "route" relays in each marker and one is assigned to each called office routing. The route relay is connected as required by the office code to which it is assigned, so that it will direct the marker to the trunks of the called office and to the office link switch frame on which these trunks appear and indicate the number of trunks in the group. The route relay also is connected to determine the type of the called office, such as Crossbar, Panel, or Manual, and to set up the corresponding circuit conditions in the subscriber sender to enable the sender to handle the connection properly after the marker has been released. The connections of the route relay contacts to the control relays in the marker are made flexible so as to permit the assignment of any route relay to any office code and to permit changes to be made from time to time in the route information, changes in trunk group sizes and location, changes in the type of terminating office, etc. The route relays and associated flexible connection facilities represent a considerable portion of the marker equipment, especially in large metropolitan offices where several hundred central offices are involved.

When the route relay is operated, the marker proceeds with the establishment of the connection as follows:

1. It connects to the office link frame on which the trunks to the

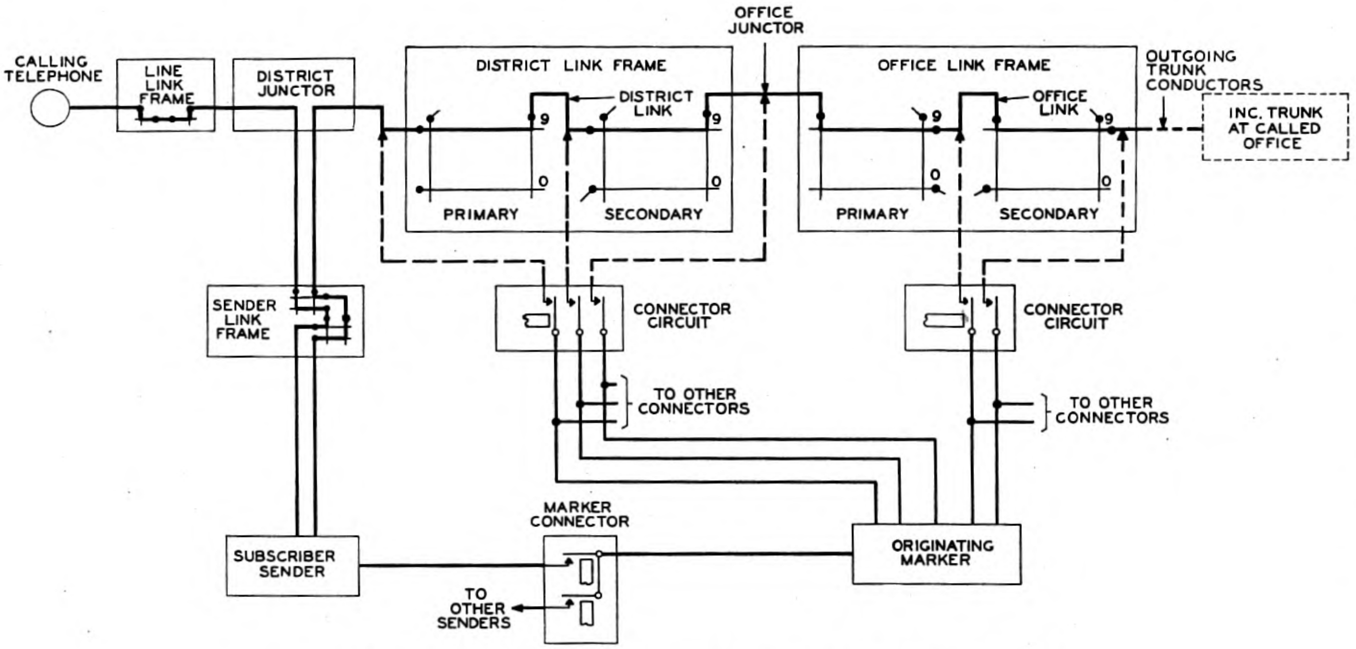


Fig. 12—District junctor and subscriber sender connected to the outgoing trunk.

called office appear. This connection is made through the office link frame connector circuit, one of which is provided for each office link frame. Through this connector the marker is extended to the test leads of any desired trunk group on the office link frame and to the crossbar switches of the frame. When so connected the marker has exclusive control of the trunks and switches of the frame and other markers which desire connection to the same frame are deprived of access until the connected marker releases.

2. The marker next tests the outgoing trunks to the called office and selects an idle one. If, as determined from the route relay, the trunks are divided over more than one frame, the marker will connect to the second group of trunks on the second office link frame in case the first group of trunks is found to be busy.

3. The marker also connects by means of a connector circuit to the district link frame associated with the district junctor to which the calling line is connected. The identification of this frame was obtained from the sender and the sender link control circuit as previously mentioned. Through the connector circuit of the district link frame the marker is extended to the control leads of the district junctor circuit and the crossbar switches of the district link frame. As in the case of the office link frame, only one marker is connected to a frame at a time.

4. The marker after selecting an idle trunk circuit which appears in a horizontal circuit path on one of the secondary switches of the office link frame, then proceeds with the selection of an idle channel through the switches of the two switch frames. A number of these connecting channels is provided between the district junctor and the outgoing trunk. Each channel consists of a "district link" on the district link frame, of an "office link" on the office link frame, and an "office junctor" connecting the district link frame to the office link frame. The marker tests a group of these channels simultaneously and selects an idle one. It then operates the switch magnets which will connect these three paths of a channel, and the district junctor and the outgoing trunk together, thereby establishing a connection from the district junctor to the outgoing trunk.

5. When the marker has completed this operation it checks the connection to insure that it has been properly established and that it is capable of being held under control of the district junctor, before releasing itself from the connection.

6. The marker performs these functions in approximately .5 second, then releases and becomes available for use on other calls.

It will be observed that the three links involved in establishing the connections between the district junctors and the outgoing trunks are used in series and are chosen simultaneously. Generally in other systems the establishment of a connection involving three such paths, is made in three successive stages with a possibility that after a selection has been made at one stage it will be found that the paths accessible to it are all busy and, therefore, the connection cannot be completed.

Before describing the next stage in the establishment of a call, it is desirable to point out other features and functions of the originating marker.

1. The marker permits wide variations in the sizes of trunk groups, permitting trunk groups as small as two and as large groups as may be required. This makes for an efficient use of the office link frame terminals and thereby tends to reduce the office link frame equipment.

2. The marker makes a second trial to establish connections over alternate trunk routes in case calls cannot be completed over the normally used groups because of busy conditions.

3. The marker makes a continuity test of the circuits over which the switches are controlled and tests them for short-circuits, crosses, opens and grounds which would interfere with the proper establishment of a call and where troubles are detected, it signals this condition to a common "trouble indicator" where an indication of the trouble and its location is recorded and a maintenance alarm given. The call is then completed over another group of circuits.

The first stage in the progress of the call through the terminating end of a crossbar office is illustrated in Fig. 13. It consists of connecting the incoming end of the selected trunk to a terminating sender for the purpose of receiving the number of the called line from the subscriber sender.

When the incoming trunk is selected by the originating end of the office equipment, the "sender link control" circuit associated with the terminating sender link frame on which the incoming trunk appears, is called into action. The control circuit then proceeds with the following functions:

1. To locate the incoming trunk circuit, which appears on one of the ten horizontal paths of a primary switch.

2. It selects an idle sender link between this primary switch and a secondary switch on which there is an idle terminating sender.

3. The control circuit selects one of the idle terminating senders reached through the secondary switch and then operates the selecting and holding magnets associated with the selected circuits, which will establish the connection from the incoming trunk to the terminating sender.

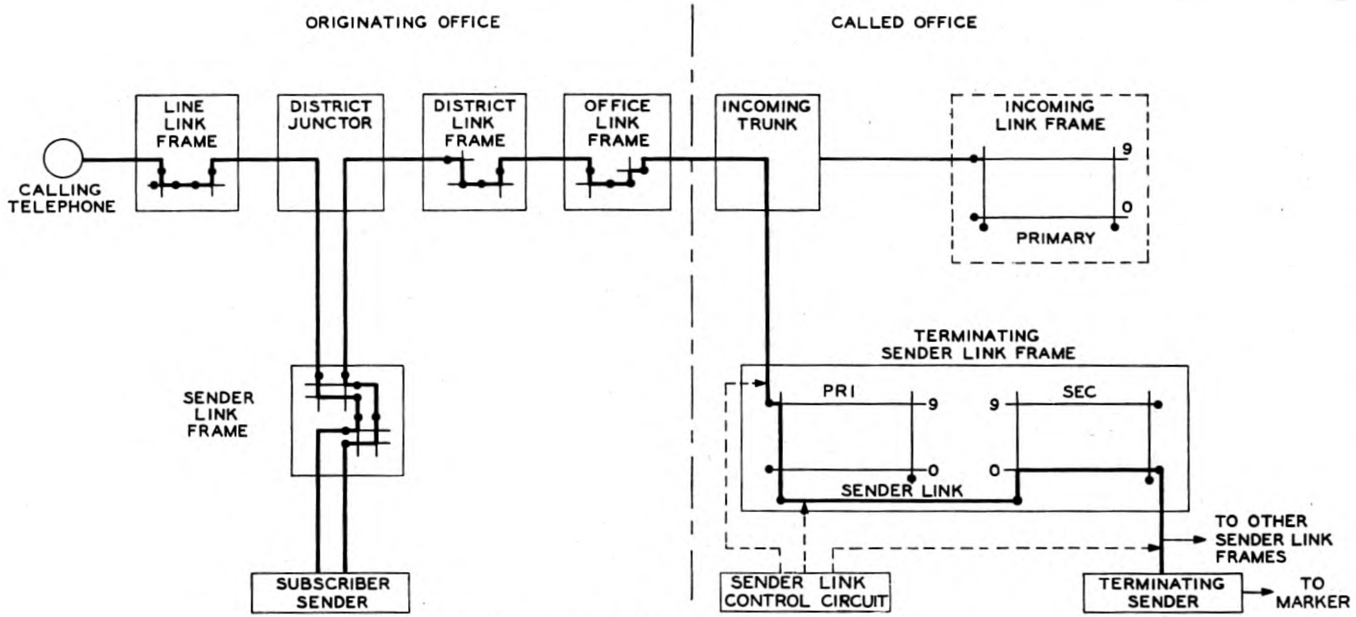


Fig. 13—Connection established to terminating sender.

4. The control circuit will signal to the terminating sender the number of the incoming link frame in which the incoming trunk appears. This frame identification will be used later in establishing the connection to the called line.

5. The control circuit will then disconnect after checking to insure that the connection to the sender has been properly established and that it will be held under control of the trunk and sender circuits after the control circuit leaves the connection.

As soon as this operation has been completed, which takes but a fraction of a second, the terminating sender will be in direct connection with the subscriber sender in the originating end of the connection. This path may be traced, by referring to Fig. 13, from the subscriber sender through the sender link frame, through the district junctor, through the district link and office link frames, over the outgoing trunk to the incoming trunk, and through the terminating sender link frame to the terminating sender.

At this stage of the connection the calling subscriber is still connected with the subscriber sender, and dialing may be still in progress. As the subscriber proceeds with the dialing of the digits of the called number, the subscriber sender will transfer them to the terminating sender. This is done by means of impulses transmitted over the circuit paths between the two senders. When the subscriber sender has completed the transfer of the called number to the terminating sender, the subscriber sender will be released and the calling line will then be connected through the district junctor to the incoming trunk.

When the terminating sender has secured the record of the called line number, the sender then connects to an idle terminating marker by means of multi-contact relays of a connector circuit.

The next stage in the progress of the call is shown in Fig. 14. The terminating marker is the principal control unit at this point in the connection. Its principal function is to provide means for establishing the connection from the incoming trunk to the called subscriber line.

When the terminating sender has connected to the terminating marker, the sender will transfer both the called line number and the incoming link frame identification to the marker. The terminating marker then proceeds to establish the connection to the called line as follows:

1. It connects itself to the particular "number group connector" circuit including the "block relay" frame in which the called line appears in its numerical sequence. All subscriber lines are provided with a set of three test terminals which appear on the block relay

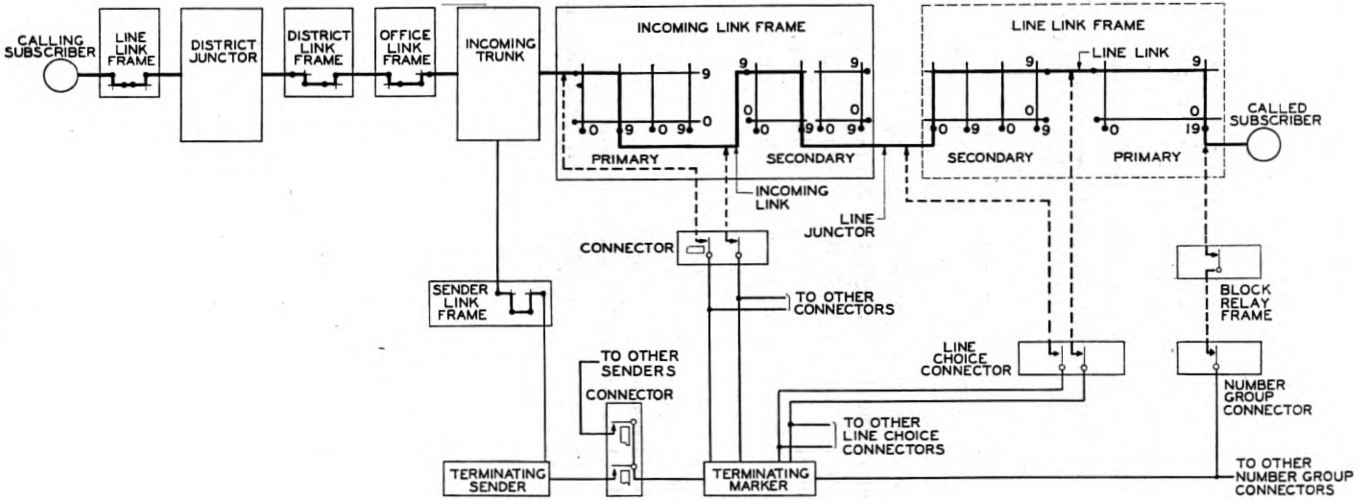


Fig. 14—Connection established to called line.

frame. These terminals correspond to the director number of the subscriber line. A number group connector generally has access to the test terminals of several hundred line numbers depending on the terminating traffic to the lines.

2. The marker will obtain a connection through the number group connector to the busy test terminal of the particular called line and determine whether the line is busy or idle.

3. It will determine from the two other test terminals, the identification of the line link frame where the called line appears, and the horizontal group of line links which has access to the called line. In addition, it determines the type of ringing to be applied to the called line from the circuit conditions on the test terminal.

4. Assuming that the called line is idle, the marker will connect, through the line choice connector circuit, to the line link frame and to the ten line links which have access to the called line.

5. It will then connect, through the connector circuit, to the incoming link frame associated with the incoming trunk to which the calling subscriber line is now connected. The incoming link frame identification was obtained from the sender link control circuit through the sender as previously mentioned.

6. The marker will then select an idle channel through the incoming link and line link frames as previously described. This channel will consist of an "incoming link," a "line junctor," and a "line link" all to be connected in series. The marker then operates the proper selecting and holding magnets of the crossbar switches in each frame which establishes the connection from the incoming trunk to the called subscriber line.

7. The marker will then cause the incoming trunk to start the proper ringing over the called subscriber line and to transmit the ringing tone signal over the trunk to the calling subscriber.

8. At this point the terminating marker and the terminating sender will have completed their functions and, together with the terminating sender link frame, will be released. The complete connection will then be established from the calling line to the called line and the conversational circuit completed when the called subscriber answers.

If the terminating marker finds the called line busy it will cause the incoming trunk circuit to transmit a busy tone to the called subscriber.

The terminating marker has the following other important functions:

1. If the call is for a PBX (Private Branch Exchange) the condition on one of the test terminals of the called line in the number group connector will inform the marker that the line is one of a group of lines. The marker will test all of the lines in the group, testing up to as many as twenty simultaneously, and will select an idle one.

The lines of a PBX may be assigned to non-consecutive numbers within the usual 10,000 series, and with the exception of the numbers dialed, they may be assigned to line numbers in a special group of 2500 outside of the 10,000 series. These features reduce the necessity for number changes due to the growth of private branch exchanges, and conserve subscriber line numbers in the office. The lines of a PBX group can be distributed over several line link frames and over two number group connectors to equalize the terminating traffic load n the case of busy private branch exchanges.

2. The marker recognizes numbers dialed which are unassigned, disconnected or changed numbers, and automatically routes such calls to an operator who will inform the calling subscribers as to the status of the numbers called.

3. In case the called number is on a party line, the marker determines from one of the test terminals which station of the line is to be rung, and signals the incoming trunk to provide the proper ringing.

4. The marker tests the continuity of the circuit paths to be used to the called line before establishing the connection, to insure that the connection is properly set up and that it will be held under control of the subscriber telephone after the marker disconnects. The marker also tests for short-circuits, crosses and grounds, and in case of a failure due to any inoperative condition it will connect itself to the common trouble indicator and leave a record of the trouble and its location and give an alarm to the maintenance force.

Figure 15 shows the complete talking connection through the various trunks and switch frames as finally established after all of the common control circuits have been released.

On a call to a subscriber served by a panel dial office, the connection is routed through the district link and the office link frames in the same manner as on a call terminating in a crossbar office, but in this case the idle trunk chosen on the office link frame terminates in an incoming panel switch in the distant panel dial office. The subscriber sender of the crossbar office causes the incoming and final selectors in the terminating panel office to select the called subscriber line without the aid of any terminating senders in either office. When the subscriber sender has completed these functions it will be released and the connection will be established from the calling line over the interoffice trunk circuit and through the terminating panel incoming and final selectors to the called line. On this type of call the subscriber sender operates in the same manner as though the called line were in a crossbar office, and the selectors in the panel office operate in the same manner as though the call had originated in another panel office.

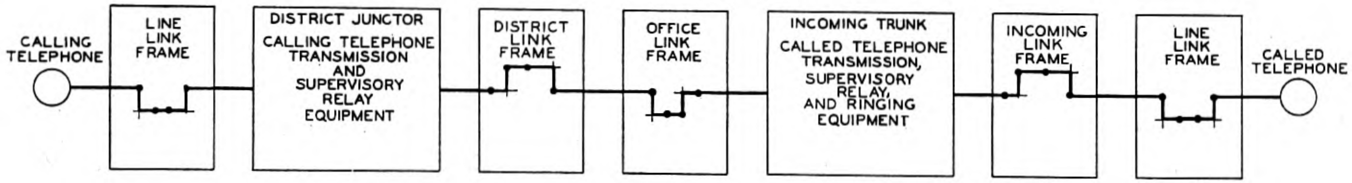


Fig. 15—Completed talking connection.

No changes are required in the panel selectors to function with the crossbar office.

On a call for a subscriber in a manual office, the call would be routed through the switches of the district link and the office link frames as previously described and connected to a trunk circuit on the office link frame which terminates in the "B" switchboard in the manual office. The subscriber sender of the crossbar office then transfers the called number by impulses transmitted over the interoffice trunk circuit to the operator's position equipment in the manual office. The called number appears in the form of visible numbers on the operator's keyshelf. The operator completes the connection by "plugging" the associated trunk circuit, which terminates on a cord and plug, into the called subscriber line jack.

A call originating in a panel dial office for a subscriber line in a crossbar office reaches the crossbar office through an incoming trunk circuit as in the case where the call originated in a crossbar office. The call from the panel office is then handled by the crossbar office terminating sender and marker in exactly the same manner as described for calls originating in the same crossbar office.

A call originating in a manual office for a line connected to the crossbar office reaches the crossbar office over an incoming trunk circuit from an "A" operator's position in the manual office. These incoming trunks in the crossbar office are similar to the incoming trunks previously described. In this case, however, the incoming trunk is connected to a terminating "B" sender and by means of this sender to a "B" board operator in the crossbar office. The "B" operator will obtain the called number verbally from the distant "A" operator and then, by means of the keyset on her position, register the called number in the terminating sender. The terminating sender will then select a terminating marker and the connection will be established in exactly the same manner as described for a call originating in a crossbar office.

MAINTENANCE FACILITIES

Automatic routine testing circuits are provided for testing all the principal circuit units, such as the district junctors, incoming trunks and senders. These test circuits automatically put each circuit, one after the other, through all of its functions on all classes of calls to insure that it performs satisfactorily. It tests the important relays of the circuits to insure that they have the proper adjustment to handle the worst circuit conditions. In case any circuit fails to meet the test conditions, the test is stopped and an alarm given to the maintenance force.

Trouble indicator circuits are provided for use in connection with the test and maintenance of the marker circuits. These circuits are arranged so that when trouble is encountered by a marker, the marker will seize the trouble indicator and operate combinations of relays and light small lamps which indicate the nature and the location of the failure and give an alarm to the maintenance force.

Equipment

Figure 16 shows a typical switch frame used in the crossbar system. This particular frame is a "line link" frame which serves a group of subscribers for both originating and terminating traffic. The frameworks on which the equipment is mounted are constructed of rolled

Fig. 16—Line link frame.

bulb angle iron sections with a sheet metal base. The bulb angle construction provides a framework which is light in weight and has the required strength, and permits an equipment mounting arrangement which conserves space and facilitates the wiring of the apparatus. The frames are welded and incorporate such features as sanitary base construction, guards to protect the apparatus and wiring against damage from the rolling ladders located between the rows of frames, and a cable duct or runway for the A.-C. power service cables with plug receptacle outlets for use with electric soldering irons, portable lights, etc.

These frame equipments are built in standardized units, which provide the required flexibility to satisfy the variations in telephone traffic and classes of service encountered in the different telephone areas. Where it has been necessary to divide an equipment assembly into several units, due to the limitations of handling, shipping and to care for different classes of service, the equipments have been designed so that the installation effort required for interconnecting such units has been reduced to a minimum.

The bays of equipment located at the right, in Fig. 16, equipped with crossbar switches, are the primary line link bays. The vertical units of these crossbar switches are wired to the subscriber lines. These primary bays are made available in units of 100 and 200-line capacities. As discussed previously the number of primary bays provided in a line link frame may be varied to fit the traffic load of the subscriber lines. The left-hand bay of this frame contains the vertical file of crossbar switches, known as the secondary switches and the vertical units of these switches are wired to district junctors and line junctors. The line link control circuit apparatus, which is common to the frame, is located at the bottom of this bay.

Figure 17 shows a group of three frame units, namely, the subscriber sender link, the district junctor and the district link frames, which are closely associated in the trunking network and have been designed as a fixed equipment group. However, for shipping reasons the group is divided into three separate equipment units. The district junctor circuits, consisting primarily of relays, are mounted in groups on the middle frame. These groups are provided in standardized units of various types, such as those required to serve coin and non-coin subscribers lines. A similar arrangement of frames is used for the combination of terminating sender link, the incoming trunk, and the incoming link frames.

Figure 18 shows a row of subscriber sender frames and a frame of "A" operator senders located at the extreme right. These frames

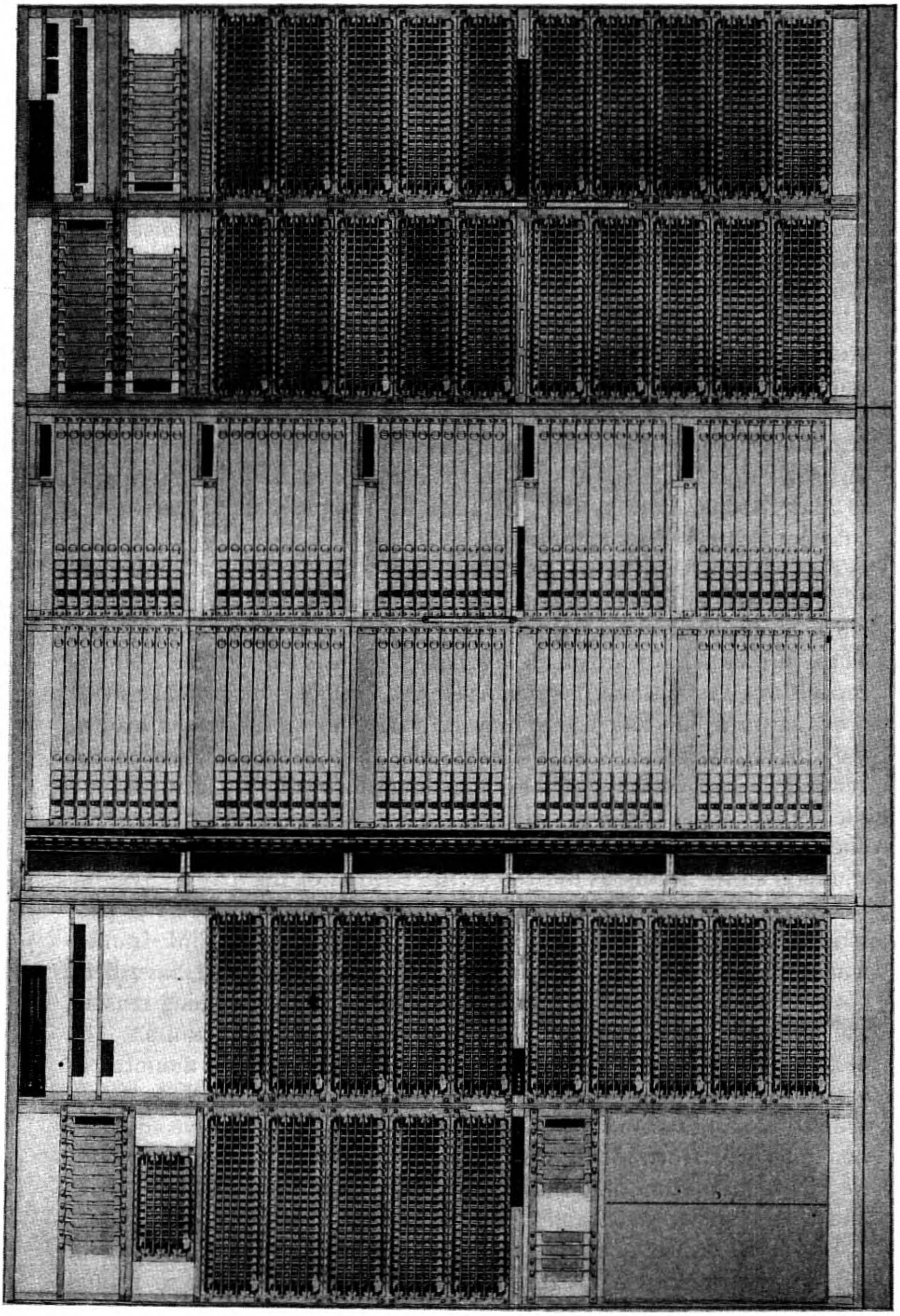


Fig. 17—Three-frame assembly of district junctor, district link and sender link frames.

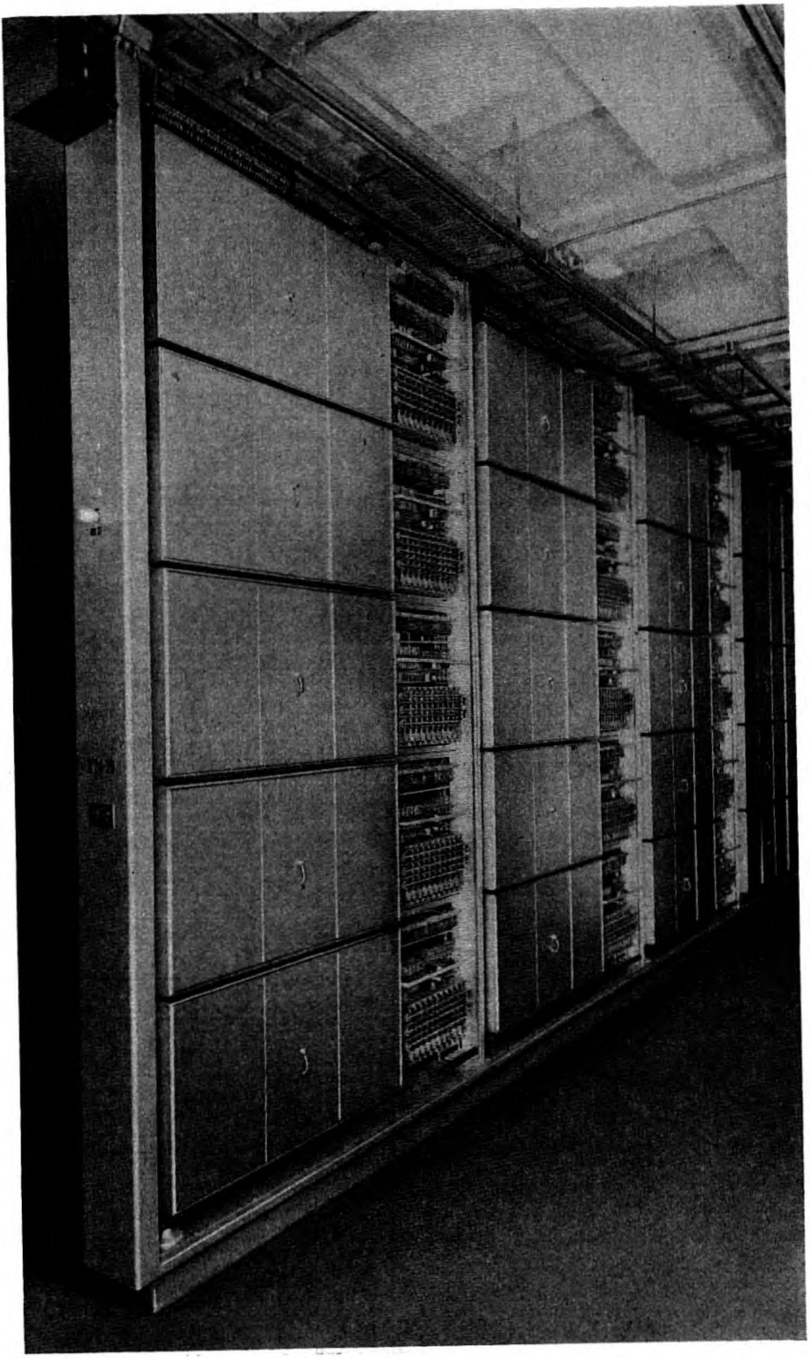


Fig. 18—Sender frames.

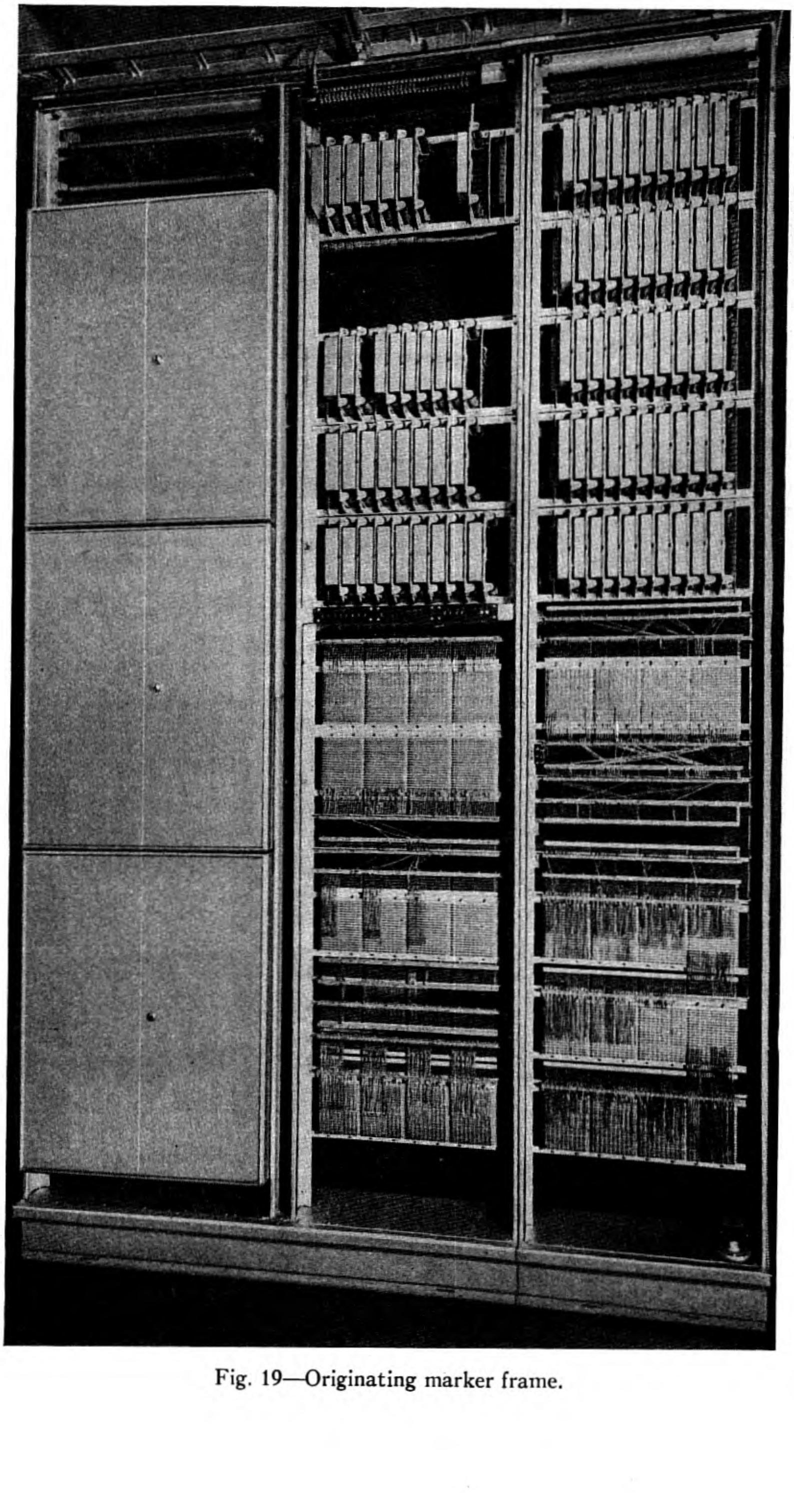


Fig. 19—Originating marker frame.

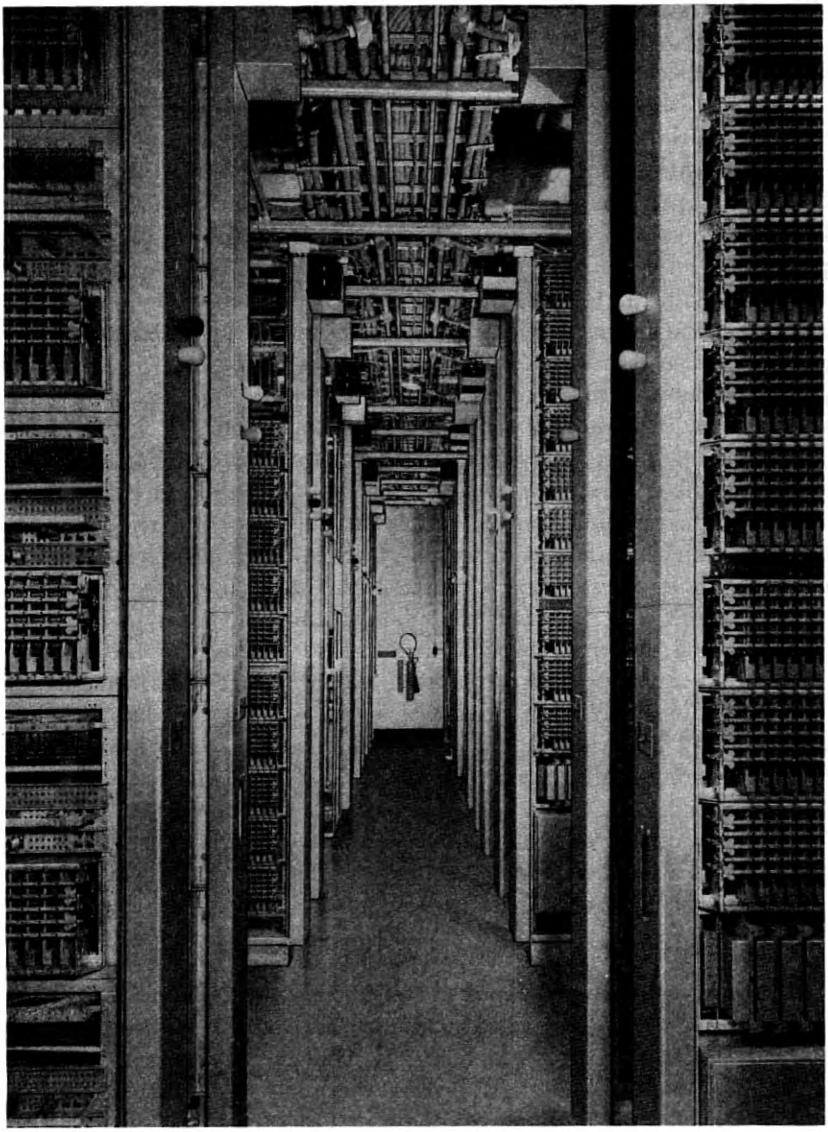


Fig. 20—Battery supply feeders, power wiring, fusing, etc.

accommodate five senders which may be of one type, or a combination of both types. The crossbar switch shown on the right of each subscriber sender unit, is a part of the sender circuit and is employed for the purpose of registering the called numbers dialed by the subscribers.

The "A" operator senders are associated with the "A" operator switchboard equipment and are used for the completion of certain classes of calls such as toll and assistance calls.

A view of the originating marker frame is shown in Fig. 19. There will be a variation in the equipment on this frame for different cities due to the variation in the number of route relays required, the number depending upon the number of central office codes that may be dialed by subscribers and operators. This variable feature is cared for by providing the route relay equipment in bays of 100 codes as shown in the right-hand bay. The terminal fields shown below the route relays on the frame provide the flexible connecting facilities which permit the use of any route relay for any office code and which readily permit changes in routings, variations in trunk group sizes and other features which are subject to change from time to time.

The power plant equipment provided for the crossbar offices is similar to the equipment now being furnished for all large dial central offices. The principal power supply arrangements provide 48-volt direct current for the operation of practically all the signaling and the telephone transmission circuits. Also several other sources of direct current are provided for miscellaneous purposes as in other standard dial systems. A new distribution scheme for the battery feeders on the frames is employed which reduces the amount of copper required. A common set of 48-volt battery feeders supplies the signaling and talking current for all frames. Individual frame filters are connected across the battery supply leads at the frames where a noise-free battery supply is required for talking circuits. Figure 20 shows a view of the overhead battery cables, conduits for the A.-C. power leads, and the fuse cabinets for the fusing of the battery supply to a row of frames.

Application

As mentioned in the first part of this paper, two crossbar dial central offices were cut into service in 1938 and these have now been in commercial operation for several months. One of these offices serves a residential area in Brooklyn, while the other serves a congested business area in the midtown Manhattan district of New York City. The operation of these offices under actual service conditions has been highly satisfactory and our expectations in regard to performance have been fully realized.

This type of system will be used for new offices in large cities instead of the panel system as rapidly as manufacturing and plant conditions permit and the apparatus which was designed for this system will be used in other fields of the telephone system.

A Twelve-Channel Carrier Telephone System for Open-Wire Lines

By B. W. KENDALL and H. A. AFFEL

A new carrier telephone system is described, together with its application in the long distance telephone plant. By its use, an open-wire pair which already furnishes one voice circuit and three carrier circuits may have twelve more telephone circuits added. Thus in all sixteen telephone circuits are obtained on a single pair. Several such systems may be operated on a pole line.

Various problems incident to the extension of the frequency range, from about 30 kilocycles, the highest frequency previously used, to above 140 kilocycles, are discussed. Among the more important of these are the control of crosstalk between several systems on a pole line, arrangements for taking care of intermediate and terminal cables, and automatic means for compensating for the effects of weather variations on the transmission over this wide frequency range.

Introduction

BARE wires supported on insulators, stretched between poles, make up the pioneer electrical communication circuit, the open-wire line. Although great advances have been made in the application of cable structures, the open-wire lines still hold their own in some sections of the country. This is because, to offset their physical vulnerability, they have several unique virtues. They are flexible and permit adding one pair of wires at a time. They are also comparatively economical where conditions favor their use. Furthermore, they are low-attenuation circuits and for this reason were the first to be used for high-frequency carrier systems.

The first carrier systems, beginning in 1918, added three or four channels to the existing voice circuit on a pair. To keep pace with this development, improvements in transposition systems were devised so that many such carrier systems might be operated on the same pole line. Such carrier systems, typified by the three-channel type C [1] system, have seen continuous growth in use in the long distance plant. Now a twelve-channel system, the type J, is being made

---

[1] "Carrier Systems on Long Distance Telephone Lines," H. A. Affel, C. S. Demarest and C. W. Green, *Bell System Technical Journal*, July 1928, and *A. I. E. E. Transactions*, Oct. 1928, pp. 1360–1387. "A New Three-Channel Carrier Telephone System," J. T. O'Leary, E. C. Blessing and J. W. Beyer, *Bell System Technical Journal*, this issue.

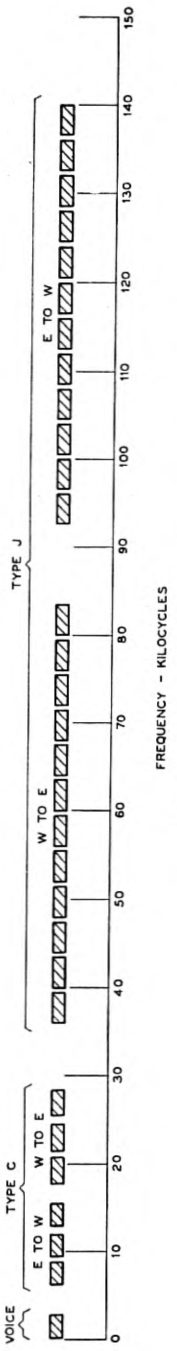


Fig. 1—Frequency allocations.

available to add to the type C system, thus giving sixteen telephone circuits on an open-wire pair in addition to the two telegraph circuits. Since there are already about 60,000 miles of pole line equipped with type C systems, the new type J system was developed to go in the frequency range above the type C system rather than to supersede it with more channels (Fig. 1).

The new system has been designed to meet high standards of transmission and reliability for distances up to several thousand miles. The frequency band transmitted by the individual derived circuits is exceptionally wide, from about 100 to 3600 cycles for a single system and has been previously discussed [2] in relation to the channel spacing in this and other new broad-band developments.

An important feature of the work on the type J system has naturally been that of making the line circuits suitable for carrying the higher frequencies. The tendency of circuits to crosstalk into one another increases rapidly with frequency. Advances in transposition design and structural improvements have now made it possible to extend the frequency range from 30,000 cycles to 140,000 cycles, which is about the upper frequency of the type J system. The problem of incidental cables in open-wire lines has also been serious, since the losses increase with frequency, and what is usually more important, there may be substantial reflection effects at junctions of the open-wire line and cable. These are serious, not only from the standpoint of the transmission loss which they entail, but from their effect on crosstalk. The increase in attenuation at the higher frequencies has also brought other problems into the picture. For example, repeaters are needed at more frequent intervals than with the lower frequency systems. Attenuation variation with frequency due to weather changes is greater than at the lower frequencies.

Figure 2 shows schematically the complete type J system, with its different major circuit elements, resulting at the terminals in the division of the single line circuit effectively into sixteen talking circuits. In no recent development is the function of the wave filter in providing essential units in a frequency dividing plan more forcefully illustrated than in the application of this new system, in combination with the type C and other facilities which exist. There are about sixty different designs of filters and networks in the terminals and repeaters. Their functions are varied,—as, for example, separating the individual channel bands, separating the opposite directional groups of channels, separating the type J frequency range as shown in Fig. 1 from the type C and other ranges, separating the different carrier frequencies

---

[2] "Transmitted Frequency Range for Circuits in Broad Band Systems," H. A. Affel, *Bell System Technical Journal*, October 1937.

of a carrier supply system in which the carriers are all derived from a common 4000-cycle source, etc.

The new system, as in the case of the type C, uses single sideband transmission with carrier elimination. Copper-oxide units are employed as translator elements of various kinds,—modulators, demodulators, and harmonic producers. Methods of mounting the equipment, and methods and apparatus for testing follow lines already worked out for the type K cable carrier system, which was described a year ago in two A. I. E. E. papers.[3]

CHANNEL TERMINALS

A terminal of the type J system changes twelve independent voice channels into a compact block of twelve carrier channels properly allocated in frequency for transmission over the open-wire line. Inversely, such a block received from the open-wire line is separated and transformed into twelve independent voice channels. The first step in transmitting the twelve voice channels is to modulate them on twelve carrier frequencies 4 kilocycles apart from 64 to 108 kilocycles and to select the lower sidebands by means of quartz crystal channel band filters. The last step in the conversion from a received twelve-channel block to the twelve independent voice channels consists in the division of the block by twelve quartz crystal channel filters and the demodulation of these messages to produce voice frequency transmissions. These two frequency changes and separations are performed by the same equipment that is used in the type K cable carrier system terminals.

Figure 3 shows the circuit of a modulator and a demodulator for the opposite directions of a single conversation with indicated connections for the eleven others which make up this fundamental twelve-channel block. The modulator consists of a bridge assembly of copper-oxide varistors and is supplied with about 0.5 milliwatt of carrier power from the carrier supply system which is described later. Of the two resulting sidebands, the lower is selected by the crystal band filter following the modulator. The line sides of twelve modulator band filters are joined in parallel and a compensating network is connected to preserve the band characteristics of the upper and lower channels.

On the receiving side, after separation by one of the twelve parallel filters the sideband is applied to a demodulator supplied with the

---

[3] "A Carrier Telephone System for Toll Cables," C. W. Green and E. I. Green, *Bell System Technical Journal*, January 1938 and *Electrical Engineering*, May 1938. "Cable Carrier Telephone Terminals," R. W. Chesnut, L. M. Ilgenfritz and A. Kenner, *Bell System Technical Journal*, January 1938 and *Electrical Engineering*, May 1938.

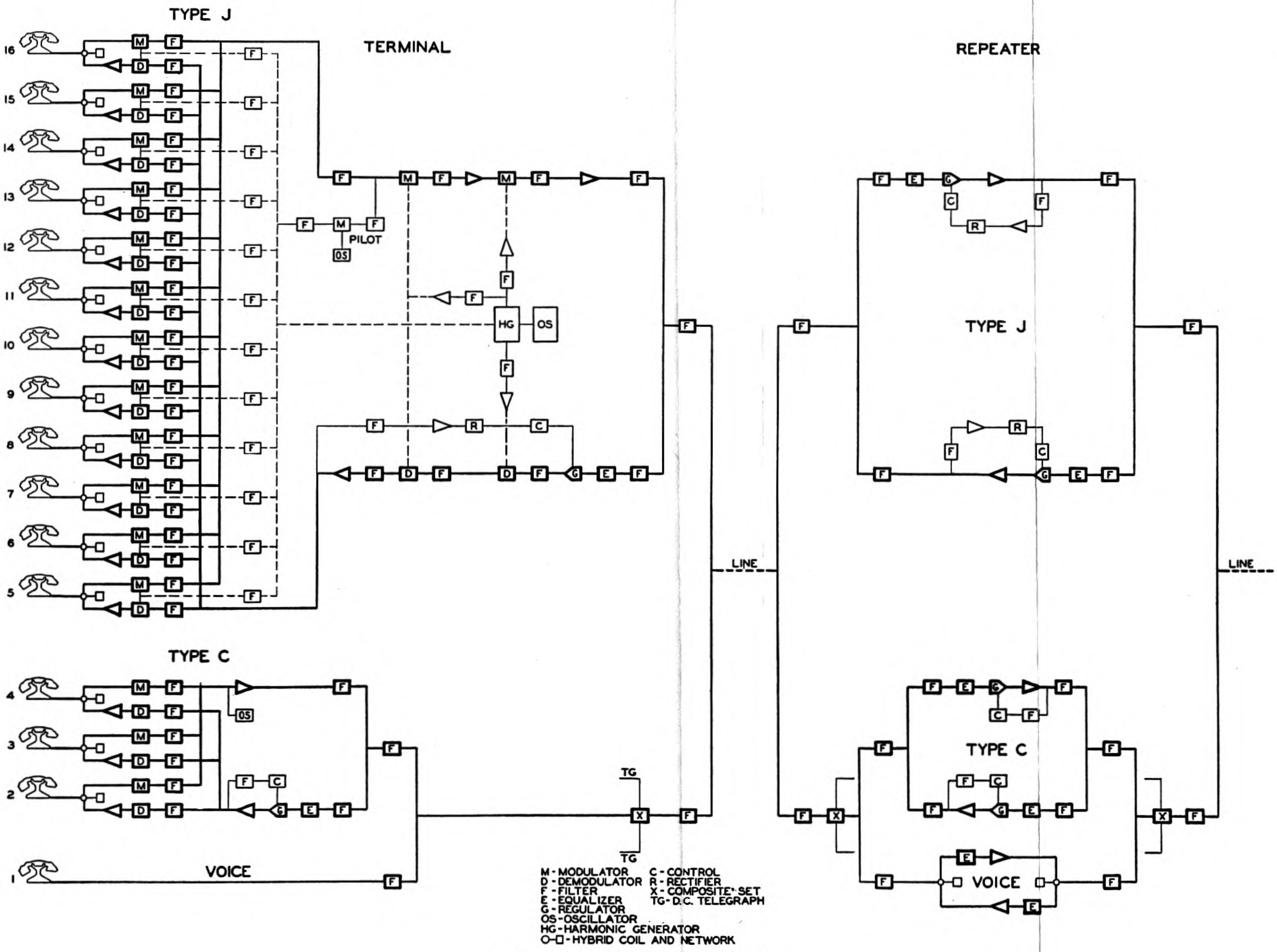


Fig. 2—Terminal and repeater layout.

proper carrier frequency to restore the voice frequency message. Because of the low level at which demodulation takes place, the demodulator is followed by a single-stage amplifier to produce the level desired in the voice frequency circuit. The gain of this amplifier is adjustable over a moderate range.

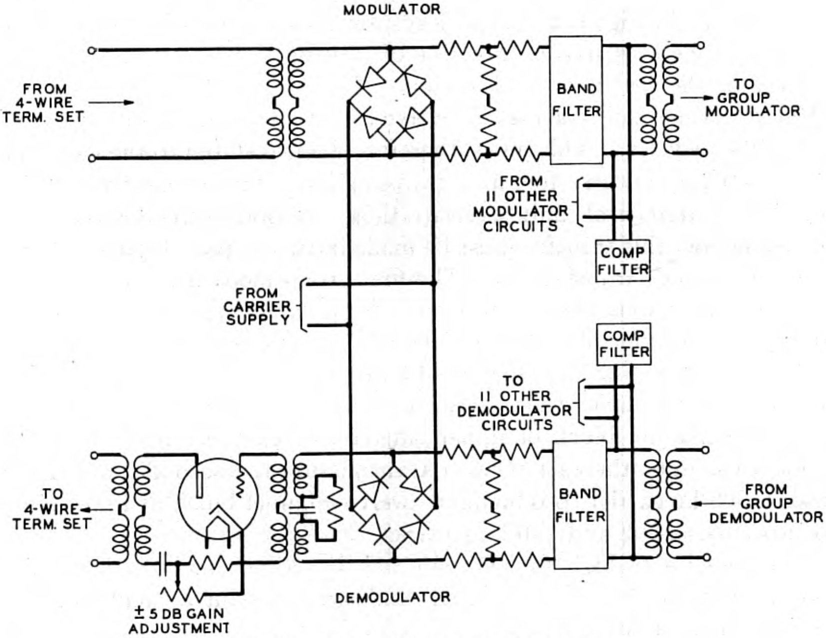


Fig. 3—Channel modulator and demodulator.

The combination of a single modulator and a single demodulator and associated equipment shown in Fig. 3 is called a "Modem" and two of these are mounted on a single equipment panel. Nine of these panels, sufficient for one and a half type J systems, or eighteen conversations, mount in a single relay rack bay of standard height.

Carrier Supply

The carrier frequencies 64–108 kilocycles are all derived as harmonics of a 4-kc frequency produced by a tuning fork controlled oscillator. This frequency is applied to an easily saturated coil to produce a sharply peaked wave which is rich in odd harmonics. Even harmonics of 4 kilocycles are obtained by rectification in a copper-oxide unit of part of the odd harmonic output. Odd and even harmonics appear in separate circuits from which each frequency desired is separated by a quartz crystal filter. Frequencies as high as the 121st harmonic, that is, 484 kilocycles, are obtained in this way from the carrier supply system. Because of the importance of the carrier supply two sources are provided, with automatic equipment to transfer rapidly from the regular to the emergency source.

Group Modulation

As shown in Fig. 1, the type J system uses a band of 36 to 84 kilocycles for the west to east direction of transmission and 92 to 140 kilocycles for the east to west direction. The output of the fundamental twelve-channel unit consists of twelve lower sidebands from carriers of 64–108 kilocycles. This must, therefore, be translated to the two type J directional groups for line transmission. Since the frequencies in the fundamental unit overlap those in both directions of line transmission, this transfer must be made in two steps. Figure 4 shows these frequency translations. The first group modulation is the same for both directions of transmission. By modulating the fundamental unit with a carrier of 340 kilocycles there is obtained a block of lower sidebands extending from 400 to 448 kilocycles. A second modulation with a 484-kc carrier then gives, for transmission from west to east, a twelve-channel block of upper sidebands extending from 36 to 84 kilocycles. For the east to west transmission the second modulation uses a 308-kc carrier, producing a twelve-channel block of lower sidebands between 92 and 140 kilocycles.

Frequencies as high as 308, 340 and 484 kilocycles are chosen for group modulation in order that undesired products shall be well separated from desired products to permit their elimination by simple filter structures.

The same group modulation processes that have been described above for adapting the twelve-channel group for line transmission are used in the opposite sequence for receiving the block from the line and preparing it for separation by the channel band filters of the receiving terminal; thus, for instance, at an east terminal the block of upper sidebands, extending from 36 to 84 kilocycles as received from the line, is first modulated with 484 kilocycles producing lower sidebands between 400 and 448 kilocycles. These are next modulated with 340 kilocycles, which produces a block of twelve lower sidebands extending from 60 to 108 kilocycles, which is the group that the fundamental twelve-channel terminal unit is designed to handle.

Figure 5 shows the essential features of the group modulating and group demodulating circuits. As in the type K system, group modulation is performed at a very low level of the message material and with a high level, about 25 milliwatts, of the group carrier supply, in order

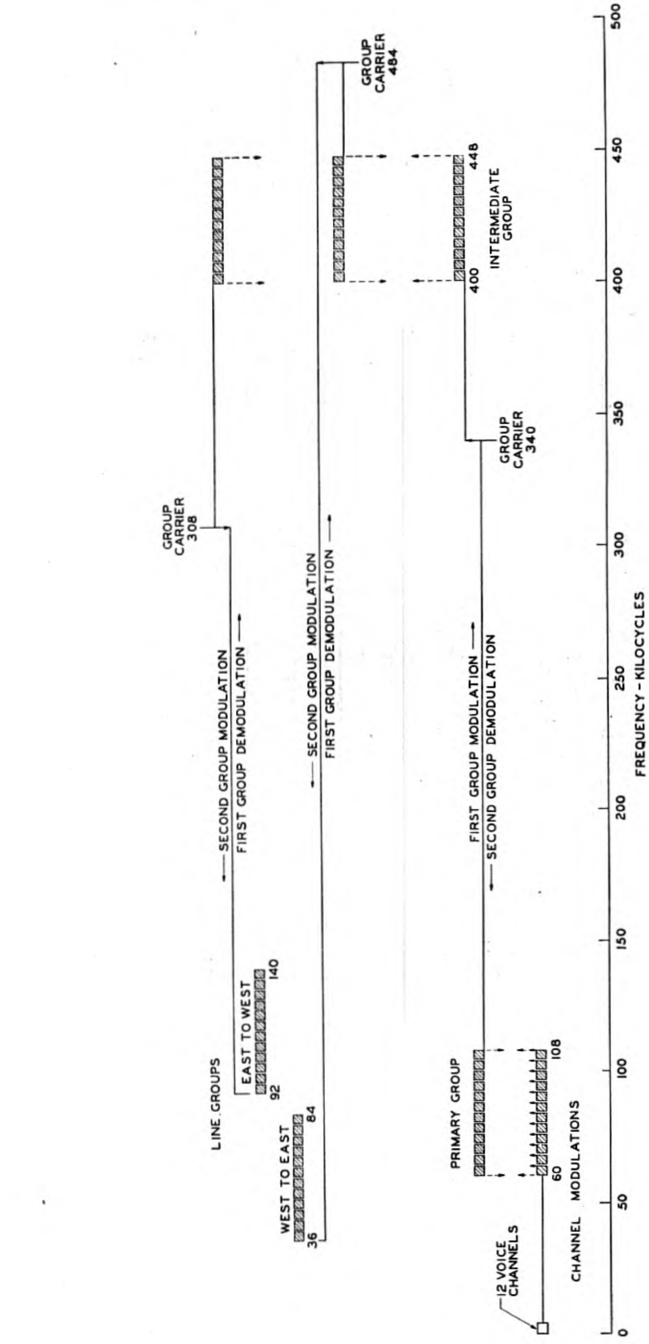


Fig. 4—Frequency translations.

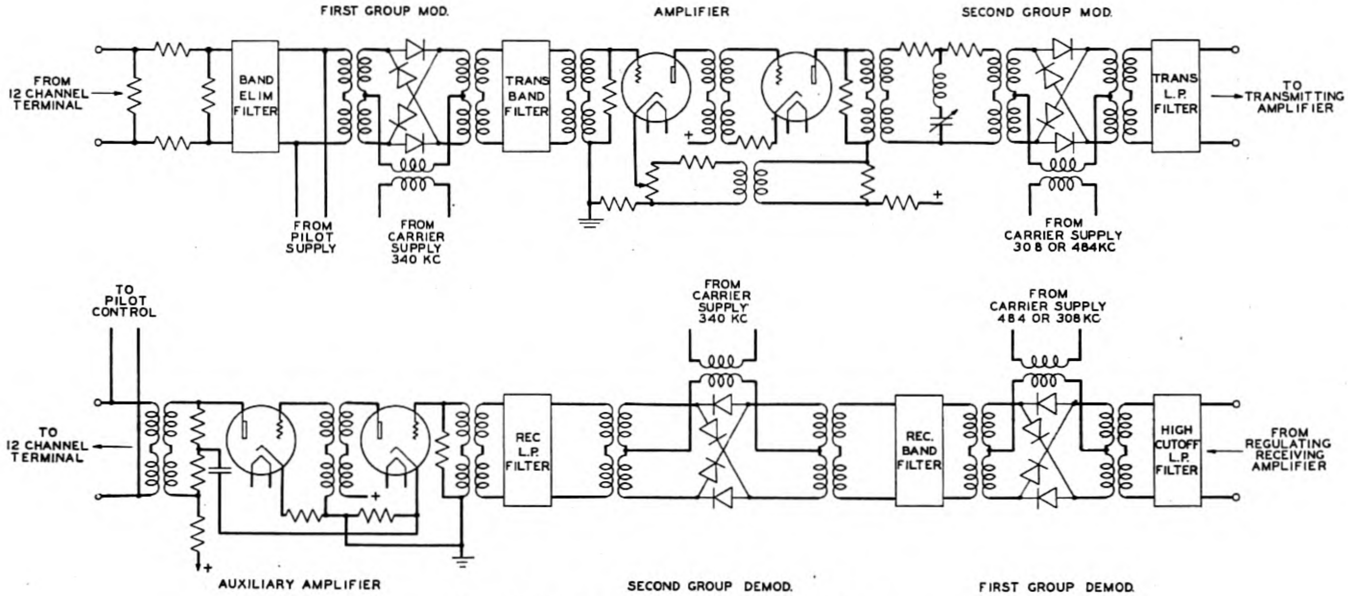


Fig. 5—Group modulator and demodulator.

to minimize interchannel crosstalk. The group modulators are of the doubly balanced bridge type which aids in suppressing some of the unwanted modulation products. Following the first group modulator and also following the first group demodulator are coil and condenser type 400–448 kc band filters which reject the unwanted products and pass the band of frequencies containing the twelve channels. Between this filter and the second group modulator on the transmitting side of the terminal, an intermediate amplifier is used in order to keep the level of the group transmission above danger of noise. Following the second group modulator and also following the second group demodulator are low-pass filters which cut off frequencies above about 160 kilocycles, to suppress unwanted modulation products. From the output of the receiving low-pass filter the twelve-channel group, 60–108 kilocycles, passes through a two-stage "auxiliary" amplifier to bring it to the desired level.

The carrier frequencies for group modulation and for group demodulation are derived from the same 4-kc tuning fork controlled oscillator that supplies carriers for the twelve-channel unit. From the circuit in which appear the odd harmonics of 4 kilocycles, the 77th, 85th and 121st harmonics, that is, 308, 340 and 484 kilocycles, are selected by carrier supply filters and separately amplified by two-stage amplifiers to produce the powers required for group modulation. Outputs from these amplifiers are fed to individual frequency busses capable of supplying the group modulators and demodulators for ten systems. An emergency carrier supply for these frequencies is also provided, with arrangements for switching rapidly from the regular to the emergency circuits.

Terminal Amplifiers

As indicated on Fig. 5, the transmitted twelve-channel group, now transferred to the proper frequency range for line transmission, goes from the low-pass filter at the output of the second group modulator to a transmitting terminal amplifier which is similar in most essentials to the amplifiers of the line repeaters. The twelve-channel group, arriving from the line, passes through a regulating amplifier arranged and controlled to compensate for variations in equivalent of the adjacent line section before passing to the first group demodulator. Similar regulating amplifiers are used at all repeater points.

Filters

At terminals and also at repeater points, two kinds of filter sets are required. One kind is used in the line to separate the type J frequency range 36 to 140 kilocycles from the type C and other lower frequencies on the line. The second kind is the directional filters of the type J system itself. These separate a twelve-channel band of frequencies lying below 84 kilocycles used for west to east transmission from the twelve-channel group lying above 92 kilocycles which is transmitted from east to west. These directional filter sets are carefully designed to equalize any non-uniformity of loss in both the directional and the line filters. As this equalization involves a considerable loss over a large part of the filter band it is provided entirely in the receiving directional filters where the transmission is at a low level and the loss can readily be made up by amplification. In this way nearly the full energy output of the transmitting or repeater amplifier is available for line transmission.

LINE CROSSTALK PROBLEMS

As noted previously, type J systems will, in general, be applied on pairs on which type C systems are already operating. Such pairs have already been arranged to transmit frequencies up to 30,000 cycles, and transposed in such a manner as to perform satisfactorily as regards crosstalk to and from nearby pairs on which similar carrier systems are operating. In addition, on most modern lines the spacing between wires of a pair has been reduced from twelve to eight inches; and, on many of the lines, in order further to reduce crosstalk by increasing the spacing between pairs, the number of pairs on a crossarm has been limited to four instead of five, omitting the pole pair. Now, by applying a new transposition system designed for type J operation up to 140,000 cycles, an eight-inch spaced four-crossarm line may be arranged to transmit type J frequencies on at least ten pairs out of sixteen. Type C systems may, of course, be used on all of the pairs. Finally by using the most advanced transposition design methods, and increasing the crossarm spacing, in addition to the features noted above, a new line may be constructed to permit the operation of sixteen channels on all pairs.

To make the pairs of wires good for type J systems, more than a fourfold increase in frequency range, was difficult. The natural tendency of the circuits to crosstalk is increased even more than the frequency ratio, so that in addition to applying a new transposition design it is necessary that the transposition poles be more accurately located, and that the sags of the two wires of each pair be kept more nearly alike. On lines which already have eight-inch spaced wires, no major structural changes are necessary. However, on lines which have only twelve-inch spaced wires and where it is desired to make available a number of pairs for type J transmission, structural changes, such as respacing the wires of the pairs concerned to six inches, are necessary in order to reduce the coupling.

One factor of extreme importance is that of reflected near-end crosstalk. In the application of transposition systems it is usually not possible to reduce the near-end crosstalk to a magnitude approximating the far-end crosstalk. It is the latter with which the carrier systems are chiefly concerned, since similar types of systems on different pairs all transmit the same frequency range in the same direction. If, however, the lines concerned do not have smooth impedance characteristics, i.e., a high degree of freedom from reflection effects, near-end crosstalk may be converted by reflection into far-end crosstalk of sufficient magnitude to be controlling over the true far-end crosstalk.

This means that lines to be used for several type J systems must be made unusually smooth electrically—impedance variations kept within a few per cent. The achievement of such smoothness consists chiefly in:

(1) Reducing the electromagnetic and electrostatic couplings to other pairs so that there are no large energy interactions, with corresponding impedance irregularities. Generally speaking, when the pairs concerned have been transposed for reduced far-end crosstalk up to the maximum frequency transmitted, this condition is also satisfied.

(2) Minimizing the effect of intermediate and terminal cables. This latter problem has caused considerable concern and is responsible for the development of several new techniques in the design and treatment of such cables, where they appear in a long line otherwise consisting chiefly of open wire.

CABLE TREATMENT

As a means of overcoming the reflection and attenuation effects of short pieces of terminal or intermediate cable, loading naturally suggests itself, as applied in type C systems, where the cable pairs involved are commonly equipped with carrier loading coils, spaced at about 700-foot intervals. This compares with the 3000-foot or 6000-foot spacings which are standard for voice-frequency loading. However, loading pairs in existing cables satisfactorily up to 140,000 cycles would mean coils at approximately 200-foot intervals. Because of physical limitations, existing manhole spacings, etc., this is highly impractical. A reasonable solution has, however, been found in the creation of a new form of low-capacitance high-frequency cable,—a disc-insulated unit which has constructional features in common with the coaxial cables and a capacitance of only .025 microfarad per mile as compared with about .062 microfarad for conventional cable pairs. This permits more practical loading coil spacings. These disc-insulated units are made up as spiral-fours, that is, two pairs (.051″ diameter wire) which form the diagonals of a square. When these cables are loaded with small coils at intervals of approximately 600 feet, they present impedance characteristics substantially equivalent to that of an open-wire pair over the desired frequency range. Accordingly, they form a desirable, although somewhat expensive, solution of the problem of intermediate or entrance cables. As shown in Fig. 6, the spiral-four units are bound together in complements of seven or less under a lead cable sheath similar to standard toll cables. It should be noted that the low-capacity disc-insulated loaded cables not only provide a satisfactory solution of the impedance matching

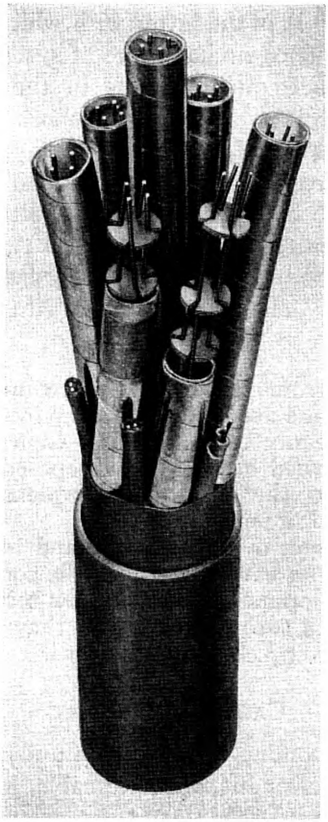


Fig. 6—Disc-insulated cable. Sheath diameter 2.3 inches.

problem, but they also give a cable circuit of low attenuation,—approximately 1.2 db per mile at 140 kilocycles.

Nevertheless, where spare pairs exist in cables, it has often been found economical to use them for type J transmission. It is possible to use them only non-loaded, in which case the attenuation is very high—4 to 6 db per mile, depending on the gauge, at 140 kilocycles, and impedance matching transformers are, of course, required at the junction of the open wire and cable. There are cases where this higher attenuation may be permitted and these pairs are used by separating the type J range from the lower frequency range, which is transmitted through pairs equipped with the older type C carrier loading. The separation is accomplished by filters which are usually housed in small filter huts at the junction of the open-wire line and cable.

In other cases it has been found economical to use the frequency separation method with filters and to install new non-loaded cables of lower attenuation to lead in the type J frequency band alone. Paper insulated 10-gauge pairs or the disc-insulated spiral-four cable of the type described above may be used for this purpose. In either case transformers are used to match the cable impedance to that of the open-wire line over the type J frequency range.

The reflection requirements are so severe and the effects of even short lengths of cable at the high frequencies so serious, that even short lead-in cables, where the open-wire line actually extends to the repeater or terminal building,—cables which are only 100 or 200 feet long, must receive special treatment. This has also been accomplished by the use of the disc-insulated spiral-four cables, loaded.

INTERACTION CROSSTALK

Because of the higher attenuation there will be many repeater points on a long line at which the type J system will be amplified but at which the other systems and wires on the line will pass through the station without amplification. In this case, even though the type J pairs are properly transposed to keep down crosstalk between themselves, there still remains the crosstalk between them and the other pairs on the line, not only pair-to-pair crosstalk but crosstalk from the type J pair to various circuit paths consisting of irregular wire combinations.

Two difficulties arise in this case: The first is that the crosstalk from the output of one J system into an irregular path may be retransferred into the input of a repeater on another type J system. The second is that the crosstalk from the irregular path may be returned to the input of the same repeater and either influence the overall transmission characteristic or, if sufficiently severe, actually cause the repeater to sing. This general situation has made it necessary to introduce in the circuits at such points "crosstalk suppression" filters in the non-J pairs and longitudinal choke coils in all pairs.

STAGGERING

In addition to the various steps which are taken in order to reduce crosstalk by improving the line conditions, the type J system may include a feature which has been used in the type C system,—the staggering of the channel bands used on neighboring pairs. The advantage of staggering results from the facts that (*a*) the sensitivity of the ear and the power of the voice vary over the audible range, (*b*) the efficiencies of transmitter and receiver also tend to vary over the frequency range, (*c*) part of a channel band may lie opposite "dead" frequency range on an adjacent pair, and (*d*) by controlling the arrangement of the sidebands the crosstalk may be made unintelligible even if not inaudible. The staggering feature is readily provided in the type J system by a suitable choice of carrier frequency for the second group modulator and first group demodulator. With the staggered systems the highest frequency used would be about 143 kilocycles.

ATTENUATION PROBLEM

In what has preceded in the discussion of line problems, the emphasis has been confined chiefly to the question of the smoothness of a line from an impedance standpoint in order to keep down reflection effects and, correspondingly, to improve the operation from a system-to-system crosstalk standpoint. There is also the problem of the higher attenuation incident to the use of higher frequencies. Between 30,000 cycles and 140,000 cycles, the normal wet weather attenuation for a 165-mil open-wire pair, for example, rises from about 0.13 to 0.28 db per mile,—an increase of approximately 2 : 1. Repeaters on the type J system, if applied on the basis of approximately the same output level and minimum level requirements, must be spaced at about one-half the interval of the type C systems. Normal spacings for type J systems would therefore be expected to range from 75 to perhaps 100 miles where no large amount of intermediate cable existed.

However, another problem, not present to a similar degree at the lower frequencies, tends in many cases to have a controlling effect on this spacing, that is, sleet or ice on the wires. With ice, frost, or snow on the wires, the wet weather attenuation may be exceeded by very large amounts. The additional attenuation is due primarily to the coating on the wires themselves rather than the coating on the insulators. It arises from the potential gradient through the ice deposit in combination with the high dielectric loss characteristic of the ice or snow coating. Figure 7 gives examples of the attenuation frequency characteristics of open-wire lines, including certain measurements with ice coating. The exact increase in attenuation due to snow and ice naturally depends on the thickness and other characteristics of the coating. Even very thin coatings of ice on the wires tend to raise the attenuation at 140 kc from the normal wet weather figure of about 0.28 db to about 1 db a mile, i.e., an increase of three or four to one. Extremes up to 5 db per mile have been measured for short lengths of line with ice nearly two inches in diameter. Such heavy ice obviously approaches the mechanical breakdown conditions for the line.

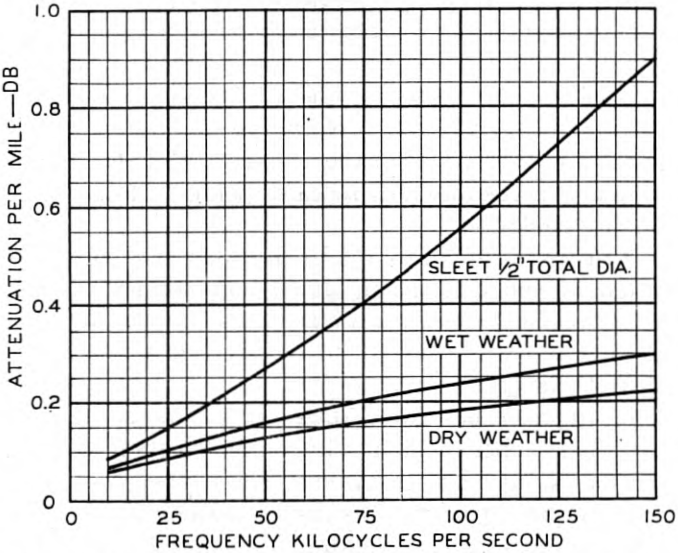


Fig. 7—Attenuation frequency characteristics of open wire lines.

Where ice and sleet occur the repeater spacings may be reduced to about fifty miles or less. The repeaters now being provided for the type J systems have gains of approximately 45 db. Repeaters are under development which are expected to raise the maximum available gain to something like 75 db. The normal dry or wet weather operation of such repeaters would be limited to gains of perhaps 10 to 25 db depending upon the amounts of cable included. The problem of obtaining automatic gain control over the extra wide range required by the high sleet attenuations is a difficult one.

REPEATERS

At each repeater point line filters and directional filters are required on both sides of the amplifying equipment to separate type J currents from those of lower-frequency services on the line and to separate oppositely directed groups for separate amplification in one-way line amplifiers. These filters have been described in connection with the terminals where they perform similar functions. Two regulating amplifiers, one for each direction of transmission, properly controlled to compensate for variations in the attenuation of the preceding line section, are also needed at each repeater point. These are described later under "Regulation."

Figure 8 shows the circuit of one of the line repeaters and indicates the location of the directional filters, and certain supplementary filters for suppressing frequencies outside the transmitted range; also the regulating amplifier circuit, and the pick-off of the pilot channel which controls the gain.

The line amplifier has three stages of pentodes. The first two stages use single tubes of high voltage amplification and low power capacity while the third stage has four power pentodes in parallel to increase the output capacity. Because of considerable heat developed by these power tubes, special precautions are necessary to dissipate the heat and to protect condensers and other elements mounted near them.

Negative feedback to improve the operation of the amplifier is supplied over two paths. The inner feedback, from the plates of the output tubes over a properly designed network to the grid of the input tube, reduces the gain at frequencies outside the transmitted band and so prevents singing at those frequencies. It has little effect at frequencies within the type J range. The outer feedback path includes the input and output transformers, which are made as hybrid coils. In each of these one pair of the conjugate windings is connected to the incoming or outgoing circuit of the amplifier while the other pair is used for the feedback connection. By feeding back through the transformers in this way, they benefit by feedback in much the same way as the tubes, and the overall characteristic of the amplifier is practically independent of the transformer characteristics. This feedback reduces the amplifier gain by over 40 db and correspondingly reduces modulation effects within the amplifier, and gives exceptionally stable transmission with respect to tube and voltage changes. It is also designed to improve and stabilize the input and output impedances.

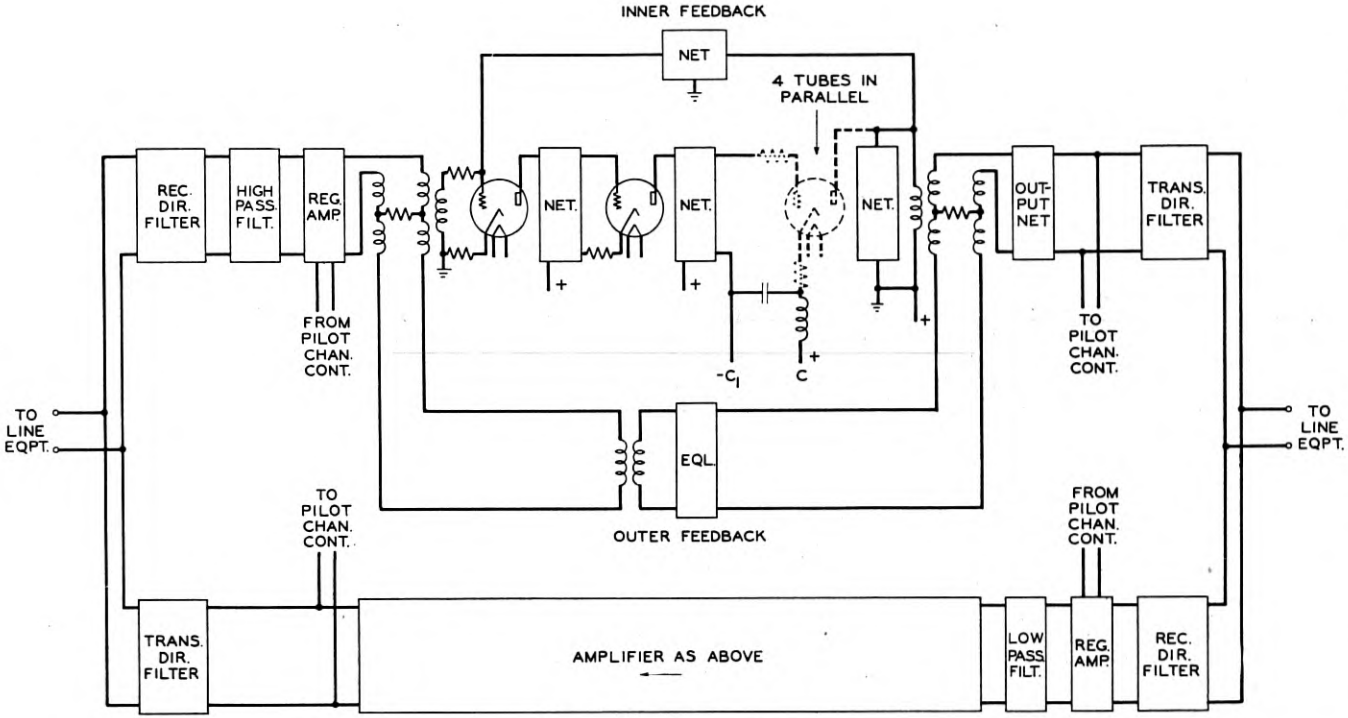


Fig. 8—Line repeater circuit.

EQUALIZATION

Equalization is necessary in each direction of transmission at a repeater point and in the receiving direction at a terminal, to compensate for frequency distortion produced by the preceding section of line. Fortunately, the attenuation frequency curves for the usual open-wire circuits, that is, 104, 128 and 165-mil wire, have nearly the same shapes for section lengths giving the same attenuation at the maximum frequencies for the two directions of transmission, so that these various circuits can be equalized alike.

As is well known, the transmission frequency characteristic of an amplifier with large feedback is almost the inverse of that of the feedback circuit itself, so that the insertion in the feedback circuit of a network having the same characteristics as a line section will provide equalized transmission over the amplifier and section combined. In the outer feedback circuit of the line repeater is included an equalizer which has a characteristic sloping with respect to frequency in the same way as the variation in loss under wet weather conditions of the longest open-wire section likely to be used. Thus, there is provided in the repeater a basic equalization for this longest wet weather line. At a receiving terminal a basic equalizer is provided which performs this same compensation, but in this case the slope of the curve must necessarily be opposite to that of the line attenuation and of the equalizer in the feedback path of the line repeater.

Line sections, however, vary in length and in the amount of entrance cable included. In order that they may be properly corrected by this basic equalization, they must be built out to equal this longest wet weather section. For this purpose there are provided flat loss pads and building-out networks whose losses have the same frequency shapes as the losses of short lengths of open-wire circuit. These pads and networks can be inserted or omitted by simple changes in strapping. They suffice to build out the shortest section which is expected to be used.

PILOT CURRENTS

For a satisfactory system, arrangements must be provided to correct automatically for the effects on line attenuation due to changes in weather, by adjusting the amplification at each repeater point and in the receiving terminal circuit. To permit measuring these effects a pilot current of fixed frequency, near the middle of the transmitted band, and of constant amplitude, is supplied from each terminal. This is applied to the transmitting side of the terminal circuit between the twelve-channel terminal and the first group modulator, where the message band lies between 60 and 108 kilocycles. The pilot frequency is 84.1 kilocycles which is obtained by modulation of 88 kilocycles, from one of the output taps of the channel supply of that frequency, with 3.9 kilocycles derived from a tuning fork oscillator. This modulation is performed in a copper-oxide bridge similar to the channel modulators and the desired product is selected by an 84-kc. carrier supply filter. The output of 84.1 kilocycles is sufficient to supply pilot current for ten terminals in the office. A sharply selective crystal band elimination filter is inserted between the output of the twelve-channel terminal and the point where the pilot source is bridged on the circuit to eliminate any current near the pilot frequency which would interfere with the small pilot current that is sent out to control the system.

The two group modulation processes alter this pilot frequency of 84.1 kilocycles so that it appears on the line as 59.9 kilocycles in the west to east directional band, and as 116.1 kilocycles in the east to west band. Correction in accordance with the magnitudes of these mid-group currents in the two directions is satisfactory over all twelve channels under ordinary conditions. In the case of ice or snow the channels at the edges of the directional frequency groups may not be properly adjusted. Additional pilot frequencies will probably be needed ultimately to care for such unusual conditions.

Regulating Amplifier

Figure 9 shows the circuit of the regulating amplifier, and above this, the circuit of the pilot channel receiving equipment which controls it. Current enters the regulating amplifier circuit from the left, coming from the receiving directional filter through a shielded transformer and the pads and building-out networks used for equalization. At the terminals the circuit includes also the basic equalizer. Last in the circuit leading from the line to the regulating amplifier is the regulating network which consists of a series of three equal networks having a total loss of 20 db at 140 kilocycles in the east to west direction and 15 db at 84 kilocycles in the west to east direction. The network loss increases with frequency in the same way as the difference between dry and wet weather attenuation of the line. The two terminals of the regulating network and the two junction points between the three networks are brought to four sets of stator plates on an adjustable condenser. The rotor of this condenser, which has about the same area as one set of stator plates, is connected to the grid in the first stage of the regulating amplifier. Rotation of the condenser therefore applies, to the grid of the first tube, a voltage which decreases continuously as the condenser rotates from left to right.

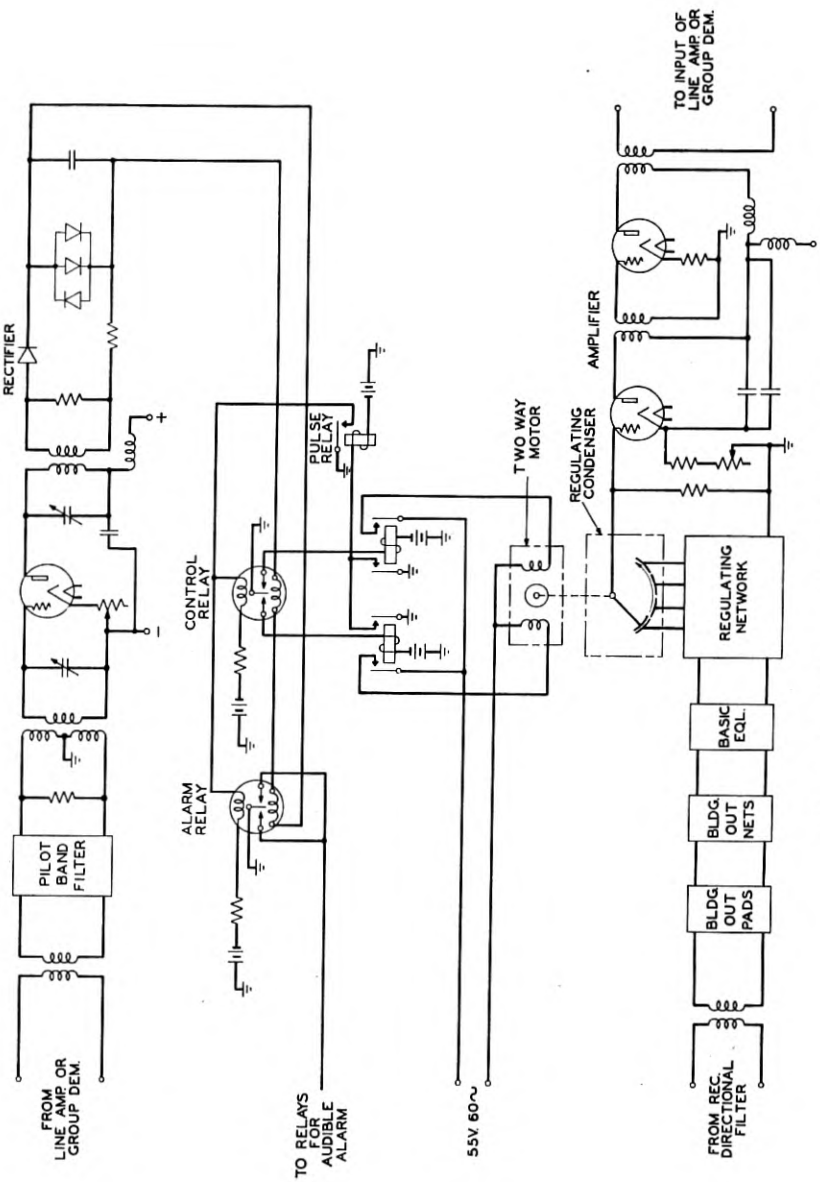


Fig. 9—Regulating amplifier and pilot control.

The regulating amplifier has two stages of pentode tubes, a high input impedance necessary for the proper operation of the condenser potentiometer, and feedback to stabilize the gain and to prevent intermodulation of the channels. Its output goes to the line amplifier at repeater stations, and to the first group demodulator at the terminals. At a west terminal there is interposed a high cut-off filter to eliminate frequencies above the upper band which may have been picked up on the open-wire line.

Pilot Control

The setting of the condenser which controls the regulating network is determined in accordance with the amount of pilot current flowing in the circuit in the direction concerned. At repeater stations the pilot current is picked off at the output of the line amplifier, being separated from the message transmissions by a quartz filter which has about a 30-cycle pass band. For control of transmission from west to east at the repeater stations, this filter selects 59.9 kilocycles and for control of the oppositely directed transmission, 116.1 kilocycles. At the terminals the pilot channel selecting filter is connected across the output of the auxiliary amplifier following the second group demodulator where the pilot frequency is 84.1 kilocycles. The pilot current from the pick-off filter is amplified in a single-stage amplifier which has feedback for constancy of operation and input and output circuits tuned to the pilot frequency. After amplification the pilot current is rectified by a temperature compensated copper-oxide rectifier.

The resulting direct current passes through the operating windings of the control and alarm relays. These Weston Sensitrol relays are, in fact, microammeters with high and low contacts made by the pointers. The mechanical bias of the moving system is adjusted so that with the normal pilot current the pointer will remain free in the middle between the two contacts. A change of about 0.5 db in this current will cause the pointer of the control relay to make contact with the terminal at the corresponding end of its swing. As the limiting contacts are magnetized and the pointer is of magnetic material, good contact is insured. When contact is made on one side a 60-cycle circuit is closed through the motor which controls the regulating condenser in such a direction as to cause the loss in the regulating network to be increased. Closure of the other contact similarly causes the loss in the regulating network to be decreased. Closure of either contact also closes a circuit containing a slow-operate "pulse" relay to release the Sensitrol relays after an interval of about four seconds. During this time the gain of the regulating amplifier will have been

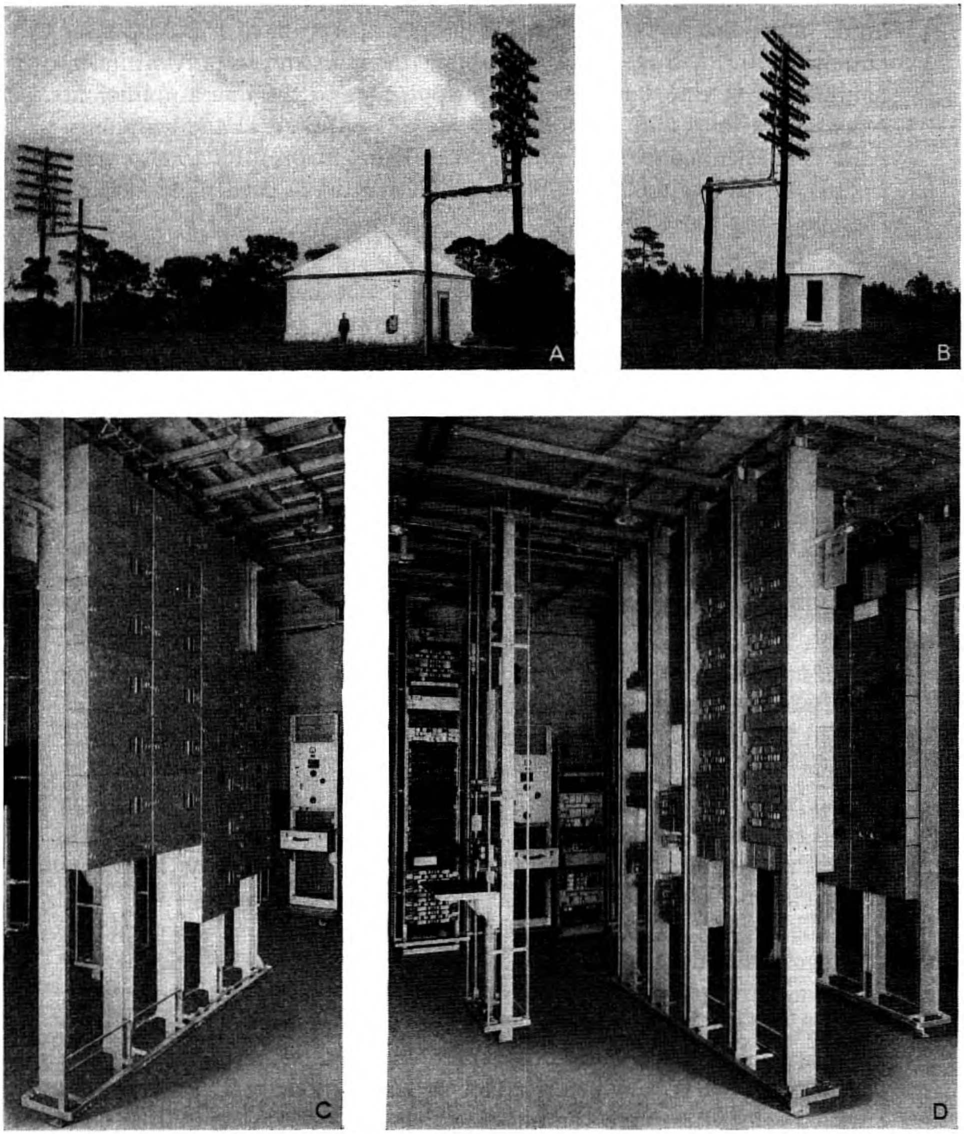
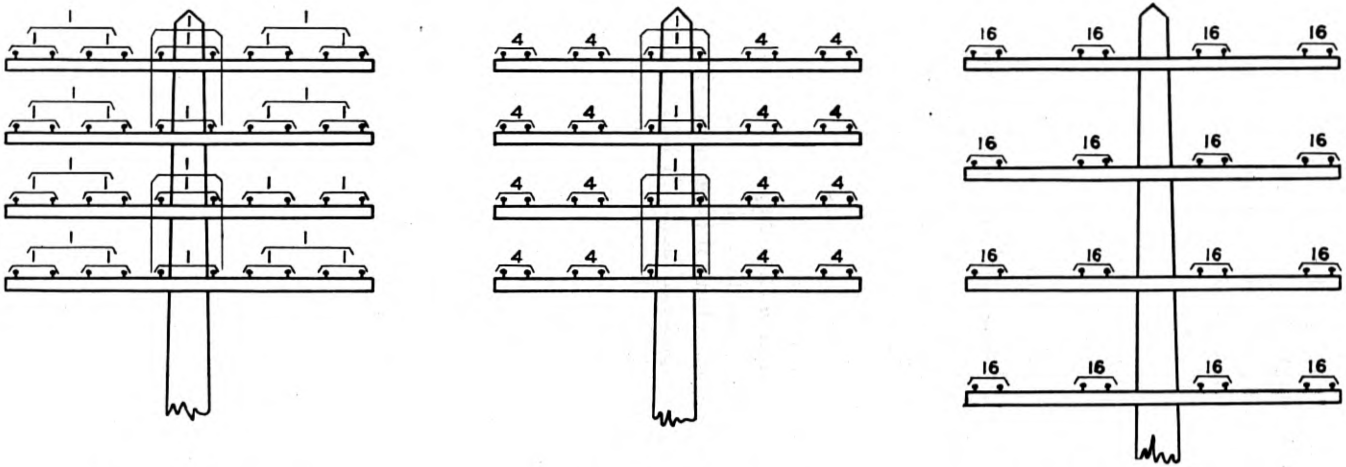


Fig. 10—Typical installations.
(A) Auxiliary repeater station.
(B) Cable hut.
(C) and (D) Terminal installations.



Facilities—30 voice circuits.

Line construction with 8″ spaced non-phantomed non-pole pairs. Type C systems on all 8″ spaced pairs. Facilities—22 voice circuits, 48 carrier circuits, total, 70

New line construction all 8″ spaced pairs, no pole pairs, crossarms spaced 36 inches instead of 24 inches. No phantoms. Facilities—16 voice circuits, 240 carrier circuits, total, 256

Fig. 11—Growth in line carrying capacity.

changed about 0.1 db. If now the pilot current level is within 0.5 db of normal the operation is complete. If not, it is repeated and the device keeps periodically testing the circuit so long as it is away from satisfactory compensation. There are also alarm circuits for attracting attention in cases of wide variations in equivalent. In severe ice conditions where a single regulating repeater has not sufficient gain to make up for the great loss in the line, the next succeeding repeater will do its utmost to make up the deficiency.

Conclusion

In what has preceded, developments have been described which are making it possible to provide a very substantial increase in circuits on open-wire pole lines without additional wire stringing. Photographs showing typical office installations of type J carrier equipment, unattended repeater stations, and filter huts are shown in Fig. 10.

Three stages in the development of the open-wire line over the past twenty years, giving successive increases in circuit capacity, are shown in Fig. 11. Prior to the application of carrier systems, a four-crossarm pole line would yield thirty voice circuits. Now, on a new line 256 circuits are potentially obtainable. Thus it is probable that the open-wire line will continue as an important factor in furnishing facilities in moderate numbers, particularly in the less densely populated sections of the country and where climatic conditions are not unfavorable. Installations of type J systems have already been made in various parts of the country.

Recent Developments in the Measurement of Telegraph Transmission

By R. B. SHANCK, F. A. COWAN and S. I. CORY

This paper describes the progress which has been made in recent years in the development of methods and apparatus for the measurement of telegraph transmission in the Bell System. Such measurements play an important part in transmission maintenance work in the field and are also necessary in development work. The changes which have occurred in service requirements, particularly the large commercial development of start-stop teletypewriter service and the effect of these changes on the technique of telegraph transmission measurement, are first discussed; then a description is given of several new measuring devices and their use.

IN keeping with advances in the telegraph transmission art, noteworthy improvement has been made in measuring devices and methods in the past few years. The faster, more accurate, and generally more dependable telegraph service now available has been made practicable not only by improvements in the telegraph systems but also by the use of improved measuring apparatus and techniques.

In the early stages of development most transmission-measuring systems were arranged to measure transmission on "looped" circuits, that is, with sending and receiving terminals at the same point, so that a comparison between the sent and received signals could be made. Although such an arrangement is quite useful for laboratory testing, it imposes serious limitations on field testing. Therefore, it is generally desirable to make tests on a straightaway basis.

For straightaway tests it is necessary to have at the receiving end certain information regarding the sent signals. This requires either the use of signals of certain known characteristics or the determination of the important characteristics of the sent signals and the transmission of this knowledge to the receiving end. The latter of these alternatives is not frequently used since it requires another communication channel, although in some instances, where tests on working circuits are desired, it represents the only practicable approach. Straightaway measurements of telegraph transmission have for this reason been generally confined to measurements in which the important characteristics of the transmitted signals were known.

Fortunately, either synchronous or stop-start teletypewriter signals fall into the last-mentioned classification and the increased use in the Bell System of stop-start teletypewriter transmission has afforded the opportunity of making telegraph transmission tests on working circuits without the need of transmitting information regarding the character of the sent signals. Of course, when sectionalized transmission measurements are desired it is necessary for proper interpretation to transmit the results of the measurements to a single point for analysis, but the communication of this information is not at all burdensome.

Aside from the ability to measure on a straightaway basis, which is primarily a field-maintenance requirement, testing equipment should preferably be direct-reading without the need of measuring adjustments on the part of the tester. Direct-reading devices in general effect considerable reduction in the time required for measurement. This feature is especially important in the field where a rapid test of possible trouble conditions is desirable and in the laboratory where the large numbers of tests which are necessary for thoroughly checking a telegraph transmission system under all of the likely operating conditions become even at best tedious and time-consuming.

The major development in telegraph transmission testing within the past few years has been the provision of instrumentalities which possess the desirable properties indicated above. They permit the rapid and direct reading of signal distortion on working teletypewriter circuits. The same instrumentalities when used with selected or miscellaneous teletypewriter test signals also permit the rapid determination of the capabilities of a telegraph circuit in the field or in the laboratory.

A paper published in 1927 [1] discussed fundamental concepts relating to signal distortion and described a number of measuring devices which had been employed in the Bell System. Another paper [2] treated the design of telegraph circuits for distortionless transmission from the standpoint of the steady-state characteristics. The fundamental ideas set forth in these papers have continued to form the basis for development of the technique of measuring telegraph transmission. However, it has been necessary better to adapt these ideas to start-stop teletypewriter operation and changing field requirements.

Operation of telegraph circuits by means of start-stop teletypewriters [3, 4] using 7.4-unit code has become of much greater importance in the Bell System in the past dozen years. The majority of private-line service is now furnished on a teletypewriter basis; also teletypewriter exchange (TWX) service,[5] inaugurated in 1931, has become an

---

[1] References are listed at end of paper.

important factor, having already grown to the point where the trunk-circuit mileage employed is a large part of the total telegraph mileage. Incidentally, there has been at the same time a general increase in operating speeds, so that the majority of the circuits now operate at a nominal speed of 60 words per minute (23 dots per second or 46 bauds).

As regards the requirements for measuring apparatus for field transmission maintenance, the desired precision and convenience have increased considerably in the last few years. This is due to several causes, chief of which are the continuing desire to give better service with greater freedom from interruptions and isolated errors, increase in speeds of operation, and the use of more complicated circuit layouts with more sections in tandem, particularly in Press and TWX service. For complicated circuits it is very advantageous to employ maintenance procedures in which each section is measured and adjusted separately to close limits, to avoid the more costly and otherwise less desirable overall line-up. Furthermore, a need has arisen for accurate transmission measuring devices for other uses such as checking the condition of receiving teletypewriters, transmitting keyboards and regenerative repeaters,[4] and use in "equalizing" of telegraph circuits, that is, the application of wave-shaping arrangements for reducing distortion. Finally, in line with improvements in main-line circuits greater emphasis has been placed on maintaining loops and circuits to outlying points so that they introduce but little distortion.

Adaptation of Measuring Technique to Teletypewriter Basis

The earlier types of measuring sets were arranged to measure the total change in the duration of signal pulses, that is, the combination of the displacements at the beginning and end of any given pulse. This method of measuring gives results which are directly indicative of the impairment for Morse operation since the interpretation of the signals depends on the total duration of pulses. This method also gives a moderately good indication of the effect of distortion for teletypewriter operation.

In start-stop teletypewriter operation there are two ways in which circuit imperfections may cause the transmission to be impaired. In the first place, imperfections other than constant delay (known as line lag), which may be neglected, may cause the start transition of any character to be displaced with respect to the time at which it should occur. This causes the starting of the receiving mechanism to be advanced or retarded and effectively displaces the succeeding transitions of the character. Secondly, other imperfections may also cause any of the succeeding transitions to be displaced in either direction. The combination of these two effects determines the effective distortion.

It is of interest to consider the case of bias alone. With uniform bias the displacement of mark-to-space transitions is in one direction and that of the space-to-mark transitions is in the other direction with respect to their positions in undistorted signals. Since the start transition is mark-to-space the result is that the effective displacement of subsequent mark-to-space transitions is zero and the effective displacement of space-to-mark transitions is numerically equal to the bias. In practice bias is seldom uniform and may vary with the signal combinations. In these cases there is an effective displacement of mark-to-space as well as space-to-mark transitions.

From the foregoing, it will be seen that it is of considerable practical value to be able to measure teletypewriter circuits on a start-stop basis in terms of displacement of transitions with respect to the start transition. (For a more complete explanation of the effect of distortion on teletypewriter operation, reference should be made to published discussions.[4, 5, 7]) In testing with miscellaneous signals, bias may for convenience be taken as the average effective displacement of space-to-mark transitions relative to mark-to-space transitions; characteristic distortion will have the appearance of a combination of fortuitous and bias effects; and the maximum total distortion will be the sum of the average effect and the variation therefrom which causes the greatest displacement.

OTHER CHANGES IN MEASURING TECHNIQUE

In measuring with normal and inverted signals[1] on circuits of the types commonly employed, the result obtained for the bias varies somewhat from pulse to pulse of the test signal. A case in which this variation is appreciable is that of carrier telegraph having level compensators (automatic devices which correct for changes in the magnitude of the received current). With these compensators, the response is fairly rapid, the result being that the bias is to some extent a function of the signal combinations of the transmitted material. This bias variation is also noticeable with open-and-close d-c. telegraph circuits having large bridged capacitance or series inductance.

On account of this bias variation, it is desirable to measure the algebraic average of distortion of the individual pulses of miscellaneous signals and take this as the bias. This may be conveniently done by measuring on the start-stop basis mentioned above. In such measurements the differences between the distortions of the individual pulses and the average distortion may be considered as due to the combination of characteristic and fortuitous effects; further measurement is necessary in order to separate these effects. A measure of characteristic distortion may be obtained by determining the difference between the systematic distortion measured with selected recurring signals and the average distortion with miscellaneous signals. Fortuitous distortions may be taken as the difference between the total and systematic distortions obtained with recurring signals. In order to distinguish between the variable bias and the true characteristic distortion in the case of level-compensated circuits measurements may be made with the compensator disabled and with it functioning.

For tests in the field where it is desired to measure systematic distortion effects, as for instance in connection with equalizing, several simple signals corresponding to certain especially selected teletypewriter characters are employed. In each of these signal combinations there are only two transitions; therefore it is convenient to observe the effect of the remnants of one transition upon the next transition. As discussed in the Appendix, the characteristic distortion obtained for miscellaneous signals is a function of the distortion obtained with the simplest characters; if there were no distortion on the simple characters no characteristic distortion effects would be expected when miscellaneous signals were transmitted. The process of equalization, therefore, consists in adjusting the transmission characteristics of the line circuit to reduce the characteristic distortion measured on six special characters to a minimum. The teletypewriter characters which are used for this purpose are Blank, *T*, *O*, *M*, *V* and Letters; the corresponding signals are shown in Fig. 1. The distortions of these signals are observed at the receiving end on a measuring set operating on the start-stop principle or a portable systematic-distortion measuring set having an integrating meter, as will be described more fully below.

In equalization testing, each of the six signals is sent repeatedly for the time required to determine the total systematic distortion—generally about 50 repetitions. In analyzing the results the bias component is assumed to be substantially the same for all of these signals and whatever difference is observed from one signal to another is taken as being due to characteristic distortion. It is found generally that the result for the *O* signals, which are practically unbiased six-cycle reversals (see Fig. 1), is not radically different from the result obtained for bias with reversals at 6 or 23 d.p.s. (shown in the lower part of Fig. 1), the results being expressed, of course, in the same terms, as for instance in per cent of a 23-cycle dot. The largest distortion is usually found on either the Blank or Letters character, this being reasonable because usually the remnants of transients practically disappear within a few dot lengths. Sometimes it is

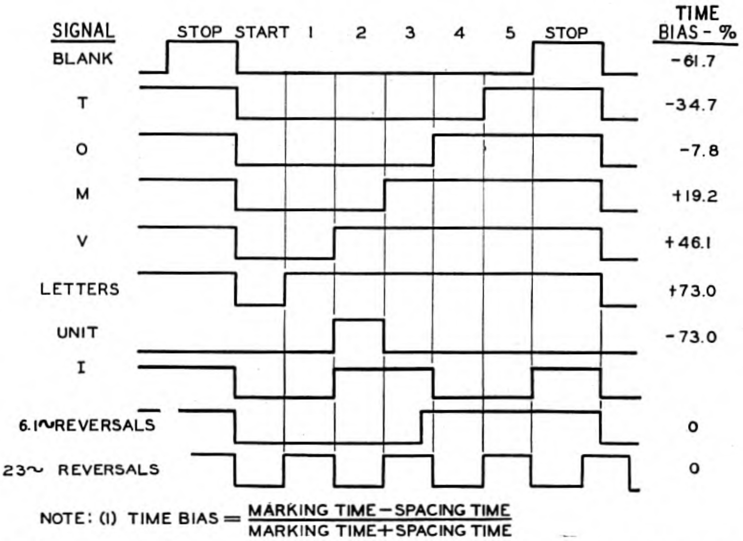


Fig. 1—Signals used in measuring systematic distortion. The first six signals and the eighth signal are teletypewriter characters.

desirable to extend the test to include biased test signals, as described below.

As an example, curves are given in Fig. 2 showing the results of tests with the six characters on a d-c. metallic telegraph[6] circuit operating on 112 miles (180 km.) of composited 19-gauge cable pair. Curves 1 and 2 show the results before and after equalization respectively; it will be noted that considerable improvement was effected. On the basis of both experiment and theory, the slope of Curve 1 is known to indicate that the received direct current is larger than it

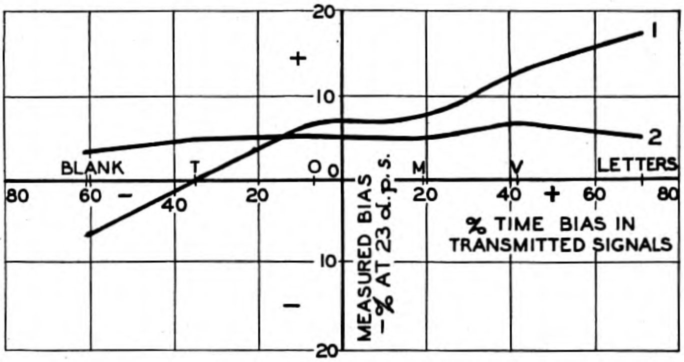


Fig. 2—Equalization test result for a 112-mile section of metallic telegraph; signalling speed 23 d.p.s. Curve 1, before equalization; curve 2, after equalization.

should be as compared to the higher-frequency components of the received current waves. The required type of equalizer in this case is one which adds loss at frequencies in the vicinity of zero, without a corresponding loss at the higher frequencies. If, however, the slope had been in the opposite direction, an equalizer which would discriminate against the higher frequencies would have been required. As a check of the equalizer setting, the total distortion and bias are usually measured using miscellaneous teletypewriter signals.

In addition to measuring with undistorted signals applied at the sending end of a circuit, the measuring technique has been expanded to include measuring with distorted signals and a device has been made available for field use by means of which reversals or teletypewriter characters may be distorted by known amounts. This kind of test furnishes additional information in that it affords an examination of the effect of signal combinations which are not included in perfect telegraph signals. It is of value because in actual operation a given telegraph section may not have perfect signals impressed at the sending end due to distortion occurring in previous sections or at the transmitter. Although such a test furnishes valuable information for line testing, it has been used in the field up to the present mainly in testing the distortion-tolerance of subscriber-station teletypewriters with signals from the adjacent central office and in maintaining regenerative repeaters.

It is necessary, of course, to make transmission measurements on the manual Morse circuits which still constitute a considerable part of the total mileage. Testing such circuits by the same methods as used for teletypewriter circuits has been found to give good results. However, due to improvement of telegraph circuits, the transmission-maintenance problem in this case consists mainly of keeping the bias within reasonable limits for which purpose simple tests with reversals can be used.

New Measuring Devices

In the following is given a description of a number of testing methods and arrangements which have been found useful in recent years both in the field and in development work.

A. Start-Stop Distortion-Measuring Set for Central-Office Use

A start-stop type of measuring set for testing teletypewriter circuits has been developed and is now used generally in maintenance work and special testing in the field and in laboratory work. This represents an outstanding advance in that it provides a quick and convenient means for reading, directly from conventional-type milliameters as illustrated by Fig. 3, the distortion of miscellaneous teletypewriter signals on working circuits. It has the distinct advantage of giving immediate indication of the occasional isolated peaks as well as the average distortion. In using this set, it is unnecessary to have a knowledge of the transmitted text.

This set employs the condenser-charging principle in the measurement of small time intervals corresponding to the distortion of the signals. The outstanding feature of this device is that the beginnings of the condenser-charging intervals are timed with relation to the start transition by a start-stop distributor which forms an integral part of the set. The charging intervals are terminated by the occurrence of transitions in the characters, at which times the operation of a receiving relay causes the condenser voltages to be compared to a reference voltage. The circuit is arranged to charge the condenser at a constant rate; hence the voltage attained is determined by the duration of the charging interval. In this way the displacements of the transitions

Fig. 3—Meters and control apparatus at telegraph board.

in the received teletypewriter characters from their proper positions are measured in terms of condenser voltages. Indications are afforded of the average distortion and the peak value of the total distortion, the latter being the sum of the bias, characteristic and fortuitous effects.

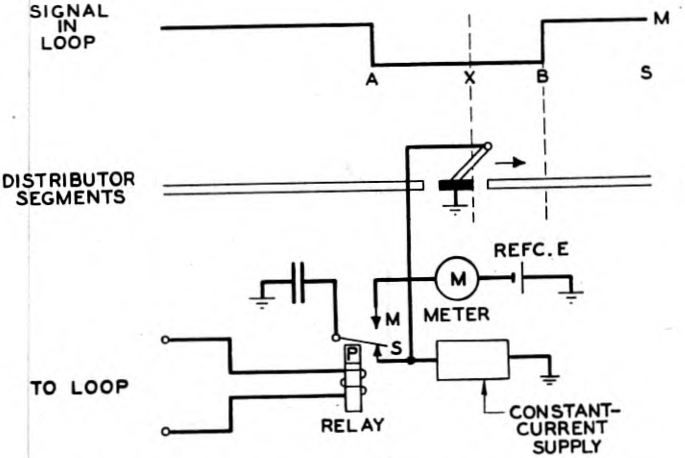


Fig. 4—Explanatory sketch of start-stop telegraph transmission measuring set.

The general features of this set will be described in the following. In Fig. 4 a condenser-charging circuit is shown with a distributor for the purpose of timing the charging intervals in the measurement of distortion occurring at transition *B*. Assume the distributor brush to be traveling in the direction of the arrow after being released by the mark-to-space transition at *A* by means of arrangements not shown. This same transition causes the relay armature to move to the spacing contact (*S*) and the condenser to begin to charge from the constant current supply, but as soon as the brush touches the grounded segment the condenser is discharged completely. After leaving the grounded segment the brush travels over an open segment and during this time the condenser accumulates a charge. At transition *B*, the armature of the relay moves to its marking contact (*M*) and the voltage of the condenser is compared with the reference voltage (REFC. *E*) which has been previously adjusted to such a value that if there is no distortion the condenser voltage and the reference voltage will be equal. However, if transition *B* does not occur at the proper time, because of distortion, the condenser voltage will differ from the reference voltage and a momentary current will flow through the indicating meter (*M*) in proportion to the amount of distortion.

Two condenser-charging circuits are provided, one for space-to-mark transitions and the other for mark-to-space transitions as is indicated in Fig. 5. Graph *A* of this figure shows an undistorted teletypewriter character having a stop pulse, start pulse, and five selecting pulses. The segments of the distributor, laid out to show their relation to the received signal, are indicated at the top of the schematic circuit (*D*). It will be noted that there are seven short grounded segments, one for each of the pulses of the character, for initiating the condenser-charging intervals referred to above.

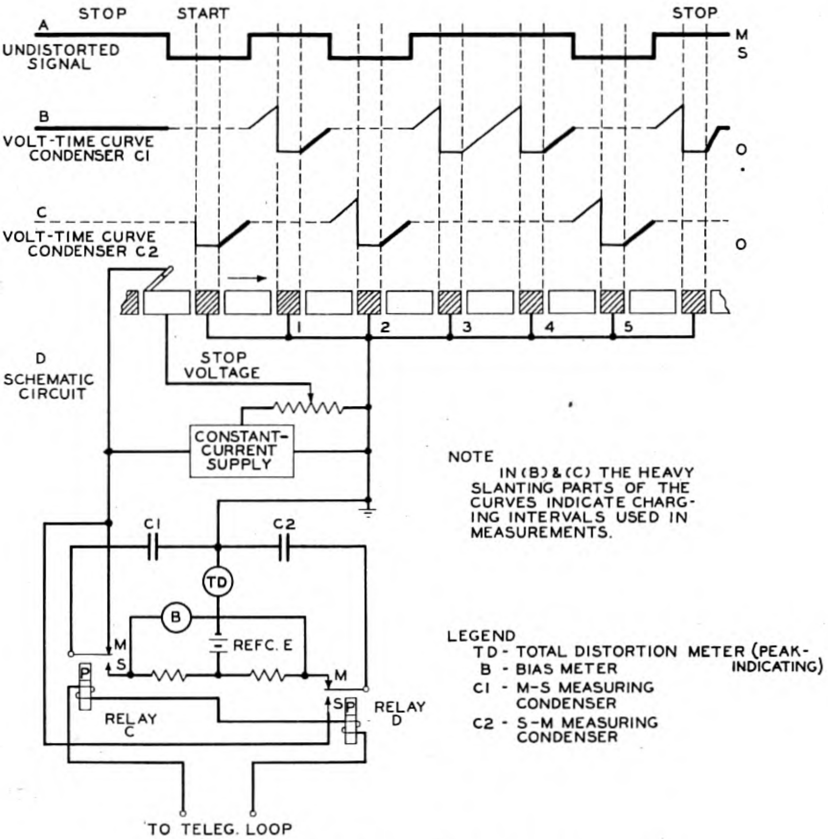


Fig. 5—Simplified diagram of start-stop telegraph transmission-measuring circuit.

Assume the distributor brush to be at rest during the stop interval as shown. The measuring condenser *C*1 is now charged to the reference voltage. As the brush leaves the stop position due to the mechanism controlled by the start pulse (not shown), it travels in the direction of the arrow and the condenser charges vary as indicated by graphs *B* and *C*. It will be seen on graph *B* that condenser *C*1 is charged during intervals between grounded segments and mark-to-space transitions. Graph *C* shows that condenser *C*2 is charged between the grounded segments and space-to-mark transitions. If there is no transition while the brush is traveling between two adjacent grounded segments the condenser continues to be charged until the brush touches the second grounded segment at which time it is completely discharged. Therefore, the useful charging interval is that between a grounded segment and a transition occurring before the next grounded segment is traversed by the brush. These intervals are indicated by the heavy-lined portions of the graphs. They amount to 37.5 per cent of a unit pulse as a maximum, i.e., the maximum distortion which can be measured is about 37.5 per cent. The currents flowing as a result of the comparison of the condenser voltages with the reference voltage are indicated on a "Total-Distortion Meter" *TD* which is, in reality, a peak-indicating voltmeter, and on a "Bias Meter," *B* which is sufficiently sluggish to give an indication corresponding to the average distortion. These meters are calibrated to indicate directly the percentage distortion with miscellaneous teletypewriter characters.

Good accuracy is obtained with these sets; when measuring distortions of small or moderate amounts with a well-adjusted set, the indication is accurate to within about 2 per cent distortion at 60 words per minute. For occasional large distortions or for higher speeds, the accuracy is not quite as good, and there are certain possible mutilations of signals, such as the dropping out of pulses, which would not be readily detected.

In addition to measuring miscellaneous teletypewriter characters these sets may be used with recurring test signals in which the spacings of the transitions are such that the maximum characteristic effects will be obtained, and with signals which experience mostly bias and fortuitous effects. In this way a measure of the components of distortion may be obtained. Such tests are commonly made in adjusting variable networks to minimize characteristic distortion, i.e., making the equalization tests, referred to above, with selected teletypewriter signals.

In special testing where it is desired to separate the total distortion into its components, this may be done by measuring the systematic distortion with the first 6 signals of Fig. 1 and then measuring the bias and total distortion using the *I* character of Fig. 1 (which is substantially the same as unbiased reversals at about 11 d.p.s.). The bias component is simply the bias measured with *I* signals, the fortuitous component is taken as the difference between the total distortion and the bias of *I* signals, and the characteristic component is obtained by averaging the results for the 6 selected characters and then selecting the result which shows the maximum departure from the average.

These sets also furnish a convenient means for the measurement of mean-square values of distortion. The current impulses flowing through the total distortion meter are proportional to the distortion and it is practicable to insert in series a specially arranged meter which is calibrated to indicate the mean-square values of the distortion. The circuit used is shown by Fig. 6, the heavy lines showing the meter circuit which is inserted in series. It contains an integrating meter whose characteristic is modified by a full-wave copper-oxide rectifier, a large condenser for increasing the damping and resistances for adjusting the amplitude and response characteristics. Measurements of mean-square distortion have been found of value in connection with producing transmission ratings of telegraph circuits, these ratings being based on the assumption that the mean-square values of distortion of component parts of a circuit may be added directly to predict the total mean-square distortion.[5]

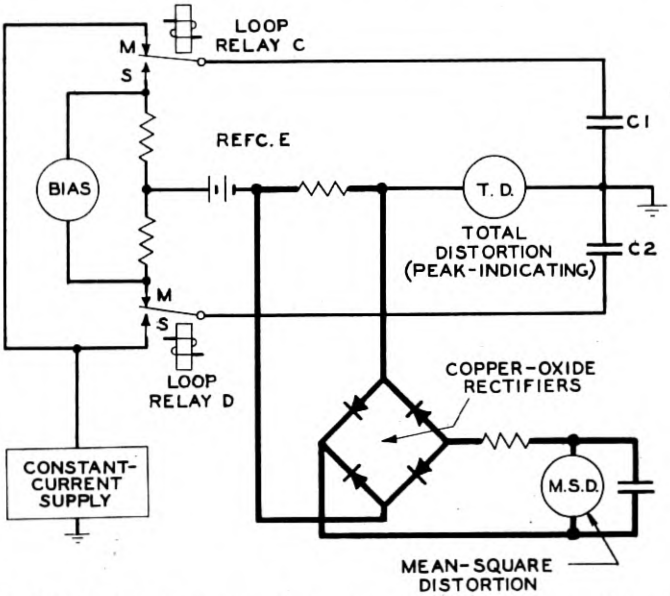


Fig. 6—Circuit of start-stop set for measurement of mean-square distortion.

In certain other special tests, where it is desired to obtain a record of the variation in distortion, recording meters have been connected in series with the meters of the set and a continuous record made using either a recurring test message or the signals from the subscriber.

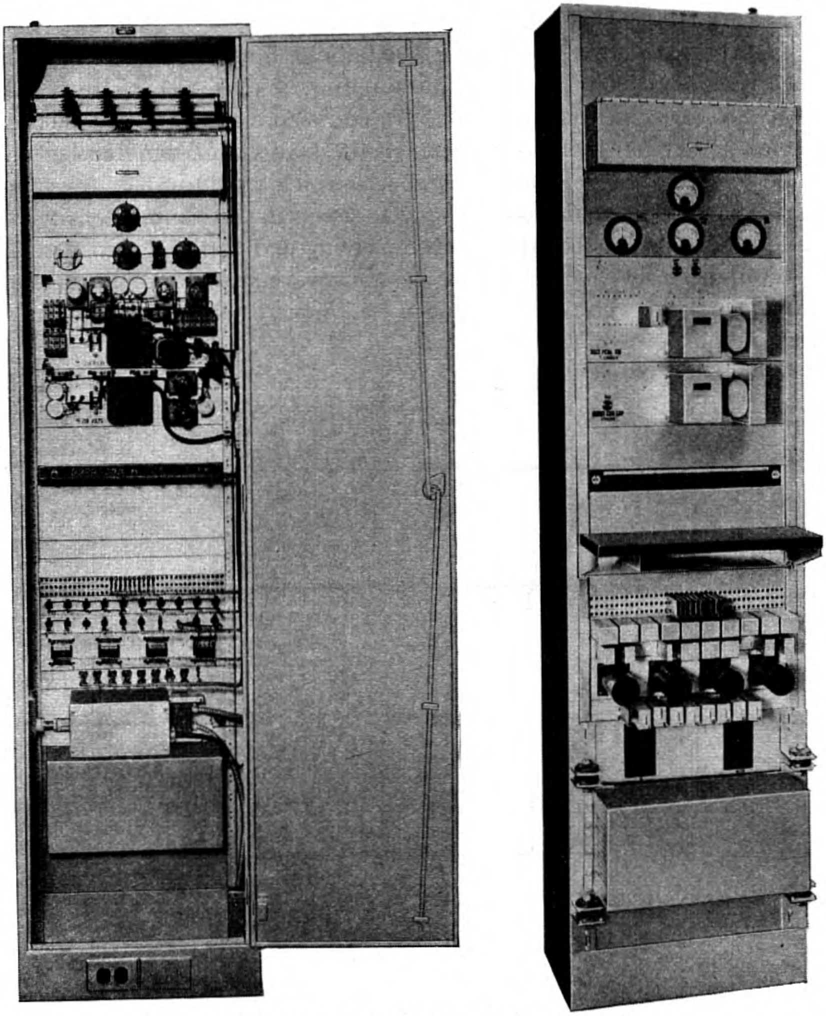


Fig. 7—Start-stop telegraph transmission measuring set.

An idea of the arrangement of the measuring set for central office use may be obtained from the front and rear views which are shown in Fig. 7. It is the practice to provide multiple appearances at telegraph boards in order that one set may be used at any one of a number of positions. The unit containing the indicating meters and associated controls for mounting at the telegraph board is illustrated in Fig. 3. The set is also provided in portable form for temporary use in cases where a permanent installation is not justified.

B. Telegraph Stability Test Set

Recording meters have been used for a number of years in the measurement of the transmission stability of telegraph circuits.[1] In such a measurement a continuous graphic record is made over as long a period as desired of the variations in the bias and of the number and time of occurrence of fortuitous effects which would impair telegraph service. For this purpose telegraph reversals are impressed at the sending end and a recorder at the receiving end makes a record of the bias of these reversals. This type of test was found to be of such utility in field work that standard stability test sets were produced for this purpose.

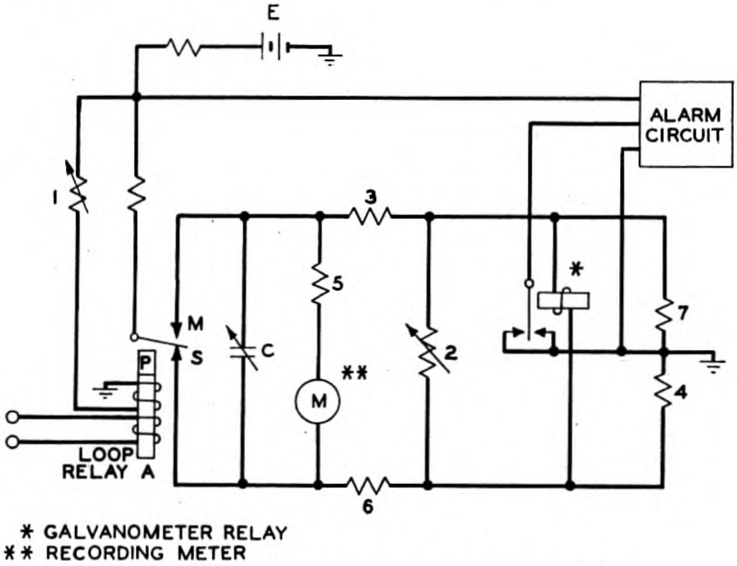


Fig. 8—Schematic circuit of telegraph stability test set.

The circuit of the telegraph stability test set is indicated in Fig. 8. This is essentially a simple bias-measuring circuit combined with an alarm circuit to indicate when given values of bias are exceeded. Movement of the armature of loop relay *A* from *M* to *S* contact or vice versa in response to the received signals causes a battery *E* to be connected first to one arm (resistances 6 and 4) and then to the other arm (resistances 3 and 7) of a bridge type of circuit containing recording meter M. The two arms of the bridge are balanced and the meter, being bridged across them, receives positive and negative current pulses of equal magnitude in response to the armature movements. If these pulses are of equal duration, as for telegraph dots or reversals having zero bias, the average meter current will be zero. Biased reversals will cause the meter current to average at other than zero by an amount directly proportional to the percentage bias. A center-zero recording meter is used and this provides a running record of the variation in bias. A damping condenser C is used to reduce the width of the trace and the amount of unsteadiness of the indication due to fortuitous effects.

The alarm circuit contains a galvanometer-relay bridged across equal resistances 4 and 7. The needle of the galvanometer-relay moves to one contact or the other when excessive values of bias are experienced, the sensitivity being adjusted for response to different values of bias by means of adjustable resistance 2. The response is made somewhat sluggish to avoid alarms being given for interruptions of short duration which are not of interest in connection with an investigation of bias stability.

A sample chart obtained by means of one of these sets is shown in Fig. 9. This chart shows slow variations in bias in the upper part and in the lower part the change in the indication due to dropping out or adding a single dot and momentary failures. Such charts do not, of course, show the characteristic distortion, since this is not present in the case of unbiased reversals.

These sets are now generally used in the field in routine checks on telegraph circuits and in special checks on circuits which have developed faults in service. Usually these checks are made with the idea of locating the cause of hits or swings which it is difficult to locate otherwise; in some cases sets are used simultaneously at several repeater points to sectionalize trouble. Charts are run for long periods, sometimes for several weeks in such tests. The stability test sets are also used to obtain data for transmission ratings of circuits, in which case it is desired to know the extent of the bias variations over long periods and the number and frequency of occurrence of hits.

Figure 10 shows a view of the portable arrangement of the set including the recording meter. The alarm buzzer and the receiving relay are located on the panel along with a row of keys for adjusting the alarm for operation on given values of bias. Jacks are provided on the side of the box for connection to circuits, batteries and the

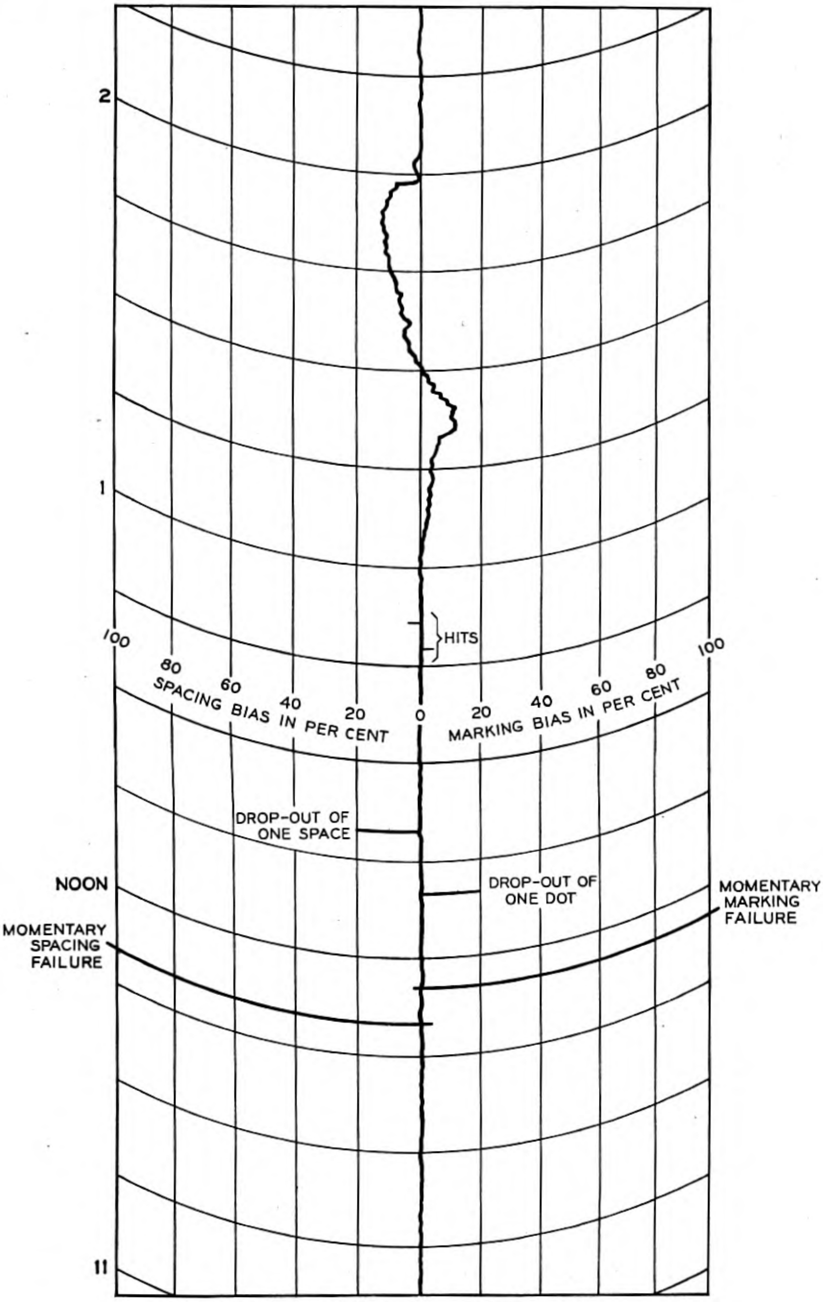


Fig. 9—Telegraph stability test—sample chart.

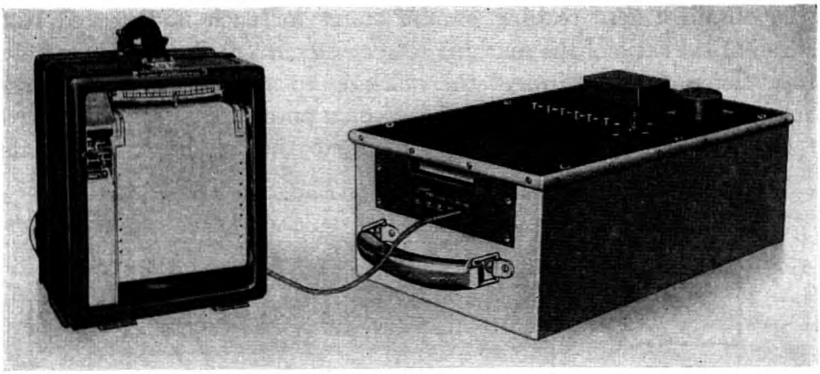


Fig. 10—Stability test set with recording meter.

recording meter. This set is also arranged for permanent mounting on a relay rack in a central office.

C. Portable Measuring Set for Use at Outlying Stations

A new measuring set has been made available for commercial use especially at outlying stations in either routine or special testing. This set is arranged for the accurate measurement of the systematic components (bias and characteristic) of distortion of recurring test signals. In addition measurements may be made of peak values of interference and of the effect of bias and variations in operating currents on the distortion. Such measurements are desirable in analyzing the causes of transmission troubles and in equalization work. By properly interpreting the results of these measurements a fair idea may be obtained of the maximum total distortion. In addition the set is arranged for the convenient measurement of the operating currents and voltages in various parts of the subscriber's circuit and in external circuits. This set may be used on either 110-volt a-c. or d-c. commercial power supply. It is mounted in an aluminum case, and weighs only about 28 lbs. (13 kg.).

The circuit employed for the measurement of systematic distortion is indicated schematically by Fig. 11. This is a simple bridge type of circuit similar to that generally used in a measurement of bias with reversals but especially arranged to indicate directly the percentage systematic distortion of the signals of Fig. 1 excepting "I" signal. The meter circuit is highly damped to prevent undesirable vibration of the meter needle in response to the 6-cycle fundamental frequency of the 60-speed teletypewriter characters. This entails the use of high resistances R, and large capacitance C. Since the voltage in the power supply is limited, a very sensitive meter M is required.

To obtain a zero reading on the meter for each undistorted test character the ratio of the marking to spacing currents is made inversely proportional to the ratio of the marking to spacing time intervals. This is accomplished by means of taps on a potentiometer as indicated in Fig. 11. With the connection made to the middle or Reversals tap, equal and opposite currents flow through the meter when the relay armature rests first on its marking and then on its spacing contact. If the relay is repeating undistorted reversals, the time bias in the signals is zero and the average meter indication will be at zero. With a given undistorted recurring character such as Blank, and with the potentiometer set at the Blank tap, the meter indication will again average at zero. The meter circuit is arranged so that the systematic distortion is indicated directly in percentage based on the duration of a unit signal element of a teletypewriter character.

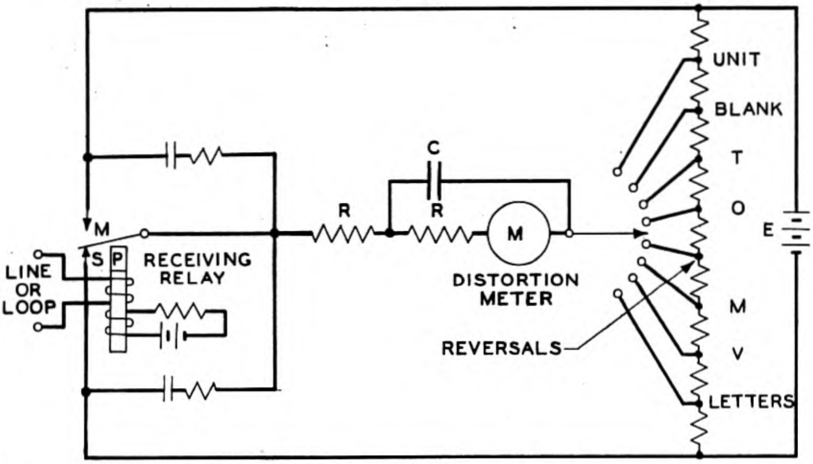


Fig. 11—Outlying-station test set. Schematic circuit for measurement of systematic distortion.

The circuit used in the measurement of interference is indicated by Fig. 12. The interfering effect is measured by noting on the meter M the biasing current which will just prevent the armature of the receiving relay from responding to the interfering currents. Movement of the armature from its contact is indicated by a response in a telephone receiver connected in the armature circuit when the switch is operated as indicated. The biasing current variation is effected by means of potentiometer P and the biasing current may be reversed by means of a switch (not shown).

The circuit of Fig. 12 may also be used to give an indication as to the amount of transmission degradation to be expected due to the interference measured and to given changes in operating currents. In this case the switch is operated to connect meter *M*1 in circuit for the purpose of measuring bias. The effect on the bias of changes in the biasing current for a given operating condition may then be

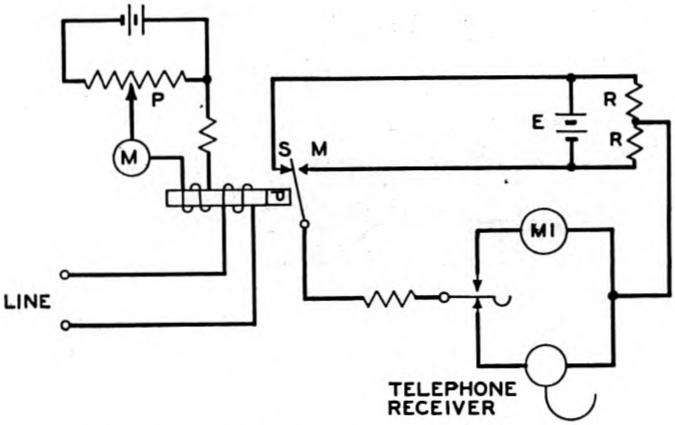


Fig. 12—Outlying-station test set. Schematic circuit for interference measurement.

noted on this meter when receiving reversals or any of the selected teletypewriter characters. The biasing current is varied by means of potentiometer *P* as before.

This set has several interesting operating features. By inspection of Fig. 13, which shows a view of the set, it will be seen that it contains its own receiving relay located in the lower left corner. This relay may be connected in series in a line circuit or in the local circuit of the subscriber set using the line jacks 1 and 2 (located in the upper right-hand corner of the set) or if desired convenient connection may be made by means of the splecial plug and adapter shown in the lower part of the figure. In the latter case the subscriber set relay is transferred to the measuring set and the special plug, and adapter if required, is inserted into the relay connecting block of the subscriber set in place of the relay. By operating the proper keys on the test set the currents in different circuits of the subscriber set may then be measured conveniently. Distortion may also be measured with this connection but in most cases it is desirable to measure with the receiving relay in its normal position in the subscriber set to obtain representative conditions.

Because of the advantages of the set, namely, the accuracy possible in the measurement of distortion, the convenience afforded in the

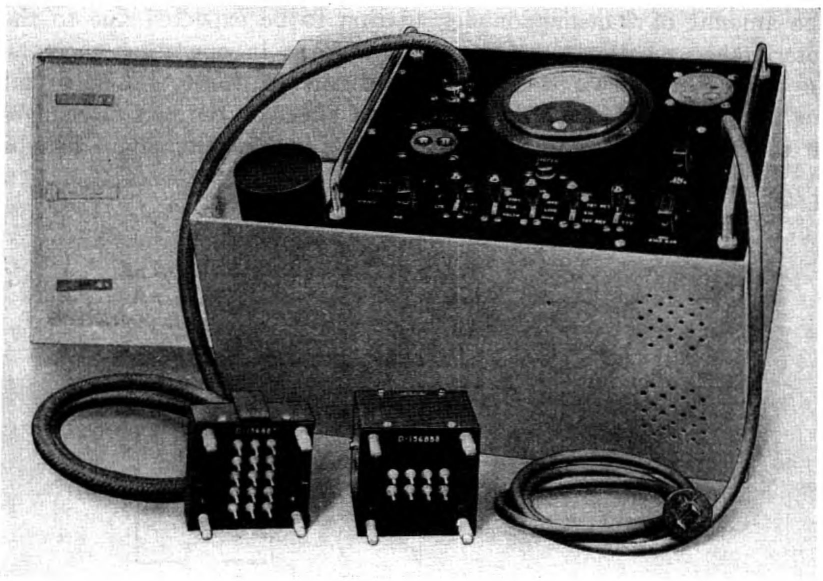


Fig. 13—Outlying-station test set.

measurement of operating and interfering currents, and the portability of the set, it is expected to be of considerable benefit in telegraph transmission maintenance work at outlying subscriber stations.

D. Distortion-Distribution Recorder

A short test to determine the peak value of the distortion existing on a circuit at any particular time is of great value but in development work it is sometimes desired to know the frequency of occurrence of different values of distortion; in other words, to obtain data from which a distribution curve of distortion may be plotted. Such distribution curves were referred to in the earlier paper and it was stated there that, in general, the distribution of distortion is in accordance with the normal law. To check this on different types of circuits in connection with work on transmission ratings of telegraph circuits and for other laboratory uses a "distortion-distribution recorder" has been devised.

The distortion-distribution recorder operates on a synchronous basis and is suitable only for testing over looped-back circuits. It may be used at any speed up to about 75 words per minute (28.5 dots per second, 57 bauds). It indicates the distortion of the transition at either end of a signal pulse, the indication being in terms of the displacement of the transition from its correct time of occurrence in the signal combination. Records of the distortion are made on message registers in ranges of one per cent for small distortions and greater ranges for larger distortions.

As shown in Fig. 14, this device contains a sending distributor face (*a*) to provide test signals and two receiving distributor faces (*b*) and (*c*). Coupled to these and running five times as fast is a large disc carrying a fine point from which a spark may be made to jump to any one of a series of stationary segments. This is referred to as "distortion-scanner" in the figure. Since the disc makes one-half revolution while the sending distributor brush is traversing one segment, one-half revolution of the disc requires the same time as one dot and is, therefore, equal to 100 per cent distortion. One hundred segments are provided as indicated, each being equivalent to 1 per cent distortion, so that this ring of segments forms a distortion scale covering the range of ± 50 per cent distortion. Distortion indicators are associated with these segments, each containing a gas-filled tube and a message register. Only the first ten segments on either side of zero are provided with individual indicating arrangements, the succeeding segments being combined in successively larger groups as shown in the figure; this affords adequate information for the usual case.

Assume the switches associated with the sending distributor face (*a*) to be operated to send the signal shown in note 2*a*. After traversing the circuit to be tested this signal operates the receiving relays, one of which is associated with a "lag-meter" and the other with a spark-producing circuit. If the time of occurrence of transition *X* (note 2-*a*) is to be used as a reference, the *M–S* switch of the lag circuit is closed and the sending segments oriented until transition *X* occurs while the brush of receiving distributor (*b*) is midway between segments 9 and 10, as is indicated by a lag-meter reading of zero. The positions of the segments of the receiving faces (*b*) and (*c*) with respect to the signal will now be as indicated in notes 2-*b* and 2-*c*, and the set is ready to measure the occurrence of a transition between any two segments of the sending face, for instance, transition *Y* or *Z*.

If for instance the displacement of transition *Y* with respect to transition *X* is to be measured, the switch associated with face (*c*) is set to connect to segment 4. When transition *Y* occurs the receiving relays operate to marking, condenser *C* is discharged through the primary of the induction coil and a spark jumps from the scanning point to a stationary segment. If transition *Y* is not distorted the spark will jump when the scanning point is opposite segment *O*. This will cause the gas tube associated with segment *O* to fire and

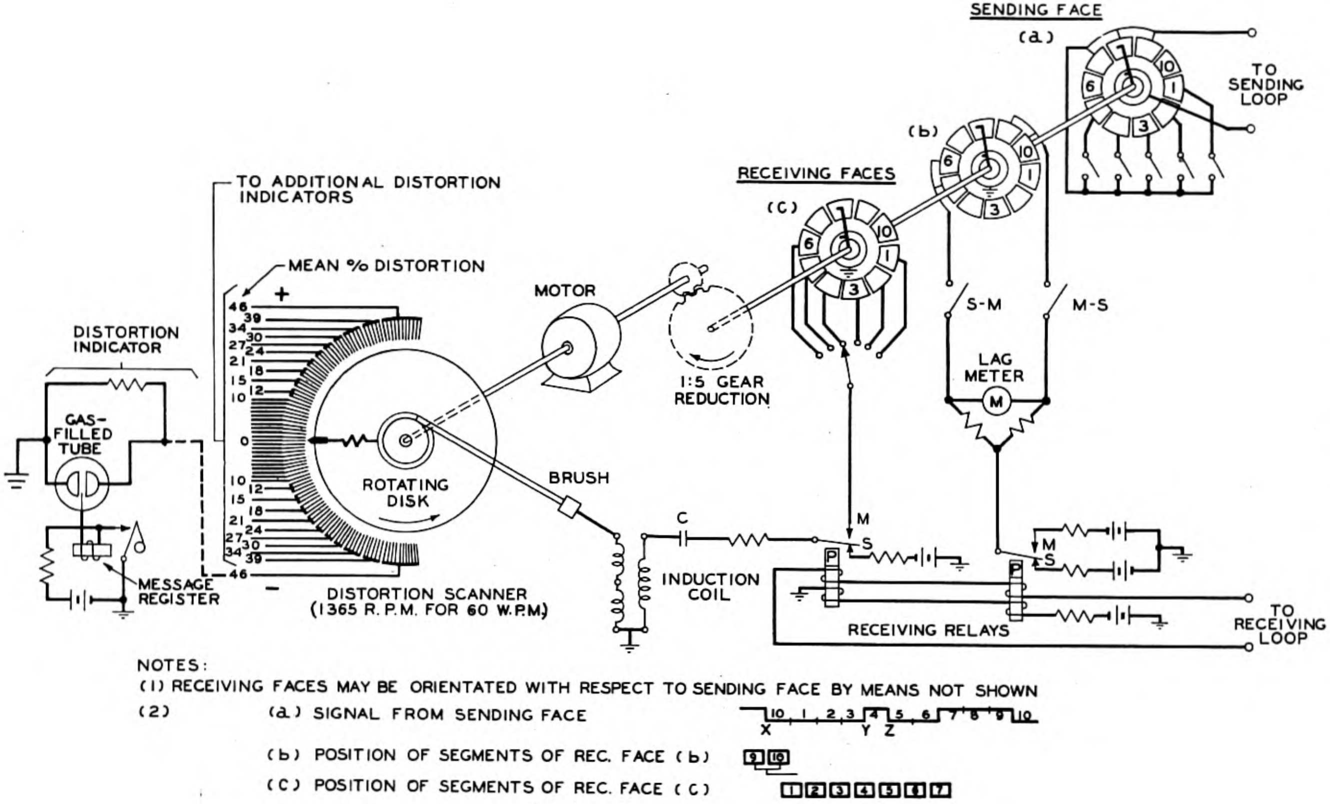


Fig. 14—Schematic circuit of distortion-distribution recorder.

operate the message register which in turn will extinguish the tube by momentarily short-circuiting the anode potential. Other transitions in the signal will not cause the condenser C to be discharged through the coil because the brush of face (c) will not be traversing the proper segment and there will be no path to ground.

The complete circuit of the device contains a selecting switch which automatically changes the signal combination for each revolution of the brush arm of face (a) and reversing switches to invert the signals and the relay connections. These have been omitted for the sake of briefness and clarity. It is thought, however, that the above description will give a good idea of how this device operates to make a record of the frequency of occurrence of distortions. A sample of such a record is plotted in Fig. 15, this being from data obtained over a short-wave radio link. The measurements covered a 15-minute period and one transition of a repeated signal was observed continuously during this period. By inspection it is seen that this circuit has bias of − 1 per cent and the r.m.s. value of distortion is about ± 3 per cent. With miscellaneous signals and with characteristic distortion present, the curve would have several peaks and would be somewhat irregular, but would have the general shape of a normal-law distribution curve.

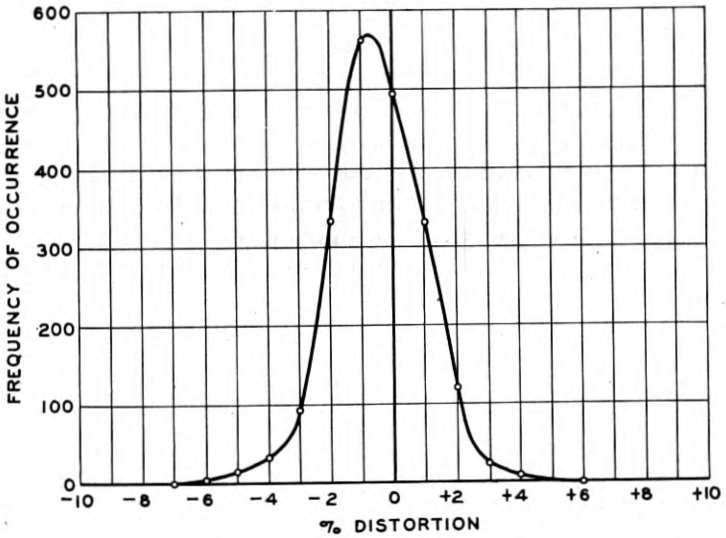


Fig. 15—Typical curve obtained with distortion-distribution recorder.

E. Sources of Test Signals

Bell System telegraph circuits are tested at present with both substantially perfect and distorted test signals. The quality of the test signals is, of course, of prime importance because the measurements are generally made on a straightaway basis and it is not practicable to make correction for accidental distortion in the test signals at the sending end of the circuit. On high-grade circuits where the distortion is generally less than about 5 per cent, distortion in excess of a few per cent in the test signals is very undesirable.

UNDISTORTED TEST SIGNALS

Substantially perfect test signals are usually supplied from motor-driven commutators. One type supplies telegraph reversals. In this case an accurately governed motor drives two brush arms, each of which is associated with two rings of segments so that four sources of signals are provided by each machine. These reversals or dot signals are used in a number of ways, their principal advantage being that since the average of the marking and spacing intervals is zero, a simple integrating meter circuit may be used for measuring bias.

When it is desired to employ miscellaneous or recurring teletypewriter characters machines known as transmitter-distributors are used. Such a device, arranged to send a standard test sentence, is illustrated in Fig. 16. This consists of a motor-driven commutator with a continuously rotating brush arm and a direct-coupled cam transmitter (on the left) which changes the connections to the segments of the commutator in accordance with the code which is cut on the cams. In another form of this device a tape transmitter [4] is used; the tape is usually of parchment although metal tapes and wheels drilled with the code combinations have been used.

When a number of sources of undistorted signals is required in a repeater station a device called a "multiple-sender" is used. This employs the distributor of Fig. 16 to operate a number of relays. The transmitting contacts of these relays are connected to jacks at convenient locations in the telegraph test board. The standard test sentence supplied by this device contains desirable signal combinations for testing transmission over lines and also for testing the operation of teletypewriters.

The signals from the commutator face traverse two groups of relay windings, a marking group and a spacing group, the circuit being as indicated in Fig. 17. This circuit effectively provides polar operation of the relays, the transmitter closing the circuit through one group of windings to operate the relays to their marking contacts and through

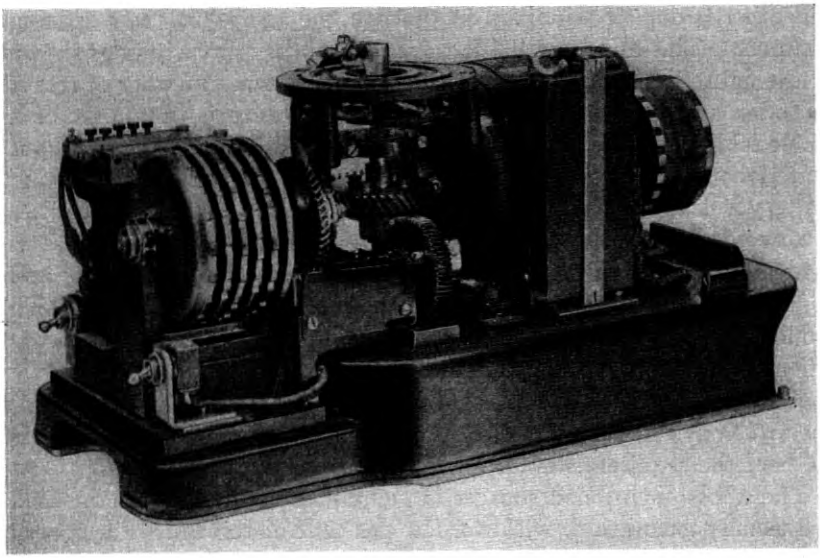


Fig. 16—Test-sentence transmitter-distributor.

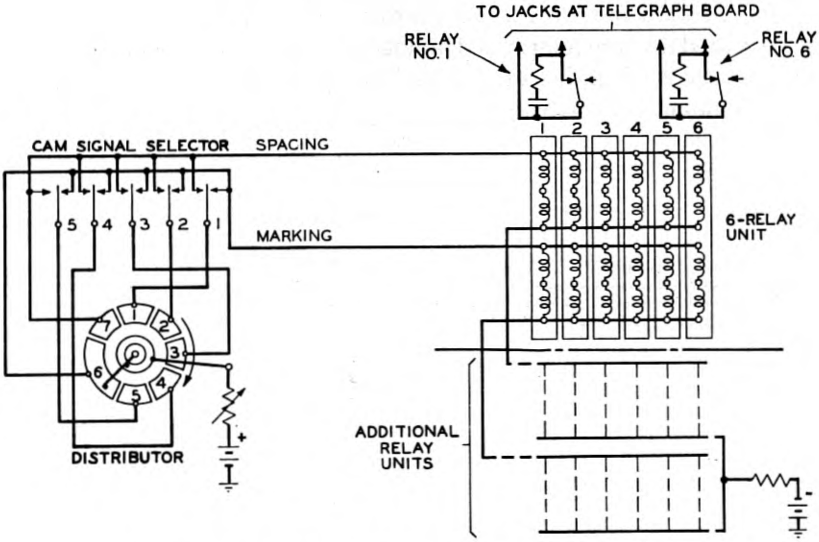


Fig. 17—Schematic circuit of automatic multiple sender.

the other group of windings to operate the relays to their spacing contacts. The series-parallel arrangement of the relay windings shown in the figure has the advantages that with the same number of relays in series and in parallel the combined inductance is only that of a single relay, and that any relay may be removed for inspection without materially affecting the operation of the others. Only one transmitting battery is used in this circuit and thus errors due to battery inequalities are avoided. As indicated in the figure, a spark-reducing circuit is associated with each output for the purpose of minimizing arcing and to neutralize the effect of travel-time of the relay armature which would otherwise cause the transmitted signals to be biased to spacing when opening and closing the circuit under test.

Each group contains 6 relays in parallel and any number of groups up to 8 may be used in series to provide a maximum of 48 outputs to meet the requirements for offices of various sizes.

The above-mentioned sources of signals as maintained in the field are usually accurate to within a few per cent distortion. For special uses it is possible to reduce this inaccuracy somewhat by additional maintenance.

DISTORTED TEST SIGNALS

A repeating device has been provided, primarily for transmission maintenance, by means of which the distortion of telegraph signals may be increased by predetermined amounts. The set is generally used with a source of undistorted signals to provide test signals having known amounts of distortion.

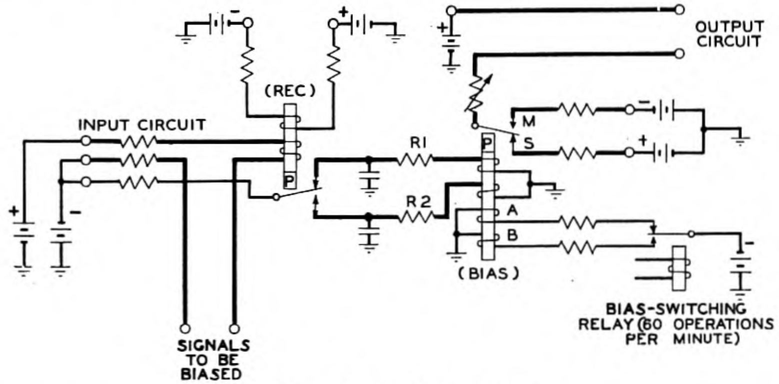


Fig. 18—Bias-producing circuit.

The schematic circuit of the signal-distorting device is shown by Fig. 18. The signals to be distorted are connected to the input and are repeated by the receiving relay (Rec) of the device into a biasing relay (Bias) through a network which modifies the wave-shape of the signals and permits them to be biased easily by changing the current through the auxiliary windings *A* and *B* of the Bias relay.

The manner in which bias is produced will be understood more fully by reference to Fig. 19 which shows graphs of the currents in the

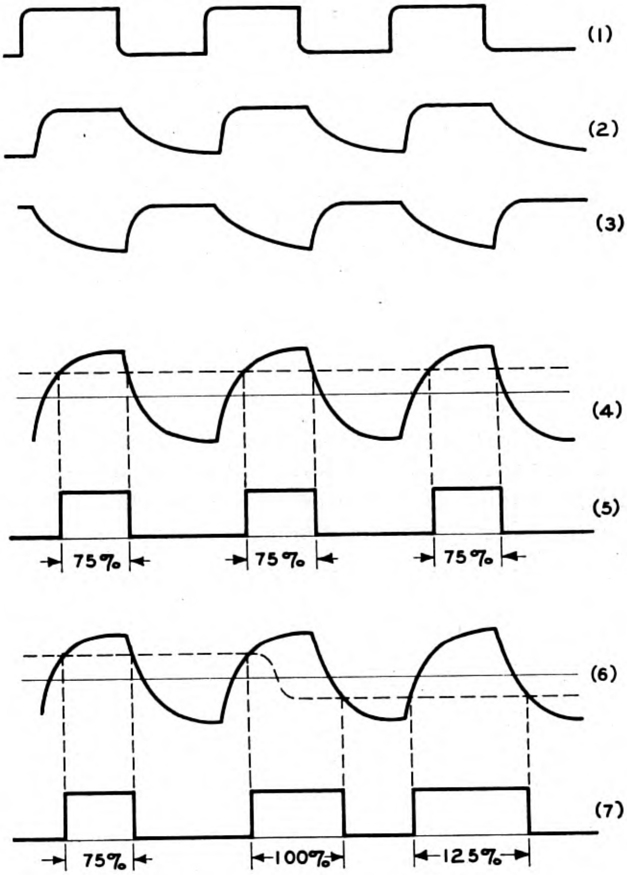


Fig. 19—Currents in bias-producing circuit.

bias-producing circuit. The current in the receiving relay input circuit is indicated by 1. These are undistorted substantially square-topped reversals. The currents flowing through the operating windings of the Bias relay in response to the marks and spaces of the reversals are shown by 2 and 3. In the case shown by 2, it is seen that at the beginnings of the marking pulses the wave fronts are steep and that the current gradually decays to zero at the beginnings of the spacing pulses. In 3 this condition is reversed so that the net operating current for the Bias relay, being the difference between the currents of 2 and 3, has a rounded wave shape for both the beginnings of marks and the beginnings of spaces as is indicated at 4. This is a symmetrical wave about the zero axis and if there is no bias effect in the Bias relay the signals will be repeated unbiased. However, if current is passed through one of the auxiliary windings of the Bias relay the operating point of this relay will be shifted as indicated by the dotted line and the repeated signals will be biased, as shown by 5. In this case the bias is 25 per cent spacing so that the repeated unit marks are 75 per cent of their original length.

One of the auxiliary windings of the Bias relay is used for introducing marking bias and the other for spacing bias, and the sign of the bias may be changed by switching from one winding to the other. Then the bias effect reverses according to the dotted line of 6, and the signal is affected as is indicated by 7. By reversing the bias periodically under the control of a commutator arrangement an effect known as "switched bias" is produced. The reversing operation occurs 60 times per minute and is not synchronized with the signals and accordingly produces a fortuitous effect on some of the signals.

The signals obtained from this device are commonly used in testing the operating margins of subscriber teletypewriters. In this case perfect teletypewriter signals are applied to the input of the device and are distorted by a predetermined amount in passing through it. These signals are then applied to the circuit extending to the subscriber station. If the receiving teletypewriter at the station responds faithfully when set at its optimum orientation point it is considered to be satisfactory for service.

Distorted signals obtained from this device are also used to determine the extent of the distorting effects in line circuits. For this purpose the set may be connected at either the sending or the receiving end of the line and a distortion-measuring set at the receiving end to give an indication of the increase in distortion caused by predistorting the signals. With the set used at the sending end, the results of the test indicate how much distortion may be applied as from preceding telegraph sections. With distortion added at the receiving end, the test may be used to show the margin in the receiving apparatus before failure.

A front view of a panel-mounted telegraph signal-biasing set of the relay type is shown in Fig. 20.

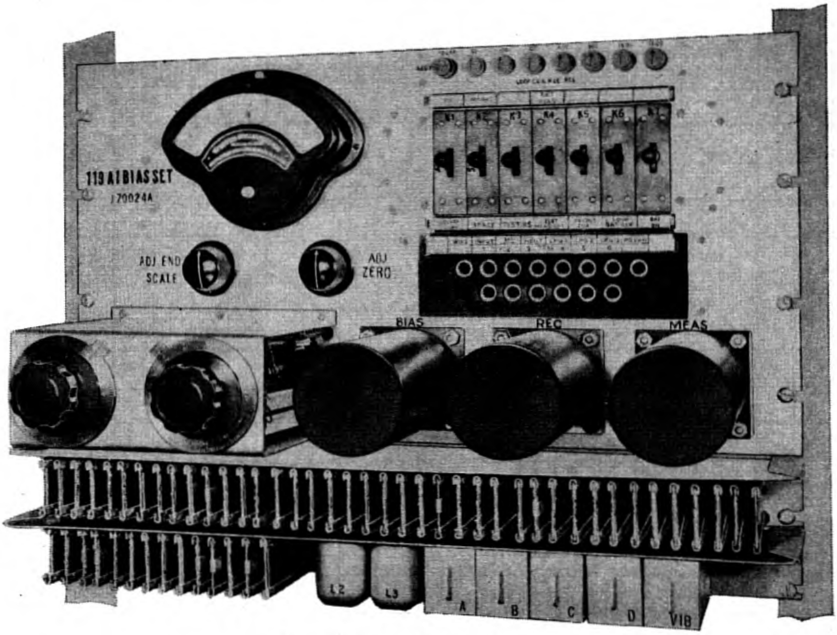


Fig. 20—Telegraph signal-biasing set.

F. Measuring with Teletypewriters

As is well known, the start-stop teletypewriter, when properly adjusted for this purpose, may be used as a transmission-measuring device. The usual procedure in field testing is to compare the range over which the orientation range finder may be shifted with substantially perfect signals to that obtained with signals from the line under test. The orientation range finder is shifted above and below the usual setting until perfect copy is no longer obtained, these limiting positions determining the margins.

At the time of the previous paper,[1] the orientation range test was the only test available to the field forces in the Bell System which utilized the start-stop principle. At that time the teletypewriter was not convenient to use and generally the results were not as accurate as desired. However, the machines have been much improved and better methods of use have been developed so that now orientation margin tests furnish a better indication of the grade of transmission. At the same time, the more convenient measuring sets described above have become available and this has, of course, reduced the field of use of the teletypewriter as a measuring device.

One of the improvements in the machines from the standpoint of

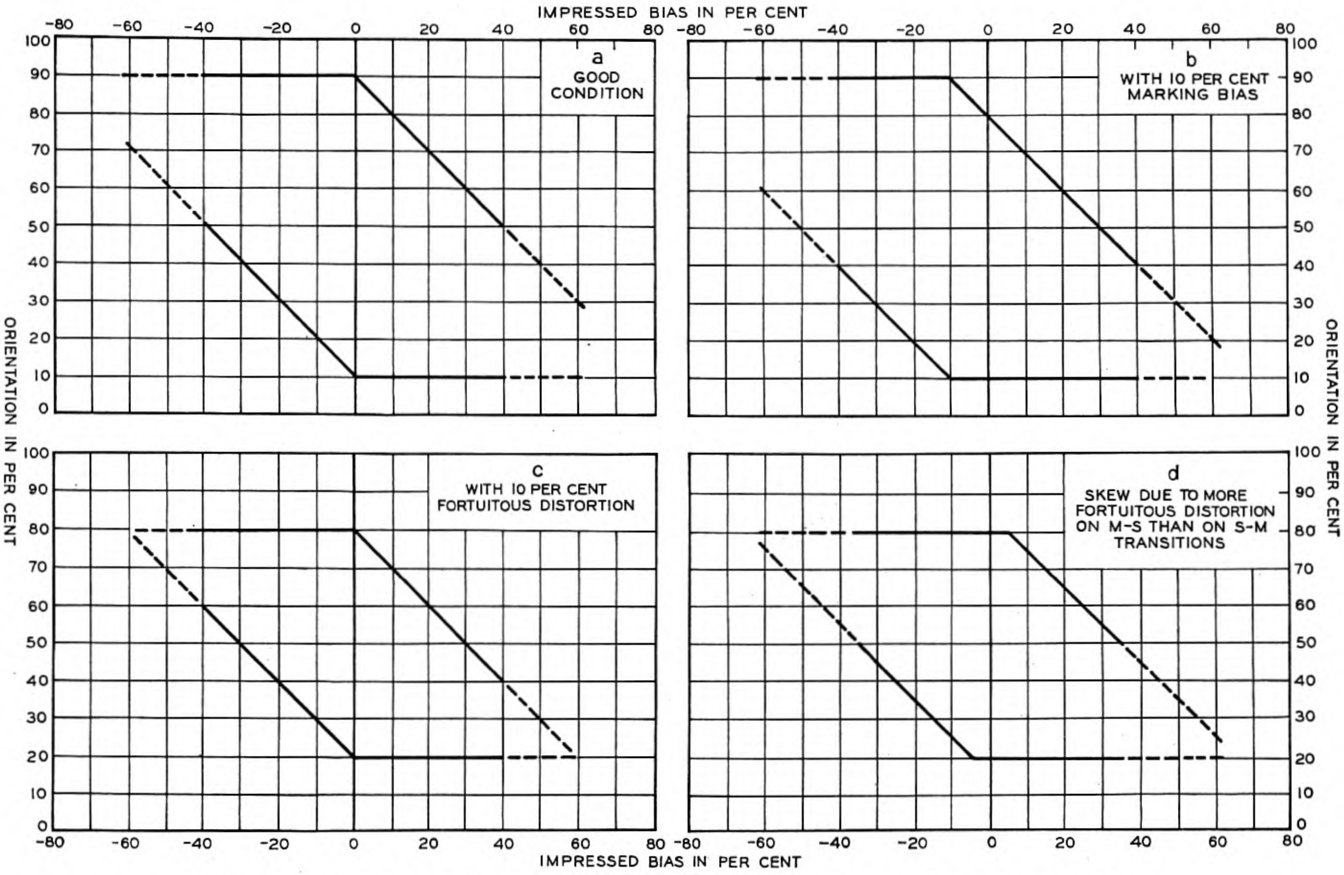


Fig. 21—Diagrams showing the effect of internal distortion on orientation limits of teletypewriters.

transmission testing consists in the addition of a small crank which extends through the cover and which is coupled to the range finder. The crank has a detent which assists in making settings to the nearest per cent, the scale being arranged to indicate directly the distortion in percentage of a unit selecting pulse. This crank and scale arrangement increases the convenience of measurement considerably.

It is of considerable importance to remove as far as practicable the internal distortion [7] of teletypewriters used for measuring purposes. These internal distortions can be identified as bias, characteristic and fortuitous effects. These effects reduce the margin from the theoretical limit of ±50 per cent to about ±40 per cent for the usual machines operating at 60 words per minute.

The presence of internal distortion can be readily determined by using signals biased by various amounts and noting the effect on the orientation range. With a machine satisfactory for measuring purposes, the results obtained would be as indicated on Fig. 21*a*. Here the range is from 10 to 90 per cent for reception of perfect signals. Marking bias reduces the upper range in direct proportion and spacing bias likewise reduces the lower range in direct proportion; thus the machine would be satisfactory for measuring purposes.

If the machine had internal marking bias, the orientation parallelogram would be as shown in Fig. 21*b*. Here the range with perfect signals is from 10 to 80 and marking bias reduces the range in proportion to the bias but spacing bias first increases the range until the internal bias is compensated and then decreases the range as the impressed bias is further increased. With internal spacing bias the parallelogram would be shifted to the other side of the zero line by the amount of the bias. It is obvious therefore that biased machines do not give accurate results in measuring.

Characteristic and fortuitous effects may also be present in teletypewriters to such an extent as to make the machines unsatisfactory for testing purposes. It will be appreciated that internal characteristic distortion changes from signal to signal and when receiving miscellaneous teletypewriter characters, it has much the same effect as fortuitous distortion and the upper and the lower margin limits would be reduced about equally as shown in Fig. 21*c*. If the machine distortions do not have the same effect on mark-to-space and space-to-mark transitions, a skewing effect is produced in the orientation parallelograms. Fig. 21*d* shows skew due to the fortuitous effect of the mark-to-space transitions, being greater than the fortuitous effect of the space-to-mark transitions. Teletypewriters showing such skew effects do not give margin reductions proportional to the impressed distortion and are, therefore, not suitable for measuring purposes.

It is apparent from the above discussion that it is necessary to adjust teletypewriters for minimum internal distortion before they can be used for measuring purposes. Where it is desired to use teletypewriters in testing, procedures have been established in the Bell System to insure that they will be in proper condition and fairly good results are obtained with them.

Although the effect of distortion on the orientation margins has been discussed previously [4, 5, 7] it may be of value to state here how distortion affects the margins at the lower and upper orientation limits in connection with the use of teletypewriters for testing purposes. In general, the maximum reduction corresponds numerically to the total distortion as indicated by the start-stop type of measuring set described above. Distortions other than bias usually affect both orientation limits equally so that the amount of bias can be estimated by subtracting the smaller reduction of the two from the larger. In addition it is possible to obtain an idea of the characteristic distortion by the indications obtained during the orientation test. If the orientation limits are fairly definite, that is, if the copy changes from good to bad when the range finder is moved only a small distance, it is likely that the distortion is due to bias. If there is a definite range over which certain characters are found to be consistently in error this is due to characteristic distortion. If the limits are not definite, that is, if there is a range over which errors occur but not consistently on certain characters this is probably due to fortuitous distortion. Although a qualitative analysis of the distortion may be made in the manner discussed above, this indirect method is somewhat laborious and may give misleading results. Moreover, it is impossible to get a measurement of isolated distortions of high value.

G. *Telegraph Service Monitoring Set*

An automatic telegraph service monitoring set has been designed for the purpose of giving an alarm at repeater stations whenever the distortion on circuits becomes abnormally high or whenever an excessive number of large distortions or "hits" is experienced. This set is still under development; however, a description of it will be given because it is thought to be of general interest to those concerned with telegraph transmission measuring.

In the interest of economy and simplicity this set contains a so-called shortest-pulse type of measuring circuit rather than a start-stop type. Measurement on this basis will of course result in a loss in accuracy for some signal combinations because of the fact, as mentioned earlier, the distortion of the stop pulse is not added to the distortion of other pulses. However for the purpose for which it is used, namely to detect trouble conditions on working circuits, the accuracy is believed to be adequate.

In this set two measuring circuits are provided, one for measuring marks and the other spaces. These are condenser-charging circuits in which the charge on the condenser is an indication of the duration of the pulse. These condenser voltages are compared to a reference voltage and only those less than the reference voltage are permitted to influence the distortion indicator. Therefore, only the shortest pulses are measured and this permits observation on working teletypewriter circuits without involving a start-stop arrangement. By adjustment of the time constant of the condenser-charging circuits, as for instance by means of continuously variable resistances, the percentage distortion for which an alarm is given may be varied at will.

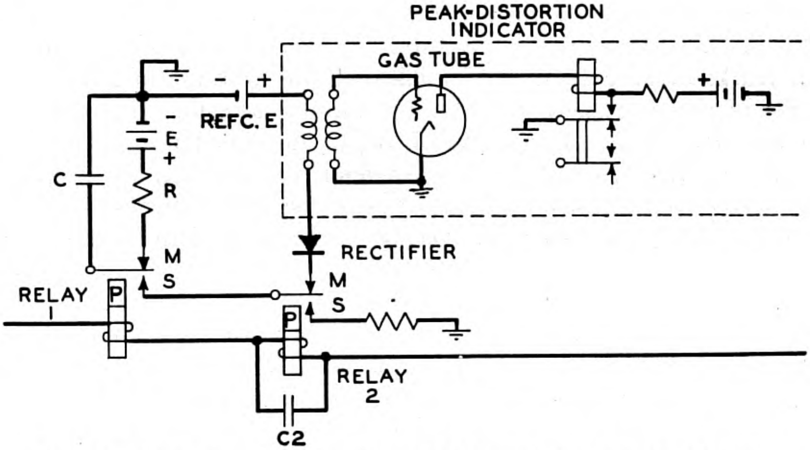


Fig. 22—Distortion-measuring circuit of telegraph service-monitoring set.

The distortion-measuring portion of the circuit used in the measurement of marks is indicated by Fig. 22. Condenser *C* is charged during marking intervals through high resistance *R* by voltage *E;* thus a voltage is produced on the condenser which depends upon the duration of the marking interval as is indicated by Fig. 23. At the time of the transition from mark to space the condenser voltage is momentarily compared to that of a reference source by way of the armature and marking contact of relay 2 of Fig. 22. Immediately afterwards the condenser charge is dissipated by the armature of relay 2 moving to its spacing contact. The small delay required in the operation of relay 2 with respect to relay 1 is obtained by connection of condenser $C2$ around the windings of relay 2.

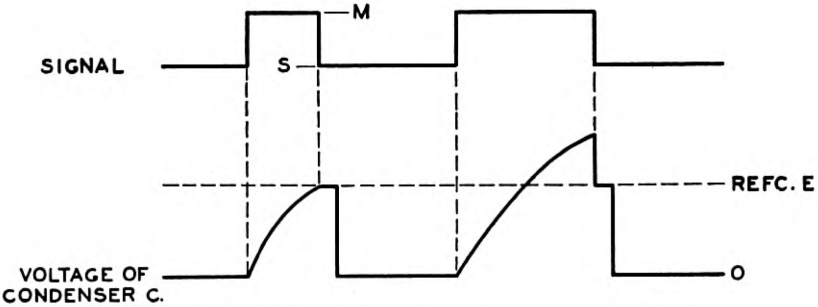


Fig. 23—Variation of charge on measuring condenser in telegraph service-monitoring set.

If at the time of the momentary comparison between the condenser and reference voltages these two voltages are equal, no current will flow in the peak distortion indicator. This condition can be obtained for marks of unit duration (for instance the duration of a selecting pulse of the teletypewriter code) by suitably adjusting the reference voltage REFC. *E*. For marks of longer duration however the condenser voltage will exceed the reference voltage but, by properly poling the rectifier in series with the peak-distortion indicator, current is effectively prevented from flowing at the time of the momentary comparison. With this poling of the rectifier the current flowing due to the marks being shorter than unit duration will affect the peak-distortion indicator and if suitably adjusted will cause a gas-filled tube to operate and thereby give an indication of distortion.

It will be apparent that the measurements of spaces may be accomplished by providing for that purpose another circuit of the same type as that of Fig. 22. For the measurement of both marks and spaces four relays are required, but only one peak-distortion indicator is necessary.

Associated with the measuring circuit is a counting circuit which counts the number of excessive distortions experienced in a given time. This circuit is indicated schematically by Fig. 24. In this circuit the charge of the condenser C is mixed with that of a larger condenser $C2$ at each operation of the gas-filled tube of the distortion measuring circuit. Thus the charge on the larger condenser becomes an indication of the number of excessive distortions. By arranging a timer to discharge this condenser through an indicating circuit suitably designed an alarm may be obtained whenever the number of excessive distortions exceeds any predetermined number up to about 7 within a given time. For this purpose another gas-filled tube is employed having a potentiometer associated with it for the purpose of adjusting the grid-biasing potential so that the tube will fire only for voltages exceeding definite amounts corresponding to definite numbers of excessive distortions. It is apparent that this type of counting circuit could be replaced by a counting-relay circuit or by a selector-switch circuit such as is used in automatic telephony.

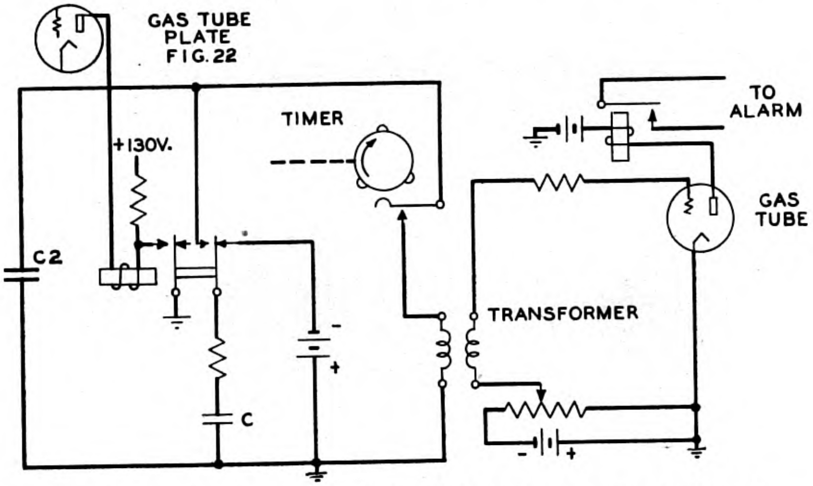


Fig. 24—Counting and alarm circuit of telegraph service-monitoring set.

A front view of the monitoring set, arranged for mounting on a relay rack in a central office, is shown by Fig. 25.

In the present arrangement these sets have jacks at a number of places along the telegraph board in the central office, for the purpose of permitting attendants to use the set conveniently. Alarm lamps and signals are provided at the board to attract the attention of the attendant. As the use of these sets is developed, it may be found desirable to employ them in conjunction with a patrol arrangement by means of which a given set may be connected in turn to each of a number of circuits for a short interval. An arrangement of this sort was described by W. Schallerer.[8]

OTHER DEVELOPMENTS

Other types of measuring apparatus have been used experimentally in the Bell System and have been found of value in laboratory work

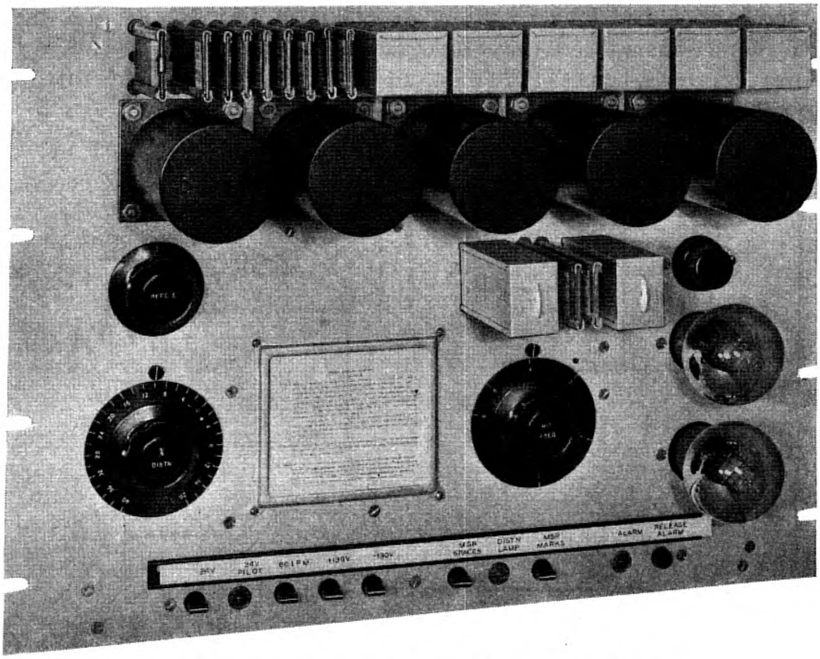


Fig. 25—Telegraph service-monitoring set, test design.

and special investigations. Among these devices are start-stop and synchronous distortion indicators employing flashes of light to indicate the position of the transitions of teletypewriter signals and a photographic recorder of teletypewriter signal transition points. A synchronous flashing indicator in combination with a distributor arranged to supply teletypewriter signals with adjustable bias is in use in shop tests of teletypewriters and in the laboratory.

ACKNOWLEDGMENT

In the development of the practical telegraph transmission measuring devices, which have been described, many members of the Bell System have contributed valuable ideas and effort. The authors wish to express their appreciation of the cooperation they have received.

APPENDIX

GENERAL

This appendix considers characteristic distortion of telegraph signals from the standpoint of the development of simple and convenient testing technique for application primarily to circuits transmitting start-stop teletypewriter signals.

A previous paper[2] developed methods of determining the correct transfer admittance for distortionless transmission under certain assumed conditions. One general and simplifying assumption was that the time interval between transitions in the telegraph signals would be an integral number of time units. If telegraph circuits were to be designed for the transmission of telegraph signals of this nature, distortionless transmission would be expected when the overall transfer admittance of the circuit was one of the many possible admittances discussed in the previous paper. Although a knowledge of the admittance requirements is a helpful guide in the design of circuits and permits the establishment of certain boundaries, the exact adjustment of transfer admittance to the proper value for satisfactory transmission on the basis of admittance measurements presents many practical difficulties and up to the present has not been generally used.

Another approach to the problem is the actual transmission of miscellaneous signals of the type required and the adjustment of the transfer admittance on a cut-and-try basis until the overall results are satisfactory. For relatively simple circuits satisfactory results can be obtained in this manner. However, for circuits which are electrically long and contain complex networks, a more orderly approach is desirable.

The problem may also be approached from the standpoint of adjusting the transfer admittance of the circuit so as to minimize the transient associated with each transition at the times at which succeeding transitions may occur. This may, of course, be done by means of oscillograph observations, but this procedure has serious practical disadvantages. An advantageous method, however, is to measure the characteristic distortion of simple signal combinations while making the adjustments. For this purpose, signals, each composed of two transitions, repeated at intervals long compared to the duration of the appreciable transient, are used. If the circuit is adjusted so as to transmit without distortion signals having respectively separations of one, two, three, etc. signal elements, between transitions, the requirements for distortionless transmission of miscellaneous signals of the nature under consideration are met, as will be shown below.

Repeated two-transition signals with varying integral time-unit intervals between transitions may be considered as telegraph reversals having a frequency determined by the period of repetition and bias determined by the interval between transitions. Therefore, telegraph reversals of varying bias may be used as a source of test signals, with a simple bias-measuring set at the receiving end and the input bias-output bias characteristic of a circuit determined for checking the characteristic distortion. The possible use and meaning of such measurements are discussed in the following:

BIAS IN-BIAS OUT

For the purpose of discussion consider the particular code indicated by Fig. 26.

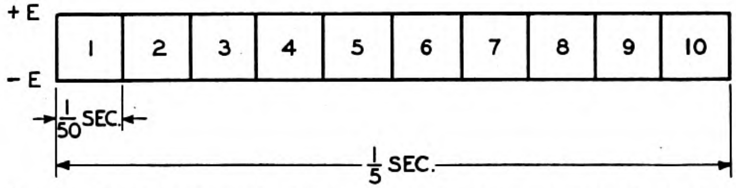


Fig. 26—Special telegraph signal. Each character consists of ten time units. Each time unit is one-fiftieth second and may be $+E$ or $-E$ depending on the character transmitted.

Each character to be transmitted is composed of ten time units, each unit being 1/50 second, so that it takes 1/5 second to transmit any of the 1024 possible characters. Although the nominal speed of signaling is 25 dots per second, the fundamental frequency of any character sent repeatedly is 5 cycles per second considered in the Fourier sense. Among the many possible signal combinations are the following which are equivalent to 5-cycle reversals with the amounts of bias indicated below.

Time Units		Time Bias $= 100\,\frac{M-S}{M+S}$
Mark	Space	
10	0	+100%
9	1	+80%
8	2	+60%
7	3	+40%
6	4	+20%
5	5	0%
4	6	−20%
3	7	−40%
2	8	−60%
1	9	−80%
0	10	−100%

It is obvious for a symmetrical circuit that if signals with any amount of positive bias are transmitted accurately signals with the corresponding amount of negative bias will be transmitted since this corresponds to a reversal of the sending polarities.

The transfer admittance for distortionless transmission of these signals will now be derived for the case in which the band width is a minimum. It has been shown in a previous paper [2] that the minimum band width required is equal numerically to the speed of signaling.

Therefore a band width of 25 cycles will be necessary and sufficient for the signals listed above; this band will also meet the requirements for transmitting the remaining characters of the possible 1024.

The Fourier series for the biased reversal is

$$E(t) = E - \frac{bE}{2\pi} + \frac{E}{\pi}\sum_{n=1}^{\infty}\frac{1}{n}[\sin n(\omega t - b) - \sin n\omega t],$$

where E and b are defined in Fig. 27, which shows the impressed voltage wave at the transmitting end of the circuit, and B is percentage bias divided by 100.

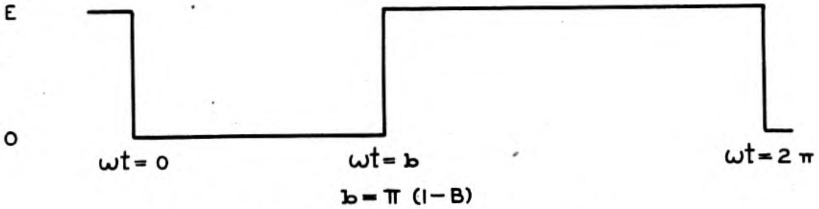


Fig. 27—Voltage wave for biased reversal.

Let the transfer admittance be Y_n at frequency corresponding to $n\omega$. Then, the requirement for perfect transmission of the reversal, assuming the receiving device to operate when the current passes through $EY_0/2$, is that

$$f(t) = EY_0/2 \text{ at } \omega t = 0 \text{ and at } \omega t = b,$$

where $f(t)$ denotes received current.

$$\text{At } \omega t = 0 \text{ or } b,\ f(t) = \left(E - \frac{bE}{2\pi}\right) Y_0 + \frac{E}{\pi}\sum_{n=1}^{\infty}\frac{Y_n}{n}\sin(-nb).$$

Hence, assuming for simplicity that $Y_0 = 1$

$$\frac{E}{2} = E - \frac{bE}{2\pi} + \frac{E}{\pi}\sum_{n=1}^{\infty}\frac{Y_n}{n}\sin(-nb).$$

Substituting the value of b and transforming

$$\frac{B\pi}{2} = \sum_{n=1}^{\infty}\frac{Y_n}{n}\sin nb.$$

In accordance with the assumption that band width is a minimum,

$Y_n = 0$ for $n \geq 5$. Therefore, the values of Y_n may be determined from the following equations wherein the values of nb are in degrees.

$$\begin{aligned}
\text{For } B&=.2 \quad .1\pi = Y_1 \sin 144 + \tfrac{1}{2} Y_2 \sin 288 + \tfrac{1}{3} Y_3 \sin 432 + \tfrac{1}{4} Y_4 \sin 576,\\
B&=.4 \quad .2\pi = Y_1 \sin 108 + \tfrac{1}{2} Y_2 \sin 216 + \tfrac{1}{3} Y_3 \sin 324 + \tfrac{1}{4} Y_4 \sin 432,\\
B&=.6 \quad .3\pi = Y_1 \sin 72 + \tfrac{1}{2} Y_2 \sin 144 + \tfrac{1}{3} Y_3 \sin 216 + \tfrac{1}{4} Y_4 \sin 288,\\
B&=.8 \quad .4\pi = Y_1 \sin 36 + \tfrac{1}{2} Y_2 \sin 72 + \tfrac{1}{3} Y_3 \sin 108 + \tfrac{1}{4} Y_4 \sin 144.
\end{aligned}$$

Solving these simultaneous equations it is found that $Y_1 = .967$, $Y_2 = .865$, $Y_3 = .685$, $Y_4 = .408$.

The computed admittance is the same as that which may be computed from the information given in Appendix III of a previous paper[2] which is:

$$Y = \frac{\pi}{2} \frac{f}{f_s} \cot \frac{\pi}{2} \frac{f}{f_s},$$

Y = transfer admittance at frequency f,
f_s = dotting speed.

Since the bias of the signals at the circuit output must equal the bias of the signals at the circuit input for the magnitudes of bias entering into the determination of the admittance, and since the transmission of additional harmonics of the fundamental frequency would be required to make the input bias equal to the output bias at additional magni-

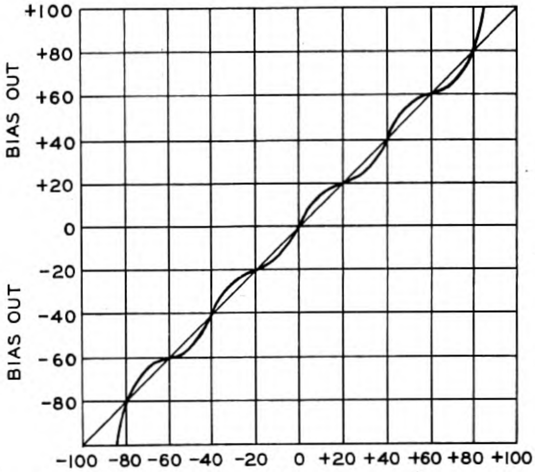


Fig. 28—Bias in-bias out characteristic for transmission of telegraph signals formed in accordance with Fig. 26 over a system having a band width equal to the dotting speed.

tudes of bias, the Bias In-Bias Out characteristic for a circuit having an admittance as computed above would have the general characteristics indicated in Fig. 28.

Considered in terms of the transient behavior, if the circuit had the transfer admittance defined by the above equation, or any of the infinite number of other prescribed admittances using higher frequencies, the transient resulting from a single transition would be such as to have zero value at each of all possible future transition points. Also if the circuit were adjusted so that the four simple characters were transmitted with negligible distortion, the transients would fulfill the conditions for the satisfactory transmission of the other 1020 possible characters.

From Fig. 28, it is obvious that a circuit which had a perfect transfer admittance and a frequency band width large compared to the character repetition frequency, would have a Bias In-Bias Out characteristic which crossed the 45-degree line at many points and approached it as a limit. It is interesting to note the relation between the deviations of the Bias In-Bias Out characteristic from the 45-degree line and the frequency band width. In the example under discussion there are five waves in the characteristic which correspond to the frequency band width divided by the number of characters per second. The band width required is numerically equal to the product of the number of waves in the Bias In-Bias Out characteristic and the number of characters transmitted per second.

Start-Stop Teletypewriter Signal

Start-stop teletypewriter systems may employ varying speeds and signal arrangements. The 60-word-per-minute (60-speed) system is the one most generally used in the Bell System and is taken as an example in this appendix. Similar methods of analyses and tests could be applied to other systems.

The 60-speed teletypewriter signal consists of a starting unit which is always spacing, five selecting units, and a stop signal which is 1.42 units in length and is always marking. The duration of a unit signal pulse is 22 milliseconds and the total length of each character is, therefore, $1 + 5 + 1.42 = 7.42$ times units or 163 milliseconds. With no pause between succeeding characters there are 368 operations per minute or

$$\frac{368}{60} = \frac{1}{.163} = 6.13 \text{ operations per second.}$$

The problem of transmitting these signals without distortion is similar to the problem just discussed except instead of uniform spacing between possible transitions of 1/50 second, the transitions may be spaced at intervals of either.022 second, .031 second, or combinations of these intervals, and the frequency of repetition is 6.13 instead of 5 per second.

As indicated in Fig. 1, there are six teletypewriter characters (Blank, *T*, *O*, *M*, *V*, and Letters) which correspond to 6.13 cycle reversals biased by certain amounts. It will now be shown that if these six characters can be transmitted without distortion, the other 26 characters will also be transmitted without distortion. The method is the same as that used in the preceding problem.

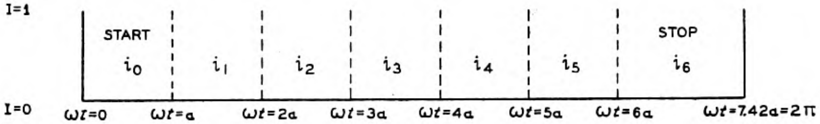


Fig. 29—7.42-Unit code start-stop teletypewriter signal.

Figure 29 indicates any teletypewriter signal where $i_0 = 0$, and $i_6 = 1$. i_1, i_2, i_3, i_4, i_5, may have values of 0 or 1 depending on the particular character. The Fourier series for the received current over a circuit with a transfer admittance Y having unit value at zero frequency, is

$$\begin{aligned} f(t) = \frac{i_1 + i_2 + i_3 + i_4 + i_5 + 1.42}{7.42} \\ + \frac{1}{\pi} \sum_{n=1}^{n=\infty} \frac{Y_n}{n} \{ i_1[\sin n(\omega t - a) - \sin n(\omega t - 2a)] \\ + i_2[\sin n(\omega t - 2a) - \sin n(\omega t - 3a)] \\ + i_3[\sin n(\omega t - 3a) - \sin n(\omega t - 4a)] \\ + i_4[\sin n(\omega t - 4a) - \sin n(\omega t - 5a)] \\ + i_5[\sin n(\omega t - 5a) - \sin n(\omega t - 6a)] \\ + 1\ [\sin n(\omega t - 6a) - \sin n\ \omega t]\}. \end{aligned}$$

Suppose that the transmitted frequency band is limited to 6 times the fundamental, i.e. $n = 1$ to 6, and Y_n adjusted so that the characters "Letters" *V*, *M*, *O*, *T* and "Blank" are transmitted perfectly.

The transfer admittance will now be determined for this case. The expression for $f(t)$ may be written for the characters just mentioned, and would have the value 1/2 at $\omega t = a$, for "letters," at $\omega t = 2a$ for *V*, at $\omega t = 3a$ for *M*, etc. Accordingly there result six equations which may be simplified as follows:

$$\text{"Letters"} \qquad \sum_1^6 \frac{Y_n}{n} \sin na = \frac{2.71\pi}{7.42},$$

$$V \qquad \sum_1^6 \frac{Y_n}{n} \sin 2na = \frac{1.71\pi}{7.42},$$

$$M \qquad \sum_1^6 \frac{Y_n}{n} \sin 3na = \frac{.71\pi}{7.42},$$

$$O \qquad \sum_1^6 \frac{Y_n}{n} \sin 4na = \frac{-.29\pi}{7.42},$$

$$T \qquad \sum_1^6 \frac{Y_n}{n} \sin 5na = \frac{-1.29\pi}{7.42},$$

$$\text{"Blank"} \qquad \sum_1^6 \frac{Y_n}{n} \sin 6na = \frac{-2.29\pi}{7.42}.$$

The values of Y_n computed by solving these equations were plotted on Fig. 30 and a curve was drawn through them. It will be understood, that, on the scale of abscissae, F is the fundamental frequency of the character and the coefficients of F are values of n.

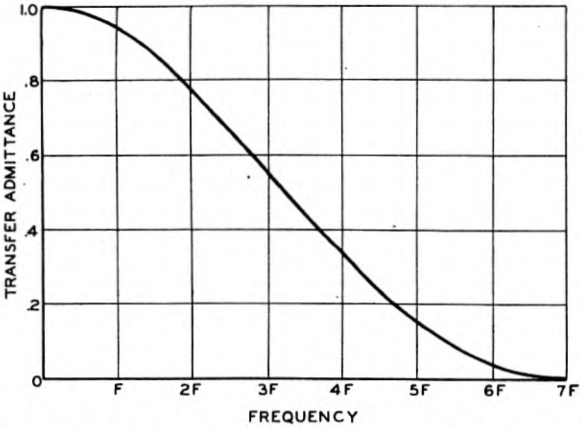


Fig. 30—Transfer admittance characteristic for undistorted transmission using frequencies up to seven times the fundamental frequency of the character.

By the methods employed in the previous paper [2] it can be shown that, for signals employing the same permissible intervals between transitions, the solutions of equations similar to the above approach the smooth curve as a limit as the period of repetition is lengthened.

These equations may be used to prove that when the six simple characters are perfectly transmitted, all other teletypewriter signals may be transmitted on a repeated basis. For distortionless transmission $f(t)$ should equal 1/2 at each transition point, assuming that the relay operates at this value of current. If the expression for $f(t)$ is written for each of the conditions which may occur, and is found to equal 1/2 at the transition points, the proof is complete. For this purpose the summation signs may be eliminated by substituting from any of the six simultaneous equations numerical values in place of summations of terms, in the equation for $f(t)$, with the following results:

Conditions

$\omega t=0 \quad i_0=0 \quad i_6=1$

$$f(t)=\frac{1}{7.42}[i_1+i_2+i_3+i_4+i_5+1.42+i_1(-1)+i_2(-1)+i_3(-1)+i_4(-1)+i_5(-1)+1(2.29)]=\tfrac{1}{2},$$

$\omega t=\mathrm{a} \quad i_0=0 \quad i_1=1$

$$f(t)=\frac{1}{7.42}[1+i_2+i_3+i_4+i_5+1.42+1(2.71)+i_2(-1)+i_3(-1)+i_4(-1)+i_5(-1)+1(1.29-2.71)]=\tfrac{1}{2},$$

$\omega t=2a \quad i_1=0 \quad i_2=1$ or $i_1=1 \quad i_2=0$

$$f(t)=\frac{1}{7.42}[0+1+i_3+i_4+i_5+1.42+0(2.71)+1(2.71)+i_3(-1)+i_4(-1)+i_5(-1)+1(.29-1.71)]=\tfrac{1}{2},$$

$\omega t=3\mathrm{a} \quad i_2=0 \quad i_3=1$ or $i_2=1 \quad i_3=0$

$$f(t)=\frac{1}{7.42}[i_1+0+1+i_4+i_5+1.42+i_1(-1)+0(2.71)+1(2.71)+i_4(-1)+i_5(-1)+1(-.71-.71)]=\tfrac{1}{2},$$

$\omega t=4a \quad i_3=0 \quad i_4=1$ or $i_3=1 \quad i_4=0$

$$f(t)=\frac{1}{7.42}(i_1+i_2+0+1+i_5+1.42+i_1(-1)+i_2(-1)+0(2.71)+1(2.71)+i_5(-1)+1(-1.71+.29)]=\tfrac{1}{2},$$

$\omega t=5\mathrm{a} \quad i_4=0 \quad i_5=1$ or $i_4=1 \quad i_5=0$

$$f(t)=\frac{1}{7.42}[i_1+i_2+i_3+0+1+1.42+i_1(-1)+i_2(-1)+i_3(-1)+0(2.71)+1(2.71)+1(-2.71+1.29)]=\tfrac{1}{2},$$

$\omega t=6a \quad i_5=0 \quad i_6=1$

$$f(t)=\frac{1}{7.42}[i_1+i_2+i_3+i_4+0+1.42+i_1(-1)+i_2(-1)+i_3(-1)+i_4(-1)+0(2.71)+1(2.29)]=\tfrac{1}{2}.$$

Hence it is concluded that if an admittance can be found such that "Blank" *T*, *O*, *M*, *V* and "Letters" are transmitted without distortion any other teletypewriter signal may be transmitted on a repeated basis, since the correct current value (1/2) will be obtained at any transition regardless of what other signal combination is used for the remainder of the character.

Computations have also been made for the case of a repeated signal combination consisting of any two teletypewriter characters. The results indicate that this may also be transmitted without distortion if the requirements have been met for the six simple characters sent repeatedly.

When distortion is present, there is deviation of the received current from the desired value at the time it should be 1/2 and the deviation to be expected at any transition for any character may be computed from equations similar to those given above. This, of course, differs from the amount of distortion which would be measured in per cent. It is of interest that the characteristic distortion on the more complicated characters may be materially greater than that measured on the simple two-transition characters.

The transient behavior of a circuit having the transfer admittance specified by the smooth curve of Fig. 30 may be determined by methods utilizing the Fourier integral. However, certain characteristics of this transient, namely, the points at which it must be zero, are known in advance from the conditions entering into the determination of the admittances. Although these conditions do not directly prescribe the magnitude of the transient at other points, additional information may be obtained from an inspection of Fig. 31.

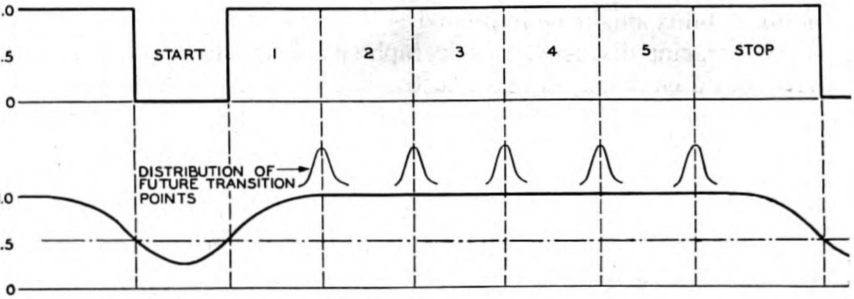


Fig. 31—Received current for "letters" character assuming the transfer admittance of Fig. 30.

This figure shows the computed received current over a telegraph circuit having the admittance shown in Fig. 30 when a repeated teletypewriter "Letters" signal is transmitted. It may be noted that the transient decays to inappreciable amplitudes at times greater than the shortest time unit of .022 second. This is significant, since it means that a particular arrangement of transitions is not of practical importance as long as transitions do not come at intervals closer than .022 second. With this admittance, therefore, no difficulty would be expected from keyboard sending, the signals from which differ from the signals previously discussed, in that the lengths of the stop signals occur on a random basis, and are never shorter than .031 second, depending on the typist.

However, if signals containing bias or other distortion such as indicated on Fig. 31 as distortion of future transition points were transmitted over the circuit, thus decreasing the minimum interval between transitions, the transient from one transition, for example, might affect following transitions. A circuit having a shorter build-up time could be made to introduce less distortion on transitions spaced at very short intervals. The rate of build-up is a function of the area under the transfer-admittance curve and from a design standpoint it is necessary to provide a suitable frequency band-width to make the slope of the received signals sufficient so that characteristic distortion will not be excessive when closely spaced transitions are transmitted, and also so that certain types of interference will not cause excessive fortuitous effects.

Where the available band width is limited the area under the admittance curve could be increased by transmitting the permissible harmonic frequencies at a greater amplitude. The transient of such a circuit would continue at appreciable magnitudes for a greater length of time and the minimum distortion would be increased but the general circuit stability might be improved.

The foregoing discussion of telegraph signal transmission has shown that when a circuit is adjusted to transmit without distortion certain selected repeated characters, the circuit can transmit on a repeated basis any of the characters possible with signals of the nature represented by the selected characters. This is true not only for the signals in which transitions are spaced at integral units of time but for signals in which transitions are spaced at predetermined non-integral units of time, such as start-stop teletypewriter signals. Incidental to the development of the proof of the later statement, the prescribed admittance for transmission without distortion was evaluated. The prescribed admittance for signals employing integral units between transitions was also similarly determined and found to be the same as that which had been determined from a somewhat different approach in a previous paper.[2]

The admittances considered have been idealized somewhat, inasmuch as no physical circuit will cut off completely at the higher frequencies and, in addition, the effective transfer admittance is not only a complex, and frequently nonlinear quantity, but is determined in part by the characteristics of transmitting and receiving relays and other terminal equipment. In the present state of the art, it is difficult to make practical use of transfer admittances in predicting the performance of a telegraph circuit.

The significant point is that the satisfactory transmission of the selected characters is an indication of the ability to transmit the desired telegraph signals satisfactorily. Also, the measurement of the distortion on the selected characters is particularly useful when it is desired to equalize individual circuits of varying length and makeup to secure a minimum of distortion.

The testing procedures suggested by the considerations of the foregoing have been incorporated into the testing instrumentalities discussed in the main paper. These methods have been used for several years in the adjustment and maintenance of telegraph circuits and found to be of considerable utility.

REFERENCES

A. References Cited:

1. "Measurement of Telegraph Transmission," Nyquist, Shanck and Cory, *Trans. A.I.E.E.*, 1927, Vol. 46, p. 367.
2. "Certain Topics in Telegraph Transmission Theory," H. Nyquist, *Trans. A.I.E.E.*, 1928, Vol. 47, p. 617.
3. a. "Telephone Typewriters and Auxiliary Arrangements," R. D. Parker, *Bell Telephone Quarterly*, July, 1929, p. 181.
   b. "Modern Practices in Private Wire Telegraph Service," R. E. Pierce, *Trans. A.I.E.E.*, Jan. 1931, Vol. 50, p. 45.
4. "Fundamentals of Teletypewriters Used in the Bell System," E. F. Watson, *Bell System Technical Journal*, Oct., 1938, p. 620.
5. "A Transmission System for Teletypewriter Exchange Service," Pierce and Bemis, *Bell Sys. Tech. Journal*, Oct., 1936, p. 529; *Elec. Engg.*, Sept., 1936, p. 961.
6. "Metallic Polar-duplex Telegraph System for Cables," Bell, Shanck and Branson, *Trans. A.I.E.E.*, 1925, Vol. 44, p. 316.
7. a. "Der Spielraum des Siemens-Springschreibers," M. J. deVries, *Telegraphen und Fernsprech Technik*, Jan., 1934, p. 7.
   b. "Der Spielraum des Springschrieber," M. J. deVries, *T. F. T.*, Sept., 1937, p. 213.

B. Additional References:

8. "Ein Neues Mess- und Überwachungsgerät für Springschreiberverbindungen," W. Schallerer, *T. F. T.*, Feb., 1935, p. 40.
9. "Versuche über eine günstige Verteilung der Trägerwellen in der Wechselstromtelegraphie," H. Stahl, *T. F. T.*, Nov., 1930, p. 340.
10. "Verzerrungsmesser für Telegraphie," A. Jipp and O. Römer, *T. F. T.*, May, 1932, p. 121.
11. "Telegraph Transmission Testing Machine," F. B. Bramhall, *Trans. A.I.E.E.*, June, 1931, p. 404.
12. "A Telegraph Distortion Measuring Set," V. J. Terry and C. H. W. Brookes-Smith, *Elec. Comm.*, July, 1933, p. 15.
13. "The Measurement of Telegraph Distortion," V. J. Terry, *Elec. Comm.*, April 1933, p. 197.
14. "Determining the Transmission Efficiency of Telegraph Circuits," E. H. Jolley, *P.O.E.E. Jl.*, April, 1933, p. 1.

Contemporary Advances in Physics, XXXII
Particles of the Cosmic Rays

By KARL K. DARROW

Even after fifteen years of intensive research following on two decades of more desultory study, the cosmic rays are still a store of new and remarkable data. The question of their ultimate origin, though by no means extinct, has been set aside by many physicists in favor of a fuller inquiry into their qualities. The distinctive mark of the cosmic-ray particles is the immensity of their energies; for, great by all previous standards as are the energy-values which physicists now can impart in their laboratories, those manifest in the cosmic rays are greater by factors not of thousands merely, but often of millions. To this remote and exalted energy-range belong the penetrating particles capable of cleaving through a metre of lead, and the wonderful and beautiful phenomenon of cosmic-ray showers. It is not to be wondered at that with energies so high, particles so familiar as electrons and photons should be invested with unfamiliar powers. So evidently they are; but some of the charged corpuscles of the cosmic rays have properties such that their strangeness cannot be ascribed to high energy alone, but apparently must be based upon some fundamental difference (perhaps a difference of mass) from all the particles thus far identified.

WHEN a new member is admitted to a small and jealously-restricted club supposedly already filled for all time, the event has a dramatic aspect. When a concept is formed in a nebulous way and rapidly gains precision with the passage of the years, the story is of philosophic interest. When physicists extend their knowledge into ranges of energy heretofore unsuspected, and find them inhabited by particles classifiable as electrons but in possession of powers ordinarily unknown, and also by particles which must be put in a class by themselves—when such things are available for telling, the tale has scientific value. When evidence comes in the form of pictures so striking as those which can here be shown, the science of lifeless matter has an aesthetic splendor such as rarely embellishes it. All of these features appear in the recent advances of the study of cosmic rays.

The small and exclusive club consists of the subatomic particles, long supposed to comprise only the negative electron and the proton and other positive atom-nuclei. Into it the positive electron had been forced in 1932, and the neutron in 1933; a vacant chair was other terminal equipment. In the present state of the art, it is difficult to make practical use of transfer admittances in predicting the performance of a telegraph circuit.

The significant point is that the satisfactory transmission of the selected characters is an indication of the ability to transmit the desired telegraph signals satisfactorily. Also, the measurement of the distortion on the selected characters is particularly useful when it is desired to equalize individual circuits of varying length and makeup to secure a minimum of distortion.

The testing procedures suggested by the considerations of the foregoing have been incorporated into the testing instrumentalities discussed in the main paper. These methods have been used for several years in the adjustment and maintenance of telegraph circuits and found to be of considerable utility.

REFERENCES

A. References Cited:

1. "Measurement of Telegraph Transmission," Nyquist, Shanck and Cory, *Trans. A.I.E.E.*, 1927, Vol. 46, p. 367.
2. "Certain Topics in Telegraph Transmission Theory," H. Nyquist, *Trans. A.I.E.E.*, 1928, Vol. 47, p. 617.
3. a. "Telephone Typewriters and Auxiliary Arrangements," R. D. Parker, *Bell Telephone Quarterly*, July, 1929, p. 181.
   b. "Modern Practices in Private Wire Telegraph Service," R. E. Pierce, *Trans. A.I.E.E.*, Jan. 1931, Vol. 50, p. 45.
4. "Fundamentals of Teletypewriters Used in the Bell System," E. F. Watson, *Bell System Technical Journal*, Oct., 1938, p. 620.
5. "A Transmission System for Teletypewriter Exchange Service," Pierce and Bemis, *Bell Sys. Tech. Journal*, Oct., 1936, p. 529; *Elec. Engg.*, Sept., 1936, p. 961.
6. "Metallic Polar-duplex Telegraph System for Cables," Bell, Shanck and Branson, *Trans. A.I.E.E.*, 1925, Vol. 44, p. 316.
7. a. "Der Spielraum des Siemens-Springschreibers," M. J. deVries, *Telegraphen und Fernsprech Technik*, Jan., 1934, p. 7.
   b. "Der Spielraum des Springschrieber," M. J. deVries, *T. F. T.*, Sept., 1937, p. 213.

B. Additional References:

8. "Ein Neues Mess- und Überwachungsgerät für Springschreiberverbindungen," W. Schallerer, *T. F. T.*, Feb., 1935, p. 40.
9. "Versuche über eine günstige Verteilung der Trägerwellen in der Wechselstromtelegraphie," H. Stahl, *T. F. T.*, Nov., 1930, p. 340.
10. "Verzerrungsmesser für Telegraphie," A. Jipp and O. Römer, *T. F. T.*, May, 1932, p. 121.
11. "Telegraph Transmission Testing Machine," F. B. Bramhall, *Trans. A.I.E.E.*, June, 1931, p. 404.
12. "A Telegraph Distortion Measuring Set," V. J. Terry and C. H. W. Brookes-Smith, *Elec. Comm.*, July, 1933, p. 15.
13. "The Measurement of Telegraph Distortion," V. J. Terry, *Elec. Comm.*, April 1933, p. 197.
14. "Determining the Transmission Efficiency of Telegraph Circuits," E. H. Jolley, *P.O.E.E. Jl.*, April, 1933, p. 1.

being reserved for the negative proton, which as yet has not turned up to claim it; few if any expected the actual applicant. The concept now hardening into the definite form of this applicant is that of the "mesotron." This is a particle presumed to be equal in charge to the electron, but in mass a couple of hundreds of times as great. In so naming it I follow (C. D.) Anderson's recent proposal, though other titles such as "barytron" and "heavy electron" are already more or less firmly rooted in the literature. The quality which marks it out, when it appears with enormous energy among the cosmic rays, is an extreme and almost incredible power of penetration. This means that the so-called mesotrons are able to traverse decimetres, nay even metres of lead (or of dense matter generally). Like electrons, mesotrons may be of either sign of charge. As for the cosmic-ray particles still classified as electrons, *they* are marked out by their power of producing one of the most magnificent phenomena of Nature, the "shower of cosmic rays," or "shower" for short. Shower-production by the supposed electrons, penetration by the supposed mesotrons, ionization along the course of either corpuscle through air: these are the three phenomena which will furnish most of the illustrations, much of the text of this article. The story of their incorporation into the structure of physical theory will furnish the remainder.

(But negative electrons and protons, not to speak of other atom-nuclei, have been identified through having their charge-to-mass ratios measured with the aid of electric and magnetic deflecting fields in elementary classical ways. Why then do I not cut this introduction short by giving the results of such a measurement upon the mesotron? The reason is, that no such measurement has yet been made. Probably one will be made ere long. Should it give something near to the result expected, the delay will not have been regrettable; for the end of the delay will mark the beginning of the time, when the story to be related in these pages will be regarded as being "of historical interest" only—which is to say, that it will then be liable to be forgotten.)

So that the reader may see at once the three phenomena which are to bulk so largely in this story, I draw his attention at once to some of the pictures which decorate this article.[1] Nearly all of them were made (of course) with the aid of the cloud-chamber or expansion-chamber of C. T. R. Wilson, that device so precious in physics and precious in so many ways.

---

[1] They decorate it with particular clarity, thanks to the kindness of Messrs. Anderson, Auger, Brode, Corson, Fowler, Fussell, Neddermeyer, Stevenson and Street in supplying me with prints of their splendid photographs.

Fig. 1—Track of a cosmic-ray particle (probably an electron) in the expansion-chamber, time having been allowed for the ions to drift apart before the expansion. (Corson and Brode, University of California)

At the beginning I place, as Fig. 1, a picture of the track of a cosmic-ray particle believed to be an electron. Anyone who has ever studied the pictures of cloud-chamber tracks will at once be impressed by seeing how distinctly the droplets stand apart. This separation was achieved by letting half a second elapse from the instant when the electron shot through, to the instant when by expansion the gas of the chamber grew suddenly cool and the water-vapor suspended in the gas condensed itself as dewdrops on the ions. These ions, formed by the passage of the electron, had been diffusing through the gas during the half-second intervening, and the diffusion-process had served in the main to carry them apart (though there must also have been cases of ions approaching and possibly even combining with each other). The counting of these droplets is germane to the question as to whether the traversing particle was or was not an electron. This question, however, we leave till later, and turn to photographs in which the droplets of the tracks lie close together and are uncountable, because the expansion took place before there had been time for much diffusion. Tracks so formed have the advantage of sharpness over what they lose in detail.

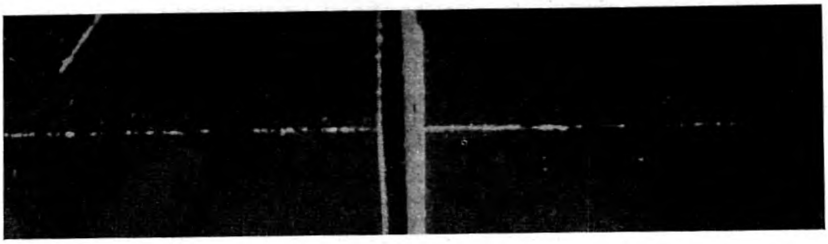


Fig. 2—Track of a particle, presumably a mesotron, traversing a metal plate without sensible deflection. (Auger; Universite dé Paris)

Figure 2 presents the track of a particle which traversed a plate of lead as it shot across the chamber. In passing through the lead, it underwent no sensible deflection; no other particle sprang from the lead; and there is nothing in the aspect of the track which differs on the two sides of the metal. It would be more impressive yet to present a similar picture for a particle traversing ten or fifty centimetres of lead, but here the practical limitations on the size of a Wilson chamber defeat the physicist, or at any rate no one has overcome them yet. Ehrenfest has lately circumvented them by the laborious scheme of setting up *two* Wilson chambers, one above the other, with as much as 9 cm. of lead or gold between them. However, the passage of single charged particles through thicknesses as great or even much greater is amply attested by the scheme of apparatus sketched in Fig. 3, even without the cloud-chamber there indicated by "Ch."

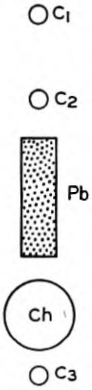


Fig. 3—Scheme of apparatus for observing very penetrative particles with counters and cloud-chamber.

In this sketch of Fig. 3, the objects C_1 and C_2 and C_3 are Geiger-Müller counters: that is to say, gas-filled discharge-tubes of a very special design, the two electrodes of each being an axial wire and a coaxial cylinder, and the electrode-size, voltage, and gas-content being very carefully adjusted. These long large cylinders, usually called simply "counters" without the prefixed names, are familiar sights in almost every laboratory where cosmic rays are studied. If a charged flying corpuscle penetrates such a tube, a momentary discharge takes place in the gas. If such discharges spring up simultaneously in all the three tubes of such a system as Fig. 3 exhibits, the event is recorded by a mechanism. ("Simultaneously" is of course a word which requires detailed exegesis; it meant at first that in all tubes discharges began within 0.01 second of each other, but this interval has been pushed down to .0001 second and lower.)

These events, the "threefold coincidences," do actually occur. Of course, since in each of the tubes a discharge occurs now and then by itself, some of the coincidences must be the result of chance; but the probable number of these meaningless ones can easily be estimated from the frequency and the duration of the individual discharges, and in the best experiments they are a small minority. For the great majority, the simplest of explanations is to attribute each of them to a single vertically-flying particle cutting through all of the counters in succession. Yet there are other thinkable causes, and confirmation of this simplest idea is needed. It was supplied when the cloud-chamber, "Ch" in the figure, was inserted. The chamber was compelled by mechanism to expand, always when and only when a three-fold coincidence happened; and at the great majority of its expansions it showed a vertical track. Figure 3 exhibits the arrangement of Street, Woodward and Stevenson at Harvard, who found the track of the traversing particle at 202 expansions out of 219. Auger and Ehrenfest at Paris had already set up *four* counters and a cloud-chamber and a block of lead in a vertical line, and found the track of the single traversing particle at fifty-five expansions out of sixty-nine. Another test is made by displacing one of the counters out of line with the others, whereupon it is found that the coincidences fall off in number sharply. And now to come to the point which most concerns us: there were 45 cm of lead between the counters in the experiment of Fig. 3, and 50 cm in the experiment by Auger and Ehrenfest, and no fewer than 101 cm in an early experiment of Rossi's with counters though without the chamber! Such is the power of penetration of some of the charged corpuscles of the cosmic rays.

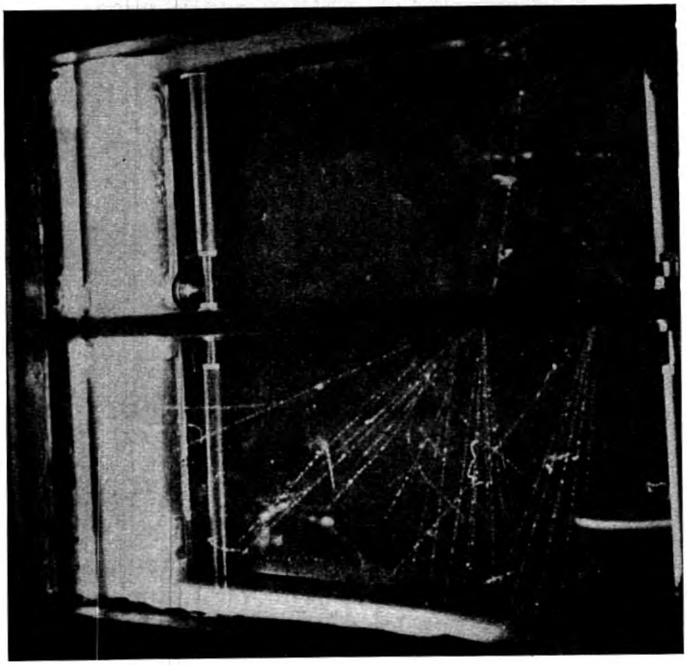


Fig. 4—Three showers, two evoked by charged particles and one presumably by a photon. (Street and Stevenson, Harvard University)

The reader has now been introduced to charged particles which bore through quantities of lead, apparently without doing or suffering anything. Next he is to be introduced to particles which begin to do something startling, when they have scarcely more than entered into a thin metal plate. This is vividly shown to him in Fig. 4, in which—after he can detach his eyes from the pretty sight beneath the transverse leaden plate—he will see that two of the "showers" beneath spring from the places where the metal was entered by two charged particles coming from above. These are accordingly called "shower-producing particles."

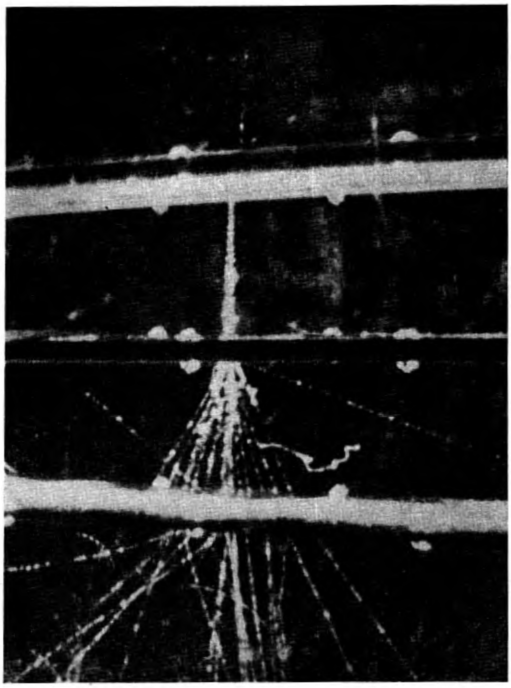


Fig. 5—Shower begun by a charged particle impinging on a 6.3-mm lead plate, and multiplied as it passes through a second such plate; in the third plate, 0.7 mm thick, only deflections occur. (Fussell, Harvard University)

Figures 5 and 6 and 7 show examples of showers even more gorgeous—regular cloudbursts, to continue with the metaphor (and indeed the term "burst" is often used as a synonym for "very large shower"). Of these, the special value of Fig. 6 is that the tracks that start in the gas itself bear witness to corpuscles of light—photons—included in the shower; for these are the tracks of electrons ejected by photons

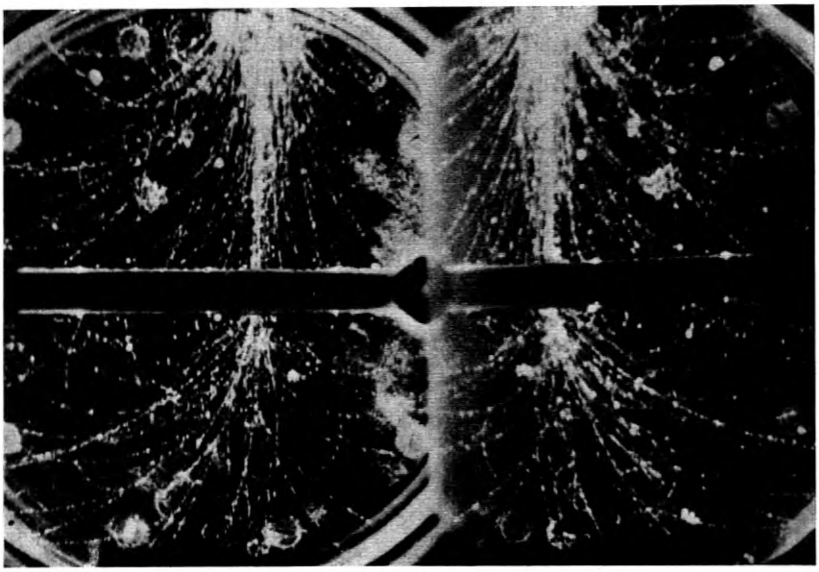


Fig. 6—Shower comprising photons attested by the (curled) tracks of slow electrons released in the gas. (Anderson and Neddermeyer, California Institute of Technology)

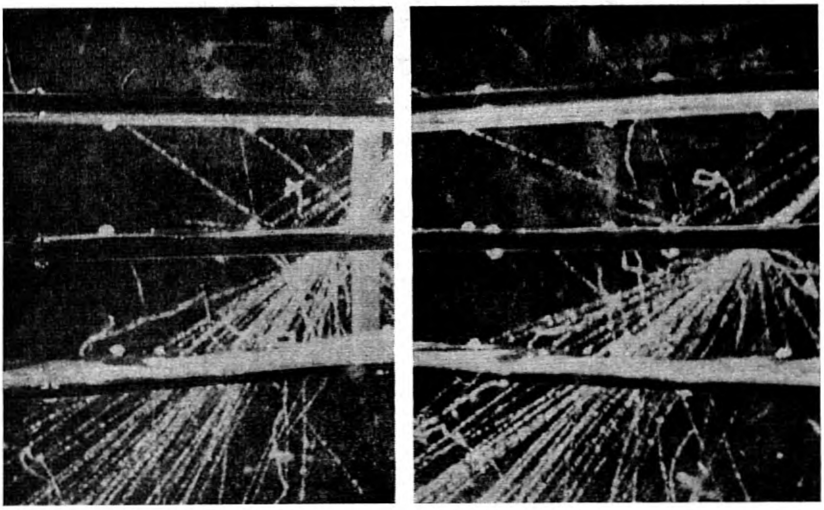


Fig. 7—Another example of a shower undergoing multiplication as it passes through metal plates. (Fussell)

from atoms of the gas. (The agent which bends them into curlicues is, of course, a magnetic field applied to the whole of the Wilson chamber.) Showers, then, comprise photons as well as charged particles. The special value of Figs. 5 and 7 is, that they show the progressive aggrandizement of showers as these pass onward through dense matter. This is called "the multiplication of showers." *Shower particles are themselves capable of being shower-producing particles.* One could not tell from these figures whether the multiplication is due to the charged particles or the photons, to either singly or to both. Here again the

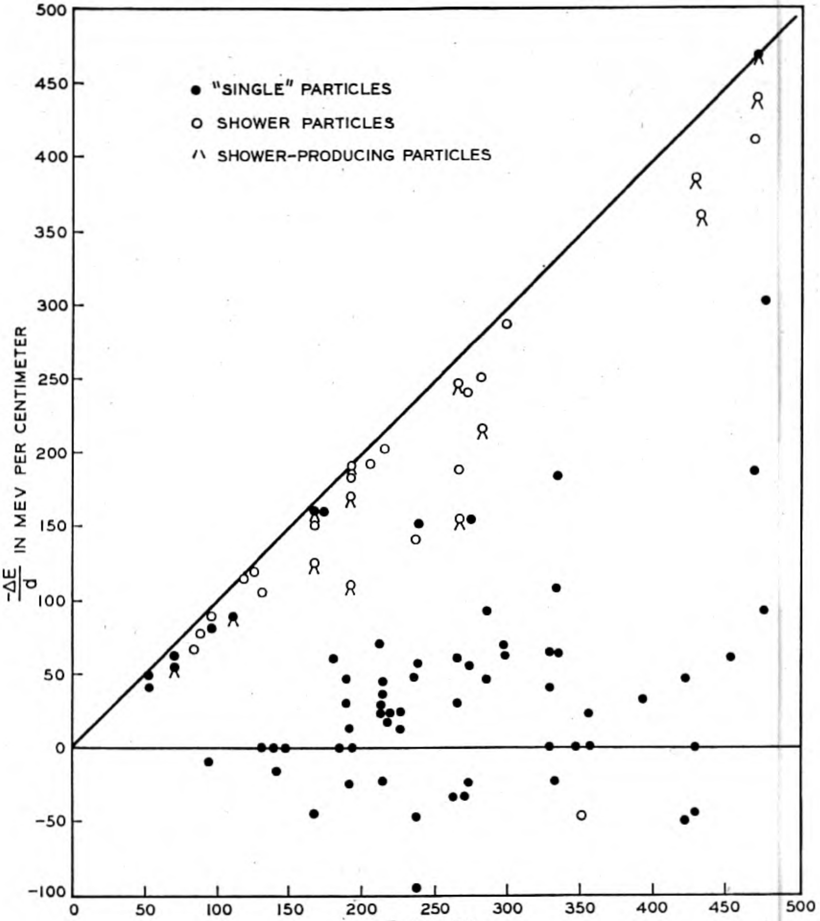


Fig. 8—Energy-losses per unit length of path (in Mev/cm) suffered by 94 cosmic-ray particles in traveling through platinum. (Anderson and Neddermeyer)

reader may consult Fig. 4, in order to notice that one of the three showers there depicted sprang from a place in the plate to which no charged particle came. This suggests that a photon may cause a shower, and that the multiplication of a shower already begun is due to the action of its charged particles and of its photons both.

Two classes of charged particles begin to take shape: the penetrating ones on the one hand, the shower particles and the shower-producing particles classified together on the other. To bring out another aspect of the distinction, I now turn to the data underlying Fig. 8.

These data are derived from cloud-chamber photographs such as Fig. 9 exemplifies. If the track of a charged particle is sensibly curved in such a magnetic field as it is possible to apply to a Wilson chamber, it may be possible to infer the momentum and the energy of the particle.[2] I digress to give the formulae, so as to make it clear just what can be deduced from what amount of knowledge. The elementary procedure consists in pointing out that the charged body describes a circle in the plane perpendicular to the magnetic field, and that consequently the force exerted on it by the field is to be equated to the product of its mass by its centrifugal acceleration. Putting ne for the charge (in electrostatic units) of the corpuscle, m for its mass, v for its speed and p for the magnitude of its momentum in the plane normal to the field, ρ for the radius of the circle and H for the field-strength, and writing down the two members of the equation, one finds:

$$Hnev/c = mv^2/\rho, \tag{1}$$

$$p = (ne/c)H\rho. \tag{2}$$

These equations remain valid when (as usually is the case with cosmic-ray electrons) the speed is so great that relativistic mechanics must be used instead of ordinary. At such high speeds equation (2) retains its aspect. Equation (1) may also be left unaltered, but one must be sure to remember that m is a certain function of v:

$$m = m_0\sqrt{/1 - v^2/c^2}, \tag{3}$$

m_0 being known as the "rest-mass" of the body.

---

[2] Curvatures of tracks being so very important in this field of research, it is necessary to examine with the greatest of care into all of the causes (apart from magnetic field) which may produce or affect them. Notable among these are currents in the gas, which are especially obnoxious if there is a metal plate in the chamber. Indeed it seems strange that the currents should not be more hampering than they are, considering the expansions which occur. Sometimes people observe that in the absence of magnetic field, there is a slight curvature of the tracks; then in the presence of magnetic field, they deduct this amount from the curvatures observed. The papers of Anderson and Blackett abound in information on these delicate questions.

Equation (2) does not involve the mass at all. In the usual loose phrasing, *$H\rho$ gives the momentum of the particle provided that its charge is known.* The like cannot be said for the energy, which is given by $H\rho$ only if both the charge and the rest-mass are known. For particles of the cosmic rays it is best to disregard the ordinary expression for kinetic energy ($\frac{1}{2}mv^2$) and adopt for good the relativistic expression mc^2, to wit, $m_0c^2/\sqrt{1 - v^2/c^2}$. Of this the portion m_0c^2 is not kinetic energy: it is the "rest-energy" associated with the "rest-mass" m_0, inseparable from the particle so long as this exists; it amounts to about half-a-million electron-volts or 0.5 Mev for the electron, to about 1000 Mev for the proton. The remainder may be called kinetic energy. For nearly all of the electrons and most of the other cosmic-ray particles, this remainder is by far the greater part. The dependence of the kinetic energy upon $H\rho$ is exhibited, for electrons and for protons, by Fig. 13 (page 213). One sees that for different masses a given $H\rho$-value leads to different energy-values, but also that the error due to an incorrect estimate of rest-mass becomes proportionately smaller as the $H\rho$-value increases. Yet the possibility of the error *is* always there, if the mass of the particle is not certainly known; and it affects many published "energy-values"

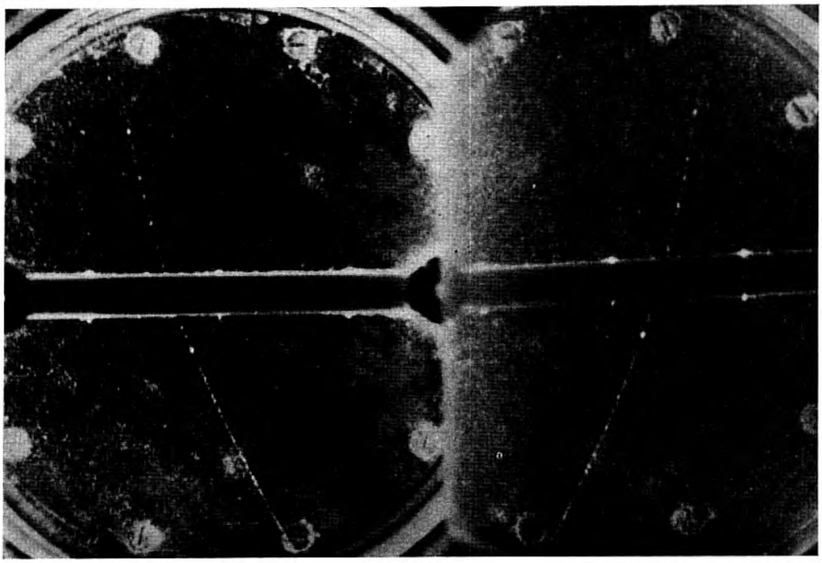


Fig. 9—Track exhibiting measurable and unequal curvatures on the two sides of a metal plate, thus indicating changes of energy and momentum suffered in the traversal. (Anderson)

based on the presumption—often admitted in the context to be more than doubtful—that the particles to which they refer are electrons. The danger might be mitigated by describing these as "quasi-energy-values" expressed in "quasi-Mev."—For actual electrons with momenta as great as those figuring in the cosmic rays, the energy-value in electron-volts is practically equal to 300 times the $H\rho$-value expressed in gauss-centimetres.

Many a cosmic-ray particle suffers no deflection that can be detected in its entire course across a Wilson chamber (diameter, 15 cm. or even more) in a magnetic field as strong as can be applied over so great a volume (field-strength, 20,000 gauss or thereabouts). One might well be tempted to think such a particle chargeless, for if this were the case, the field would have no grasp at all upon it; but if it were chargeless it could not ionize the molecules of the gas and therefore could not form the chain of ions on which the droplets are founded. In some of the finest of the experiments (those in Pasadena and those in Paris) a detectable curvature of the track would be shown if this were made by an electron of energy so enormous as $2 \cdot 10^{10}$ electron-volts (20,000 Mev!). The uncurved tracks accordingly speak of electrons of energies greater than 20,000 Mev, if these particles are electrons; and the inference is not much less drastic, if they are more massive than an electron.

We, however, are more interested, for the present, in the tracks which are sensibly curved; and most of all, in the tracks which are intersected by a metal plate and which show a curvature on one side of the plate and a larger curvature on the other (Figure 9). From the two ρ-values one can deduce the momentum-loss Δp and the energy-loss ΔE suffered by the particle in passing through the plate. (Yet I emphasize again that Δp is computable only if the charge is correctly guessed, and ΔE only if the rest-mass is correctly guessed in addition to the charge.) With this ambition Anderson inserted such plates for the first time into a Wilson chamber, in 1931. The idea had a wonderful and unforeseen result, some years ago recounted in these pages. Notice that above I spoke of the momentum-*loss* and the energy-*loss* suffered by a particle in going through a plate. In so doing I was making the assumption that it is a loss and not a gain which happens. If this highly plausible assumption is correct, then the sense in which the particle is traveling its path is knowable; it is from the side of the plate on which the curvature is less, to the side on which the curvature is greater. If the sense of the motion is knowable, so also the sign of the charge of the particle is knowable, being positive or negative according as the track is bent with its concavity toward the left or toward the right of an observer looking into the chamber from the north-seeking pole of his magnet. Without the plate, neither sense nor sign would be knowable except in the rarest of cases.[3] Anderson in August 1932 found on one of his photographs the track of a particle which by this criterion was positive, and which by the density of droplets along its track (we take up this topic later) he identified as an electron. He thus became the discoverer of the positive electron.

Concentrating on the measuring of ΔE after the excitement of the positive electron had subsided, Anderson presently found that its values are very fluctuating. Thus in 1934 he published the details of nine traversals, made by particles assumed to be electrons, through thicknesses of lead from 7 to 15 mm. (Even with a single metal plate the effective thickness varies, since corpuscles traverse the plate with varying degrees of obliqueness.) These were by no means identical in initial energy, this ranging from 38 to 240 Mev; nevertheless one might have expected the energy-loss per unit length of path in lead to be about the same for all, and yet the nine values thereof were scattered all the way from 18 to 120 Mev/cm! Such fluctuations suggest that the energy is lost in great amounts at a few events, and not in driblets at many. They did not deter Anderson and Neddermeyer from making such measurements on hundreds of later particles, classifying the particles into groups according to their energy-values, and averaging the energy-losses within each group. What then was found has a bearing upon the problem; but we pass over it for the time being, and consider in Fig. 8 the record of ninety-four particles which, during a later experiment, passed through a plate of platinum one centimetre thick.[1]

Plotted horizontally are the energy-values of the particles while above the plate, vertically the energy-changes divided by the lengths of path in the platinum. The axis of abscissæ is the locus of energy-losses imperceptibly small; the line slanting at 45° is the locus of energy-losses which are total, the particles shown on this line having been stopped by the plate. The fact that some of the representative points lie below the horizontal axis means only that for every particle the observers subtracted its energy below the plate from its energy above, irrespective of its direction of motion. Suppose that these

---

[3] One might be misled by the adjective "cosmic" into believing that all cosmic-ray particles come from above, their sense of motion making an angle of less than 90° with the downward-pointing vertical. Many, however, including Anderson's first positive electron, have been found by this criterion to be moving upward (i.e. at more than 90° to the downward-pointing vertical). The showers of Figs. 6 and 7 show that this is not a forced interpretation.

[1] I am indebted to Dr. Anderson for a plate exhibiting data thus far unpublished.

subjacent points correspond to upward-going corpuscles, and transfer them across the horizontal axis. Then, the sprinkling of points extends all the way from axis to slanting line; and this is the sign of fluctuations such as Anderson from the start had observed. Notice however that the representative points are of four aspects: solid dots and hollow circles, with or without downward-pointing barbs. The dots refer to tracks which were seen in the chamber singly; the circles, to particles which "entered the chamber accompanied by other particles." The lonely particles are prevailingly able to pass through matter without suffering energy-losses nearly so great as those which the others incur! Thus by itself and without any theory, Fig. 8 establishes a distinction between the singly-appearing corpuscles on the one hand, and those which appear in company on the other. Moreover the barbs are often attached to the hollow circles, bearing out the inference from Figs. 5 and 7 that shower particles are likely to be shower-producing particles; but rarely are they attached to solid dots, never to those which lie far off from the slanting line.

(This seems the best place for mention of the similar work now being done in England by Blackett and (J. G.) Wilson, in France by Ehrenfest. The Englishmen have set plates of gold, lead, copper and aluminium, of various thicknesses from 3.3 mm to 2 cm, into the middle of an expansion-chamber in Anderson's fashion; Ehrenfest, using a pair of cloud-chambers one over the other, was able to put between them a block of gold no less than 9 cm thick! Their way of reducing their data for plotting is not the same as that employed at Pasadena, and their diagrams therefore look very different[1] from Fig. 8. Their energy-range runs much further upward, as far as 5000 Mev, and the great majority of the particles which they plot lie beyond the limit of Fig. 8. Many of Ehrenfest's particles got through the great thickness of gold without losing anywhere nearly the whole of their energy, and are therefore to be classed as much more penetrating than electrons should be. So did nearly all of the particles of energy greater than 250 Mev observed in England, but there were a few of these which lost most of their energy in 0.33 cm of lead, and of these few about half seemed to belong to showers.

---

[1] For the benefit of those who may consult the original papers, I give the difference. Let E_1 and E_2 stand for the (quasi) energy-values of a particle before and after passing through a thickness d of metal; ΔE for $(E_1 - E_2)$; x for $\frac{1}{2}(E_1 + E_2)$. What is plotted by Anderson and Neddermeyer (Figure 8) is $\Delta E/d$ as ordinate and E_1 as abscissa. Blackett (in all his papers but the earliest), Wilson and Ehrenfest begin by subtracting from ΔE a quantity sd which is supposed to be the amount of energy spent by the particle in detaching electrons from atoms while traversing the metal (Blackett assigns the value 15 Mev/cm to s in lead, Ehrenfest takes 28 for gold); they then plot $(\Delta E - sd)/xd$ as ordinate and x as abscissa. Their ordinate (denoted by them as R) is then more nearly ready for comparison with theory.

At energy-values below 200 Mev Blackett finds almost no penetrating particles, a singular contrast with the Pasadena observations; he suspects that the penetrating particles become ordinary electrons when they are slowed down into this energy-range. I mention also the measurements made on some twenty penetrating corpuscles by Leprince-Ringuet and Crussard, leading to the exceptional conclusion that positives suffer smaller energy-losses than negatives.)

But granting that there are two sorts of particle with a right to different names: has either a right to the name "electron"? To settle this question, and for several other reasons, it is time to call upon theory.

It is now some thirty years since there entered into physics a German word, *Bremsstrahlung*, which can be translated literally into English as "braking radiation," and would no doubt be so translated if "braking" did not sound like another English word of entirely different meaning. This is chiefly observed emerging from X-ray tubes, being emitted from their metallic targets when these are struck by the stream of bombarding electrons. It consists of photons or corpuscles of light, each containing at least a part of the kinetic energy of one of the incident electrons. The distribution-in-energy of the photons makes it clear that the electrons frequently lose large fractions of their initial energy *en bloc*, throwing it off in individual parcels which are these photons (indeed it sometimes happens that the entire kinetic energy of an incident electron is shed in the form of a single corpuscle of light). This radiation forms the so-called "continuous X-ray spectrum" or "X-ray continuum" emerging from targets of X-ray tubes. With the spectrum-lines which are sometimes seen superposed on this continuum we have nothing here to do.

By the classical theory of thirty years ago this continuous spectrum is attributed to the slowing-down of the electrons as they penetrate into the metal, whence the name *Bremsstrahlung*. By the quantal theory of today it is still ascribed to the slowing-down, which must now be conceived as taking place in instantaneous jerks, occurring probably in the close vicinity of atom nuclei. At each of the jerks, the electron-speed is suddenly reduced and the kinetic energy goes forth in the form of light. The later theory in its quantitative form gives a competent account of the continuous X-ray spectrum as it springs from the tubes of the laboratory, with their bombarding electron-streams energized by voltages of a few tens or hundreds of thousands. For a long time nobody seemingly troubled to extend it to voltages of the order of thousands of millions; a futile extension indeed this would have been, so far as X-ray tubes are concerned.

When finally the extension was made by people interested in the cosmic rays, it turned out that according to the quantal theory the liability of electrons to these "radiative energy-losses" goes up so greatly with increasing speed, that electrons of even the cosmic-ray energies should not be able to bore their way through as much as five centimetres of lead!

After the meaning of this inference sank in, there ensued a period lasting for months (in 1935 and 1936) in which several eminent theorists were willing to concede that Nature must have set a limit to the scope of quantal theory. It was beginning to be believed that somewhere between the energy-range attainable in the laboratory and the energy-range manifest in the cosmic rays, there is a critical energy-value beyond which the electron escapes from the sway of the quantal laws, and is exempted from losing its energy by the process of *Bremsstrahlung*. This belief was an artifice for permitting the penetrative particles of the cosmic rays to be called by the name of electron. It might have remained a credible artifice, if the penetrative particles had been the only ones—if, that is to say, there had never been any evidence for the existence of particles among the cosmic rays having the properties required of electrons by the quantal theory. Such a situation may have seemed to exist at the time when the belief was dominant. It exists no longer, as the description of Fig. 8 has just suggested; but before considering further the data, I must introduce something more of what the theory has to say.

Since 1934 it has been known that a photon of energy greater than about one million electron-volts is capable, when in the vicinity of an atom-nucleus, of converting itself into a pair of electrons of opposite sign. About one million electron-volts—1.02 Mev, to be somewhat more precise—becomes "rest-energy" of the twin electrons, being incorporated with their rest-masses; the remainder ($h\nu - 1.02$, if by $h\nu$ we denote the photon-energy in Mev) becomes kinetic energy of the electrons. The process may be produced at command and exhibited to the eye, by projecting the photons known as gamma-rays against metal targets contained in expansion-chambers. The gamma-rays originally used for this purpose proceeded from natural radioactive substances; mostly they were those emitted by a certain substance (thorium C'') with a photon-energy of 2.62 Mev. Nowadays gamma-rays of energy several times as great can be produced by effecting certain transmutations, in the course of which (or afterward) they emerge from the new-born nuclei. Figure 10 shows an admirable example of an electron-pair formed out of such a photon. Moreover, the converse process is well-known: positive electrons falling against

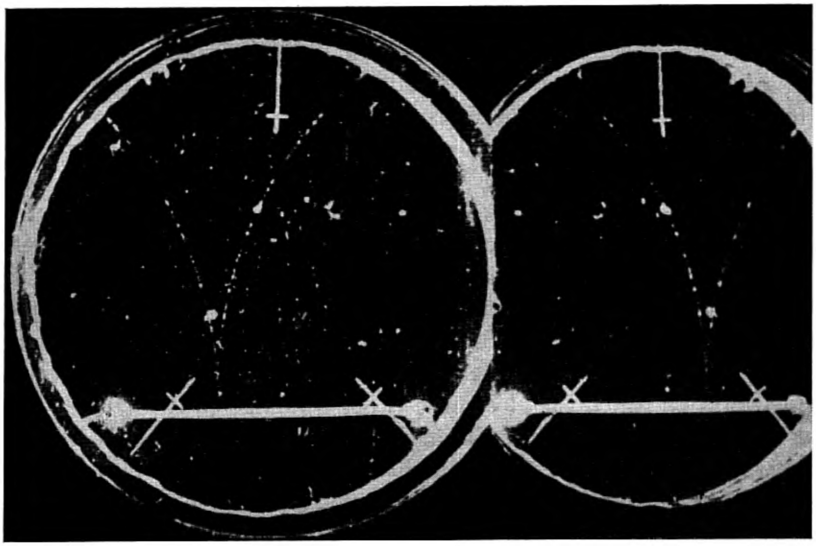


Fig. 10—An electron-pair born from a photon. (W. A. Fowler, California Institute of Technology)

a plate of dense matter bring about the emission of photons of energy 0.51 Mev, and these are just what are to be expected if the positive electrons (after being slowed down) unite with some of the innumerable negative electrons already in the plate and produce, at every such union, a pair of equal photons.[4] Much too abundant to be here described is the evidence for the ability of electron-pairs to pass into light and light to pass into electron-pairs, making it permissible to imagine a continual alternation of energy between these two so sharply contrasted forms.

Formation of the photons of *Bremsstrahlung* by electrons of enormous energy, and formation of electron-pairs out of such photons: these reciprocal processes engaged the attention of several theorists (Bethe, Heitler, Sauter, Weiszaecker, Oppenheimer) in the years 1933 and 1934. The problem was, to evaluate by quantal theory the chance that electron or photon would spend its energy in producing photon or electron-pair, while traversing given thickness of given element.

---

[4] Evidently this is not quite the converse of the process previously described, which if reversed would consist in the merger of a positive and a negative electron with the formation of a single photon bearing away all of their energy. Some evidence exists for the occurrence of this process. There is no sign of the fourth conceivable process (the meeting and merger of two photons to form two electrons) which must obviously be very rare in practice owing to the feeble concentration of photons in actual beams of gamma-rays. Nevertheless this last is the process first predicted by the theorist Dirac.

Approximations had to be made in the calculation, as nearly always in quantal problems; but they are supposed not to affect the rightness of the main result. To quote Oppenheimer's description of this result: "a beam of high-energy electrons should have a good part of its energy converted into photons in a centimetre of lead; in an equal distance these photons will be largely reconverted into pairs."

Such was the result from which, in 1935, it was inferred that quantal theory must be wrong because it was predicting something which could not be found in Nature; and from which, in 1936 and thereafter, it was concluded that quantal theory not only was correct but had made a splendid triumph, in explaining the phenomena of showers! It is not altogether clear why the later conclusion was not drawn at the start; perhaps the reason is, that as lately as the summer of 1936 fine photographs of showers were still rather rare, while such pictures as Figs. 5 and 7 with their examples of self-augmenting showers had not as yet been made. On the other hand it would be premature to say and misleading to imply that the process which the theory describes is in exact and quantitative accord with the observations on showers. There are at any rate good grounds for hoping that as the mathematics of the theory is more fully worked out and the art of the experiments refined, the agreement will grow better and better. The most that seems safe to say is, that now we have a general scheme for the interpretation of showers of a certain type, and a very hopeful prospect that this general scheme will be converted into a detailed and quantitative explanation as the mathematics of the theory on the one hand, the aptness and precision of the observations on the other hand are gradually improved.

By inserting the words "of a certain type" in the foregoing sentence, I leave open the possibility that showers may be classified into more than one type, and all of these but one be ascribed to other processes. This is no mere possibility but already almost a certainty. Certain showers which include "heavy tracks" due to protons or still more massive particles are ascribed to nuclear explosions provoked by cosmic rays. If a shower fails to undergo the "multiplication" illustrated in Figs. 5 and 7, it is taken as belonging to this other type. Exception made for such cases, it is strongly plausible to say that shower particles and shower-producing particles are electrons; that accordingly high-energy electrons exist among the cosmic rays, behaving as the quantal theory says that they should; and that consequently the other particles, setting themselves apart from electrons by their penetrative power and their failure to make showers, are of another sort.

Ability to penetrate matter, inability[5] to make showers: these are the complementary aspects of the property which distinguishes this other type of particle, the mesotron. If one wishes to contrive a particle having this property and differing otherwise as little as possible from the electron, how must it be done? The electron has the qualities of charge and mass; also those of spin and magnetic moment, but these are considered (perhaps wrongly) to be little or not at all concerned with shower-production. If we imagine the mass to be increased while the charge remains the same, the liability to *Bremsstrahlung* will diminish; for *Bremsstrahlung* occurs when sudden sharp deflections or decelerations occur, and these are less sharp and sudden the more massive the particle is. Now *Bremsstrahlung* is the prelude to the entire manifold process of the forming of a shower, and hence a mere increase in the mass of the hypothetical particle leads in the desired direction. The theory indicates that a particle with the electronic charge and a few dozen times the electronic mass will be penetrating enough. We do not need, however, to be contented with such vague intimations, for there is yet another phenomenon in respect of which the mesotron differs from the electron, and from this the mass can be deduced more sharply.

So far, we have been considering the passages of particles through solids. There, the paths are concealed, the adventures of the particles can only be inferred—from the difference between energy before and energy after traversal, or from the photons and the secondary electrons which are driven out of the solid. Now we are to consider the passages of charged particles through the gas of the Wilson chamber, which, unlike the scriptural way of the eagle through the air, are preserved for our inspection by the droplets. Figure 1 has shown to us a track in which the number of droplets in unit length of path can rather readily be counted. What does this number signify? And is it truly an indication of the mass of the traveling particle, as I hinted on an early page?

The latter question might perhaps be sufficiently answered without reference to the former; but for completeness, and for the sake of its own interest, the former ought to be treated more fully than it was in that brief earlier mention. In the voyage recorded in Fig. 1, nothing so drastic happened to the traversing particle as would have been the losing of a large part of its energy in the form of a photon of *Bremsstrahlung*. It lost its energy in driblets, spent in detaching electrons from molecules and giving them a small extra bonus of kinetic energy

---

[5] It is better to say "relative inability" since occasional showers are attributed to mesotrons, which perhaps operate by making a violent impact on an electron and so giving it the energy needful for starting the process.

with which to go wandering around in the gas. They had not speed enough to wander far, even in the half-a-second afforded them before the condensation. Probably they had already adhered to molecules before the condensing water immobilized them. One speaks of the droplets as being condensed partly on negative, partly on positive ions; the last-named are the molecules from which the electrons were reft. (If, during the half-a-second, an electric field of suitable strength is applied, the ions of the two signs drift in opposite ways, and when the water-vapor comes down there are seen two parallel trails of droplets with an empty space between.)

The simplest idea is that the traversing particle tears off one electron from each of many molecules through or near which it passes, and that half of the droplets are formed on these electrons and the other half upon the molecules bereft. This is too simple to be true. It is likely that sometimes the particle removes two electrons or more from a single molecule, so that there well may be more negative ions than positive. Much more serious is the certain fact that often when an electron is thus released by the direct action of the traversing particle, it shoots away with speed and energy enough to enable it to release one or several more from neighboring molecules. Now and then one comes on a cloud-chamber photograph in which there appears a track with branches (Fig. 11); each of these is the trail of an electron which has received a truly abnormal and extraordinary amount of energy. Much commoner, in fact universal, is the "beaded" appearance of such trails as appear in most of the pictures of this article: it is presumed that each of the beads is an unresolved cluster of droplets formed on a cluster of ions, all but one pair of them made in the indicated way. Occasionally one sees a picture in which the interval allowed for diffusion has been so happily chosen that the droplets in the clusters are far enough apart for counting, and yet consecutive clusters do not overlap. In making Fig. 1 the interval allowed was a little too long, and yet perhaps it is possible to think that the ions are denser in some parts of the trail than in others, as though they had been formed in clusters which have broadened almost but not quite to the point of losing their identity.

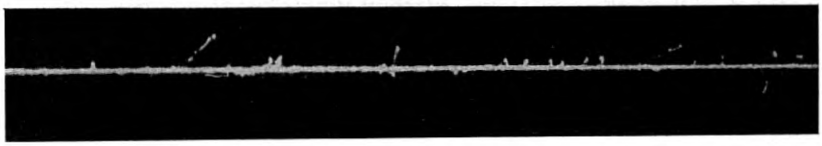


Fig. 11—Tracks of a charged particle bristling with short branching tracks, made by electrons ejected from atoms with energy sufficient to ionize. (Auger)

It is therefore necessary to distinguish, in mind if not in fact, between the "primary ionization" consisting of the electrons and the molecules torn apart from each other by the direct immediate action of the traversing particle,[6] and the "entire ionization" (sometimes called "probable ionization") consisting of these together with all the ions formed by the directly-ejected electrons. Under ideal conditions it is presumed that the measure of the former would be the total number of droplet-clusters,[7] the measure of the latter would be the total number of droplets, in unit length of path. Not many physicists have tried to evaluate both of these numbers. Of those who have, the data have been scanty, but the consensus of opinion is that the latter is about or not quite twice as great as the former. It is, however likely that the value of the ratio of the two is not important when one wants only to distinguish between electron and mesotron, as we shall presently see.

The problem of the primary ionization is one of the major tasks of theoretical physics. Classical and quantal theorists alike have spent great labor on the question: given a charged particle of specified charge and mass and speed traversing air (or any other gas), how many electrons will it set free from the molecules in unit length of path? At this point I will give only one of the results—or rather, something which is not a result at all, but a part of the assumptions. It is assumed that as the traversing charged particle flies along through or close to a molecule, it operates upon the electrons thereof by virtue of the ordinary electric forces between its charge and the charges of the electrons. It follows, then, that *whatever expression finally may be derived for the primary ionization must depend only upon the charge and the speed of the traversing particle, and not upon its mass.* (Mass and momentum of the particle must indeed be great enough to hold it on a sensibly straight course as it plows onward through the gas, despite its losses of energy as it detaches electrons; but this condition is always realized, with the corpuscles of the cosmic rays.)

I seem to have said that the primary ionization gives no power of distinguishing between an electron on the one hand, a particle of equal charge and different mass on the other. However, it *does* confer on us this power, for the reason that the curvature of a particle-track in a known magnetic field is a measure not of particle-speed but of

---

[6] Unluckily called "secondary ionization" by some of the German theorists.

[7] Best to observe the droplet clusters as individual entities, one would wish the expansion to occur before the ions have any time at all to diffuse. To attain this, Williams and Pickup caused the chamber to expand at moments taken at random, and trusted to luck for the appearance of cosmic-ray tracks formed at just the right instants. Luck served them with no fewer than four tracks betokening particles of a distinctive mass.

particle-momentum (equation 2). If by luck an experimenter should happen upon two tracks having the same curvature but made by particles having masses[8] standing to one another in the ratio (say) 100 : 1, the speeds would stand to one another in the ratio 1 : 100, and this might well entail a perceptible difference in the primary ionization. It would come to the same thing, if someone should take the data for a large number of tracks, and plot primary ionization as function of curvature: if there are really two kinds of particle differing in mass, there should be two sets of points lying along two curves, and from the ordinates of these curves at any abscissa the ratio of the masses would be derivable.

Perhaps the last sentence suggests that someone already has made this correlation, and has found that the points for all of the single or penetrating particles lie upon one curve, and all the points for shower-particles and shower-producing particles lie on another. This has not been done. The reason is, that many of the penetrating particles exhibit no perceptible curvature of track at all, and most of the others a very small curvature. The former are moving so fast that their momentum cannot even be estimated, except as being beyond a certain critical value. As for the latter, the speeds of even these are so great as to approach the speed of light; for a given momentum-value the speed varies only a little with the mass, and the primary ionization varies too little to serve as an index of mass. To make a profitable correlation, one must use only the particles of which the tracks are notably curved. Nearly all of these are shower-particles, which already are presumed to be electrons. To find a penetrating particle with a highly-curved track, one must find it when it is near to the end of its course and its energy wellnigh gone. Such is the principle which directed some of the recent successful searches for particles proclaiming themselves by their ionization to be more massive than electrons.

Before looking at the track of one of these particles, we ought to notice a couple of questions concerning ionization. One of them is: is the distinction between primary and entire ionization—or rather, our lack of perfect ability to make it in practice—likely to lead to trouble? Many observers are far from clear in reporting whether what they observe is more like the one or more like the other; but it seems probable that the second like the first is dependent only upon the speed and the charge of the traversing particle, not on the mass thereof; and this diminishes the dangers from confusing the two. The question is implicated with the second: to what extent do experi-

---

[8] Allowance being made for the relativistic dependence of mass on speed.

ment and theory aid us in identifying the shower-particles with the electrons? As to experiment, there exist the records of a few studies made by the Wilson chamber upon particles acknowledged to be electrons, of energy-values ranging from about 2 Mev downward to some 25000 electron-volts. In respect of the trend with energy, they agree fairly well with the assertions of the quantal theory; but when one inquires whether the absolute value for the number of clusters of ions in unit length agrees with the absolute value of the quantal expression for the primary ionization at any particular energy, one is confronted with the fact that the quantal expression contains a multiplying factor which depends on intimate details of the structure of the molecule, and is not exactly known. The quantal theory, however, predicts a minimum in the curve of primary ionization vs. energy, at an energy of about 2 Mev. Such a minimum (Fig. 12) was actually found by Corson and Brode in their study of some fifty particles of the cosmic rays, and probably is to be ranked as evidence for the electronic nature of these particles quite as forcible, as would be an absolute agreement between the observed ionization and the predictions of a reliable theory.

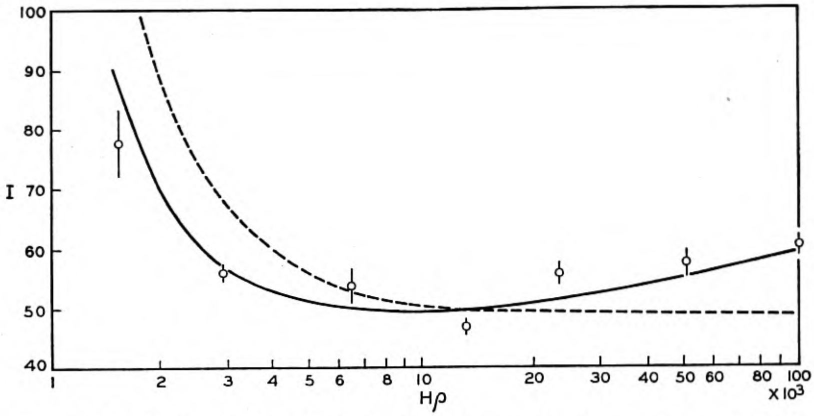


Fig. 12—Ionization-density (entire) along the tracks of cosmic-ray particles, plotted as function of $H\rho$. The continuous curve is that of a theoretical function containing a multiplying factor which has been adjusted to get the best fit to the data. (Corson and Brode)

Street and Stevenson, with a row of counters and an interposed cloud-chamber such as appeared in Fig. 3, adjusted their counters in such a way that the chamber expanded only when the counters above the chamber had simultaneous discharges and the counter below did *not*. A thousand photographs yielded to them the track of one particle having a notable curvature and displaying an ionization six times as great as that attributable to an electron; they inferred a "mass 130," *i.e.* a rest-mass one hundred and thirty times as great as that of an electron. Neddermeyer and Anderson transposed the bottommost counter into the very centre of the cloud-chamber itself,

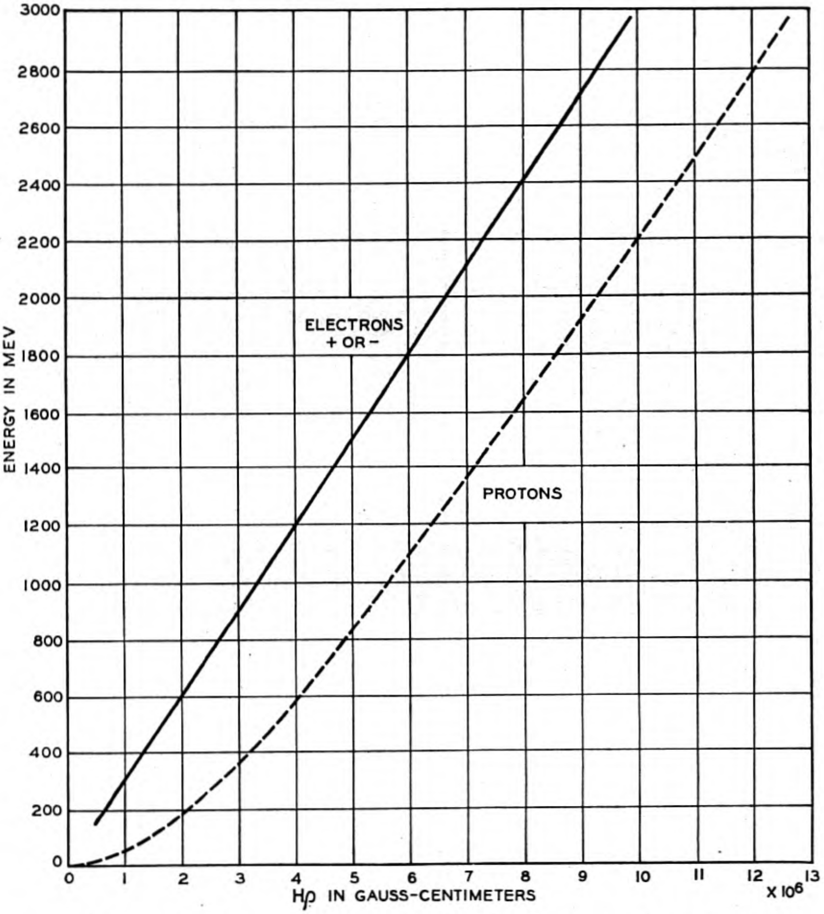


Fig. 13—Relation between energy and $H\rho$-value for electrons (of either sign) and protons. (Anderson)

and there it appears in Fig. 14, neatly intersected by the course of a particle which above it made a track lightly curved and thinly studded with droplets, and beneath it made a track sharply curved and densely congested. Comparing ionization with curvature along the track above and the track below, they found 240 to be a satisfactory ratio of the mass of the traversing particle to the electron-mass. Williams and Pickup, to whose technique I have already alluded (footnote 7 on page 210), observed four tracks of which three were compatible with a rest-mass of about 200, the remaining one requiring a mass-value between 430 and 800. A few more such tracks have appeared in the literature, but instead of describing them I turn for the climax to another and an exacter way in which Fig. 14 furnishes the desired value of mass.

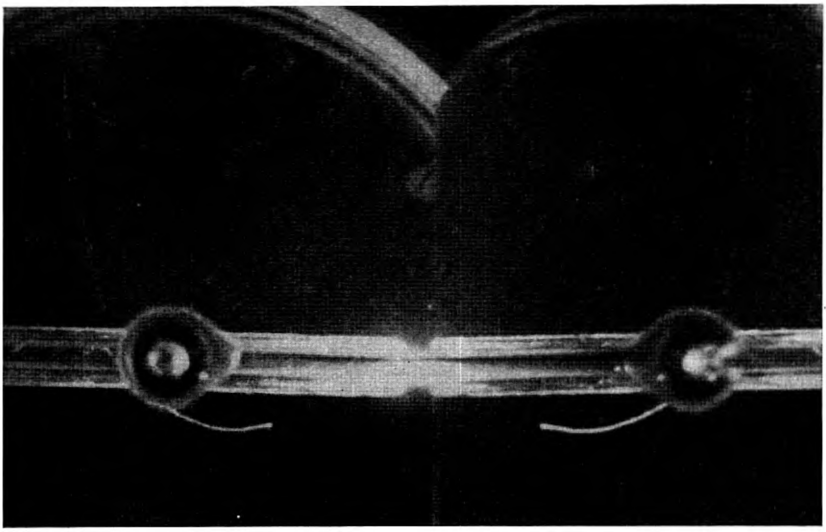


Fig. 14—Track of a mesotron slowed down by an obstacle in a Wilson chamber and finally brought to a stop in the gas of the chamber itself. (Neddermeyer and Anderson)

In Fig. 14, the track beneath the counter comes to a sudden end. One could take a sheet of coordinate-paper, and plot along the horizontal axis the curvature of the path as it emerges from the counter, and along the vertical axis the length of the path from that point of emergence onward to its end. This would give a single point of what is known as a "range-vs.-curvature relation" or a "range-vs.-momentum" relation. A second point can be found by measuring the thickness of the glass counter-wall twice traversed by the particle, converting it into an equivalent thickness of gas, adding this to the length of the path beneath the counter, and correlating the sum with the curvature of the path at the point where the particle enters the counter. Now, range-vs.-curvature relations are among the best-studied of the features of the charged particles already known—

electrons, protons, alpha-particles. These two points pertaining to the particle of Fig. 14 lie far from the curves appropriate to any of the three. An electron departing from the counter in a path of such a curvature as there is shown would have traveled 2000 times as far before reaching the end of its course! a proton, on the other hand, only one seventy-fifth as far! This at the moment is deemed the sharpest and most clear-cut evidence for the existence of a particle intermediate in mass between proton and electron, to which Anderson now assigns a mass of 220 (± 35) times the electron-mass.[9]

It is fitting to end this article by mention of several other kinds of evidence which have bearing on the question of the mesotron; mainly they are relatively indirect, and would require much space to describe and assess. Inferences have been drawn from the number of electrons ejected with high energy from metal plates by penetrating particles traversing these: J. G. Wilson derives a mass-value greater than 100. A curious inference has been drawn from the deflections suffered by these particles in traversing metals: the magnitude of these should by theory be independent of the mass of the particle—since it *does* appear to be the same for penetrating particles as for electrons, it is deduced that the mesotron and the electron can differ only in mass. Inferences have been drawn from the trend of cosmic-ray intensity with elevation in the atmosphere, and from the trend of cosmic-ray intensity beneath metal screens as function of the material and thickness of these last (it was thus that Auger as early as 1934 was led to suspect the existence of two kinds of charged particle among the rays).

Inferences have also been drawn from nuclear theory. To enter adequately into this difficult field is impossible here: it must suffice to say that Yukawa conceived, as a constituent of nuclear structure, of a particle possessing the charge of an electron and a mass of about the magnitude which the mesotron appears to have, and possessing in addition the quantity of *instability*. The "Yukawa particle," that is to say, has the qualities demanded of the mesotron, and in addition is liable to emit an electron; what is left behind is then a neutral particle which could elude observation. The emission is expected to follow the law familiar in radioactivity, the durations of individual Yukawa particles being distributed according to the law of chance about a mean value. Is there evidence that the mesotron behaves in this way?

---

[9] Values diverging from this by more than the estimated uncertainties have been published by other observers of other particles, and may betoken an underestimate of the uncertainty or the existence of particles of several masses. A "nomograph" for facilitating the evaluation of mass from curvature of path combined with ionization-density or range is given by Corson and Brode.

For this there *is* some evidence, of the following kinds. First let us compare (in imagination) the number (per unit time per unit area) of penetrating particles flying vertically downward and the number flying obliquely downward. The comparison can be readily made with such an apparatus as that sketched in Fig. 3, the cloud-chamber being superfluous and the lead absorber reduced to the least thickness sufficient to stop electrons; the axis is oriented first at 90° and then at various lesser angles θ to the horizontal plane. Even the whole of the atmosphere is insufficient to stop such mesotrons as the cloud-chamber discloses; and yet the observations show a marked decline of the number thereof as θ decreases. But the particles which travel obliquely traverse a greater distance from the top of the atmosphere than those which come vertically down, and take a longer time in doing so; the decline of number with decrease of θ may therefore be ascribed to the perishing of the mesotrons *en route* to the apparatus as the route grows longer and longer. Second: Let us compare the effect of the obliquely-traversed atmosphere with that of a sheet of lead in cutting down the number of particles arriving at the apparatus. One must make a guess as to the thickness of lead which would be required to produce a falling-off of the number of particles equivalent to that observed in the atmosphere, if the falling off were due to actual stopping of mesotrons in air and lead respectively, and the impermanence of the mesotron did not enter in at all. It is commonly conjectured that the equivalent thicknesses of lead and air would stand to one another inversely as the densities of these materials. When, however, the effects of such "equivalent" thicknesses are compared, it is found that the falling-off beyond the lead is decidedly less than that beyond the air. Now the mesotrons take very much less time for traversing the sheet of lead than the wide expanses of the atmosphere; and the "anomaly," as it has been called, is tentatively explained by assuming that few of them perish in the lead, many in the long journey through the atmosphere.

Estimates of the mean life of the mesotron thus made yield values of the order of a millionth of a second. It is supposed by many that the mesotrons are born in the upper layers of the atmosphere. Such conjectures, however, lead beyond the scope of this article, which must be confined to these few recent fruits of the seemingly exhaustless cornucopia of the cosmic rays.

SELECTIONS FROM THE LITERATURE

ENERGY-LOSSES OF PARTICLES TRAVERSING METALS: Anderson and Neddermeyer, *London Conference on Nuclear Physics* (1934); *Phys. Rev.* **50**, 263 (1936); Neddermeyer and Anderson, *Phys. Rev.* **51**, 884 (1937); Blackett and Wilson, *Proc. Roy. Soc.*

160, 304 (1937); Blackett, *ibid.* **165**, 11 (1938); Wilson, *ibid.* **166**, 482 (1938); Leprince-Ringuet and Crussard, *Comptes Rendus* **204**, 112, 240 (1937); Ehrenfest, *Comptes Rendus* **207**, 573 (1938).

PENETRATING PARTICLES DETECTED WITH COUNTERS, WITH OR WITHOUT CLOUD-CHAMBERS: Street, Woodward and Stevenson, *Phys. Rev.* **47**, 891 (1935); Street and Stevenson, *Phys. Rev.* **51**, 1005 (1937); Auger and Ehrenfest, *Comptes Rendus* **199**, 1609 (1934).

TRACKS OF PARTICLES WITH CHARACTERISTIC IONIZATION DENSITIES: Electrons: Corson and Brode, *Phys. Rev.* **53**, 773 (1938). Mesotrons: Street and Stevenson, *Phys. Rev.* **52**, 1003 (1937); Ehrenfest, *Comptes Rendus*, **206**, 428 (1938); (E. J.) Williams and Pickup, *Nature* **141**, 684 (1938); Maier-Leibnitz, *Naturwiss.* **26**, 677 (1938); Neddermeyer and Anderson, *Phys. Rev.* **54**, 88 (1938), and literature there cited.

THEORY OF SHOWERS: Oppenheimer, *Phys. Rev.* **50**, 389 (1936); Carlson and Oppenheimer, *ibid.* **51**, 220 (1937); Bhabha and Heitler, *Nature* **138**, 401 (1936), *Proc. Roy. Soc.* **159**, 432 (1937); Bhabha, *Proc. Roy. Soc.* **164**, 257 (1938); Montgomery and Montgomery, *Phys. Rev.* **53**, 955 (1938).

DEFLECTIONS OF PARTICLES: Blackett and Wilson, *Proc. Roy. Soc.* **165**, 209 (1938).

ELECTRONS RECOILING FROM IMPACTS OF MESOTRONS: Wilson, *Nature* **142**, 73 (1938).

INSTABILITY OF MESOTRON: Blackett, *Nature* **142**, 992 (1938); Rossi, *ibid.* 993; (T. H.) Johnson and Pomerantz, *Phys. Rev.* **55**, 104 (1939).

GENERAL REVIEWS: Euler and Heisenberg, *Ergebnisse d. exakten Naturwiss.* **17**, 1 (1938); Froman and Stearns, *Rev. Mod. Phys.* **10**, 133 (1938).

Hurricane and Flood—September 1938

By **W. H. HARRISON**

Editor's note: The following was presented by Mr. Harrison as the closing address of a symposium on the effects of the hurricane and floods of September 21, 1938 on transportation, power and communication utilities. The symposium was held in New York at the Winter Convention of the American Institute of Electrical Engineers, Thursday, January 26, 1939. After the close of the meeting a motion picture on the hurricane prepared by the Bell System for the information of its own employees was shown.

THE experiences of the telephone companies are naturally much the same as those already described. The aftermath tally showed that more than one-half million telephones were put out of service—in the New England States about thirty per cent of the telephones in that area. Through the destruction of toll lines, the storm temporarily cut off telephone communication with the outside from over two hundred towns. The total damage to telephone plant was in the neighborhood of ten million dollars.

The story of restoration—the immediate provision of emergency services—the handling of emergency supplies in unprecedented quantities—the augmenting of forces locally to supplement the normal forces—and the mobilization of forces from other areas—all are replete with engineering interest and are very intriguing, but it would not be appropriate to take the time to tell the story here. A few facts will give you a sketchy idea of the situation.

As to materials:

3,500,000 feet of lead covered cable
54,000,000 feet of paired wire
7,000,000 feet of steel strand for guys and supporting cables

As to mobilization of forces:

Local construction forces were expanded from 3,000 to 5,000. In addition, 2400 highly skilled linemen, cable splicers and installers and over 600 fully equipped construction trucks and other special motor vehicles were brought from fourteen other telephone companies as far south as Virginia and as far west as Nebraska and Arkansas.

Of striking significance in the prompt restoration of service was the traditional Bell System background of standardization of materials and methods. This standardization greatly facilitated the collection of large quantities of suitable supplies and made possible maximum effectiveness of the men who came from many parts of the country. The striking effectiveness of these measures is a great tribute to the engineers who long ago by their recognition of the value of standardization laid the broad foundation for this effective work.

In every disaster much is learned with regard to formulating plans and caring for specific situations. Of interest in this specific situation, there had been serious floods in much of this territory in 1936. The experience at that time pointed to certain precautionary measures and we know of no case where these did not prove effective in the present situation. For example, while the water rose five feet above the ground floor level of the main telephone building in Hartford, it was successfully kept out of the building by bulk-heads about the doors and windows, provided since the 1936 flood. Also, at various places where lines had been carried away due to the failure of bridges or other forms of river crossings the restored lines did not fail.

Over and above all of these more or less specific points, which I might say are somewhat routine, lies a broad engineering fundamental vividly illustrated by this whole experience.

Engineers by their work have made a pattern of life which has come to make individuals and communities dependent to a large extent in their day-to-day activities and mode of living, on the proper functioning of the services of power, transportation and communication.

Having done this, they have seen their works fall before the fury of nature—have seen the utter disruption of the organized scheme of life, with all the anguish that goes with such disruption.

It is in the light of this experience that an engineering fundamental of first magnitude presents itself, and one which offers a long range problem that is going to call for nicely balanced judgment, both on the part of the engineer and the management. This fundamental stands out clearly—dependability of service, and specifically the degree to which dependability can soundly and wisely be built into the physical plant.

It is trite to say that dependability is fundamental to good service, that it is of prime consideration in the design, construction and operation of all communication, power and transportation facilities. On the other hand, it would be foolhardy to assume that any man-made structure could completely withstand the fury of the elements, as typified by this storm.

Consider the circumstances. For four days rain was progressively heavier. It totaled between five and ten inches at many New England points. At some places more than six inches of rain fell in one day. As a result large rivers were brought to flood stage and small brooks and streams became raging destructive torrents. And then came the hurricane—then the seas. Wind velocities as high as from 120 to 180 miles per hour have been reported. Raging flood and tidal waters inundated important sections of many communities. Our services extended over the entire band of the storm and we can definitely trace the relationship of high wind velocities and resultant damage.

Another important circumstance, and bearing particularly on engineering consideration, is that nothing like this had happened in this area since the year 1815.

Obviously, to build plant to be unyielding to the sea and to be hurricane tight against such occurrences at century intervals would be as unsound as to ignore them altogether. Thus a challenge is presented to the engineer, taxing his best judgment. On the one hand, not failing to take every reasonable precaution in the future design of the plant, such as the avoidance of known exposures, the provision of alternate routes, the use of emergency restoration facilities of every conceivable character, adequate emergency operating routines; and on the other hand, not to be led by the tragedy of the storm to recommend extreme construction and operating procedures such as wholesale substitution of underground for aerial plant, which would obviously not be in the public interest.

This, it seems to me, is the broad lesson that we draw from this experience and the challenge presented to the engineer.

It was my good fortune to have been in the midst of the restoration work. It was comforting and inspiring to see how the men and women of all service agencies responded to the call, each presented with a trying problem of his own but ever ready to lend helpful and effective cooperation to those in other utilities, and all motivated with the common objective of maximum service to the community in this period of great distress. I know we in the telephone end could not have done our job had we not had the help of others, including the highway and other public agencies.

The final measure of any man's work is, has it been for mankind? A grateful public has put the mark of approval on the work of the men and women of the utilities and transportation groups in the stricken area. My admiration for them knows no bounds. Frequently when we fail of expression we turn to the pens of immortals.

Two lines in one of Kipling's poems—"Sons of Martha"—beautifully express the work of these men and women:

> "Not as a ladder from earth to Heaven, not as a
> witness to any creed,
> But simple service simply given to his own kind
> in their common need."

A Terrain Clearance Indicator*

By LLOYD ESPENSCHIED and R. C. NEWHOUSE

There is described a radio altimeter that gives continuously on the plane a measurement of the separation between the plane and the earth's surface or projections therefrom. There is projected from the plane and reflected from the earth back to it a very short radio wave, the frequency of which is continuously swung back and forth. The returned wave is thereby made to differ from the outgoing wave in frequency by an amount that is proportional to the echo path; and the difference or "beat" frequency is indicated on a frequency meter calibrated in feet of separation. The paper outlines some of the early efforts in this field, some of the technical problems involved, the theory of the system and the practical experimental results that have been obtained.

INTRODUCTION

THE problem of an altimeter for aviation has engaged the attention of many inventors and experimenters for twenty years or more. As a result, about every conceivable fundamental method of attacking the problem, by the utilization of acoustic or electric phenomena, is disclosed in the art, including the many U. S. patents on the subject.

The familiar aneroid altimeter has reached a high degree of perfection and enables the pilot to maintain level flight at any desired altitude but it gives no clue as to the variation of the elevation of the terrain beneath. The pilot has to know his position at all times and perform a mental calculation, in order to know his height above the ground at any given moment. A number of airplanes have drifted off their normal courses and have crashed on higher ground.

An altimeter based upon the use of a sound echo is subject to two fundamental limitations. The first of these limitations is the extremely high noise level produced by the airplane's motors and propellers, which tends to submerge the relatively weak echo at heights of more than a few hundred feet. The second is that the speed of sound is not enough greater than the speed of airplanes. At a height of one thousand feet approximately two seconds are required for a sound to travel to the ground and return. In this time interval a modern airplane would travel six hundred feet and the clearance may have changed materially.

---

* Read before the Institute of Aeronautical Sciences at the Chicago meeting, November 19, 1938, and to be printed in the Journal of the Institute.

There is in radio the corresponding phenomenon of an echo, an electric-wave reflection. The velocity of a radio signal is so great that an echo from the earth's surface is almost instantaneous; in fact, the time interval is so small as to give rise to a problem in measuring it. For instance, for heights less than a thousand feet the time to be measured is less than two millionths of one second.

The method used in the present instrument is extremely simple in theory. A radio transmitter is provided on the airplane which sends toward the earth a signal, the *frequency* of which changes at a definite rate with respect to time. The signal is reflected by the earth and returns as an echo after a time delay equal to twice the height, divided by the velocity of propagation. During this interval the frequency of the transmitter has changed and now differs from that of the echo by an amount equal to the product of the rate of change of frequency and the time of transit. The reflected wave is combined in the plane receiver with some of the outgoing wave energy and the difference or "beat" frequency is measured by a frequency meter. Since the reading of the meter is that of the "beat" frequency, it is proportional to the time delay of the echo and, hence, to height and thus can be calibrated directly in feet.

EARLY EFFORTS

The evolution of this method is interesting because it illustrates how one art is built upon another, and also the familiar story of separate inventors arriving at the same answer almost simultaneously, actually somewhat in advance of the existence of instrumentalities having the characteristics required to make the invention practically serviceable.

Many systems employing electromagnetic waves for the purpose of indicating altitudes of an aircraft have been proposed.[1] Among early workers in this field who independently of each other were concerned with methods involving frequency modulated waves were J. O. Bentley[2] of the General Electric Company; Professor W. L. Everitt[3] of Ohio State University and certain students in his department of Electrical Engineering including the junior author[4] and M. W. Hively; and the senior author.

Under the direction of Professor Everitt, some experimental work on the frequency modulation method, using wire lines, was undertaken in the school year 1928–29. On the basis of this work a grant was made by the Guggenheim Fund for the promotion of aeronautics and an investigation was continued with experimental tests, during the following school year under the auspices of the Ohio State Engineering Experiment Station. The experiments were reported upon in the bulletins of the Station, and in a graduate thesis [5] of the junior author and J. D. Corley.

As early as 1920, the senior author proposed the use of electric wave reflection in railway safety systems [6] and entertained the idea of frequency-modulated transmission with beat-tone detection for measuring distance along a track. Radio wave reflection for aircraft altitude determination was considered at times from 1926 to 1930 when a patent application was filed [7] for an arrangement similar to that which has been worked out, including the use of a frequency meter to give continuously a visual indication of the altitude.

At that time, however, a really practical terrain clearance indicator could not be built due in large part to the lack of suitable radio instrumentalities. Vacuum tubes capable of operating on frequencies approximately fifty times higher than those generally available were indicated as necessary before a satisfactory system could be built.

A long-range program, however, of vacuum tube development for high frequencies was under way in Bell Telephone Laboratories. This resulted in the production of suitable tubes, and they were described by A. L. Samuel [8] to the Institute of Radio Engineers in October, 1937. One of these was capable of providing a stable output of between five and ten watts at a frequency of approximately 500 megacycles, so it became feasible to undertake the development of a practical terrain clearance meter.

The Japanese have been experimenting recently with apparatus operating upon the same basic theory and a paper [9] was published in Japanese in 1936. A later paper [10] was published in English in 1938 by the same author, which describes the apparatus and the results of tests made on the ground over short distances with the equipment at rest.

TECHNICAL PROBLEMS

At the time this development was undertaken a number of questions presented themselves as to what the earth's surface would do to the incident wave in reflecting it. It seemed possible that the signal might be so scattered and broken in reflection by small irregularities that the echo would be more like static than a useful signal.

Even if the reflected signal proved satisfactory over the smoother surfaces, it was hard to predict what would happen when flying over timber land or over very irregular mountainous terrain. There was also the question of what would happen when the surface happened to be that of a city where an airplane flying at 250 to 300 feet per second passes over several buildings and streets with abrupt altitude changes of possibly hundreds of feet several times in the course of one second.

Even with the most directive systems that can be devised, the beam radiated from the airplane is so spread that echoes can be expected to arrive simultaneously from several surfaces, for instance from both the leaves on the trees and the ground between the trees, or from the top of a building and from the adjacent street.

Several problems were anticipated in the apparatus itself. The theory is based upon a frequency-modulated signal free from any amplitude modulation, and it was questioned whether a transmitter could be built to operate on ultra-high frequencies which would be sufficiently free from amplitude modulation, when subjected to the vibration of the airplane, to be satisfactory. Since the receiver utilizes both the direct and reflected signals in making the altitude measurement, it is necessary that some signal be picked up directly from the transmitting antenna but not enough to overload the receiver and thus prevent reception of the echo. It was expected that difficulty would be encountered in sufficiently reducing the direct signal.

After considering all these problems, it was decided that the cheapest and easiest way of determining the answers was to build the apparatus and try it out to see if correct operation could be obtained, first, under the more or less ideal conditions of flying over smooth water and, then, over less favorable surfaces.

Most of the measuring equipment available for radio frequency test work is useless at ultra-high frequencies. Hence, it was necessary to get the system functioning as a whole before any means were available for determining the best adjustment of the radio-frequency parts of the system. Because of the difficulty of providing, while on the ground, an adequate reflector at distances of from a few feet to thousands of feet from the apparatus, it was necessary to install the equipment in an airplane very early in the development and make most of the tests during flights. Nearly a hundred airplane flights were made in one of the Bell Telephone Laboratories' airplanes during the development period of seven months which preceded the public demonstrations made in the United Air Lines Flight Research Airplane.

Operation and Theory

The fundamental parts of the altimeter in relation to their application are shown in Fig. 1. An ultra-high frequency oscillator is provided, whose frequency is varied up and down by a modulator which consists of a small rotating variable condenser driven by a motor. The oscillator is connected through a coaxial transmission line to a transmitting antenna which is located on one of the lower surfaces of the airplane. The signal is radiated downward by this antenna. A radio receiver is connected through a similar coaxial line to a second antenna similarly located but arranged in such a way that a minimum of direct signal is received from the transmitting antenna and as much echo as possible from the ground. The direct and reflected signals are applied to a detector circuit in the receiver. The output of this detector is a signal of a frequency equal to the instantaneous difference existing between the direct and the reflected signals and is proportional to the height of the plane above the terrain. This signal is amplified by the receiver and applied to a frequency meter or counter circuit which is so designed that a current proportional to the frequency and, hence, to the height flows through a meter calibrated in feet and located on the airplane's instrument panel. A number of types of

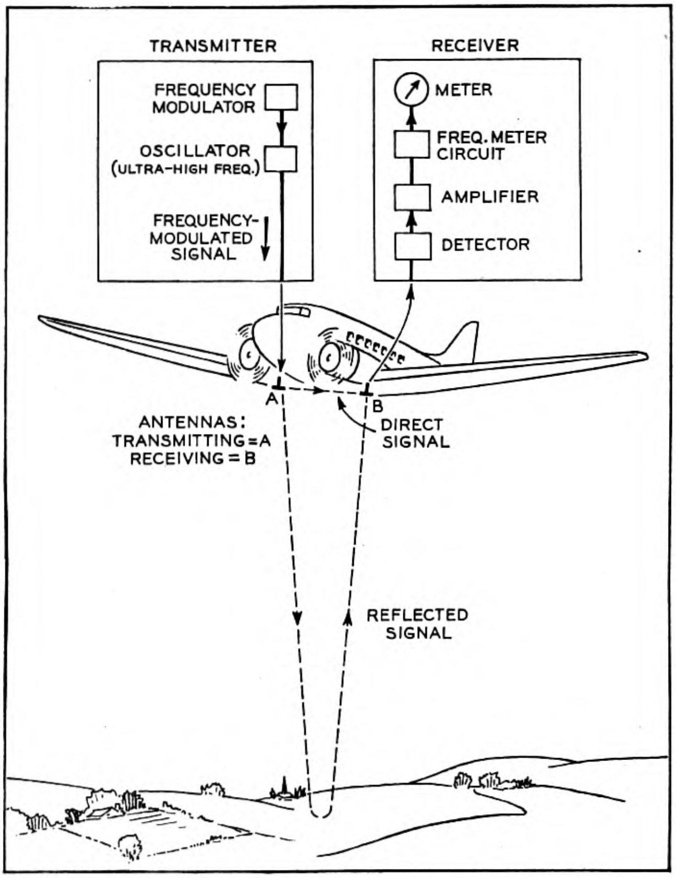


Fig. 1—Overall system.

indicating frequency meter circuits[11] of the condenser charge and discharge variety have been described in the technical literature.

The operation of the system can be understood more easily by reference to Fig. 2. The variation of the transmitter frequency with time is indicated by the solid sawtooth line.* The value of the ordinate of this curve at any point is the transmitter frequency for the corresponding time. The frequency is varied from $F_{\text{MIN.}}$ up to $F_{\text{MAX.}}$ and back F_m times per second, so the rate of change of frequency is $2\Delta F\ F_m$ when ΔF is substituted for $F_{\text{MAX.}} - F_{\text{MIN.}}$ The linear frequency variation shown, while ideal, is not essential for the successful functioning of the apparatus. The dashed sawtooth line represents the variation with time of the frequency of the echo signal from the earth's surface. This curve is displaced to the right by a time equal to twice the height divided by the velocity of propagation, or, in other words, the time it took the radio signal to go down to the earth and

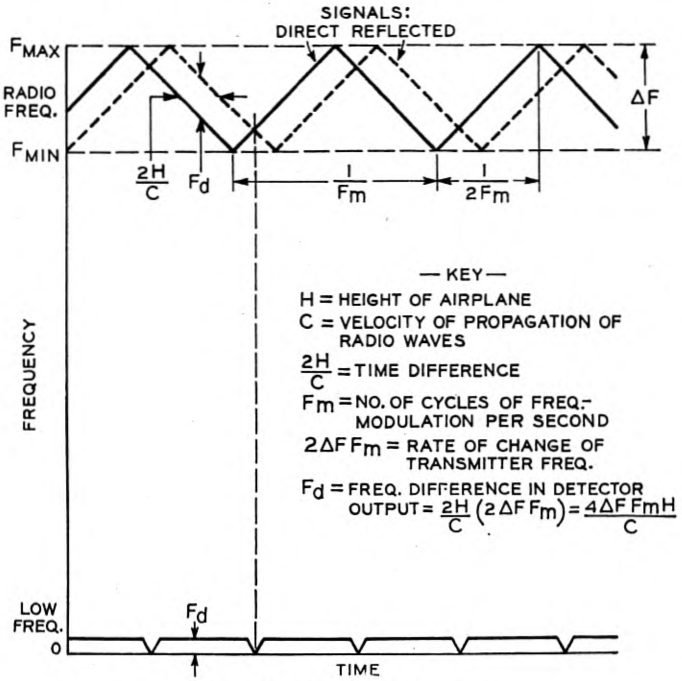


Fig. 2—Operating theory.

---

* A simple harmonic wave that changes in frequency from instant to instant is no longer a single frequency but a series of discrete frequency components. In the present instance, the number of cycles of frequency modulation per second is small compared to the transmitter frequency swing, so the spectrum occupied by the signal is substantially that of the swing itself.

return. This results in a frequency difference between the direct and reflected signals which is equal to the product of the time delay $2H/C$ and the rate of change of frequency, and is given by the equation,

$$F_d = 4\Delta F\ F_m H/C \text{ cycles per second.}$$

The difference is plotted again at the bottom of the diagram and appears as a series of trapezoids of height F_d. The time delay, $2H/C$, has been greatly exaggerated in comparison with $1/F_m$, the time interval corresponding to one cycle of frequency modulation, in order to make the difference, F_d, large enough to show on the diagram. F_d is actually only a few cycles in hundreds of millions. It will be noted that F_d drops momentarily to zero twice for each complete sawtooth variation of the transmitter frequency. This is due to the necessity of varying the transmitter frequency first up and then down, instead of forever in one direction. Hence the theory must be considered from the standpoint that one altitude measurement is made for each upward and another for each downward sweep, ΔF, of transmitter frequency so that a total of $2F_m$ measurements are made per second. The number of cycles of frequency F_d, occurring during one frequency sweep, is

$$F_s = F_d \times \frac{1}{2F_m} = 2\Delta F\ H/C,$$

since $\frac{1}{2F_m}$ is the time of one sweep, ΔF. F_s is directly proportional to both the height and to the amount of transmitter frequency change, ΔF.

The fact that $2F_m$ separate measurements are made per second is important only when considering small altitudes. The height which gives a value of unity for F_s corresponding to a frequency meter signal of $2F_m$ cycles per second is the minimum height which can be indicated since lower altitudes give the same reading. Lower altitudes cause only a fraction of a cycle of frequency, F_d, to be generated per sweep, but since this fraction is repeated $2F_m$ times per second, it constitutes a signal of the same frequency $2F_m$ and is so counted by the frequency meter. In order to make this minimum altitude small, it is necessary that ΔF be large, since they are inversely proportional to each other. A frequency sweep of approximately 25 megacycles is required to provide measurements down to the present minimum of about twenty feet. If a high antenna efficiency is to be obtained over a band 25 megacycles wide, it is necessary that the percentage variation from the average frequency during the modulation cycle be small. This

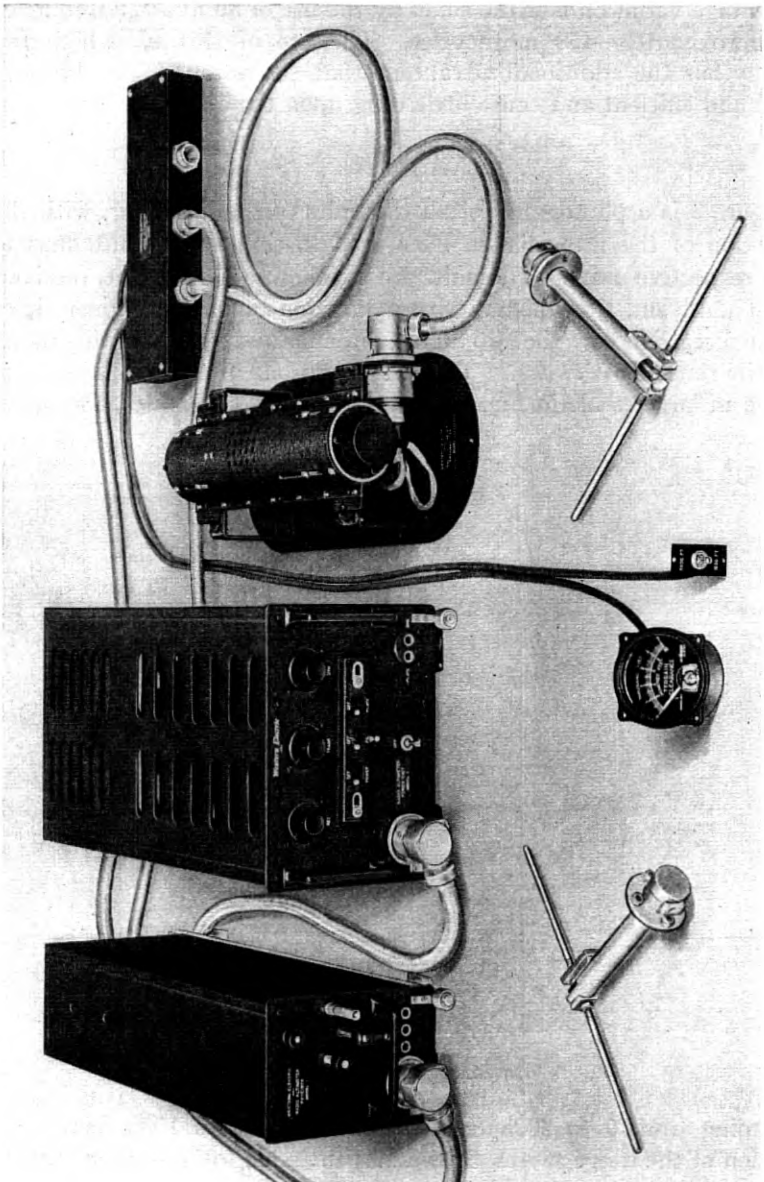


Fig. 3—Terrain clearance indicator units.

percentage variation is made small by the use of an average frequency of approximately 450 megacycles. The use of this ultra-high frequency has the additional advantage that the antennas can be both small and efficient and cause little drag upon the airplane.

APPARATUS

Figure 3 is a photograph of all the units of the altimeter, with the exception of the transmission lines used to connect the antennas to their respective units. The units are as follows: left to right, receiver, power unit, and transmitter, with a junction box in the upper right. In the foreground are the two dipole antennas and the indicating meter with its range-shift switch. The meter and one of these antennas are shown in larger scale in Fig. 4. The meter has two scales, the upper extending from 0 to 5000 feet and the lower 0 to 1000 feet. The position of the range switch determines the scale to be used in reading the meter.

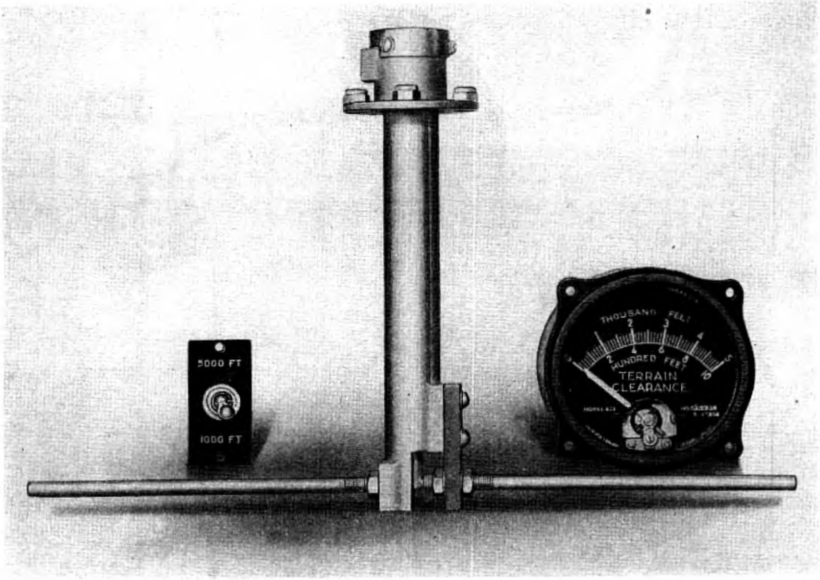


Fig. 4—Antenna, meter, and range switch.

Figure 5 shows an assembly of the various units located approximately as they would be installed in an air transport. The transmitter, power unit, receiver and a junction box are installed in the baggage compartment just aft of the cockpit with cable connections

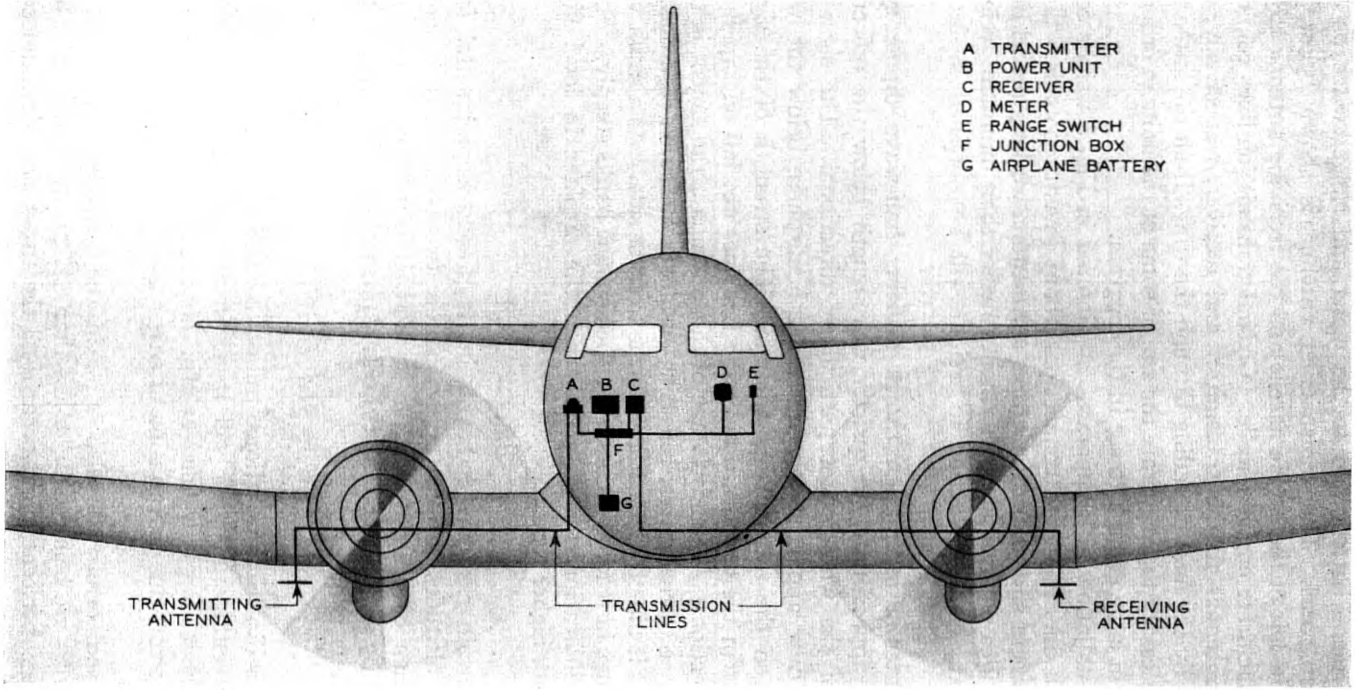


Fig. 5—Airplane installation.

to the airplane battery and to the meter and range switch on the instrument panel. The transmitting antenna is shown below the wing to the left of the engine nacelle and the receiving antenna to the right of the other engine nacelle. Coaxial transmission lines connecting the antennas to the transmitter and receiver, respectively, are indicated by the lines extending through the wings from the antennas. It was necessary to exaggerate the size of some of the units in order to make them large enough to see in the diagram.

The installation with apparatus as pictured in Fig. 3 weighs complete with all cables and connections about seventy pounds. Since the equipment shown in the pictures represents a working model built with the idea of attaining performance rather than minimum weight, undoubtedly some reduction in weight will be obtained in future models.

The antenna installation shown utilizing half-wave dipole type antennas approximately a quarter wave-length below the reflecting surface of the wing is not particularly directional. The signal is radiated over approximately the whole hemisphere below the wing centered on the transmitting antenna. The strength of the signal is greatest in the downward direction but does not fall off rapidly in other directions. The advantage of this antenna arrangement is that the distance to the nearest reflecting surface is measured regardless of whether it is directly beneath, or to the front or side. As a result very little change in reading occurs when the airplane banks steeply. Some advance indication also is given when the airplane in level flight approaches higher terrain.

PERFORMANCE

The terrain clearance indicator in its present experimental form indicates altitudes between approximately twenty and five thousand feet. When over smooth water or land, it is subject to errors as indicated by a consideration of the fundamental equation upon which the altimeter is based,

$$F_d = 4\Delta F\, F_m\, H/C.$$

Since F_d is directly proportional to both ΔF and F_m, any variation of a given percentage in either will result in a corresponding percentage error in the reading of the meter. It is believed from the data available that the errors due to variation of either ΔF or F_m do not exceed ±1 per cent.

Additional errors can also occur in the frequency meter circuit. These errors are believed to be less than ±7 per cent, so that a total error of ±9 per cent might occur if all the errors were simultaneously in the same direction. Fortunately, all these are of a percentage nature, so that the error in feet becomes smaller as the ground is approached. An absolute error in the indication is still possible because of the limitations of the milliammeter used on the instrument panel. The Weston aircraft meter used is guaranteed to be correct to within one per cent of its full scale reading at any point on its scale, which permits maximum errors of ten feet on the 1000-foot scale and fifty feet on the 5000-foot scale.

When flying over rough water, wooded terrain or cities, reflected signal is received from surfaces at different distances simultaneously, resulting in addition and subtraction interference effects, thus sometimes momentarily reducing the echo signal below the minimum required for accurate indication. In such a case, the meter hand may swing down momentarily as much as 10 per cent. For the present limited transmitter power and receiver sensitivity, at altitudes above 2500 feet, these momentary signal reduction effects become progressively more serious when flying over irregular surfaces so that for a substantial part of the total time the echo signal may be below the minimum required for correct meter reading. This is indicated by a reading fluctuating between 3000 and 5000 feet when flying at 5000 feet over a surface dotted by buildings, timber, etc. The meter swings up to the correct reading every time the airplane passes over a smooth field or body of water of any size. Up to 2500 feet the echo signal has proved to be sufficient for steady operation over all kinds of terrain.

Tests have been made over New York, Raritan, Newark and San Francisco Bays, Great Salt Lake, Lakes Erie and Michigan, the timbered mountains of Washington and Oregon, the deserts and mountains of the southwest and the cultivated areas of the midwest during the period of the recent demonstration flights made with the equipment installed in the United Air Lines Flight Research Airplane.

An indication of the character of the surface over which the airplane is flying is given by the variations in the meter reading. A city usually causes rapid fluctuations of the order of fifty feet, depending, of course, upon the height and the spacing of the buildings. Cultivated farmland causes fluctuations of lower frequency and amplitude. An isolated high object such as a skyscraper or a chimney is indicated only by a slight meter kick as the airplane passes over it, which may not be noticed by the observer. If the airplane passes over only a few feet above the object and the top is large enough to contribute momentarily most of the echo signal received by the airplane, the indication is unmistakable and the correct distance to the object is indicated by the meter. For instance, the gas storage tank near the Chicago airport is an excellent object upon which to demonstrate the altimeter performance. The instrument is useful as a position indicator when approaching an airport on a course which crosses an obstruction of appreciable height and size since the moment of passage over the obstruction is clearly indicated. In fact, use as a position indicator may be one of the altimeter's most valuable applications.

A study of the circumstances in connection with a number of crashes in the west during recent years has revealed that in most of the cases the airplanes crashed after having been within a few feet of the ground without the pilot knowing it for several minutes before they struck. In such a situation the terrain clearance indicator should be capable of warning the pilot in ample time to avert a crash.

The writers wish to express their appreciation of the contributions of a number of other members of the technical staff of the Bell Telephone Laboratories to the success of this project.

REFERENCES

1. H. Loewy, *U. S. Patent* 1,492,300.
   H. Loewy, *U. S. Patent* 1,585,591.
   C. F. Jenkins, *U. S. Patent* 1,756,462.
   E. F. W. Alexanderson, *U. S. Patent* 1,969,537.
   F. H. Drake, *U. S. Patent* 1,987,587.
2. *U. S. Patent* 2,011,392 issued August, 1935 to J. O. Bentley.
3. Page 29 of "Solving the Problem of Fog Flying," a publication of the *Daniel Guggenheim Fund for the Promotion of Aeronautics*, 1929.
4. "Altitude Measurements by Reflected Electro-Magnetic Waves" by Murray Hively and R. C. Newhouse, Ohio State University Library, 1929.
5. "An Electro-Magnetic Altimeter" by R. C. Newhouse and J. D. Corley, Ohio State University Library, 1930.
6. *U. S. Patent* 1,517,549 issued December, 1924 to Lloyd Espenschied.
7. *U. S. Patents* 2,045,071 and 2,045,072 issued June, 1936 to Lloyd Espenschied.
8. "A Negative Grid Triode Oscillator and Amplifier for Ultra-High Frequencies," A. L. Samuel, *Proceedings of Institute of Radio Engineers, 25*, Oct. 1937 (1243).
9. "A Research of Direct Reading Altimeter for Aeronautical Use by Radio Reflection Method," Sadahiro Matsuo, *Journal I. E. E.* Japan, No. 571, February, 1936.
10. "A Direct-Reading Radio-Wave-Reflection Type Absolute Altimeter for Aeronautics," Sadahiro Matsuo, *Proceedings of Institute of Radio Engineers*, Vol. 26, 1938 (848).
11. *Trans. A. I. E. E.* 49, 1930 (1331).
    Proc. I. R. E. 19, 1931 (659).
    Review of Scientific Instruments 6, 2, January, 1935.
    Journal of Scientific Instruments 14, 1937 (136).

Transcontinental Telephone Lines *

By J. J. PILLIOD

Late in 1937 a large construction project was completed which added 16 telephone circuits to the transcontinental layout, and the work was so planned that 48 additional circuits can be obtained by the addition of equipment but without stringing additional wire. A brief description of some features of this project and the general development of the transcontinental telephone routes since the first one was opened for service in 1915 is given in this article. Although most of the discussion relates to transcontinental lines, the methods described are generally applicable to other similar situations.

LESS than twenty-five years ago, it was impossible to talk by telephone from coast to coast across the United States. Furthermore, it was impossible to talk between points separated by any such distance anywhere in the world. By 1915, technological advancement had reached a point such that telephone service could be established across the country, and three telephone circuits had been built which connected San Francisco and the Pacific Coast with points in the East. Four telegraph circuits were also provided by the new wires. An improved loading system and especially the successful development of the vacuum-tube telephone repeater were outstanding factors which made telephone connections of such length possible for the first time in history.

Open-wire lines played the major role in the early transcontinental telephone circuits. The transmission losses caused by cable were so great that it was avoided whereever possible. The steady improvement of telephone repeaters, types of loading for use on cable circuits, and carrier telephone systems for use on open-wire lines made it possible to provide rapidly and economically more telephone circuits across the continent as use of the service grew. In the cross section

---

* This paper has been prepared from an address given before the Communications Group of the A. I. E. E., New York Section, March 22, 1938, and published in *Electrical Engineering* for October, 1938. Since the paper was written, three type J 12-channel carrier systems have been placed in service on the new line. Two of these systems operate between Oklahoma City and Whitewater, 1200 miles, and the third between Oklahoma City and Albuquerque, N. M. Twelve additional intermediate repeater stations have been constructed. Three of these are located at such remote distances from primary power that experiments are being made in generating by means of wind-mill power plants part of the power required. One such station is shown in Fig. 8. *Editor.*

just west of Denver there are today one hundred and forty through telephone circuits and about the same number of telegraph circuits carried by four open-wire routes, the last of which was completed during 1937. While open wire was used almost exclusively as a matter of necessity in the first transcontinental telephone lines, cable is now used for about half of the circuit mileage. This is a striking illustration of the large-scale changes which have taken place in the interest of more reliable toll telephone service.

Continued Importance of Open-Wire Lines

The open-wire line seems destined to continue to play an important part in long-distance telephone communication, particularly where distances are great and circuit requirements on any one route are relatively small. Improvements in the usage to which the wires may be put have made this increasingly so. The three circuits on the first transcontinental line were operated at voice frequencies and were obtained from two pairs of line conductors, the third circuit being derived by means of phantom circuit arrangement of these two pairs. The development of carrier telephone systems made it possible to obobtain three additional circuits on some pairs of wires, using frequencies above those required for existing voice-frequency circuits. Carrier telephone systems were first installed on a transcontinental route in 1926 and were quickly followed by others, so that today ninety-six of the one hundred and forty circuits mentioned earlier are obtained by means of these three-channel carrier telephone systems. Development work, however, has been continued, and it is now expected that it will be possible, by means of carrier telephone systems using still higher frequencies, to obtain as many as twelve more telephone circuits on some pairs of wires. It has been with a view toward using such systems and obtaining a total of sixteen telephone circuits on a pair of wires that the latest of the four transcontinental routes has been designed.

Construction of New Transcontinental Line

Early in 1937, it became clear from a study of loads carried on existing transcontinental routes that additional circuits would be required in the near future. Circuits in cable were available as far west as Omaha, Kansas City, Oklahoma City, and Dallas. After consideration of all the factors, it was decided to construct the new facilities west from Oklahoma City to Los Angeles on the route shown in Fig. 1. It was also decided to carry out the work in such a way that the route could be utilized for the future addition of a relatively

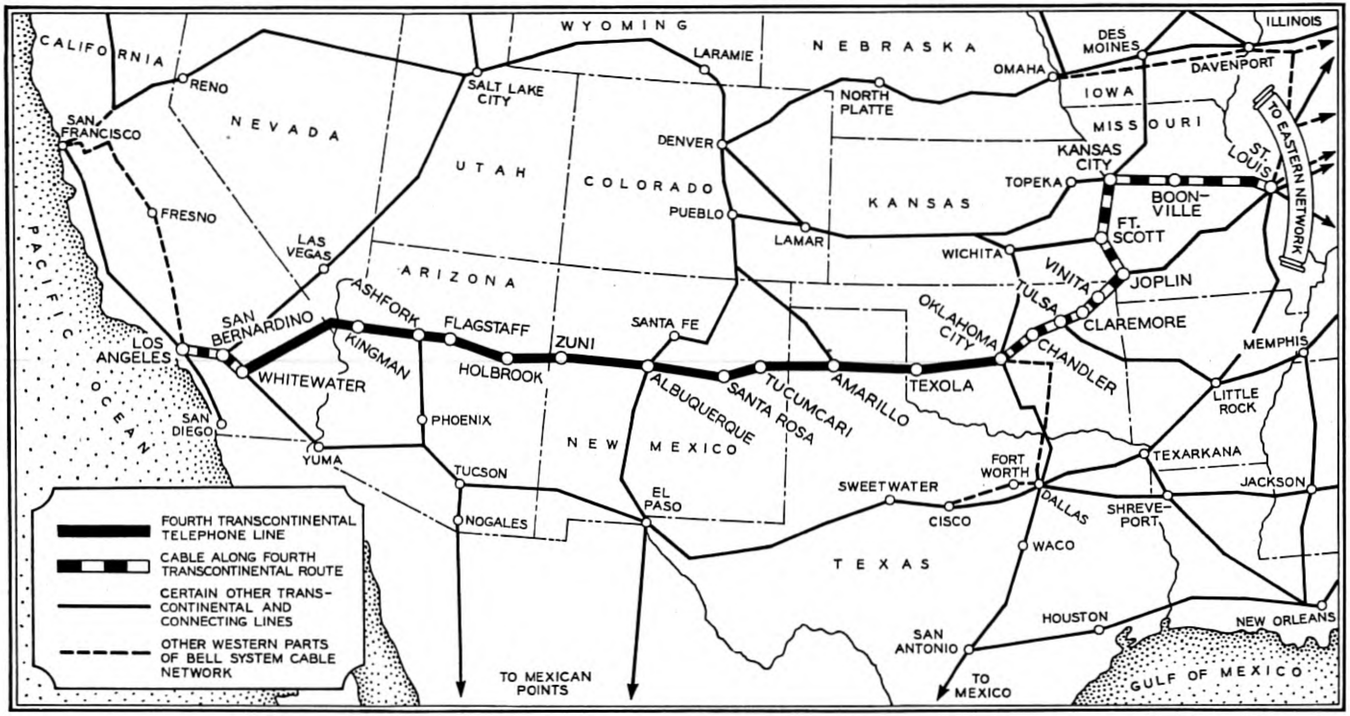


Fig. 1—Route of new transcontinental line across western states.

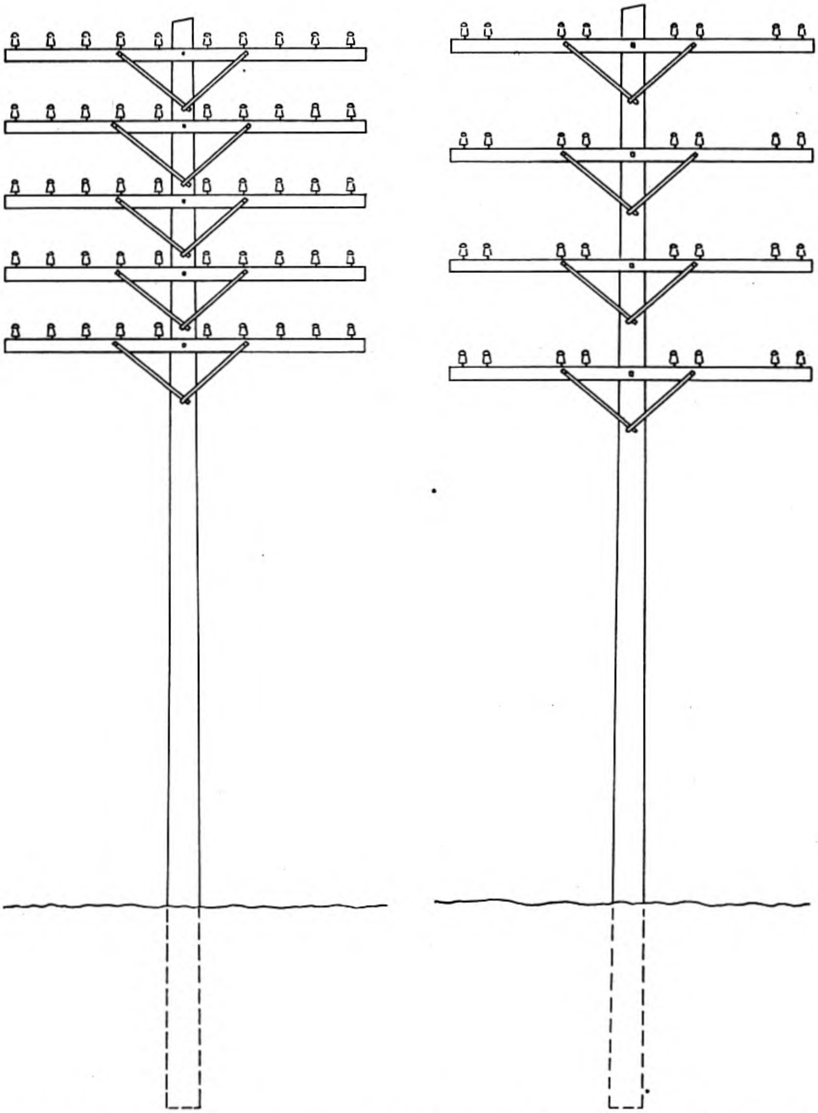


Fig. 2—Cross-arm arrangements. Left—50-wire phantomed line, capacity 77 telephone circuits. Right—32-wire non-phantomed line, capacity 256 telephone circuits.

large number of circuits through the application of the twelve-channel carrier-current telephone system then under development. Among the conditions favoring this particular route is freedom from winter storm hazards throughout most of the distance, which, looking ahead, is particularly important to the future application of twelve-channel carrier telephone systems. The work done in 1937 consisted of building a length of nearly three hundred miles of new pole line and stringing four pairs of wires throughout most of the section from Oklahoma City to Whitewater, California, a distance of 1,200 miles. Initially the voice channel and three-channel carrier telephone systems have been developed on these four pairs, providing a total of sixteen telephone circuits.

WIRE SPACING AND TRANSPOSITIONS

Open-wire telephone lines designed to carry frequencies up to 140 kilocycles per second, as used in the operation of the twelve-channel carrier telephone systems, have structural requirements substantially more stringent than those designed to carry only three-channel systems, which use frequencies up to 28 kilocycles. The usual type of open-wire toll telephone line has ten wires on each crossarm, spaced at about one-foot intervals, five on each side of the pole and with the crossarm spaced twenty-four inches apart. In the case of the line designed to conduct high carrier telephone frequencies, this configuration has been changed and is illustrated by Fig. 2. Eight wires are strung on each arm, grouped as four pairs, two on each side of the pole. The wires of the pair are spaced eight inches apart, and the nearest wires of the two pairs on each side of the pole are spaced twenty-six inches, while the spacing at the pole is thirty inches. Crossarms are spaced thirty-six inches apart.

These new wire spacings reduce the coupling between pairs on the same line or between pairs on this line and pairs on other lines which may parallel it. New transposition systems are used further to reduce this coupling. Transpositions are closer together and a transposition bracket of the type shown in Fig. 3 is used to turn the wires completely over at as nearly a given point as possible. Transpositions in one or more pairs are installed on every pole with an occasional exception, and certain pairs are transposed at every other pole. The wires of a pair must be adjusted to the same sag within close limits. These sag variations are held to a fraction of an inch, and a check of the completed work indicated that fifty per cent of the spans had been adjusted to within one-quarter inch. Telescopes are used to help obtain these close sag adjustments, and a final check is made by oscillating the wires in a span and observing the periods at which they oscillate.

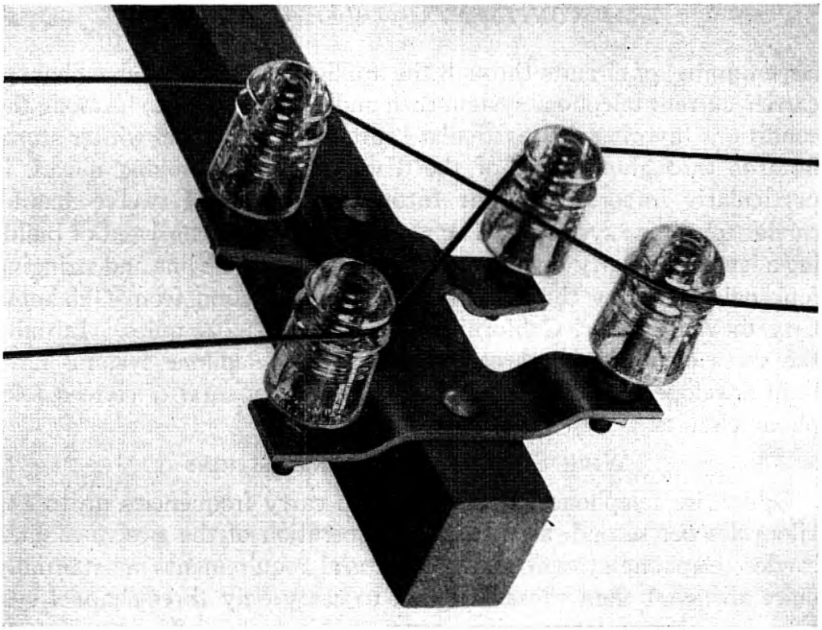


Fig. 3—Transposition bracket shown as installed. Tie wires are not used with this type bracket.

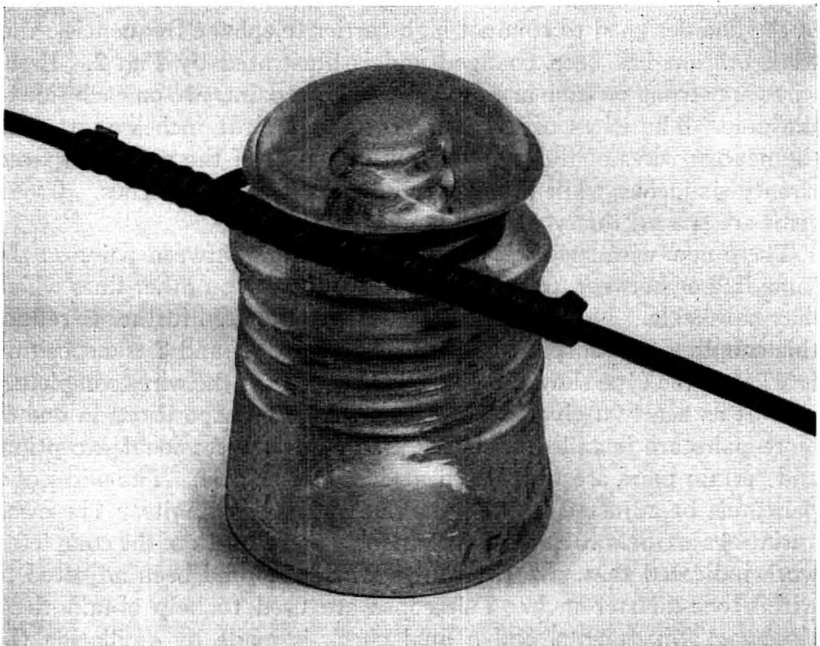


Fig. 4—Flat tie wire shown as installed.

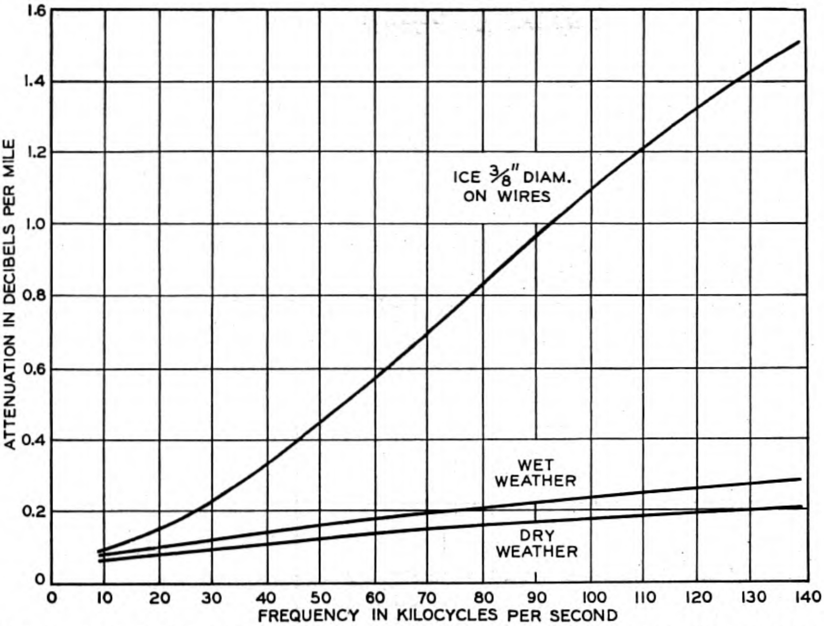


Fig. 5—Attenuation of open-wire pairs, 165-mil copper, 8-inch spacing, CS insulators.

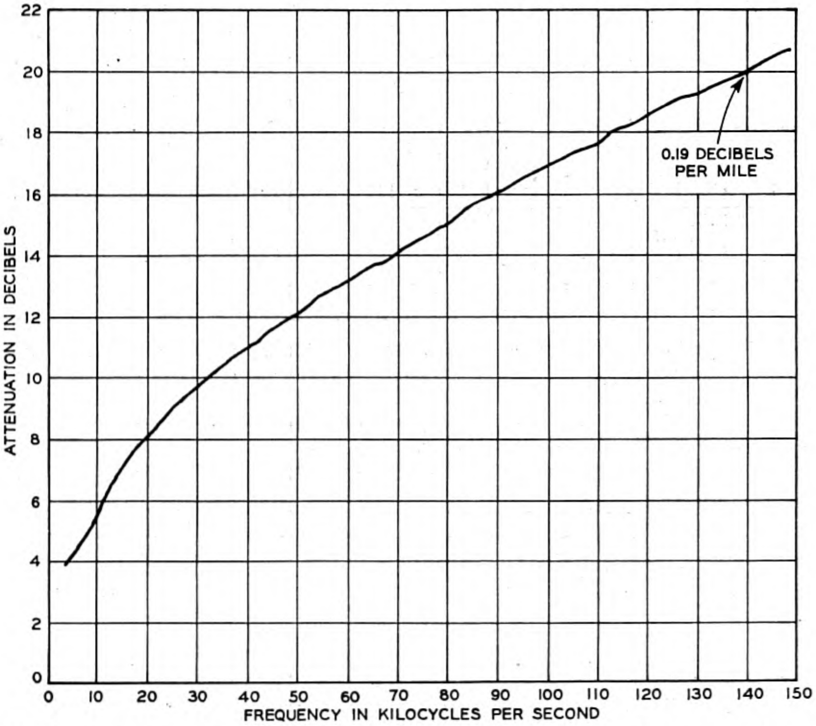


Fig. 6—Attenuation of 165-mil copper, 8-inch spaced pair of wires measured on a 105-mile section of the Amarillo-Albuquerque line in clear weather.

Pole Spacing and Insulators

Poles must be spaced uniformly in order that the transpositions may be most effective, and an occasional deviation of only thirty-five feet is the maximum permitted. Where it is impossible to locate poles within this limit, such as is the case at long-span crossings, special fixtures are suspended from steel cables at the proper points to permit making the transpositions.

New types of insulators on steel pins, each pair of which is electrically bonded, are used to improve the stability of the transmission characteristics.

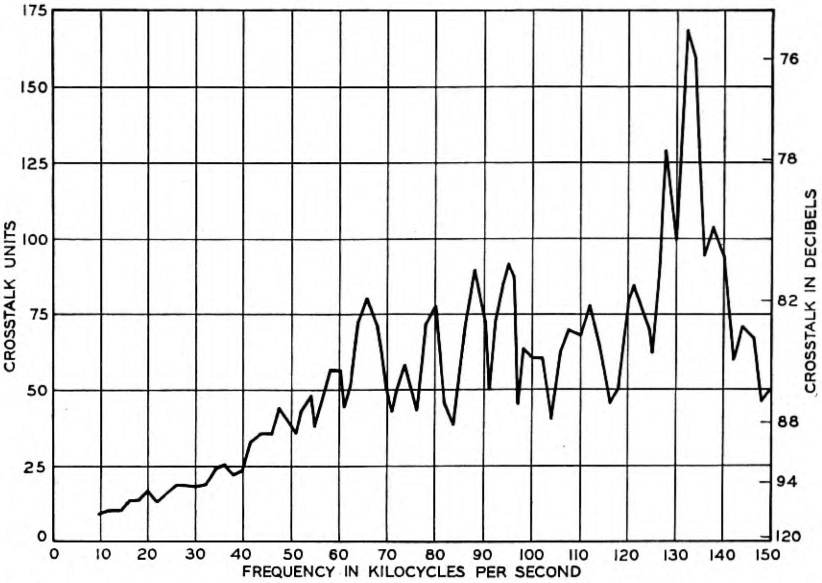


Fig. 7—Far-end crosstalk between wires 7–8 and 9–10 of Amarillo-Albuquerque line, measured from pole 1 to pole 4236, a distance of 105 miles.

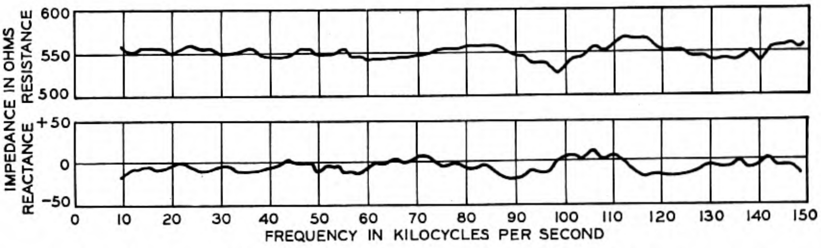


Fig. 8—Impedance of 165-mil copper, 8-inch spaced, CS insulated pair of wires on Amarillo-Albuquerque line, measured from pole 1 to pole 4236.

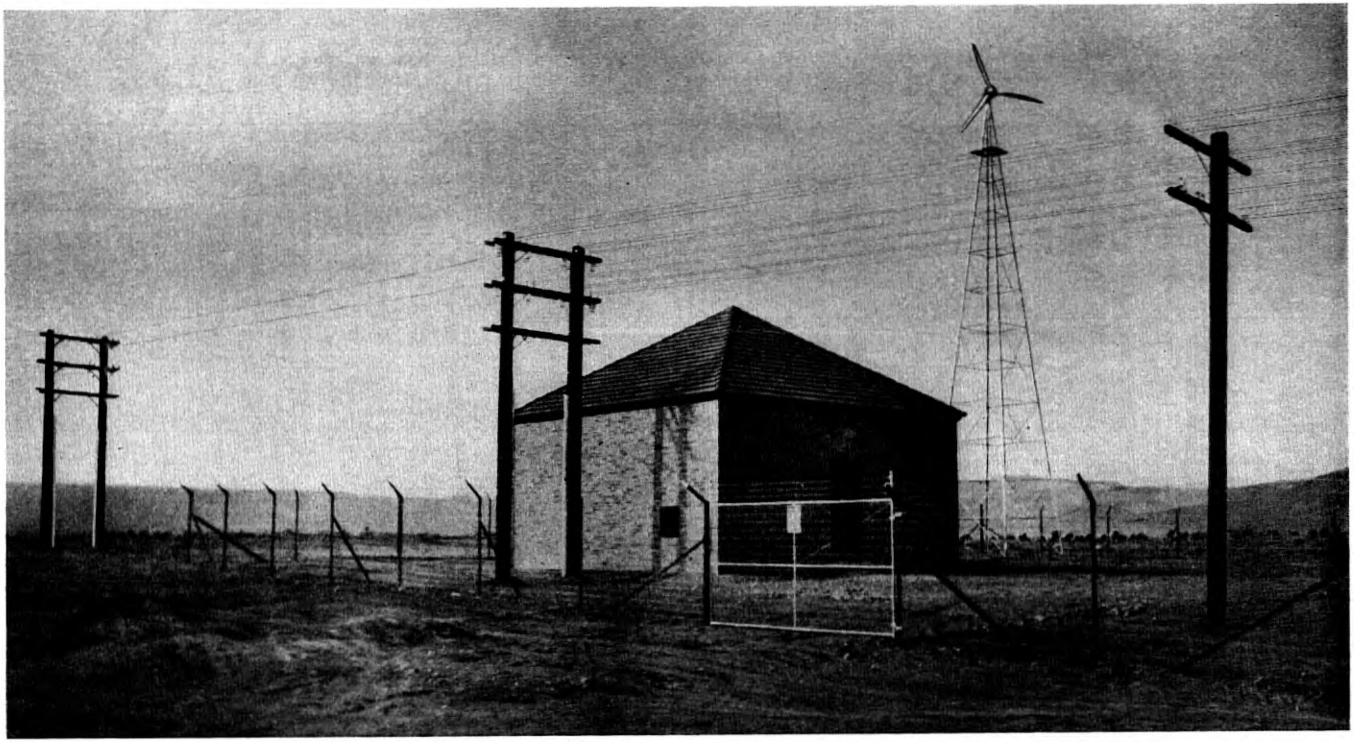


Fig. 9—Type J carrier repeater station at San Fidel, New Mexico (about sixty miles west of Albuquerque), showing wind mill power generating plant. This station is unattended.

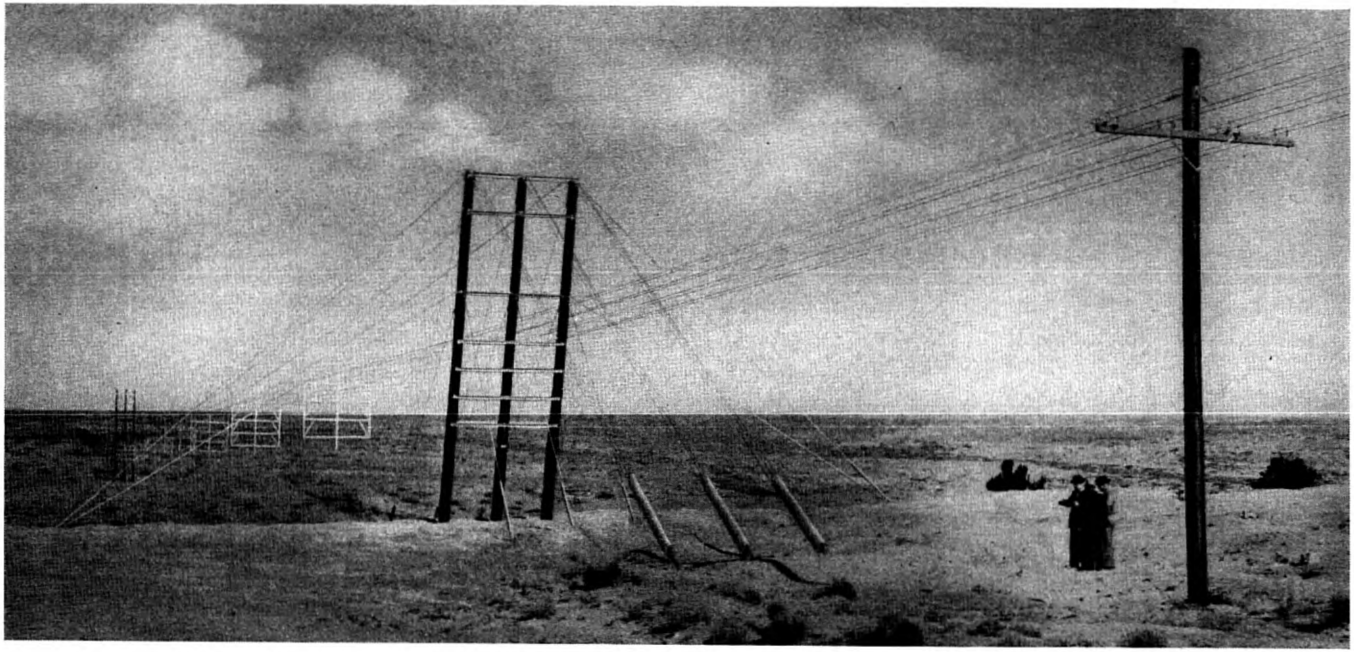


Fig. 10—Long span crossing at Tucumcari Creek, New Mexico. This span is 1080 feet between supporting towers. Normally, the creek carries very little water and the long span is to avoid damage at times of flood.

Wire 165 mils in diameter was used on this construction west of Oklahoma City because of its strength and resultant relative freedom from interruptions. Transmission losses were also a factor in this case. Rolled sleeve joints were used in splicing the wire because of their strength and good electrical characteristics. Flat tie wires, shown in Fig. 4, were used to reduce chafing of the line wire when the wires vibrate. Tie wires are not used at transposition points, as may be seen in Fig. 3.

Repeater Stations

The computed losses on open-wire pairs of this type at high-carrier frequencies and under different weather conditions are shown by Fig. 5. Field tests confirm these data. To offset the line losses it will be necessary to locate repeater stations at intervals of from fifty to one hundred miles, depending upon the weather conditions which may be expected. Between Oklahoma City and Whitewater, California, in order to operate the twelve-channel carrier telephone systems, it will be necessary to equip sixteen intermediate repeater stations. Most of these will be unattended and maintained from other offices.

It is not practicable, of course, to bring the open-wire pairs directly into all repeater stations and in some cases entrance cables several miles long must be used. Although ordinary non-loaded cable pairs may be used for this purpose, their usage involves transmission difficulties, and except where other factors dictate the use of this type of facility, it is planned to use low-loss cable conductors of a new design. These cable pairs have more favorable impedance characteristics as well as lower losses.

With the building and the further equipping of this latest open-wire line across the western states, open-wire facilities have played one more important part in the development of long-distance telephony. Although cable is being found more and more useful, there still remain many important links in the nation-wide telephone communication network where, for the present at least, the open-wire line can serve best and the development of it toward maximum usefulness is still being carried on.

Abstracts of Technical Articles from Bell System Sources

Paper as a Medium for Analytical Reactions.[1] B. L. CLARKE and H. W. HERMANCE. Absorbent paper has long been used in chemical laboratories for filtering suspensions. Paper has also found a special use as a container or holder for certain testing reagents; litmus paper is a common example. In this article and a preceding one (*Indus. & Engg. Chem., Anal. Ed.*, June 15, 1937), are reported exploratory investigations on the extension of the use of absorbent papers in chemical analysis.

Rapid identification "spot tests" have been described by Feigl in which, by successively placing drops of unknown and reagent solutions on filter paper, characteristic color changes are produced. Methods and apparatus are described in the present articles whereby some of the variables in such tests are controlled. Chief among these innovations is the use of semi-soluble instead of soluble reagents, and the precipitation of these compounds directly on the paper fibres to form a more or less permanent test paper. By these changes in technique the sensitivity—the smallest amount of a given metal detectable—is decreased from ten to one-hundredfold. For example, 0.002 microgram of copper can be detected by the new method, as compared with 0.2 by the old.

In another application a very dilute solution of some metal ion is slowly siphoned through a small circular piece of reagent paper suitably mounted. The metal is entrapped on the paper in an insoluble form strongly adsorbed by the paper. Theoretical analysis indicates that copper, for example, may be removed from a solution in this way so completely that only 8×10^{-12} microgram will be left in a liter.

Neutral Particles in Physics.[2] KARL K. DARROW. During the early days of science, the elementary particles which scientists and philosophers alike saw fit to postulate were always imagined as chargeless. With the remarkable growth of the understanding of electricity during the nineteenth century, and with the invention of instruments for detecting small charged particles during the twentieth, it became the custom to suppose that the fundamental particles of

---

[1] *Indus. & Engg. Chem., Anal. Ed.*, October 15, 1938.

[2] *Amer. Phil. Soc. Proc.*, September 30, 1938.

matter all bear charges and that the forces exhibited in Nature are all electrical (exception being made for gravitation). A noted and serious objection to this view was temporarily met by the adoption of quantum mechanics. Since 1930 a reversal of trend has set in, heralded by the discovery of the neutron as a subatomic chargeless particle capable of independent existence; and at present there is a strong tendency to develop the view that neutral as well as charged particles of subatomic size, and non-electrical as well as electrical forces, exist together in Nature.

Electrical Networks for Sound Recording.[3] F. L. HOPPER. Electrical networks are employed in sound recording for modifying and limiting the frequency-response characteristic. The necessity for their use, application, and design is described. Particular emphasis is placed upon the constant-resistance type of structure.

Sound Pictures in Auditory Perspective.[4] FRANKLIN L. HUNT. Soon after sound reproduction in auditory perspective was demonstrated over telephone circuits between Philadelphia and Washington in 1933, experimental sound pictures in auditory perspective were made at the Bell Telephone Laboratories' sound picture laboratory. Listening tests showed that they distinctly enhanced the illusion that the sound originated at its apparent source on the screen and they strikingly improved the feeling of spaciousness and reality. The auditory perspective effect is not primarily dependent upon perfect synchronism of the two sound-tracks required, nor on frequencies above the present commercial range. Existing equipment can be converted to project sound pictures in auditory perspective without great difficulty.

Composition and Colloidal Properties of Balata Latex.[5] A. R. KEMP. This paper reports the composition and colloidal properties of two types of balata latex from Dutch Guiana. The white variety is shown to be superior to the red, owing to its higher content of hydrocarbon.

It is shown that balata latex is very stable owing to the presence of a highly protective water-soluble substance in its serum. It cannot be coagulated by acids or salts, but is readily coagulated by alcohol or acetone.

The balata latex particles are spherical and vary in diameter from about 0.1 to 2.5 microns with an average diameter of about 0.5 micron.

The balata latex particles are shown to enclose the resins, which

---

[3] *Jour. S. M. P. E.*, November 1938.

[4] *Jour. S. M. P. E.*, October 1938.

[5] *India Rubber World*, December 1, 1938.

appear to be present in a dispersed state in the hydrocarbon. The particles are shown to contain about 18% of water, determined as water of retention in pressed coagulum.

The "resins" have been separated from both types of balata latex as water-white viscous liquids which deposit crystals of β-amyrin acetate on standing. The red balata latex resin is shown to be more viscous than the white and to differ from it as regards its iodine value, refractive index, and solubility in cold 95% ethyl alcohol.

The serum constituents have been separated into four main fractions: protein, carbohydrate, gummy substance, and ash. Minor constituents such as tannin and amino-acid have also been noted.

A complete analysis of balata ash has been made and compared with the analysis of ash from *Hevea* latex by Bruce. Balata ash was found to contain higher contents of CaO, Na_2O, and MgO and lower contents of K_2O and P_2O_5 than *Hevea* latex ash.

New data are presented on the density, refractive index, dielectric constant, and heat of combustion of balata hydrocarbon which are believed to be more reliable than similar data previously available in the literature. Data on the effect of temperature on the refractive index of balata and gutta percha hydrocarbon are presented, showing the crystallization of the gutta hydrocarbon on cooling, which starts at about 37° C. resulting in an abrupt increase in refractive index occurring between 37° and 35° C.

A Short-Wave Single-Sideband Radio Telephone System.[6] A. A. OSWALD. There is described briefly a short-wave single-sideband system which has been developed for transoceanic radio telephone service. The system involves the transmission of a reduced carrier or pilot frequency and is designed to include the testing of twin-channel operation wherein a second channel is obtained by utilizing the other sideband.

The paper indicates the reasons which led to the selection of this particular system and discusses at some length those matters which require agreement between the transmitting and receiving stations when single-sideband transmission is employed.

The Oxide-coated Filament. The Relation between Thermionic Emission and the Content of Free Alkaline-earth Metal.[7] C. H. PRESCOTT, JR. and JAMES MORRISON. The oxide-coated filament had its beginning in the sealing-wax era of vacuum technique. The obscure accident of its origin is not recorded, but all of our older physicists knew

---

[6] *Proc. I. R. E.*, December 1938.

[7] *Jour. Amer. Chem. Soc.*, December 1938.

that an enhanced emission of electrons could be obtained by smearing sealing-wax on a platinum ribbon and burning it off in air. The first authentic study is recorded by Wehnelt, who investigated the voltage-drop in a gas discharge tube with cathodes coated with various metallic oxides. Its further evolution and development to the status of a cathode in Western Electric vacuum tubes has been described by H. D. Arnold. A comprehensive treatment of its history, the various modifications in current use, and divergent theories of its preparation and behavior has been given by Saul Dushman in a treatise on "Thermionic Emission." A later review is given by J. H. deBoer.

The present work is devoted to a quantitative determination of the relation between thermionic emission and the content of free alkaline earth metal. To this end we have employed a filament which is a platinum rhodium core coated with barium, strontium, and nickel carbonates. On heating in a reducing atmosphere this coating becomes a grossly homogeneous colloidal mixture of barium oxide, strontium oxide, and free nickel. After a thorough preliminary clean-up of the experimental tube, the requisite amounts of free alkaline-earth metal are generated by reaction with methane. The electrical measurements are summarized by the use of the Richardson equation for thermionic emission. Free alkaline earth metal has been determined by oxidation with carbon dioxide and analysis of the gaseous reaction products.

Using a filament coated with a colloidal mixture of barium oxide, strontium oxide, finely divided nickel, and free alkaline earth metal, we have investigated the quantitative relation between thermionic emission and the content of active metal. A high level of activity was found from 15 μg./sq.cm. to 60 μg./sq.cm. of equivalent Ba, with a slight apparent maximum at 30 μg./sq.cm. where the thermionic current at 1050° K. is 600 m. a./sq.cm. The electron work function is 1.37 v.

The radiant emissive power at 0.66 μ is approximately 64%, independent of the content of active metal.

The free alkaline earth metal was determined by oxidation with carbon dioxide and analysis of the gaseous reaction products.

A Single-Sideband Receiver for Short-Wave Telephone Service.[8] A. A. Roetken. A new radio telephone receiver has been developed for the reception of reduced-carrier single-sideband signals in the frequency range from 4 to 22 megacycles. This receiver employs triple detection in which the first beating oscillator is continuously

---

[8] *Proc. I. R. E.*, December 1938.

variable and the second is fixed in frequency. The first oscillator is a very stable tuned-circuit type, the proper adjustment of which is maintained through the use of an improved type of synchronizing automatic-tuning-control system. The second oscillator is crystal controlled. Separation of the carrier and sideband is accomplished in the receiver by means of band-pass crystal filters which provide extremely high selectivity. Unusually high stability and selectivity characterize the performance of the receiver.

Dielectric Constant and Dielectric Loss of Plastics as Related to their Composition.[9] W. A. YAGER. Data are presented for the frequency variation of the dielectric constant and dielectric loss factor of various plastics over a broad frequency band extending from 1 kc. to 35 mc. The extremely low loss of polystyrene compared to that of polar plastics confirms the theory that a hydrocarbon is inherently more satisfactory from a dielectric point of view. Of the several possible mechanisms of dielectric loss which might account for the high-frequency dielectric absorption observed in polar plastics, the rotation of polar units in the chain and of polar side groups appears most probable. The fact that the loss factor maxima of phenol fibers, phenol fabrics, and phenol or urea formaldehyde molding compounds containing cellulosic fillers occur at essentially the same frequency is viewed as evidence that this dielectric absorption is an intrinsic property of cellulose. Substitution of mineral fillers for cellulose reduces the high-frequency loss to that residing in the polar resin binder. Furthermore, the dielectric loss of mineral-filled molding compounds is less moisture-sensitive. The large increase in dielectric loss at low frequency always found in materials of relatively high free-ion conductivity manifests itself in Duprene, and the humidified phenolic plastics containing cellulose fillers or laminations.

[9] *Electrochemical Society Preprint* No. 74-24, October 12–15, 1938.

Contributors to this Issue

H. A. Affel, S.B. in Electrical Engineering, Massachusetts Institute of Technology, 1914; Research Assistant in Electrical Engineering, 1914–16. American Telephone and Telegraph Company, Engineering Department and the Department of Development and Research, 1916–34; Bell Telephone Laboratories, 1934–. As Assistant Director of Transmission Development, Mr. Affel is concerned with toll transmission problems, including the development of carrier telephone systems.

J. W. Beyer, B.S., State College of Washington, 1915; Westinghouse Electric and Manufacturing Company, 1915–16. Instructor in Electrical Engineering, State College of Washington, 1916–19. Western Electric Company, Engineering Department, 1919–24. Bell Telephone Laboratories, 1924–. Mr. Beyer is concerned with the development of carrier telephone terminals and pilot channels for open-wire circuits.

E. C. Blessing, B.S. in Electrical Engineering, Purdue University, 1922. Western Electric Company, Engineering Department, 1922–25; Bell Telephone Laboratories, 1925–. Mr. Blessing has been engaged in the development of carrier systems.

S. I. Cory, B.E.E., Ohio State University, 1916. American Telephone and Telegraph Company, Engineering Department and Department of Development and Research, 1916–1934; Bell Telephone Laboratories, 1934–. During this entire time Mr. Cory has been engaged in transmission development work, chiefly on telegraph systems and transmission-measuring methods.

Frank A. Cowan, B.S. in Electrical Engineering, Georgia School of Technology, 1919. American Telephone and Telegraph Company at Atlanta, Ga., 1920; Long Lines Engineering Department, Special Service Group, New York, 1922; appointed Division Transmission Engineer, Division No. 1, New York, 1926; appointed to present position, Engineer of Transmission, Long Lines Department, 1928.

Karl K. Darrow, B.S., University of Chicago, 1911; University of Paris, 1911–12; University of Berlin, 1912; Ph.D., University of

Chicago, 1917. Western Electric Company, 1917–25; Bell Telephone Laboratories, 1925–. Dr. Darrow has been engaged largely in writing on various fields of physics and the allied sciences.

Lloyd Espenschied, Pratt Institute, 1909, has taken an important part in much of the Bell System's invention and development in the field of radio and carrier currents; also in technical contacts and international conferences abroad. Previously, in the Department of Development and Research of the American Telephone and Telegraph Company; now, as Research Consultant of the Bell Telephone Laboratories, participates in a broad front of electric-communications development.

Frank Gray, B.S., Purdue, 1911; Ph.D., University of Wisconsin, 1916. Western Electric Company, Engineering Department, 1919–25. Bell Telephone Laboratories, 1925–. Dr. Gray has been engaged in work on electro-optical systems.

William H. Harrison, Vice President and Chief Engineer, American Telephone and Telegraph Company. Pratt Institute, graduate in Industrial Electrical Engineering, 1915. New York Telephone Company, repairman, apparatus inspector, 1909–14. Western Electric Company, engineer, 1914–18. American Telephone and Telegraph Company, 1918–33, Department of Operation and Engineering, engineer; Equipment and Building Engineer; Plant Engineer. Bell Telephone Company of Pennsylvania, Diamond State Telephone Company, Vice President in charge of Operations, 1933–37. American Telephone and Telegraph Company, Department of Operation and Engineering, Assistant Vice President, 1937–38; Vice President and Chief Engineer, 1938–.

L. W. Hussey, A.B., Dartmouth, 1923; M.A., Harvard, 1924; B.S. in E.E., Union College, 1930; Mathematics Department, Union College, 1924–29. Bell Telephone Laboratories, 1930–. Mr. Hussey has been engaged principally in work on the stability of regenerative systems and on modulation in non-linear resistances.

B. W. Kendall, S.B., Massachusetts Institute of Technology, 1906; Instructor in Physics at Massachusetts Institute of Technology, Barnard College, and Columbia University, 1906–13. Engineering Department of the Western Electric Company, 1913; Bell Telephone Laboratories, 1925–. As Toll Development Director Mr. Kendall has charge of the development of carrier, voice frequency, and telegraph circuits. His early work was on repeaters in connection with the transcontinental line; he has also been connected with carrier-current development since its inception.

R. C. **Newhouse**, B.E.E., Ohio State University, 1929; Guggenheim Fellow, Ohio State University, 1929–30; M.Sc., Ohio State University, 1930. Bell Telephone Laboratories, 1930–. During most of this period Mr. Newhouse has been engaged in the design and development of aircraft radio transmitters. In recent months his efforts have been confined to the development of the terrain clearance indicator, for which he has been given the 1938 Lawrence Sperry Award by the Institute of the Aeronautical Sciences.

J. T. **O'Leary**, B.S. in Electrical Engineering, Villanova College, 1918. American Telephone and Telegraph Company, Department of Development and Research, 1919–34. Bell Telephone Laboratories, 1934–. Mr. O'Leary has been concerned with the transmission aspects of carrier systems.

E. **Peterson**, Cornell University, 1911–14; Brooklyn Polytechnic, E.E. 1917; Columbia University, A.M. 1923; Ph.D. 1926. Electrical Testing Laboratories, 1915–17; Signal Corps, U. S. Army, 1917–19. Western Electric Company, Engineering Department, 1919–25; Bell Telephone Laboratories, 1925–. Lecturer in Electrical Engineering, Columbia, 1934–. As circuit research engineer, Dr. Peterson's work has been largely in theoretical studies of non-linear circuits and circuit elements.

J. J. **Pilliod**, E.E., Ohio Northern University, 1908. American Telephone and Telegraph Company, Long Lines Department, 1908–11; General Engineering Department, 1912–13; Long Lines Department, Division Plant Engineer, 1914–17; Engineer of Transmission, 1918–19; Engineer, 1920–. Mr. Pilliod is the head of the Long Lines Engineering Department.

J. N. **Reynolds**, B.S. in Electrical Engineering, Purdue University, 1904; E.E. 1907. Western Electric Company, Engineering Department, 1904–25. Bell Telephone Laboratories, 1925–. Mr. Reynolds has been continuously associated with the development of machine switching apparatus. As Special Studies Engineer, he is now engaged in the development of improved forms of crossbar switch and allied apparatus.

Frederick J. Scudder, New York Telephone Company, 1905–10; Western Electric Company, Engineering Department, 1910–25; Bell Telephone Laboratories, 1925–. Mr. Scudder has been engaged in the development of machine switching systems since 1910, and in his present capacity as Systems Development Engineer is in charge of fundamental studies and circuit development of panel and crossbar systems.

R. B. Shanck, B.E.E., Ohio State University, 1915. Railroad telegraph service, 1909–10. American Telephone and Telegraph Company, Plant Department, 1910–11 and summers 1912–14; Engineering Department and Department of Development and Research, 1919–34; Bell Telephone Laboratories, 1934–. As Telegraph Transmission Engineer, Mr. Shanck is engaged in development work on the transmission features of telegraph systems and the measurement of telegraph transmission.