The Bell System Technical Journal

July, 1927

Measurement of Inductance by the Shielded Owen Bridge

By J. G. FERGUSON

SYNOPSIS: The study described in this paper shows that the Owen bridge is well adapted to the accurate measurement of inductance and effective resistance to above 3,000 cycles. The construction of a shielded bridge for audio frequencies is described and a theoretical discussion is also given. It was found possible to measure inductances ranging from 0.1 to 3 henrys with an error of measurement less than 0.1 per cent, and for 10 henrys the accuracy is better than 0.25 per cent. As a means of measuring effective resistance the bridge shows an accuracy of about 2 per cent. The sources of error and method of eliminating or correcting them are discussed.

INTRODUCTION

THE accurate measurement of inductance and capacitance is essential to the correct design of practically all precision electrical apparatus. Particularly is this so in the field of electrical communication where the successful introduction of new circuits and equipment, such as the carrier telephone and the telephone repeater, depends largely on the accuracy with which the elements can be adjusted to the nominal values, this accuracy in turn depending on the accuracy with which the electrical measurements can be made.

Owing principally to the ease with which a telephone receiver may be used to indicate a balance at audio frequencies, bridge measurements are very generally used for the measurement of capacitance and inductance in telephone work. The simplest type of bridge and the one used most for the comparison of like impedances is the equal ratio arm bridge described by Shackelton.[1] This bridge requires standards of the same kind and magnitude as the impedances which are to be measured. The calibration of these standards is a separate problem, for which a distinct type of bridge is required.

Either capacitance or inductance may be measured by a bridge method in terms of time and resistance, both of which are fundamental quantities. However, since condensers may be obtained with very low losses and small changes with frequency, this type of measurement is usually made with capacitance,[2] inductance measurements being

[1] W. J. Shackelton, "A Shielded Bridge for Inductive Impedance Measurements," *Bell System Technical Journal*, January, 1927.
[2] J. Clerk Maxwell, *Electricity and Magnetism*, Vol. 2, pp. 776–7.

made by comparison with capacitance and resistance or with capacitance and frequency. The resonant method is adapted to the comparison of inductance with capacitance and frequency. However, this method demands an accurate measurement of the frequency used, which is not always convenient. It is therefore evident that a bridge which furnishes a comparison of inductance with capacitance and resistance serves a very useful purpose in the calibration of standards of inductance for use in simple comparison bridges.

A bridge circuit due to Owen [3] furnishes a very good example of this type, the balance conditions being independent of frequency and the equations of balance giving a relation between inductance, capacitance, and resistance. The circuit is shown in Fig. 1. It consists of a fixed resistance r_1 in the arm BC, a fixed capacitance C_3 in the arm AB, a fixed capacitance C_4 in series with a variable resistance R in

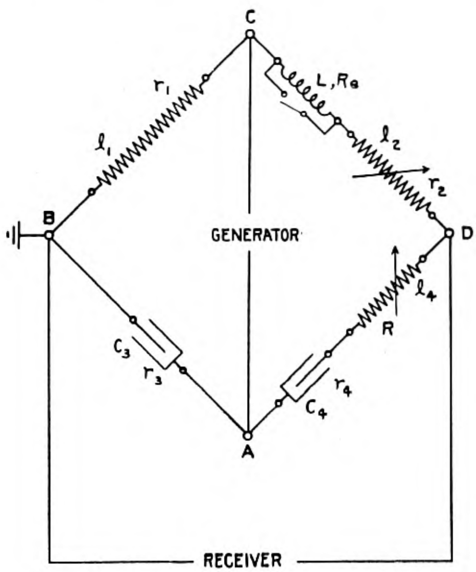


Fig. 1

AD, and a variable resistance r_2 in series with the inductance to be measured in CD. The adjustments for balance are made with R and r_2. These two adjustments are independent of each other. The relations between the quantities at balance, as will be shown later, are such that the bridge may readily be made direct reading for

[3] D. Owen, "A Bridge for the Measurement of Self Induction," *Proceedings of Physical Society of London*, Oct. 1, 1914.

inductance, and these advantages make this bridge superior to practically all other bridges for this type of comparison.

This paper contains a discussion of the theoretical relations of this bridge circuit, its possibilities and limitations for the accurate measurement of inductance and effective resistance, and the sources of error and methods of eliminating them. A shielded bridge, constructed for use in calibrating inductance standards, is described and sufficient measurements are given to show the accuracy of which it is capable.

The maximum frequency at which measurements were given by Owen is 530 cycles. For the measurement of telephone apparatus considerably higher frequencies are used, and it is desirable that the bridge be capable of measurements up to 3,000 cycles without loss of accuracy. It is in the upper part of this range that the greatest difficulties are encountered, requiring special precautions not so necessary for the lower frequency measurements.

While in the following discussion the maximum frequency considered is 3,000 cycles, this is not meant to indicate a maximum limit to this type of bridge.

Equations of Balance

Taking into consideration the phase angle of the resistances and the loss in the condensers, the complete network is shown in Fig. 1, the reactive component of the resistances being shown as series inductance, and the condenser losses as series resistance. Let

L and R_e = Inductance and effective resistance of coil to be measured,
r_1 and l_1 = Total resistance and inductance in arm BC,
r_2 and l_2 = Resistance and inductance in CD exclusive of R_e and L,
R and l_4 = Total resistance and inductance in AD including the equivalent series resistance of C_4,
r_3 = Equivalent series resistance of C_3.

The inductance in the arm AB may readily be reduced to a negligible amount and will not be considered.

We may now balance the bridge with the inductance terminals short circuited, that is, take a zero reading, and then balance again with the inductance inserted.

Writing the equations of balance in each case, subtracting one from the other, and separating reals from imaginaries, we get the following equations:

$$C_3 r_1(R - R') = L + (l_2 - l_2') + C_3 r_3(r_2 + R_e - r_2') + p^2 C_3 l_1(l_4 - l_4') \quad (1)$$

and

$$\frac{r_2' - r_2 - R_e}{C_3} = p^2 l_1(R - R') + p^2 r_1(l_4 - l_4') - p^2 r_3(L + l_2 - l_2'), \quad (2)$$

where l_2', r_2', l_4', and R' are the values of l_2, r_2, l_4, and R at balance with L short circuited, and p is 2π times the frequency. These are practically identical with Owen's equations (10) and (12).

In equation (1), each of the third and fourth terms contains two factors of second order, namely r_3 and $(r_2 + R_e - r_2')$, and l_1 and $(l_4 - l_4')$ respectively.

We may therefore write

$$C_3 r_1(R - R') = L + (l_2 - l_2'). \quad (3)$$

In equation (2), let

$$\frac{1}{pC_3} = -X_3, \quad pl_1 = x_1, \quad pl_4 = x_4,$$
$$pL = X, \quad pl_2 = x_2, \text{ and } pl_2' = x_2'.$$

Then we may write

$$-(r_2' - r_2 - R_e)pX_3 = px_1(R - R') + pr_1(x_4 - x_4') - pr_3(X + x_2 - x_2'). \quad (4)$$

But from (3)

$$R - R' = \frac{L + l_2 - l_2'}{C_3 r_1} = \frac{-(X + x_2 - x_2')X_3}{r_1}.$$

Substituting in (4),

$$r_2' - r_2 - R_e = \frac{(X + x_2 - x_2')x_1}{r_1} + (x_4 - x_4')\frac{(X + x_2 - x_2')}{R - R'} + \frac{r_3(X + x_2 - x_2')}{X_3}$$
$$= (X + x_2 - x_2')\left[\frac{x_1}{r_1} + \frac{x_4 - x_4'}{R - R'} + \frac{r_3}{X_3}\right]$$

and

$$R_e = r_2' - r_2 - (X + x_2 - x_2')\left(q_1 + q_4 + \frac{1}{Q_3}\right), \quad (5)$$

where q_1 = ratio of reactance to resistance of arm BC,
q_4 = ratio of reactance to resistance of *change* in arm AD,
Q_3 = ratio of reactance to resistance in arm AB.

From equation (3) we see that, if we take a zero reading first, the inductance is given by the expression

$$L = C_3 r_1 (R - R'), \tag{6}$$

the percentage error due to neglecting $l_2 - l_2'$ being

$$\frac{100(l_2 - l_2')}{L}. \tag{7}$$

From equation (5), the effective resistance of L is given by

$$R_e = r_2' - r_2, \tag{8}$$

the percentage error due to neglecting corrections being

$$\frac{100(X + x_2 - x_2')}{R_e}\left(q_1 + q_4 + \frac{1}{Q_3}\right). \tag{9}$$

The error in L is approximately, from equations (7) and (8),

$$\frac{x_2 - x_2'}{R_e} \cdot \frac{R_e}{X} = \frac{q_2}{Q},$$

where q_2 = ratio of reactance to resistance of change in arm CD, and
Q = ratio of reactance to resistance of the inductance being measured.

This error is usually negligible and may be approximately corrected for when appreciable. Dr. Owen has pointed out that this type of error is not peculiar to the Owen bridge, but is present in practically all methods of inductance measurement.

The error in R_e is a function of the Q of the coil measured, and of q_1, q_4 and Q_3. It is greatest for coils of high Q.

It is possible to make $q_1 = -\frac{1}{Q_3}$ for a given frequency, in which case the error reduces to approximately Qq_4 and the two errors are of the same order of magnitude for $Q = 1$,—the error in R_e becoming greater, and in L less as Q is increased.

However, in the general case we cannot cancel q_1 against $\frac{1}{Q_3}$ over any appreciable range of frequencies, and they are normally additive. Also for ordinary inductance coils Q is considerably greater than one, sometimes as large as 100. For such cases the error in R_e becomes large and difficult to determine without an accurate knowledge of the reactances of the resistances used and the losses in the condenser.

From the above relations we see that a method of this type is capable of measuring inductance with a high degree of accuracy and may be made to measure effective resistance with fair accuracy, provided that there is no coupling between any of the four arms nor any between them and the input and output circuits. This is in practice a difficult result to realize, and this difficulty in obtaining a simple but adequate system of shielding is one of the most serious limitations to the bridge.

Shielding

Since the bridge contains no inductances of appreciable magnitude, it is a comparatively simple matter to eliminate electromagnetic coupling by using input and output transformers in toroidal form, the input transformer being so designed that the core will not be saturated when using the maximum input to the bridge.

The elimination of the electrostatic coupling is not so simple, as any electrostatic shielding introduced adds capacitance which, unless due care is taken, will involve errors in the bridge. This means that such capacitances must be limited to the corners BD and AC where

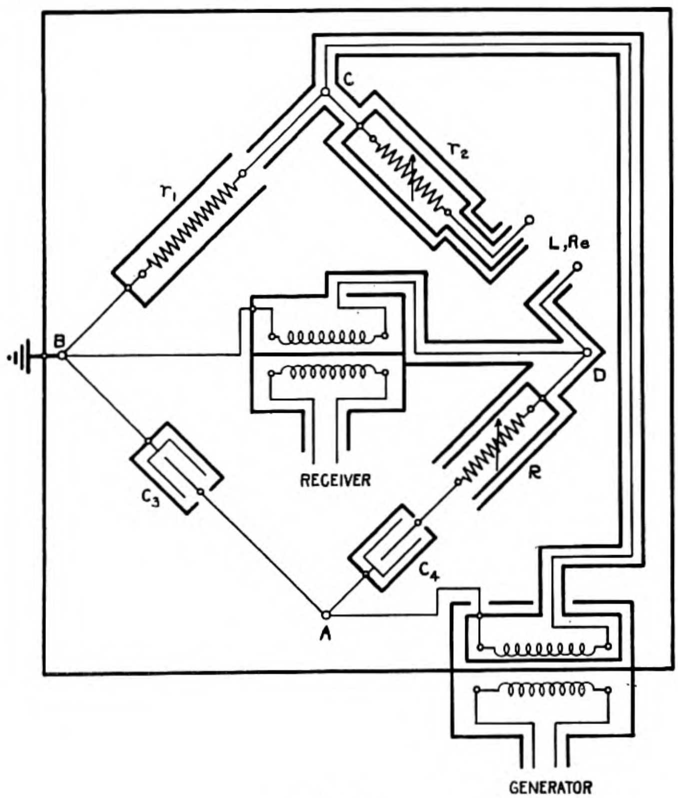


Fig. 2

they simply shunt the input and output circuits, and to AB where they shunt the capacitance C_3 and may be included in the assumed value of C_3. If such shielding is not used, the balance of the bridge will be affected by external conditions such as body capacitance, and the position of the bridge arms with respect to each other and to other apparatus, with the result that accurate results can be obtained only by the use of the greatest precautions.

A shielding scheme which satisfies the above requirements is shown in Fig. 2. In this system all capacitance between shields is limited to the diagonal corners of the bridge and the arm AB. However, this system of shielding, while about as simple as can be designed where complete shielding is required, is rather difficult to carry out in any practical bridge construction.

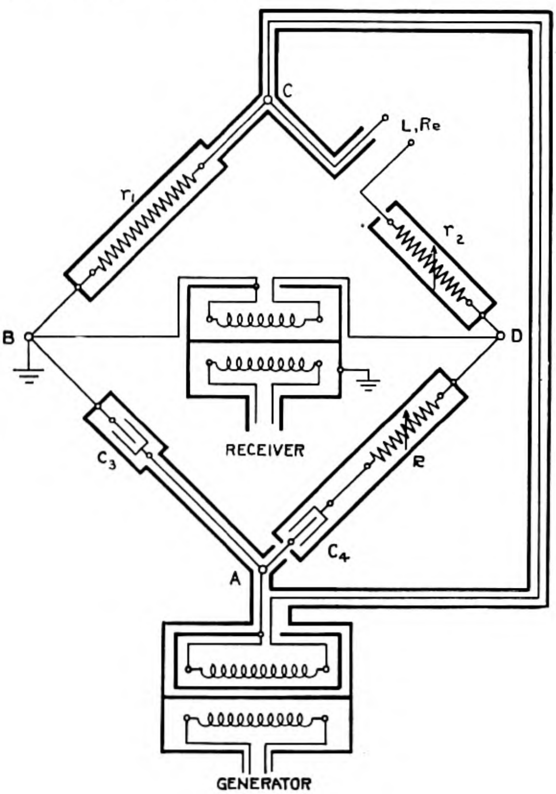


Fig. 3

The question of reducing the amount of shielding and still retaining a high degree of accuracy has been investigated and the modified scheme shown in Fig. 3 has been developed. In this circuit the shielding is complete insofar as it limits the electrostatic coupling to specific points in the bridge, and eliminates coupling between the bridge and the input and output circuits. However, in addition to capacitance across the diagonal corners and across arm AB, capacitances are introduced across r_1, across R, and across arm AD. The capacitance across AD may be made small enough to neglect since it consists of the capacitance of one condenser lead to the shield. Capacitances across r_1 and across R do not enter as first degree errors in the value of L but do directly affect the measurement of R_e. However, where the bridge is used primarily for the accurate measurement of inductance this compromise is justified. Even for the measurement of effective resistance, although the corrections may be larger due to the presence of the shielding, the bridge will give more consistent results and the corrections may be fairly well estimated.

The method of shielding shown requires one transformer having two shields between the windings and one transformer with a single shield between windings. It is essential that these shields be as perfect as possible. The other shielding shown is comparatively simple, no equipment requiring more than a single shield. The ground is shown at the point B simply because grounding at this point results in the simplest shielding. It would be desirable to have the ground at C in order that one terminal of the coil under test would be grounded, but at the time of balance the points B and D are at the same potential, and provided that r_2 is only a small fraction of the total impedance of the coil under test we may consider that one terminal of the coil is practically at ground potential. However, it should be noted that for a coil having a considerable capacitance from intermediate points in its winding to ground, a ground at B cannot be considered exactly equivalent to a ground at D. This difficulty is only appreciable in the case of very large inductances of large physical size when measured at high frequencies, and in such cases the effective inductance will be dependent on external conditions, whatever bridge circuit it is measured in. In the case of shielded coils, the ground should in all cases be connected to D rather than to B. In spite of the slight disadvantages noted, this method of shielding appears to be the most satisfactory, and a bridge has been constructed in accordance with it.

Construction of the Bridge

From the equation giving the value of L, it is seen that we may obtain an additional range for the inductance by having either r_1, C_3, or both, variable in steps. In the present bridge we have used two steps for C_3 and five steps for r_1. It is possible by choosing the correct values for r_1 to make the bridge direct reading for inductance. The actual values used for the capacitance were .6 mf and .06 mf. The values used for r_1 were 1,000/.6 or 1,667 ohms and multiples or submultiples of this value. In this way the bridge was made direct reading in millihenrys.

The capacitance C_4 has only one requirement to meet. It must be small enough so that the ratio of resistance to reactance of arm AD shall always be less than the ratio of reactance to resistance of the coil.

Taking 3,000 cycles as the maximum frequency, 10,000 ohms as the maximum resistance in arm AD, and 200 as a maximum value for the Q of the coil measured, then

$$2\pi fC < 200,$$

and

$$C < 1 \text{ mf}.$$

We have accordingly used a value of .6 mf in this arm to correspond with the value of C_3.

Resistances R and r_2 are dial type completely shielded resistance boxes which can be varied from 0 to 10,000 ohms in .01 ohm steps. The resistances are all of the reversed layer type, wound on impregnated wood spools and designed to give low phase angle and high stability.

The condensers are of the paraffine impregnated mica type, about ten years old, thus ensuring high stability, and having temperature coefficients less than .003 per cent per degree C., over the ordinary range of working temperatures.

The transformers are of a special type described by Shackelton.[1]

Accuracy—Measurement of Inductance

As previously stated the shielding, while increasing the stability of the bridge, introduces capacitances across R and r_1 which increase the corrections necessary in computing the effective resistance and may also require corrections in the measurement of inductance if sufficiently large. Accordingly, measurements were made on the bridge to determine the magnitude of this error. By shunting R and r_1 respectively, it was readily shown that capacitances as high as 200 mmf would not change the indicated inductance reading by as much as .01 per cent for all settings of r_1, for the whole range of R, over the whole audio frequency range. This conclusion is in accordance with equation 1. Since the shielding introduced capacitances across these points of the order of 25 to 50 mmf, this source of error may be neglected in the measurement of inductance.

Table I gives the exact values for C_3 and r_1, and the corresponding constant K by which the indicated value of R must be multipled to give the true inductance. This table shows how accurately the resistance r_1 has been adjusted to make the bridge direct reading. K is a simple number within .02 per cent in all cases when using the large condenser. The two condensers might have been made to have a ratio more nearly 10 to 1 by adding an auxiliary condenser to the larger one.

TABLE I

$K = C_3 \times r_1$ = Millihenrys per Ohm

r_1 (Ohms)	82.785	165.59	828.04	1656.1	8280.9
C_3 (mf)					
.60381	.049987	.099985	.49998	.99998	5.000
.06052	.0050103	.010022	.050113	.10023	.50117

A check was next made on a single inductance having a nominal value of .1 henry to determine the relative accuracy of different values of K at different frequencies. These values are given in Table II. It will be noticed that the value of L obtained is approximately

TABLE II

Comparison of Different Values of K Using a Single Inductance

Nominal Inductance, Millihenrys	K	Frequency, Cycles	R Ohms	R' Ohms	$L = K(R - R')$ Millihenrys
100	.099985	1,000	1,006.64	.03	100.65
"	.49998	"	201.34	.00	100.66
"	.050113	"	2,009.4	.45	100.67
"	.049987	"	2,013.4	.09	100.64
"	.099985	3,000	1,022.0	.03	102.18
"	.49998	"	204.40	.00	102.20
"	.050113	"	2,040.3	.00	102.24
"	.049987	"	2,044.1	.09	102.17

independent of K but the highest value obtained is for the value of K corresponding to the highest value of r_1. Since the reactance of this coil is only approximately 600 ohms at 1,000 cycles and the largest value of r_1 used was 828 ohms, it is evident that the potential of the coil with respect to ground varies considerably for different values of K. This is sufficient to account for the increased inductance value obtained for values of K using r_1 = 828 ohms. Keeping this in mind the different values of K agree with each other very closely. It has already been stated that r_1 should be small compared with X and therefore the values of K using $r_1 = 828$ ohms would not normally have been used for the measurement of this coil.

Table III gives a comparison of the inductance of several coils as measured on the Owen bridge and by a resonant method, the last column giving the difference between the two methods in per cent.

TABLE III

COMPARISON OF OWEN BRIDGE WITH RESONANCE BRIDGE

Nominal Inductance, Henrys	Frequency, Cycles	Measured Inductance		Difference, Per Cent
		Owen Bridge, Henrys	Resonance, Henrys	
.1	1,000	.10065	.10066	− .01
.1	2,000	.10124	.10118	+ .06
.15	1,000	.15072	.15082	− .04
.15	2,000	.15112	.15111	+ .01
1.0	2,000	1.0143	1.0144	− .01
2.9	1,000	2.918	2.918	.0
2.9	2,000	2.976	2.974	+ .07
10.0	2,000	11.295	11.27	+ .22

The resonant method was a highly accurate one in which frequency errors were negligible. The accuracy was probably of the same order as the measurements on the Owen bridge. The agreement between these two methods does not in itself indicate the accuracy of either method. However, the resonant measurements were made on a completely shielded equal ratio-arm bridge,[1] in terms of frequency and capacitance, using entirely different equipment from the Owen bridge in which the inductance is measured in terms of resistance and capacitance. Accordingly it is very improbable that these two methods had any errors in common and we may assume that the agreement obtained is a fair measure of the combined error of the two methods. Consequently from this table we see that for a range of .1 to 3 henrys and for frequencies up to 2,000 cycles the error in the measurement of the inductance by the shielded Owen bridge is less than .1 per cent and for 10 henrys is less than 1/4 per cent.

ACCURACY—MEASUREMENT OF RESISTANCE

The measurement of effective resistance in the case of an impedance of low reactance practically consists of the substitution of the unknown for the known resistance. In this case the accuracy of the measurement is high. However, the usual case we have to consider is the measurement of the effective resistance of coils of high Q. It is in such measurements that the greatest corrections are necessary, and it is also in such measurements that the greatest errors in effective resistance are produced by incomplete shielding in the bridge. Consequently it is in the measurement of effective resistance that shielding is most essential, and although this shielding may introduce a necessity for larger corrections due to the capacitance it introduces, these corrections may be made with a certain degree of precision and having made them the value obtained will be more reliable than in the case of a complete absence of shielding.

Table IV gives the figures for the measurement of effective resistance of three coils having a high Q. Referring to equation 5 we see that q_1 and q_4 are positive when the reactance is inductive and that Q_3 is

TABLE IV

Inductance, Henrys	Frequency, Cycles	r_2' Ohms	r_2 Ohms	q_1	q_4	$\frac{1}{Q_3}$	$X\left(q_1 + q_4 + \frac{1}{Q_3}\right)$	R_e Ohms	R_e' Ohms	Diff., %
1	1,000	1,670.66	1,358.30	− .0004	.0000	+ .0023	− 17	329.4	326	1
1	3,000	1,670.15	1,373.14	− .0013	"	"	− 68	365	378	3
.1	1,000	1,670.66	1,641.46	− .0004	"	"	− 1.7	30.9	30.1	3
.1	3,000	1,670.15	1,643.34	− .0013	"	"	− 6.8	33.6	32.5	3
.02	1,000	1,670.66	1,659.30	− .0004	+ .0003	"	− .30	11.66	11.47	2
.02	3,000	1,670.15	1,659.15	− .0013	+ .0011	"	− .94	11.94	11.8	1

always negative. The column headed R_e is obtained from equation 5. The column headed R_e' is obtained from a resonant method of measurement which has the same order of accuracy as the present method. Consequently the last column of differences gives the combined error in the two methods. In these measurements covering the most used range of inductance and a frequency range of 1,000 to 3,000 cycles, the largest difference between the two methods is 3 per cent. The total corrections to be made are in some cases extremely large, especially for the higher inductances and frequencies. This correction may amount to 30, or 40 per cent in some cases, and this means that an effective resistance obtained by the Owen bridge when not corrected may be in error by this amount. However, after allowing for the necessary corrections we can say that the bridge is capable of an accuracy for the measurement of effective resistance of about 2 per cent over the greater range of inductance and frequency.

Determination of Electrical Characteristics of Loaded Telegraph Cables

By J. J. GILBERT

SYNOPSIS: The use of permalloy for continuous loading has introduced a number of new factors of importance in the study of transmission of signals over long submarine telegraph cables. Data to check the theoretical assumptions that are used in the design of permalloy loaded cables can be obtained by measuring on such cables the attenuation and time of propagation of sinusoidal currents of various frequencies in the telegraph range. By combining the results of these measurements with data obtained on the cable during process of manufacture, the resistance, inductance, capacity and leakance of the cables can be determined.

This paper describes the experiments that were performed on three laid cables and discusses in a general way the methods of computing the cable parameters.

WITHIN the last few years the art of telegraphing over submarine cables of transoceanic length has been revolutionized by the development of effective means of applying to such cables the principle of inductive loading. By surrounding the copper conductor of the cable with a thin layer of permalloy, a material of high magnetic permeability, the range of signal speeds attainable over cables of the order of 2,000 n.m. in length has been multiplied eight to ten times.[1] In place of the low frequency band extending from zero to about 15 c.p.s., which represents the range of frequencies which can be efficiently transmitted over the usual type of non-loaded cable, we are concerned in the case of the loaded cable with a transmission band extending from zero to about 120 c.p.s. Largely because of this comparatively high speed of operation, a number of factors, which were of negligible influence in the case of non-loaded cables, have become of primary importance in affecting the speed of signalling, and it has been found necessary, in order to establish a definite basis of estimating the performance of loaded cables, to make a thorough study of these factors by theoretical analysis supplemented by experimental work in the laboratory, and by measurements on laid cables.

PRINCIPLES OF CABLE TRANSMISSION

The theory of transmission of signals over submarine telegraph cables[2] and the principles governing the design of permalloy loaded

[1] O. E. Buckley, *B. S. T. J.*, Vol. IV, No. 3, July 1925; Electrical Communication, Vol. 4, No. 1, July 1925; *Jour. A. I. E. E.*, Vol. XLIV, No. 8, August 1925.

[2] H. W. Malcolm, "The Theory of the Submarine Telegraph and Telephone Cable," London, 1917.

J. W. Milnor, *Jour. A. I. E. E.*, Vol. 41, p. 118, 1922.

cables [1] have been fully discussed elsewhere and only a brief summary will be given here for the purpose of indicating the importance of the measurements that will be described. On account of the fact that for a given value of sending voltage the amplitude of the signals received over a submarine cable diminishes rapidly as the speed of signalling is increased, there is a practical limit to the speed of operation of any cable. This limit depends on the electrical characteristics of the cable and the magnitude of extraneous interference encountered at the receiving terminal. The criterion for legibility of signals is, in general, that the attenuation constant of the cable at a value of frequency which may be termed the critical frequency shall not exceed a given value, the attenuation constant αs being defined by the relation

$$\frac{|V_R|}{|V_S|} = \epsilon^{-\alpha s}, \tag{1}$$

where $|V_R|$ is the amplitude of voltage arriving at one end of the cable when a sinusoidal voltage of amplitude $|V_S|$ is impressed at the other terminal. The value of this critical frequency depends mainly upon the method of operation, and it usually lies somewhere between the signal frequency and one and one half times the signal frequency.

Given the values of the four fundamental parameters of the cable, resistance (R), inductance (L), capacity (C) and leakance (G), the attenuation constant at the frequency $p/2\pi$ can be computed by means of the formula

$$2\alpha^2 = \sqrt{(R^2 + p^2L^2)(G^2 + p^2C^2)} + RG - p^2LC, \tag{2}$$

which to a close approximation reduces to the form

$$\alpha = \sqrt{\pi f C R} \tag{3}$$

in the case of a non-loaded cable, where R is large compared with $2\pi f L$, and to the form

$$\alpha = \frac{1}{2}\left(R + \frac{G}{C}L \right)\sqrt{\frac{C}{L}} \tag{4}$$

in the case of the loaded cable, where R is small compared with $2\pi f L$ at the critical frequency. In all cases it is assumed that G is very small compared with $2\pi f C$, which is strictly true for the insulating materials employed on submarine cables.

The manner in which the attenuation constant varies with frequency for typical loaded and non-loaded cables is shown in Fig. 1, the signal frequencies at which they are designed to operate being as indicated. In the case of the non-loaded cable the resistance and capacity are practically constant over the frequency range and the attenuation

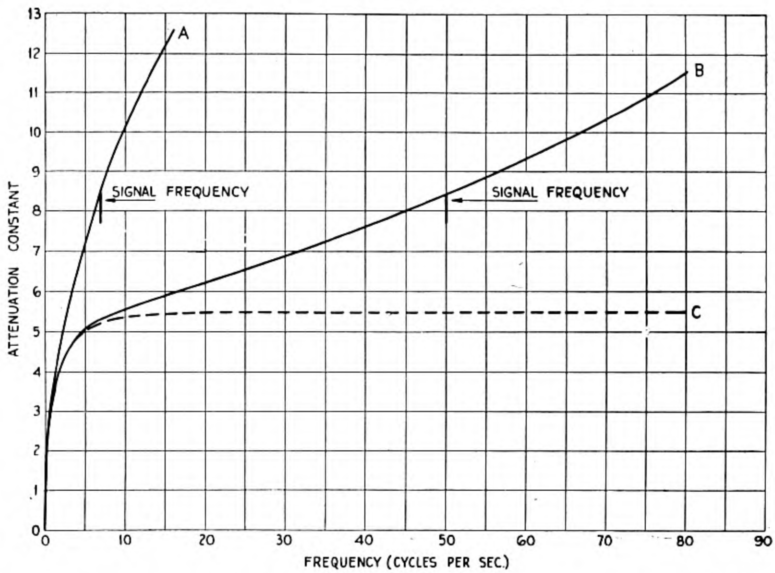


Fig. 1—*A*—non-loaded cable; *B*—actual loaded cable; *C*—ideal loaded cable

curve is approximately a parabola as indicated by formula (3). The curve for the loaded cable for small values of frequency is similar to the curve for the non-loaded cable, since for such frequencies the loading inductance has very little effect upon transmission. As soon as $2\pi fL$ becomes appreciable compared with R the beneficial effect of the inductance becomes apparent and the attenuation constant increases at a less rapid rate. If the cable parameters were constant throughout the frequency range, as in the case of the ideal cable, the attenuation constant would, at a value of frequency considerably below the signal frequency, attain a constant value, as represented by the dotted curve. On account of the fact, however, that R and G increase rather rapidly with frequency, the attenuation-frequency characteristic of an actual cable merely inflects, then increases, and at some frequency will actually cross the attenuation curve of the non-loaded cable.

To insure that legible signals will be obtained at the desired signal frequency the amplitude of the extraneous interference must be accurately determined. If, for example, the interference in the case of the cable having the attenuation-frequency characteristic shown in curve B were found to be twice as great as had been anticipated, the amplitude of received signal would likewise have to be doubled, which would mean a reduction of 0.7 in the allowable attenuation constant. This, as can be seen from curve B, would correspond to a reduction in speed of 8 to 10 c.p.s. Also since the value of attenuation constant is considerably affected by variations of the electrical parameters, it is desirable that the values of these parameters in the laid cable be capable of predetermination to a degree of accuracy comparable with that obtained in the case of non-loaded cables. Methods of estimating the value of extraneous interference to be expected at the terminals of a projected cable have been described in a previous paper.[3] The present paper will be devoted to a discussion of methods of predetermining the electrical parameters of cables.

Measurements during Manufacture

In the case of a non-loaded cable the attenuation constant, as indicated by formula (3), is determined solely by the dielectric capacity and the conductor resistance. For the values of frequency involved in the operation of such cables, the latter consists almost entirely of the direct current resistance of the copper conductor. The values of capacity and copper resistance of a considerable part of the cable can be measured during the process of manufacture, and, by reducing these values to sea bottom conditions, an accurate estimate of the resistance and capacity of the laid cable is obtained.

In the case of the loaded cable the problem of predetermining the electrical parameters of the laid cable is much more difficult, since a number of the quantities involved in computing the attenuation are influenced by conditions which are not entirely known and which are difficult to simulate in laboratory experiments. The dielectric leakance, for example, is affected by pressure as well as by temperature, and since the hydrostatic pressure to which the cable is subjected may be as high as 10,000 pounds per square inch, it is evident that measurements of this characteristic of the cable, on any but a very small scale basis, will be very difficult and costly. The permeability of the loading material and consequently the inductance of the cable may be affected by mechanical strain and by superposed magnetic fields. An estimate of the average inductance of the laid cable can be obtained by bridge measurements in the factory on pieces of core about 1

[3] J. J. Gilbert, *B. S. T. J.*, Vol. 5, p. 404, and *Electrician*, Vol. 97, August 6, 1926.

nautical mile in length, selected at intervals during manufacture, the effect of strains and of superposed fields being estimated by means of experiments on short lengths of cable. However, there are ordinarily small unavoidable variations in electrical characteristics from point to point along the cable and it is not entirely certain that the average inductance obtained from measurements on a fraction of the core lengths entering into the cable structure will represent the average inductance of the entire cable. The resistance of the laid cable is likewise difficult to estimate. This parameter comprises, in addition to the copper resistance, the resistance of the return conductor consisting of the armor wires and sea water in parallel, components resulting from eddy current and hysteresis losses in the loading material and other components of lesser importance, the nature of which will be discussed later. The losses in the loading material depend upon the average permeability obtained in the laid cable, and their predetermination from factory measurements may be uncertain for reasons that have been pointed out. As regards the sea return resistance, rigorous methods of computation are available,[4] but there is some uncertainty regarding the conditions that should be assumed as existing at the ocean bottom.

Measurements on Laid Cables

For the purpose of placing the design of loaded cables upon a definite basis, it has appeared desirable to measure the parameters of a number of cables of this type that have been laid, and to compare the values so obtained with the estimates based on analytical methods and upon factory measurements. In order to simplify the problem, attention will be devoted mainly to determining the values of the parameters corresponding to a very small value of current in the cable conductor. Under these conditions the hysteresis component of resistance is negligible and the inductance and eddy current resistance can be considered constant at any frequency. This is entirely consistent with the method employed in the design of loaded cables, in which the attenuation constant is computed, first on the assumption that the current is very small throughout the cable, and then corrected for "head end losses" due to the effect of hysteresis losses which are present under actual conditions of operation.

The usual method of determining the parameters of a transmission system consists in measuring the propagation constant, Γ, per unit

[4] J. R. Carson and J. J. Gilbert, *Jour. Franklin Institute*, Vol. 192, p. 705, 1921; *Electrician*, Vol. 88, p. 499, 1922; *B. S. T. J.*, Vol. 1, No. 1, p. 88.

length and the characteristic impedance, K, which quantities are defined at the frequency $p/2\pi$ by the formulas

$$\Gamma = \sqrt{(R + jpL)(G + jpC)}, \tag{5}$$

$$K = \sqrt{\frac{R + jpL}{G + jpC}}. \tag{6}$$

Knowing these two quantities at any frequency, the values of the four parameters can be readily computed.

The propagation constant and the characteristic impedance of *telephone* cables 100 miles or less in length have been determined by measuring the input impedance of the cable with the distant end in turn insulated and grounded. These two impedances are determined for a cable of length s by the formulas

$$Z_I = K \coth \Gamma s$$
$$Z_G = K \tanh \Gamma s,$$

and given the values of Z_I and Z_G it is an easy matter to compute the corresponding values of propagation constant and characteristic impedance, the accuracy of this determination depending upon the difference between Z_I and Z_G. In the case of a submarine telegraph cable of the order of 2000 miles in length, the value of Γs is so large that Z_I and Z_G differ by less than one part in 10,000 in the frequency range in which we are interested. This means physically that the remote parts of the cable have little effect upon the terminal impedance of the cable and the values of input impedance are determined almost entirely by the parameters of the 400 or 500 miles of cable adjacent to the terminal. It is true that by going to extremely low frequencies, perhaps fractional cycles per second, the method above described could be used to determine the characteristic impedance and the propagation constant of long cables, but at such frequencies these quantities are determined almost entirely by the d.c. resistance and capacity of the cable and no information regarding the quantities in which we are particularly interested would be obtained.

The method that has actually been employed to determine the parameters of several of the continuously loaded cables which have recently been laid is to measure separately at a number of frequencies the real and imaginary parts of the propagation constant, the capacity of the cable at various frequencies being determined by correlating the results of laboratory tests with d.c. measurements of capacity made on the laid cable.

As can be seen from formula (4), the real part of the propagation constant, αs, the attenuation constant of the cable, involves all four of the cable parameters, but on account of the fact that the inductance, leakance and the various components of the effective resistance predominate in influence at different points in the frequency range it is possible, by methods of successive approximations, to obtain a reasonably good set of values of these quantities.

The imaginary part of the propagation constant, βs, is to a close approximation, given by

$$\beta s = sp\sqrt{CL}. \tag{7}$$

From this it follows that the time of propagation of a sinusoidal wave of voltage or current over the cable is given by

$$T = s\sqrt{CL}, \tag{8}$$

and knowing the time of propagation and the capacity at any frequency the inductance of the cable at this frequency can be easily computed. Since the resistance and leakance have only a slight effect upon the time of propagation, this is the most direct method of determining the average inductance of the cable.

Measurement of Attenuation

The attenuation constant of the cable is determined by measuring the values of voltage received at one end of the cable, due to various values of voltage of constant frequency impressed at the other end. The impressed voltage may be either sinusoidal or square top in shape, the latter being preferable for the reason that, at the low frequencies and high voltages required, it is difficult to obtain a wave form from an oscillator sufficiently free from harmonics to enable an accurate determination of the fundamental component to be made. Square top reversals of any frequency and amplitude can be easily obtained by means of a relay actuated by an oscillator, and the amplitude of the fundamental component can be accurately computed.

At the receiving end, for the frequencies of particular interest, the arriving voltage is practically sinusoidal, since the harmonic components are eliminated by the higher attenuation of the cable for such frequencies. This voltage is measured by terminating the cable in an impedance which is very large compared to the characteristic impedance of the cable, and measuring the potential drop across all or part of this impedance by means of a vacuum tube amplifier in the output of which is a thermocouple and meter. The advantages of the high impedance termination are, first, that by reflection it doubles the amplitude of the arriving voltage, thus giving larger quantities to work with, and second, that it eliminates the necessity of taking into account the characteristic impedance of the cable and the impedance of the balanced type of sea earth which is usually employed as the earth connection of the amplifier. By means of a string oscillograph in the output of the amplifier, the wave shape of the received voltage and the nature of the extraneous interference can be determined. The amplifier is calibrated by impressing on it a measured voltage of the same frequency as that of the received voltage.

Knowing the values of received voltage and the corresponding transmitted voltage, the values of attenuation constant can be readily computed. By plotting the values of attenuation constant corresponding to various values of frequency and transmitted voltage as functions of the latter quantity and extending these curves to the axis of zero transmitted voltage, the values of attenuation constant corresponding to a very small current in the conductor can be obtained for various frequencies.

Assuming that all the parameters have been accurately predetermined, there are three sources of error which might possibly cause a difference between the measured value of attenuation constant and that computed from the average values of the cable parameters by means of formula (2). In the first place, the parameters are not uniform throughout the cable as is assumed in deriving this formula. In particular, the inductance may vary from point to point. At each point where the capacity or inductance changes value reflections of voltage and current will take place and the effect of these reflections should be to increase the attenuation constant of the cable. For variations of the parameters of the order that is to be expected in loaded cables, the increase in attenuation constant is quite small, and the magnitude of this increase can be computed approximately by a method due to Carson.[5] Another source of error is the presence of extraneous interference superposed on the received voltage. This factor is usually troublesome only at the highest frequencies and lowest voltages employed, and in this case measurements of the oscillograms of received voltage and of calibrating voltage will give a value of the received voltage independent of interference. The third source of error is due to the presence in the transmitted voltage of harmonics of the fundamental frequency. These harmonics are attenuated in transmission over the cable to a much greater degree than is the fundamental, so that they constitute only a small per-

[5] *Electrician*, Vol. 86, p. 272, 1921.

centage of the received voltage and are practically negligible in their effect upon the thermocouple.

Measurement of the Time of Propagation

The time of propagation of a steady state sinusoidal voltage over a loaded cable of transatlantic length is of the order of 0.3 second. It is measured by means of the circuit shown in Fig. 2, which is

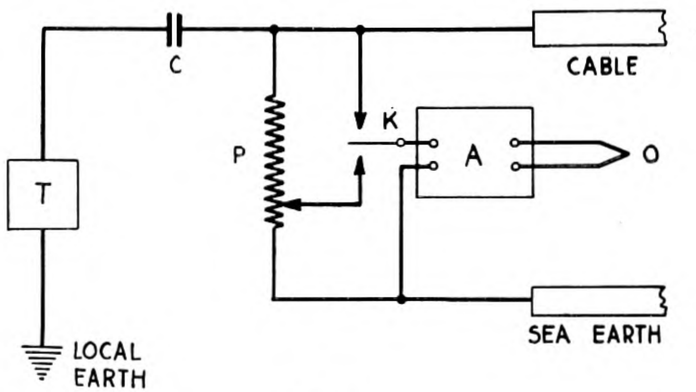


Fig. 2.

operated simultaneously at both ends of the cable. At each end a perforated tape is prepared which when inserted in the high speed transmitter *T* will cause a train of about ten reversals to be sent out over the cable. The potentiometer *P* is adjusted so that a measurable record of either transmitted or arriving trains, depending upon the position of the key *k*, will be obtained on the string oscillograph *O* after amplification by the vacuum tube amplifier *A*. The condenser *C* is inserted between the cable and transmitter in order to remove the low frequency components of the transient part of the train, which would otherwise overwhelm the steady state component at the distant end of the cable. The oscillograph, shown in Fig. 3, gives a continuous record of the current in a fine wire, which is free to respond to the interaction between the current and the strong magnetic field in which the wire is placed. The displacement of the wire, and hence the amplitude of current in it, is recorded on a long strip of sensitized paper, which is developed and fixed within the camera by a continuous process immediately after exposure. By this means it is possible to obtain a continuous record, over a period of several minutes, of voltages transmitted and received over the cable. A second wire can be used to give simultaneously a record of any other current which may be desired for comparison. An arrangement is provided for superposing on the records vertical timing lines at intervals of one hundredth of a second. Short pieces of record are shown in Fig. 4.

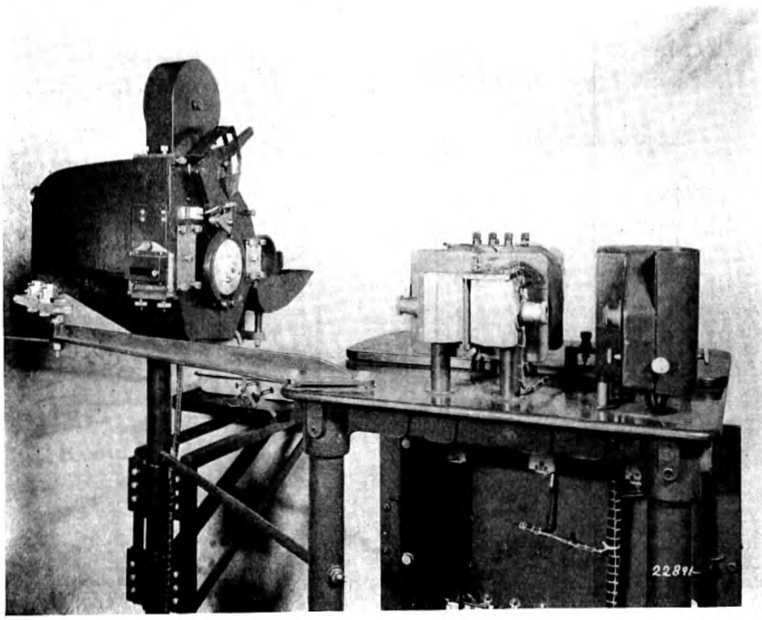


Fig. 3.

At a prearranged time the oscillographs at both ends are started. A train of reversals is transmitted from one end, a record being taken on the oscillograph at that end, and received at the distant end, where a record is also taken. Both stations quickly change potentiometer connections from send to receive or vice versa, and the distant station transmits a train of reversals, recording it on the same tape as was used for reception. Similarly at the first station the arriving train is

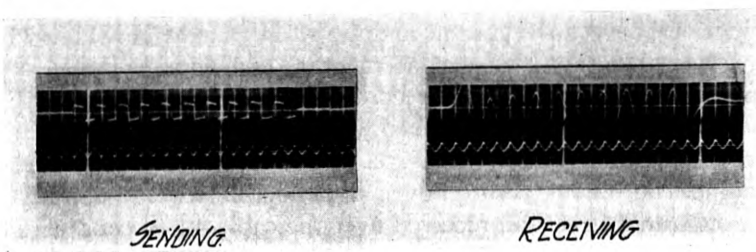


Fig. 4.

recorded on the strip containing the record of their transmitted train. Station 1 measures on its oscillogram the time elapsing between its transmitted train and its received train, and at Station 2 the time elapsing between the received train and the transmitted train is measured. After making suitable corrections, which will be described later, the difference between the interval measured at Station 1 and that measured at Station 2 will be equal to twice the time of transmission of the train of reversals over the cable.

A typical record such as would be obtained at Station 1 is shown in Fig. 4. It will be observed that in the record of received voltage the first few cycles are somewhat distorted because of the fact that the steady state has not yet been reached. Because of this fact the time of arrival or departure of a train is referred to a later cycle in the series, say the fifth. The times of departure and arrival of the various zeros following this cycle are measured, and the average of the values so obtained is defined as the time of arrival or departure of the train. In this way the possible errors due to interference or to distortion in the sent record due to improper functioning of the transmitter are eliminated.

It will be observed that, mainly on account of the presence of the condenser C, the voltage reversals impressed on the cable are not flat-topped and the zero phase of the fundamental component which we are measuring occurs somewhat ahead of the point in the transmitted voltage which we have used as the zero of reference in measuring the oscillograms. Since we are interested in the time elapsing between zero phase of the fundamental frequency in the transmitted voltage and the zero phase of the corresponding cycle in the received train, it is necessary to compute this interval, either by graphical analysis of the oscillogram or by computation from the constants of the circuit, and add the corresponding time to the time which has been measured.

Although the mechanical arrangement by which the timing lines are obtained on the oscillogram is adjusted as accurately as possible so that the interval between lines is very nearly one hundredth of a second, the very slight variations which occur in such a system are apt to introduce considerable error into the measurement of time of propagation. This is due to the fact that the time of propagation is obtained from the difference of two intervals each of which may be as much as ten times the time of propagation. An error in either interval will therefore result in a tenfold error in the final result. To guard against this condition a record is taken during the experiment of a periodic voltage obtained from a standard oscillator or fork, and the peaks of this oscillation serve as a check on the timing lines. As a final check, records similar to Fig. 4 are taken with various times elapsing between reception and transmission at the second station. If an error exists in the timing arrangement its effect on the time of propagation will be greater the greater the interval between receiving and sending, and the time of propagation corresponding to negligible error in the timing system can be easily obtained by graphical methods. The error of measurement of the time of propagation is probably less than 1 per cent.

The inductance of a loaded conductor is an increasing function of current for the range of current values used in cable practice because of the increase of permeability of the permalloy, and since with finite transmitting voltage the current at the sending end may be quite large, the inductance of this portion of the cable under such conditions will be larger than the value it would have for very small current in the conductor. Accordingly the time of propagation at a given frequency will be a function of voltage. The value of inductance corresponding to very small current in the conductor can be derived from the time of propagation corresponding to zero transmitted voltage, which is obtained by extrapolation from measurements of the time of propagation at several values of transmitted voltage.

Measurement of Capacity

The dielectric capacity of submarine cables in the telegraph range of frequencies is in general comparatively insensitive to changes in temperature and hydrostatic pressure, so that it is possible to estimate this quantity rather accurately at various frequencies by means of measurements made in the factory, the factors required to reduce the results of the measurements to sea bottom conditions being relatively easy of determination. In order to check these values, however, the d.c. capacity of the laid cable is measured by the method of mixtures, employing a charging time of 10 seconds or more and a mixing time of equal duration.

Computation of Cable Parameters

The inductance of the cable can be computed at any frequency from the measured values of capacity and time of propagation by means of equation (8), proper allowance being made for the rather small effect of resistance.

Having computed the inductance and the capacity of the cable, only the resistance and the leakance remain undetermined. The direct current resistance can be computed from factory measurements and checked by measurement on the cable. The resistance component due to eddy currents in the loading material can be computed from the resistance measurements obtained in the factory in the process of determining the inductance of sample core lengths. The eddy current resistance is proportional to the square of the product of frequency and permeability, and corresponding reduction factors must be employed in computing the eddy current resistance of the laid cable from the factory measurements. Since we are dealing with values of the parameters corresponding to very small current in the cable conductor, the hysteresis resistance is zero. In addition to the losses in the loading material there are other losses peculiar to continuously loaded cables due to currents induced in the cable structure. The loading material is ordinarily applied to the conductor in the form of a tape or wire of finite width, so that it has a definite lay, and since the magnetic flux in the loading material tends to follow the convolutions of the latter there is a component of this flux parallel to the axis of the central conductor. Consequently as the flux changes with signal current, electromotive forces are induced in those portions of the cable structure which link with it—the teredo tape and armor wires, for example. The resulting energy loss has in most practical cases comparatively small effect on the performance of the cable, and the magnitude of the corresponding resistance component can be estimated by theoretical methods and by measurements in the factory. The various components of resistance having been estimated, the total resistance at any frequency can be computed. Likewise the value of dielectric leakance of the laid cable at any frequency can be estimated from tests made during manufacture. These values of resistance and dielectric leakance should be considered merely as first approximations, since they are based in part on assumptions that cannot be directly verified.

Formula (2) is then employed to determine the effect upon the attenuation constant of departures from the approximate values of resistance and leakance, and by comparing these results with the measured values of attenuation constant, mutually consistent sets of values of resistance and of dielectric leakance can be computed at various frequencies. A choice of the best sets of values can then be made, due weight being given to the evidence available from computations and laboratory measurements regarding the manner in which these quantities vary with frequency.

From the curves relating the values of measured attenuation constant and the transmitted voltage, a check can be made of the method of computing the increase in attenuation due to hysteresis and to variation of inductance with current.[6] Since this method employs

[6] See Buckley, loc. cit., and U. Meyer, *E. N. T.*, Vol. 3, No. 1, 1926.

the inductance-current and resistance-current characteristics of the loaded conductor, as determined in the factory, the attenuation measurements also afford a check on these characteristics.

Conclusions

Measurements of attenuation, time of propagation and dielectric capacity of the laid cable at various frequencies, supplemented by measurements of eddy current resistance in the factory and by information regarding the manner in which sea return resistance and dielectric leakance vary with frequency are sufficient for determining the values of the four parameters of a loaded cable and for dividing the resistance into its component parts. A quantitative comparison of the results so obtained with the values of parameters that would be predicted from factory measurements alone would require a detailed discussion of the methods involved in such measurements, and is outside the scope of the present paper. A general conclusion that can be drawn from the results of measurements made on three cables of somewhat different characteristics is that the method of estimating the characteristics of laid cables from measurements made on short lengths of core during process of manufacture is capable of considerable accuracy. The values of inductance and dielectric leakance obtained from factory measurements are close enough to the actual values in the laid cable to give a value of attenuation constant within a few per cent of the actual value. The value of resistance obtained from the cable measurements appears to be about three to five per cent higher than the estimated value. This may in part be due to latent errors in measurement or in the method of allowing for the effect of reflections along the cable.

The greater part of the discrepancy between the estimated and measured values of resistance is perhaps due to erroneous assumptions involved in computing the value of sea return resistance employing the method described in the paper by Carson and Gilbert. In this work it was assumed that the cable is surrounded by a homogeneous medium, the sea water. For values of frequency higher than the telegraphic range this assumption appears to be sufficiently close to the truth, since only a comparatively small region around the cable plays any part in the phenomena. In the telegraph range, however, the return current is distributed through a comparatively large cross-section and more exact specification of the electrical characteristics of this region is required. To determine by rigorous methods the sea return impedance in the case where the cable lies in a plane separating two different media is a problem of considerable difficulty. An approximate method, which gives results which are sufficiently accurate for purposes of cable design, consists in computing the combined impedance of the three parallel conductors, namely, the armor wires, the sea water, and the earth, the impedances of the latter two conductors being determined by the methods outlined in the aforementioned paper. The physical interpretation of the sea return resistance as obtained by this method is that the high value of reactance of the sea water and earth, due to the large cross-section of the conducting area, forces the return current to flow in the armor wires even though the resistance of this path is much higher than that of the paths through the sea water and earth. It appears probable that the electrical conductivity of the earth is very much less than that of sea water which would result in a larger cross-section of conducting area external to the armor wires and larger inductance of this path. This leads to higher values of sea return resistance than are obtained on the assumption that the cable is surrounded on all sides by sea water and thus gives a result more nearly consistent with the observed facts.

Automatic Printing Equipment For Long Loaded Submarine Telegraph Cables

By A. A. CLOKEY

SYNOPSIS: The introduction of the permalloy loaded submarine cable has presented the possibility of telegraph transmission at speeds several times those obtainable on non-loaded cables and has made practicable the operation of printer telegraph equipment. The present paper presents the various factors which affect the design of operating equipment and describes the apparatus which has been developed and used for a considerable period of time under service conditions. The transmission speed attained may exceed 2,400 letters per minute. To a certain extent, the detailed design of the terminal apparatus is controlled by the electrical characteristics of the particular cable to which it is to be applied and this type of equipment cannot, therefore, be completely standardized.

GENERAL

AT the time the development of the loaded submarine telegraph cable was undertaken, non-loaded cables were generally being operated duplex at signalling speeds ranging from 5 to 8 cycles per second (160 to 260 letters per minute) in each direction. The transmitting apparatus consisted of transmitters of the reciprocating contact type controlled by perforated tapes and the signals were received and recorded by the delicate moving coil type of amplifiers (generally referred to as magnifiers), relays and siphon recorders which produced a received signal record of such a character as to require the employment of highly skilled operators to translate and type the messages in final form. Except for a few trials, automatic printers had not been applied commercially to the operation of submarine cables, although the highly successful results which had been previously obtained with multiplex printing telegraph equipment on land lines coupled with the increasing demands made upon the cable systems as a result of the World War had directed the attention of telegraph and cable engineers to the need for applying automatic printing telegraph methods to submarine cables.

Preliminary studies of the characteristics of permalloy as a loading material for long telegraph cables indicated that, through its use, transmission speeds many times that of non-loaded cables could be readily attained. As the then existing apparatus was incapable of operation at the high speeds thus obtainable and the operating methods in use were not suited to handling the greatly increased volume of traffic over a single cable, it became apparent that new operating methods and equipment would have to be developed if the full advantage afforded by the use of permalloy loading [1] was to be realized. The development of the permalloy loaded cable was, therefore, paralleled by a study of the newly presented operating requirements and the development of suitable operating methods for high speed loaded cables. It is the purpose of this paper to present the various factors which affect the design of operating equipment for use on long loaded cables and to describe the apparatus and principles of operation which have been developed. The system which is to be described is similar to the multiplex system now in use on American land lines [2] but has been modified in several important respects in order to adapt it to the requirements of cable transmission.

The Cable and Amplifier as a Transmitting System

Submarine cables have heretofore been thought of as transmitting media which greatly distorted the signals and so reduced them in amplitude as to require the use of a sensitive siphon recorder for reception. The effects of signal distortion were, to a certain degree, compensated for by the addition of sending and receiving condensers, magnetic shunts at the receiving end, and, to a slight degree, by the inherent characteristics of the siphon recorder itself. A separate instrument termed a cable magnifier, of which there are several different types, was inserted between the cable and the siphon recorder to "magnify" the signal delivered to the recorder and thus partially offset the effects of attenuation. No two cables are identical as regards the distortion and attenuation of the signals and the means which will effectively provide for the correction and amplification of the signals on one cable will not necessarily be suitable for use on another cable of different length or construction. The apparatus provided for the correction of distortion in and amplification of the received signals is therefore an essential part of a signal transmitting system which includes the cable, and, except for the necessary switching arrangements, is independent of the means employed for impressing the signalling impulses upon the cable and for producing a permanent record of the corrected signals. Thus the development of terminal equipment for loaded cables comprised two separate and distinct developments, viz. the study of signal distortion and design of suitable signal shaping amplifiers (described in a separate paper by Mr. A. M. Curtis which appears elsewhere in this issue), and the development of apparatus for delivering signals to the system comprising the cable

[1] O. E. Buckley, "The Loaded Submarine Telegraph Cable," *Jour. A. I. E. E.*, June 26, 1925.

[2] J. H. Bell, "Printing Telegraph Systems," *Trans. A. I. E. E.*, Vol. 39, Part 1, 1920.

and amplifier and for converting the signals delivered by that system into a permanent printed record.

With the combination of cable and signal shaping amplifier, signals which are transmitted into the sending shaping network and the cable as square topped impulses as shown in Fig. 1 emerge from the amplifier

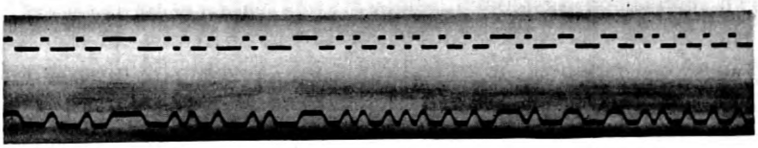


Fig. 1

as rounded impulses from which the high frequency components have been removed as a result of the attenuating effect of the cable. The receiving and printing system must therefore be capable of accurately translating rounded signals of this nature into printed characters.

Requirements of Operating System for Loaded Cables

The outstanding characteristic of the loaded cable is the enormously increased speed of transmission which may be as high as 2,400 letters per minute or more. For practical utilization of such high speeds the operating system must include some means for dividing the line time to provide a number of traffic channels. This is necessary in order to facilitate the distribution of the work of preparing the perforated transmitting tapes and checking the received message records among the required number of operators. The system must also provide for efficient two-way operation to avoid delay in the transmission of traffic from either terminal and should be capable of being joined with other cables or land lines through automatic repeaters to avoid the delay and expense introduced by manual methods of repetition.

As the shape of the received signal is determined by the cable-amplifier system and also by the character and amount of interference present which cannot be eliminated by the distortion correction networks, the operating system must be able to take the partly corrected signals delivered to it by the amplifier and accurately restore them to the form in which they were originally transmitted before using them to control the final recording mechanism.

The apparatus associated with loaded cables will in practically all cases be installed in the same offices as the equipment in use on non-loaded cables and, in order to avoid the necessity for duplicating the operating and maintenance staffs, it should be of such a nature as to permit of its being operated and maintained by men familiar with the operation of apparatus in use on land lines and ordinary cables.

Codes

The signalling speed attainable on any telegraph circuit, the effect of interference upon the received signals, and the design of the operating equipment depend to a certain extent upon the telegraph signal code used. A great variety of codes have been devised from time to time with a view to effecting greater economy of line time or greater freedom from the effects of interference, but only a few of them have been generally adopted in commercial practice. These may be divided into two general groups: the two-element codes which are composed of various combinations of positive and negative current impulses, of which the continental Morse and the Baudot codes are well-known examples, and the three-element codes in which a zero or no-current interval is employed to separate individual pulses of a group or as a third element in the combinations. The cable code and three-unit code are examples of the three-element type.

The codes in each of these two groups may be subdivided into two classes, those known as uniform codes in which all characters are composed of the same number of equal time units and those known as non-uniform codes in which the impulses forming the characters vary in length, number or both. The non-uniform codes are well adapted for use where the received signals are translated manually, but are not so well suited to automatic translation as the uniform codes on account of the mechanical and electrical complications introduced by the necessity for distinguishing between signal combinations of varying length.

Of the uniform codes which have been used in automatic printing telegraph systems, the Baudot or two-element five-unit type of code possesses advantages over the three-element three-unit type of code which make it much better suited for automatic operation of submarine cables. The three-element three-unit code employs, as does also the cable code, a zero interval of unit length in forming the signal combinations representing each character or letter and the shape of the received signals must be sufficiently refined to make this zero interval easily distinguishable (see *A*, Fig. 2) in order to prevent confusion in translation. Even with the best shaping obtained to date on long non-loaded cables operated at high speeds the presence of this zero interval may be indicated only by a difference in slope of the recorded curve as shown in *B*, Fig. 2. Interference currents, which are present to some extent in all cables, are superposed upon the

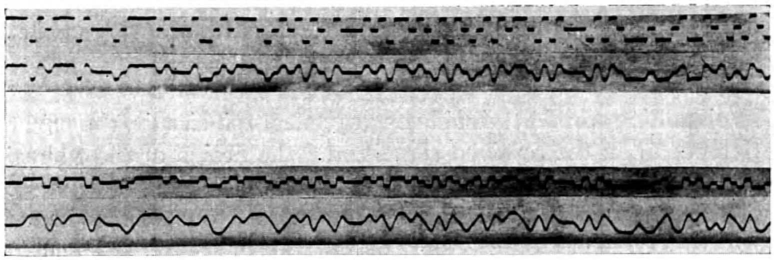


Fig. 2

received signals and cause troublesome distortion in the zero intervals and the length of the sustained pulses. The absence of these zero intervals of unit length in the two-element five-unit code, combined with the fact that only the middle portion of each received signalling impulse is used to operate the selecting mechanism of the printers, considerably reduces the effect of interference upon the accuracy of translation and makes it unnecessary to secure such refined signal shape.

The accurate evaluation, in terms of transmission speed, of the relative merits of the various telegraph codes is a highly complex problem which does not readily lend itself to solution through purely theoretical methods since it involves consideration not only of the total number of separate combinations which must be provided to represent the letters of the alphabet and all other characters to be transmitted, the frequency of occurrence in traffic of the various characters, and the average number of unit impulses required to form the combinations, but also depends upon the characteristics of the line or cable and the nature and distribution of the interference encountered and its effect upon the shape and definition of the received signals. The application of a code to any specific case also involves the more practical considerations of the type and operating characteristics of the apparatus employed. Practical experience therefore probably forms the best guide to the choice of a code.

In consideration of the conditions referred to above and the experience previously gained through the extensive use of the Baudot type of code on automatic telegraph circuits both in the United States and Europe, it was concluded at an early stage in the development that the multiplex code used on American land line multiplex circuits would be the most suitable for high speed automatic submarine cable transmission. Subsequent experience has indicated that the original conclusion was amply justified and has shown that the Baudot type of code, when used in connection with terminal apparatus of suitable design, is probably faster than any of the other types of codes which have been considered for high speed loaded cable operation.

Outline of System

The multiplex system [3] used on land telegraph lines was in many respects well suited to the requirements of loaded cable operation. It was capable of operation at high transmission speeds, was more economical of line time than other methods which were considered, and provided for the division of traffic between a number of traffic channels in a manner which afforded great flexibility in the handling and routing of traffic and permitted the channel speeds to be fixed at values which would allow the operating staff to work at maximum efficiency. Its long continued use on land telegraph lines had resulted in bringing the apparatus, operating methods and routines to a high degree of perfection and the development of a thoroughly trained staff skilled in the operation and maintenance of the equipment, all of

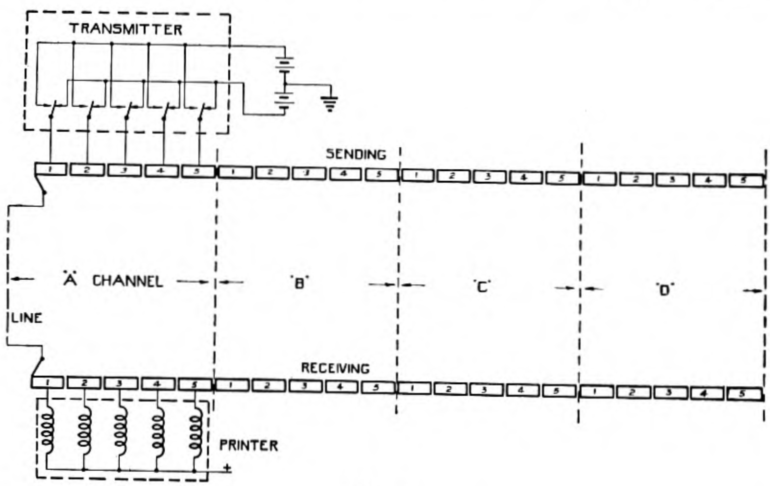


Fig. 3

which was of inestimable importance in the successful application and operation of printing telegraph methods to submarine cables.

The multiplex system provides for associating the line at the sending

[3] J. H. Bell, loc. cit.

27

end with each one of a number of transmitters in rotation by means of a rotating brush which passes over a segmented commutator to which the transmitters are connected as illustrated in Fig. 3. In this figure the commutator segments are shown developed for sake of simplicity. At the receiving end, the line is similarly associated in rotation with each one of a corresponding number of printers by means of the receiving brush and commutator. The commutator brushes at the two ends of the line are maintained in nearly exact synchronism by short correcting impulses which are derived from reversals in the received signalling currents, and their phase relation is such that each of the five segments connected with the "*A*" channel transmitter will be connected in rotation through the line to the corresponding segments of the "*A*" channel printer once during each revolution of the brushes, and the impulse transmitted from any one sending segment will pass through the corresponding receiving segment and operate the printer selector magnet which is connected to it. Similarly the transmitter on each of the other channels will be connected to its corresponding printer once during each revolution of the brushes. The commutators and the associated brushes together with the mechanism provided for correcting the phase relation of the brushes are usually referred to as distributors.

In the operation of this system a transmitting tape is prepared in which the characters to be transmitted are represented by combinations of holes perforated in the tape by means of a keyboard perforator which resembles a typewriter. The tape thus prepared is drawn through a transmitter which is arranged to apply to its associated distributor segments, positive and negative battery in the proper combination to form the five-unit impulses corresponding to the perforations in the tape. The received signal combinations control the operation of an automatic telegraph typewriter or printer which converts the signals into printed characters. Detailed descriptions of the perforating, transmitting and printing apparatus and the various methods for maintaining synchronism used in the multiplex system are given in the paper by Mr. J. H. Bell, previously referred to, and also in an excellent book by Mr. H. H. Harrison entitled "Printing Telegraph Systems and Mechanisms."

On account of the many advantages which this system embodied, it was chosen in principle as a basis for the development of the new system, although in several important respects much of the apparatus and operating methods employed were entirely unsuitable for loaded cable operation. The multiplex had been employed almost entirely in the operation of duplexed circuits and therefore was not applicable to simplex operation of cables. The character of the received signals and interference on long cables is such as to require the use of entirely different methods of reception in order to utilize the line time most efficiently, and the higher transmission speeds expected on cables necessitated departure from standard land line practices in the matter of apparatus design and number of channels employed. The system as finally developed embodies the following important improvements over previous methods.

1. An entirely automatic means for quickly reversing the direction of transmission on a simplex circuit at short intervals which can be altered as required to accommodate varying traffic loads in the two directions.

2. A synchronous vibrating relay which corrects for the residual distortion in the signals delivered by the amplifier, and practically doubles the speed of transmission.

3. A high degree of precision and refinement in the design and construction of apparatus which is justified by the great cost of the cable relative to that of the terminal apparatus.

The inclusion of these improvements in a modified multiplex system involved, of course, the solution of a number of important incidental problems such as the provision of Morse "talking circuits" which could be made instantly operative, and the development of suitable arrangements for linking two simplex cable sections together through repeaters.

Two Directional Working

The use of duplex methods in the operation of non-loaded cables enables communication to be carried on simultaneously in both directions and usually effects an increase of from 60 to 90 per cent in the total traffic capacity of the cable. As only a moderate capital expenditure is required to equip a non-loaded cable for duplex operation practically all cables of this type are now equipped in this way as a matter of economy. The characteristics of the loaded cable, however, are such as to require the use of highly complicated and extremely expensive artificial lines and balancing equipment for duplex operation and it is quite doubtful whether the total duplex traffic capacity thus secured would equal that obtainable by the use of simplex methods. Duplexing the loaded cable therefore appeared to afford no certain economic gain over simplex operation and the extremely high cost of duplexing could hardly have been justified merely for the sake of securing simultaneous transmission in both directions.

The apparatus and methods formerly employed for reversing the direction of transmission on manually operated simplex cables were so time-consuming that it was impracticable to reverse direction oftener than once every quarter or half hour. The delay in transmission which would result from the adoption of the older methods could not be permitted on the loaded cable and it therefore became necessary to develop special apparatus for automatically reversing the direction of transmission at comparatively short intervals in order to approximate simultaneous transmission in both directions and reduce traffic delays to an absolute minimum.

The design of suitable switching arrangements which would permit stopping transmission on a long cable operated with multichannel printing equipment and almost immediately starting transmission in the opposite direction presented several difficult problems. On account of the lack of uniformity in the lengths of the messages to be transmitted and the number of channels employed, it rarely happens that the transmitters on all channels complete the transmission of their respective messages at exactly the same instant, therefore it was necessary to arrange for making the change in direction of transmission at more regular and frequent intervals even though the transmitters on all channels had only partly completed the transmission of their respective messages at the time the change was made. To accomplish this without introducing any errors or other evidence of the interruption into the final printed message necessitates first stopping the transmitters on all channels at precisely the right instant, then allowing an interval equal to the time of signal propagation over the cable to elapse before cutting off the printers at the distant end, and finally upon resumption of transmission in that direction starting all of the transmitters and printers at the proper time and in the correct sequence to avoid the loss, repetition, or mutilation of any character.

The last signals transmitted into the cable before changing to the receiving position result in leaving the cable charged to a potential which would paralyze or "block" the amplifier were it to be immediately connected. Part of this charge must be dissipated and the current due to the residual charge and the presence of any interference or earth currents must be allowed to attain its steady value in the shaping network and input transformer elements of the amplifier before connecting any of the actual amplifying elements to the cable. The switching operations involved in applying the amplifier to the cable must be effected in the proper sequence and at precisely timed intervals in order to leave the amplifier in the proper condition to avoid mutilation of the first signals received from the distant end.

The required degree of accuracy in timing the various switching operations involved in reversing the direction of transmission was secured by utilizing the rotating shafts of the distributors at both stations to control a timing mechanism which determined the lengths of the transmission intervals in the two directions.[4] This timing mechanism is essentially an electrical revolution counter which can be set to count any desired number of revolutions of the distributor shaft and close within a fraction of a revolution of that number the circuit which controls the operation of the various contacts which do the actual switching. As the distributor shafts at the two ends of the cable are maintained in exact synchronism in a manner previously described, the timing mechanisms will therefore also operate in synchronism, and if at the time of setting up the circuit they are started in the proper phase relation the correct phase relationship will be maintained as long as the operation of the circuit continues without interruption. The timing mechanisms are driven from the distributor shafts through the medium of an electrically operated clutch which when disengaged permits the timing mechanisms at all stations to be manually set in their proper positions and started together in this relationship by means of a starting impulse sent over the line which causes the clutches to engage.

In order to provide for transmission intervals of various lengths in the two directions, the timing mechanism includes a number of timing elements each representing a different division of the line time, any one of which can be quickly selected at will by the movement of an indicating lever to control the length of the transmitting and receiving periods.

Upon the completion of the predetermined number of revolutions of the distributor the timing mechanism operates a direction control relay, see Fig. 7, the contacts of which are arranged to operate and cut off the transmitters, discharge the cable, and connect the amplifier and the printers in properly timed sequence. The actual time consumed in making all of the circuit changes necessary to reverse the direction of transmission, measured from the time of transmission of the last signal combination to the time of printing the first character on the printer at the same station, is of the order of five seconds but will vary somewhat on different cables according to the length of the cable and the magnitude and character of the interference and earth currents encountered.

During the interval in which the actual switching operations are taking place no signals are being transmitted in either direction so

[4] A. A. Clokey, U. S. Patent No. 1,601,941.

that neither of the distributors will receive any correction impulses and as a result the sending and receiving brushes may depart considerably from their normal phase position. This would cause errors to occur in the first signals received upon resumption of transmission if means were not taken to bring the brushes back into proper phase relationship before the transmission of actual signals was begun. This is provided for by arranging to have the distributors transmit, at the close of each switching period, a number of "spacing" signals which do not affect the receiving printers since they are not connected in circuit until a sufficient number of reversals have occurred in the line current to correct the receiving brush into the proper position. The transmission of signals which must be recorded by the printer is then started.

As the length of the interval allowed for these switching operations is determined by a definite number of revolutions of the distributor shaft, which may be set to rotate at various speeds, the gearing between the distributor shaft and the timing mechanism is designed to allow for a five-second switching period when the distributor is rotating at a speed which corresponds to the maximum transmission speed of the circuit.

Although this system lacks the advantage of absolutely continuous communication in both directions, it possesses another feature which goes far toward offsetting, if it does not entirely outweigh, the advantages afforded by the duplex method. Almost all of the long cables of the world run in an east and west direction and the difference in time between the terminal stations of those cables results in an unequal distribution of traffic in the two directions except perhaps during a comparatively short time each day. The provision of the selective timing mechanism permits the total traffic capacity of the cable to be divided between eastward and westward transmission in about the same proportion as the eastward traffic load bears to the westward load and thus permits efficient utilization of the entire traffic carrying capacity of the cable.

The Synchronous Vibrating Relay

The vibrating relay principle was first suggested by Gulstad [5] who applied it to short cables for overcoming the effects of distortion. As originally used, it consisted of a sensitive polarized relay provided with a line winding, upon which the received signals were impressed, and two auxiliary windings included in a local vibrating circuit adjusted to cause the relay armature to vibrate continuously when

[5] K. Gulstad, "Vibrating Cable Relay," *Elec. Rev.*, London, Vol. 42, 1898; Vol. 51, 1902.

the line winding was de-energized. The rate of vibration was adjusted to be approximately the same as the frequency of the transmitted signals and the amplitude of the vibrating current was adjusted to be approximately equal to the received signalling current so that the latter, if of one polarity, would neutralize the effect of the vibrating impulse and prevent the movement of the relay armature and if of the same polarity would aid the vibrating impulse. The effect of this combined action of the vibrating and received signalling impulses is to reproduce, in the local circuit, signals of approximately the same shape and duration as the original transmitted signals.

The frequency attenuation characteristic of a system comprising a long telegraph cable and its signal shaping amplifier and networks when the latter are adjusted for the maximum transmission speed is such as to cause the impulses of unit length, which represent half cycles of the fundamental signalling frequency, to be received in considerably smaller amplitude than the impulses of two units (or more) length which represent half cycles of one half (or less) the fundamental signalling frequency.[6] The highest signalling speed obtainable on a given cable is therefore determined by the length of the shortest impulses which must be received in sufficient amplitude to exercise control over the receiving apparatus and at that speed the two-unit and longer impulses will be received in much greater amplitude than is necessary for operation of the receiving apparatus. Gulstad pointed out [7] that as the received impulses of unit length always occur in the proper direction to aid the vibrating impulses they may therefore be greatly reduced in amplitude without impairing the accuracy of reception. On account of this fact the speed of signalling may be increased to a point where only the two-unit and longer impulses are received in sufficient amplitude to overcome the effect of the locally generated vibrating impulses and control the movement of the relay armature. At this increased speed the impulses of unit length will be either greatly diminished in amplitude or entirely removed by the attenuating effect of the cable and at such times the armature of the vibrating relay will be operated by the locally produced impulses.

As the rate of vibration of the Gulstad relay was determined entirely by the values of the resistances and capacities in the local vibrating circuit, the vibrations of the relay armature did not exactly coincide either in frequency or phase relation with the signals sent by the distant transmitter so that complete restoration of the incoming signals to their original form was impossible and full advantage of the speed

[6] The fundamental signalling frequency is defined as the fundamental frequency of a train of alternate positive and negative impulses of unit length.

[7] K. Gulstad, loc. cit.

possibilities of the device could not be realized. For these reasons its use was limited almost entirely to comparatively short cables where the strength and shape of the received signals were sufficiently good to control the relay directly with only a small improvement in shape and with no amplification. The original arrangement was later modified [8] to adapt it to the operation of longer unloaded cables.

One of the principal features of the cable multiplex herein described is the synchronous vibrating relay [9] which was developed particularly for high speed operation on long cables, and is a great improvement over the Gulstad device. The vibrating impulses, instead of being derived from an adjustable vibrating circuit, are generated by a segmented commutator located on the receiving head of the distributor. As the brushes on the receiving head of the distributor rotate in nearly exact synchronism with the transmitting brushes, it is evident that the rate of vibration of the relay will coincide exactly with the frequency of the transmitted signals and by properly adjusting the angular position of the vibrating segments, the time of closure of the relay contacts with respect to the incoming signalling impulses can be accurately fixed. The accuracy with which the missing impulses of unit length in the received signals are reinserted by this means makes it possible to realize the full speed possibilities of the vibrating relay principle and obtain faithful reproduction of signals on a given cable at almost double the speed obtainable through the use of ordinary non-vibrating relays.

Another important advantage gained through the use of the synchronous vibrating relay is greater freedom from the effects of extraneous interference. The amplitude of the received signals is sufficiently great to permit of its being reduced by the effects of interference to approximately half of the normal value before the distortion becomes sufficiently great to cause errors in printing. Likewise interference occurring during the zero intervals in the received signals must attain a value of approximately half of the normal received signal amplitude before causing errors. Interference occurring during the intervals between vibrating impulses will, of course, produce no effect upon the relay unless the amplitude of the interference attains a sufficiently large value to operate the relay directly. This ratio of interference to received signal amplitude represents the absolute limit of operation and some margin must obviously be allowed. It has been found that continuous satisfactory operation can be maintained so long as the interference does not

[8] W. Judd, British Patent No. 9,768, April 25, 1913; G. R. Benjamin and Herbert Angell, U. S. Patent No. 1,579,999, April 6, 1926.

[9] A. A. Clokey, U. S. Patents Nos. 1,521,870 and 1,522,865.

exceed one third of the normal signal amplitude. The presence of a proportionate amount of interference in the received signals in ordinary cable code operation would cause the recorder record to be so mutilated as to render it entirely illegible.

A detailed description of the operation of the synchronous vibrating relay is given in the appendix.

Refinement

The extreme speed at which the apparatus must operate to utilize the entire capacity of a loaded cable precludes making any adjustments while in operation and the importance of maintaining uninterrupted service for long periods demands that the apparatus shall be absolutely reliable in its operation and as free as possible from any variation in adjustments which would require occasional correction. This degree of reliability is secured through a refinement in mechanical and electrical design and a precision in construction which might be considered uneconomical for ordinary land line operation. The extra expense incurred in the design and construction of such highly refined apparatus is well justified by the resultant large increase in traffic capacity of the cable and the small cost of even the most refined apparatus relative to that of the cable on which it is used.

A general idea of the type of apparatus and construction used can be gained from one of the terminal distributors which is illustrated in Fig. 4. The greatest permissible variation in the phase relation

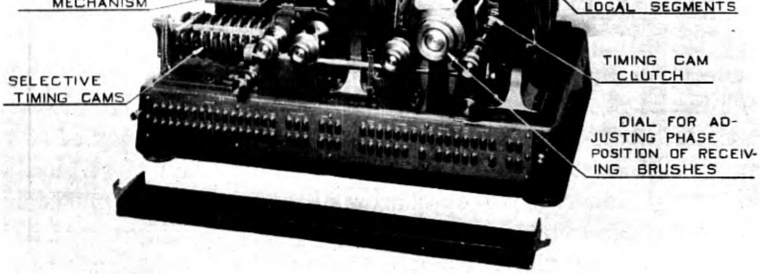


Fig. 4

between the brushes on the distributors at the two ends of the cable is only about one and one half degrees of revolution and in order to hold within this limit it was necessary to design a driving unit in which the phase shift, resulting from variations in the line voltage, was reduced to a minimum, and to arrange the gearing and coupling between the driving motor shafts and the various rotating brush arms so as to reduce to a minimum any lost motion or back lash. The driving unit consists of two motors: the one which supplies the power for driving the brushes is a dynamotor in which the DC side is used as a motor to supply the power and the AC side is included in a circuit with an electrically driven tuning fork which controls the motor speed within very close limits; and the other is a phonic wheel or La Cour motor driven from the same driving fork. This motor normally supplies little if any power for driving the distributor but by increasing or decreasing the load it prevents the occurrence of any appreciable phase shift in the DC motor due to variations in the driving voltage. In order to prevent slight shifting in the phase of the brushes due to vibration and axial twisting in the shafts and gears, it was necessary to employ much heavier construction in the rotating parts than is actually required to transmit the small amount of power used. The cutting of the gears, the distributor segments, and timing cams was done with the utmost precision to eliminate mechanical errors. The distributor segments included in the vibrating relay circuit are heavily faced with coin silver to reduce variation in the resistance of the contact between them and the rotating brushes.

The satisfactory operation of the system depends upon the accuracy with which the various relays in the system follow and repeat the signals. None of the available types of relays were found to be sufficiently reliable to permit of use in the system and it became necessary to develop for the purpose a new type of high speed relay shown in Fig. 5. The size and inertia of the parts comprising the moving system of this relay were reduced as much as possible in order to secure quick response and freedom from contact chatter at the highest operating speeds. A magnetic circuit was designed in which the effects of magnetic hysteresis are practically negligible, which results in the relay always operating upon the same value of current irrespective of its previous magnetic history. Permanency of adjustment, which is essential in relays used in this class of service, was obtained by adhering to standards of accuracy and precision of manufacture heretofore considered unnecessary in relay construction. The accuracy with which relays of the new type will operate at high speeds and the entire freedom from contact chatter is illustrated in the oscillogram reproduced in Fig. 6, which shows 200- and 600-cycle sine waves applied to the relay windings and the character of the reversals repeated by the contacts. At these and lower frequencies

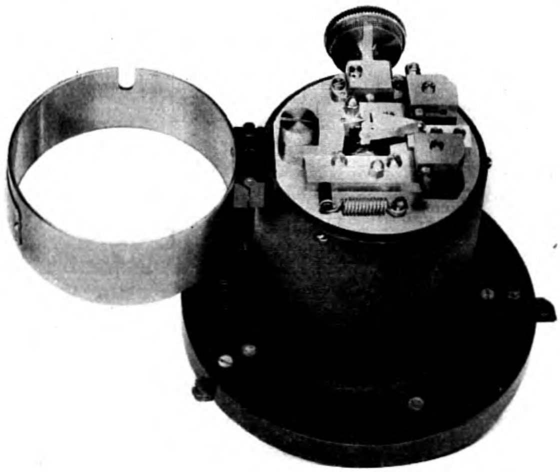


Fig. 5

the adjustment of the relay is sufficiently stable to permit of its being operated continuously for long periods without requiring readjustment or other attention.

Apparatus of this nature is frequently installed in isolated stations where materials or parts needed for making repairs cannot be obtained promptly, and the climatic conditions at some of these stations often

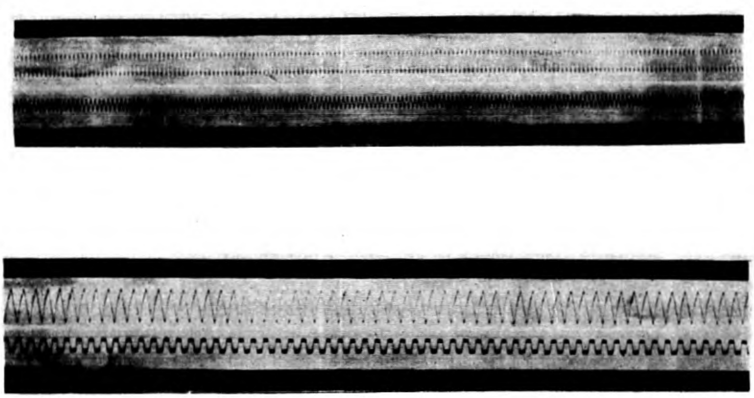


Fig. 6

impose quite severe requirements upon the mechanical as well as the electrical portions of the apparatus. In designing the apparatus, the greatest care was therefore exercised in selecting, for the construction of even the smallest details, materials which would withstand the most severe usage and be unaffected by the most severe climatic conditions.

Summary of System

The inclusion of these newly developed features in the multiplex system and the application of the modified system to a long loaded cable presents a number of interesting aspects and new possibilities. The entire system is shown schematically in Fig. 7.

The use of an amplifier containing no mechanical moving parts, in which all adjustments are made by alteration of the constants of electrical circuits, makes it possible to determine at the time of installation the proper amplifier adjustments to give satisfactory signal shape at a number of different transmission speeds and thereafter the amplifier may be quickly set for any speed by duplicating the adjustments that were previously found suitable for that speed. As the operation of the correcting relays and circuits and the vibrating relay depends to some extent upon the shape of the signals delivered by the amplifier, the ability to reproduce accurately a signal shape which has been previously found satisfactory is of considerable importance in the operation of the system.

Although the amplifier and shaping networks are considered a part of the cable system rather than an element in the transmitting and receiving system, their operation must be controlled by the direction control switching mechanisms. The relays included in the direction control system which switch the amplifier circuits are built in the amplifier to simplify wiring and maintenance. The speed of the distributors is controlled as in the multiplex system by vibrating tuning forks, but in order to secure under certain conditions greater stability and freedom from speed variations due to alteration in the fork contact adjustment and changes in room temperature and voltage of the power supply there was developed a constant temperature vacuum tube driven fork. The distributor, with its driving fork, the relays included in the direction control and vibrating relay circuits, and the apparatus usually provided in land line multiplex equipments for phasing and lining up the circuit, including the Morse talking circuit, are mounted in accessible positions on a table which is separated from the operating tables on which the printers and transmitters are located.

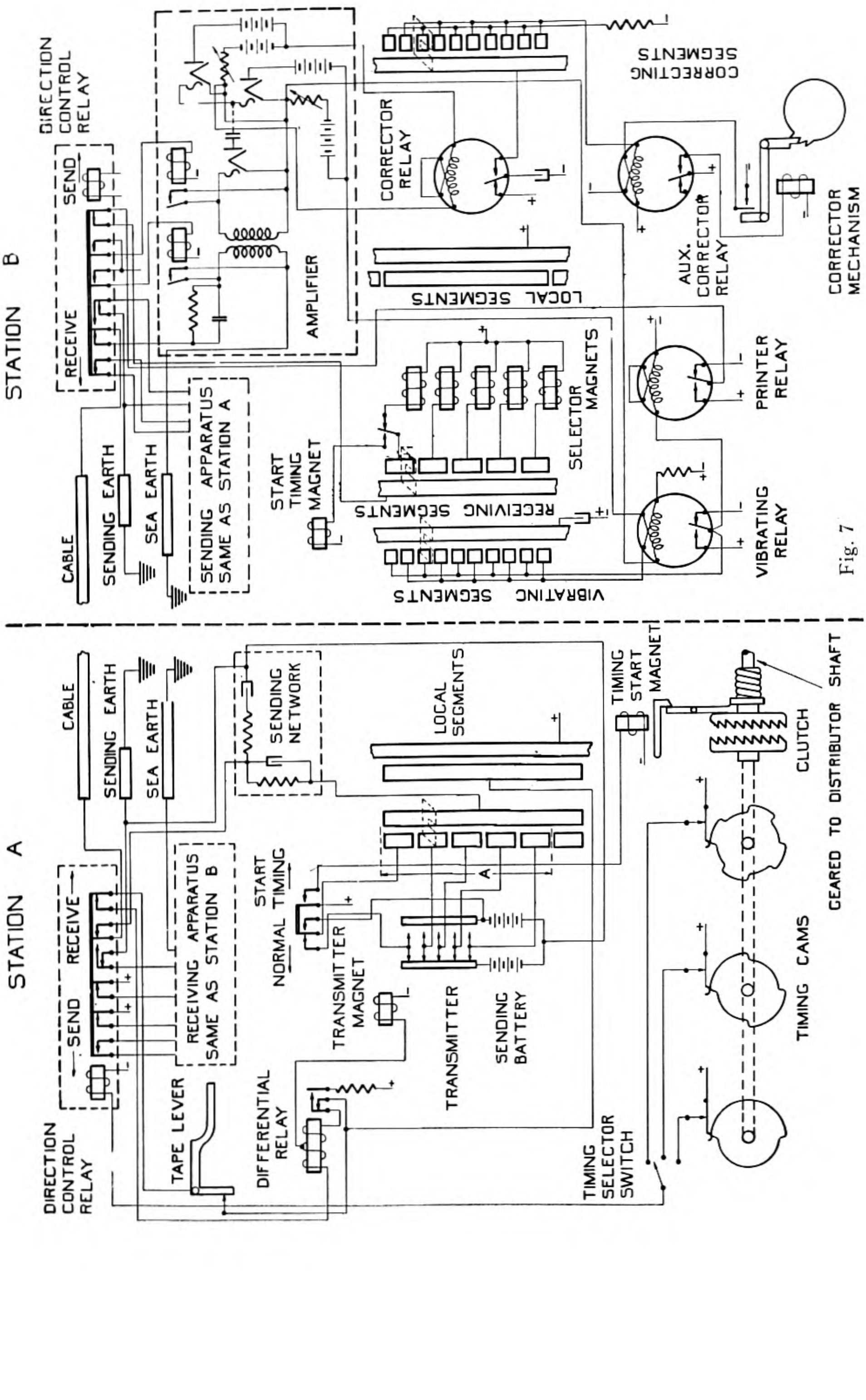

STATION B
DIRECTION CONTROL RELAY
SEND
RECEIVE
CABLE
SENDING EARTH
SEA EARTH
SENDING APPARATUS SAME AS STATION A
AMPLIFIER
CORRECTOR RELAY
CORRECTING SEGMENTS
LOCAL SEGMENTS
AUX. CORRECTOR RELAY
CORRECTOR MECHANISM
START TIMING MAGNET
SELECTOR MAGNETS
RECEIVING SEGMENTS
VIBRATING SEGMENTS
PRINTER RELAY
VIBRATING RELAY
STATION A
DIRECTION CONTROL RELAY
SEND
RECEIVE
CABLE
SENDING EARTH
SEA EARTH
RECEIVING APPARATUS SAME AS STATION B
SENDING NETWORK
TAPE LEVER
DIFFERENTIAL RELAY
TRANSMITTER MAGNET
NORMAL
START
TIMING
TRANSMITTER
SENDING BATTERY
LOCAL SEGMENTS
TIMING START MAGNET
TIMING SELECTOR SWITCH
TIMING CAMS
CLUTCH
GEARED TO DISTRIBUTOR SHAFT

Fig. 7

The adaptation of the multiplex to cable operation does not involve any modifications which affect the design of the perforators, transmitters, or printers, so that it is possible to employ in this system the same type of instruments used in land line operation.

In cases where it is desirable to link two cable sections together automatic repetition is provided for by the provision of additional sending and receiving commutators on the repeater distributor for transmitting and receiving on the second section. A photograph of such a repeater distributor is shown in Fig. 8. The incoming signals

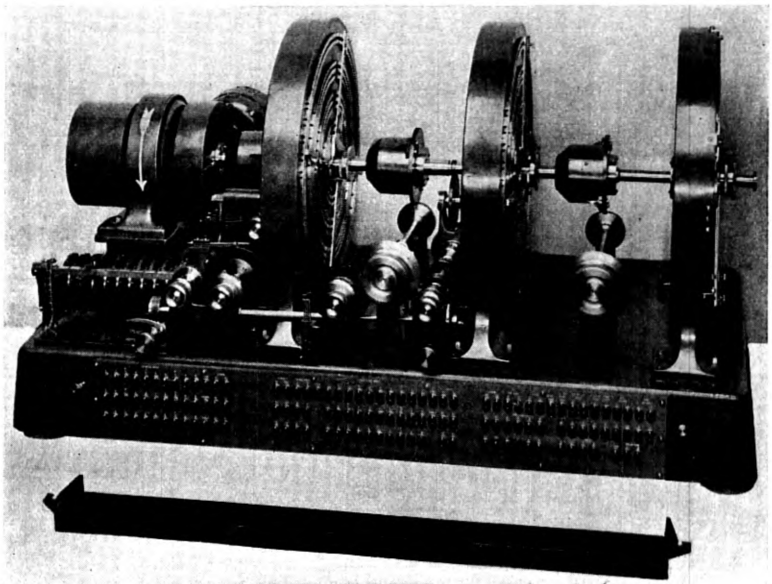


Fig. 8

at the repeater are received in the regular way and operate the vibrating relay which applies the completely corrected impulses to the receiving segments. The receiving segments, instead of being connected to the selector magnets of a printer, are connected to the windings of storing relays which are operated by the incoming signal combinations and set up the identical combinations on the corresponding sending segments associated with the next section of cable. The storing relays thus perform the same function as a tape transmitter except that they are controlled directly by the received signals instead of a perforated tape. With this method [10] of repetition it is possible to replace the storing relays on any channel with a printer and trans-

[10] E. P. Bancroft, et al., U. S. Patent No. 1,541,316.

mitter on each section so that one or more channels on both cable sections may be terminated at the repeater station without interfering with automatic repetition of traffic on the remaining channels.

Not only is it possible to link two or more simplex cable sections together through automatic repeaters, but it is also possible to link such a system through repeaters with a duplexed land line multiplex without introducing serious complications. The printer on the receiving side of each channel of both the land and cable circuits at the repeater station may be replaced with an automatic reperforator which will prepare from the incoming signals a perforated tape for retransmission. As this tape leaves the reperforator it is automatically drawn through a standard transmitter which will transmit the signals into the corresponding channel of the next section of line or cable. In moving between the reperforator and transmitter the tape passes under a contact closing lever arranged to stop the operation of the transmitter when the slack in the retransmitting tape drops below a predetermined minimum as the result of a difference in transmission speed on the two sections or the stoppage of the reperforator on the simplex section during the transmitting periods. This avoids the possibility of mutilation of the transmitted signals or tearing the tape.

The provision of a comparatively large number of traffic channels and automatic repeaters by means of which traffic on any or all channels may be automatically repeated into the other cable sections or land telegraph lines affords a high degree of flexibility in handling and routing traffic and permits the several channels to be terminated at the two ends in widely separated points.

Conclusion

Although the general principles of the system and the general design of the apparatus described herein are applicable to all loaded cables irrespective of length or construction, it is quite obvious that the detailed design of the various pieces of apparatus required will be determined to a great extent by the electrical characteristics of the particular cable to which they are to be applied and by the operating and traffic requirements which that system must fulfill. Equipment of this type can not therefore be standardized to the degree possible in the case of similar equipment for land line service, and the provision of apparatus for each cable becomes a special engineering problem which must be worked out with the cooperation of the engineers of the operating company in order to make the apparatus capable of satisfactorily meeting all of the conditions which will obtain in subsequent commercial use.

A complete operating equipment embodying the general principles described has been designed with the cooperation of the engineers of the Western Union Telegraph Company for the New York-Azores permalloy loaded cable and has been in actual commercial operation for many months. Provision has been made in the design of this apparatus for the extension of the circuit to Emden, Germany, over the Azores-Emden cable of the Deutsch-Atlantische Telegraphengesellschaft, and automatic repeaters for the Azores station and terminal equipment for the Emden station have been installed and are now undergoing tests preliminary to the establishment of through-operation between New York and Emden.

Appendix I

Synchronous Vibrating Relay

There are several methods which may be employed to obtain the synchronous vibrating feature, one of which is shown schematically in Fig. 9. The relay is of the polarized type having two separate

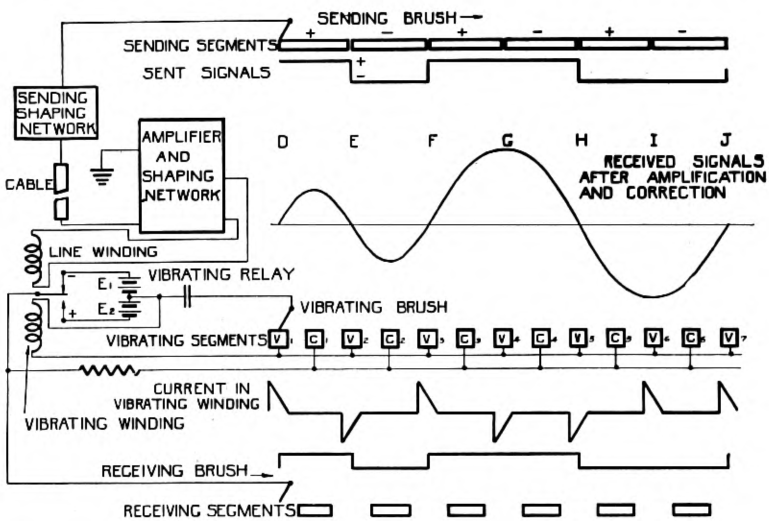


Fig. 9

windings, one of which, termed the line winding, is connected directly in the output circuit of the amplifier, and the other, or vibrating winding, is included in a circuit comprising the vibrating condenser, the vibrating segments V and the vibrating brush of the distributor. The distributor segments are shown developed for the sake of clearness. Disregarding the line winding for the moment, the passage of the vibrating brush over segment C_1, when the relay armature is resting against the negative contact, causes the vibrating condenser to be negatively charged by the battery E_1, and as the brush continues its rotation and passes upon segment V_1, the charged condenser is disconnected from the charging circuit and is connected to the vibrating winding through which it immediately discharges in the proper direction to cause the relay armature to be moved against its opposite or positive contact. This change in the position of the relay armature connects all of the "C" segments to positive battery, so that the passage of the brush over segments C_2 and V_2 in succession causes the condenser first to be positively charged, then to discharge through the vibrating winding which restores the relay armature to its former position against its negative contact. This cycle of operations will be repeated as long as the brush rotates and the rate of vibration can be made to coincide exactly with the frequency of the transmitted signals by suitably arranging the vibrating brush so as to be corrected from the incoming signals in a manner similar to that employed in the standard multiplex system.[11] The armature of the vibrating relay, in addition to controlling the polarity of the charge upon the vibrating condenser, controls the polarity of a battery applied to the receiving brush which distributes the received and corrected signalling impulses to the selector magnets of the receiving printer. In practice an intermediate relay, not shown in the figure, is employed between the armature of the vibrating relay and the receiving brush.

The line winding of the vibrating relay is connected in the amplifier circuit in the direction which will cause its armature to move toward its positive contact in response to incoming signalling impulses of positive polarity and vice versa, and as the amplitude of the current in that winding is adjusted to be approximately equal to that of the vibrating impulses, the effect of impressing the amplified and partially corrected positive impulse D upon the line winding at the time the vibrating brush is passing over segment V_1 would be only to aid the condenser discharge current in reversing the position of the relay armature. Every received impulse of unit length and the current during the first interval of unit length in every sustained pulse will produce the same result as shown at D, E, F and H, but the effect of current in the line winding due to the second and all succeeding time units of every sustained pulse will be neutralized, as shown at G, I

[11] J. H. Bell, loc. cit.

and J, by a pulse of approximately equal amplitude and opposite polarity in the vibrating winding and the position of the relay armature will therefore remain unchanged. Thus as the relay is actually operated by the energy supplied by the locally generated vibrating impulses and the received signalling current is employed only to neutralize the effect of the vibrating impulses during the second and succeeding time units of the sustained pulses, a considerable amount of distortion may be present in the signals without causing errors in reception, in fact the impulses of unit length may be entirely missing without interfering in any way with the accuracy of reception.

Ordinary polarized relays not provided with the vibrating feature are operated by the energy supplied by the amplified signals, which must consequently be quite free from distortion in order to insure faithful reproduction of the transmitted signal. The speed of transmission on long submarine cables employing non-vibrating relay reception cannot exceed that frequency at which the received signalling impulses of unit length are reduced by the attenuation of the cable to an amplitude which is only enough greater than the interference currents present to cause positive operation of the relay. The vibrating relay, however, does not require the pulses of unit length to be present at all, so that its limiting speed is that at which the impulses of two units length are received in just sufficiently large amplitude to prevent the relay from vibrating and allow a reasonable margin for overcoming the effects of any interference currents that may be present. This limiting speed is approximately double that obtainable with the use of non-vibrating relays.

The Application of Vacuum Tube Amplifiers to Submarine Telegraph Cables

By AUSTEN M. CURTIS

SYNOPSIS: Vacuum tube amplifiers have been developed for use in submarine telegraph reception and at present are in successful operation on four high speed permalloy cables. There is no limit to the speed at which vacuum tube amplifiers may be operated and in the present stage of development, the rate at which messages may be passed over loaded cables of the length used in the Atlantic Ocean is determined by the cable itself and the mechanical transmitting and receiving apparatus. In regard to maintenance, vacuum tube amplifiers have a great advantage in that they do not require any delicate mechanical adjustments.

THE laying of the new permalloy loaded cable between New York and Fayal (Azores) in September 1924 marked the most radical change in construction and operation of submarine cables that has taken place in many years. During 1926 three additional cables of this type were laid, the four sections being arranged to provide a line of communication between New York and England and another between New York and Germany. The traffic handling capacity of these cables when ultimately developed to its maximum by suitable terminal apparatus, will be nearly equal to that of all the other cables between North America and Europe combined.

The construction of these cables and the principles underlying their operation have already been described by O. E. Buckley[1] before the American Institute of Electrical Engineers in June 1925. The speed of operation of loaded cables of this type is many times that of the older non-loaded cables, and new apparatus has had to be developed to realize the full advantage offered by permalloy loading. One of the most important of these new developments has been the signal shaping vacuum tube amplifier, which is now in use on the four North Atlantic loaded cables.

The purpose of this paper is to point out the requirements which must be met by cable amplifiers, particularly those used on high speed loaded cables, and to describe how these requirements have been fulfilled in the present signal-shaping amplifier of the Western Electric Co. It will not be necessary to consider in any detail the general principles of operation of telegraph cables as these have been discussed with reference to non-loaded cables by Mr. J. W. Milnor[2] and are

[1] *Bell System Technical Journal*, Vol. 4, July 1925; *Journal of the American Institute of Electrical Engineers*, Vol. XLIV, No. 8, August 1925.
[2] "Submarine Cable Telegraphy," *Journal of the American Institute of Electrical Engineers*, Vol. 41, February 1922.

considered with particular reference to loaded cables in the paper by Mr. A. A. Clokey published in this issue of the *Bell System Technical Journal.*

The Necessity of Cable Signal Amplifiers and Their Requirements

Telegraph signals passing over a long submarine cable are distorted so severely that only a small fraction of the ultimate speed would be possible if extraordinary means were not taken to compensate for this distortion. It may be found that a certain cable attenuates very low frequencies to only one half of their original voltage while the higher frequencies may be received at less than one ten-thousandth of their initial strength. The transmission of an ideally perfect signal requires that its components of all frequencies be received at amplitudes proportional to those transmitted, consequently the reshaping of a signal received from a cable involves equalizing the strength of all its component frequencies by reducing the amplitude of the lower frequencies and amplifying the higher frequencies.

As the voltage which may be impressed upon a cable is limited by considerations of the safety of its insulation, the sensitivity of the receiver to currents of the highest frequency necessary in a properly defined signal is one of the limitations of the speed at which a cable may be operated. Unfortunately this is not the only limitation or it would be possible to increase the speed indefinitely by simply increasing the sensitivity of the receiver. All cables are exposed to extraneous interfering currents, some natural in origin and some the result of human activities, and while a great deal may be accomplished by proper design of the cable and the associated apparatus the speed is ultimately limited by interference. A large proportion of this interference is similar in nature to "static" and the bane of radio communication is also, but to a lesser degree, the bane of cable communication.

Experience has shown that continuous communication of the high standard of accuracy required in the transmission of code and cypher messages cannot be maintained unless the voltage received at the nominal signaling [3] frequency is between two and five millivolts. A receiver must therefore be capable of responding to voltages of this order at the signaling frequency in order to utilize the cable efficiently. On an average cable the power available at this voltage is of the order of 2×10^{-8} watts and there is at present no signal recording device

[3] This is defined in the case of the Morse cable code as the fundamental frequency in a series of alternate dots and dashes.

known to the art which will operate on so small a power except at uneconomically low speeds. For this reason it is necessary to insert between cable and recorder an instrument which will amplify the received signal.

A cable signal-shaping amplifier must fulfill many severe requirements. With its associated apparatus it must be capable of correcting the attenuation of the cable by equalizing the strength of all important component frequencies of the signal and it must also be capable of controlling in its output circuits a power many times as large as it receives.[4] It must be as insensitive as possible to interfering currents not included in the band of frequencies necessary to the signal and it is very desirable that overloading, which may be caused occasionally by these currents, should not permanently influence its adjustment or destroy any of its elements. The strength of its output current should be readily adjustable. It should be mechanically rugged, as otherwise its maintenance will require too large a proportion of the time of the staff at the cable station, and delays to traffic will be caused. Finally it should be protected as well as possible against local electrical fields and mechanical vibration and its operation should not be affected by conditions of extreme humidity.

Comparison of Mechanical Amplifiers and Vacuum Tube Amplifiers

In recent years several satisfactory mechanical amplifiers (called magnifiers in cable parlance) have been invented and their use has led to radical improvements in the speed of transmission over non-loaded cables. Most of these magnifiers utilize a sensitive moving-coil galvanometer, which moves some device a small distance in order to control a much greater power than that which caused the original motion of the coil. We may consider as typical of these the selenium magnifier which causes a beam of light to move over one or the other of two groups of selenium cells and thus varies their resistance, the Heurtley hot wire magnifier which changes the resistance of two pairs of almost microscopic heated wires by causing them to move relatively to each other, and the electrolytic magnifier which changes the resistances of a group of immersed electrodes. With all of these devices the controlled power is obtained from a local battery, but it is so small that it can do little more than operate a sensitive siphon recorder or a delicate moving coil relay. The latter may of course control a larger power which may in turn cause the operation of a comparatively rugged electromagnetic relay and thus indirectly a considerable power

[4] In practice the power amplification factor of the various types of amplifiers may range between five thousand and one hundred million.

may be controlled. With any of these magnifiers the suspended coil forms a mechanical oscillating system which is of great assistance in correcting the distortion of the cable, and allows signals to be shaped properly with the aid of a simple network of inductance, capacity and resistance. The inertias of the suspended coil and of the controlled devices make these magnifiers insensitive to high frequencies, and while this has some advantages in discriminating against high frequency disturbances, it also sets a rather definite limit to the speed at which they may be used. In order to utilize them as efficient signal shaping devices the natural frequency must be not far from one and one half times the nominal signaling frequency.[5] On this account and because their sensitivity decreases roughly as the square of the natural frequency to which they are adjusted, the moving coil magnifiers are rarely operated at signaling speeds of more than fifteen cycles per second. As they are easily damaged by relatively small overloads it is not safe to keep them in circuit when the approach of a thunder storm to a cable terminal makes the reception of induced surges in the cable likely. This sometimes results in keeping a cable out of operation for several hours, although the surges would only occasionally cause the loss of a letter if the magnifiers were not subject to damage by overloading.

A vacuum tube amplifier is free from many of the disadvantages of the mechanical amplifiers. It contains no delicate parts which require skilled manipulation, and once adjusted it maintains its adjustment indefinitely. There is no inherent limitation to the speed at which it may be operated; this being determined only by the requirement that the signal be sufficiently stronger than the interference. There is no practical limit to the amount of power which may be controlled and at the same time it is easy to limit this power and insure that momentary overloading shall not damage the amplifier or the associated apparatus. A multi-stage vacuum tube amplifier possesses still another important property in that there is practically no reaction between its various stages at telegraph frequencies. For this reason a number of interstage shaping networks may be used, and it will be found that the adjustment of one network is entirely without influence on the effects of the others.

History of Development of Vacuum Tube Amplifiers in Bell Telephone Laboratories

The signal shaping amplifier now in use is the outgrowth of studies of the applications of vacuum tubes begun in the laboratories of the

[5] Milnor, *A. I. E. E.*, February 1922.

American Telephone and Telegraph Company and the Western Electric Company in 1912. The vacuum tube amplifier appeared to offer important advantages for use on submarine cables because of its lack of distorting effects which are a function of the frequency of the current amplified, and also because of the ease with which signal distortion correcting circuits could be associated with the vacuum tubes. The initial studies on amplifier circuits suited to currents of the low frequencies involved in submarine cable telegraphy were made by Mr. R. V. L. Hartley and Mr. B. W. Kendall.[6] One of the difficulties which loomed quite large at that time was that most cables were operated duplex and the connection of an amplifier to a duplex circuit would involve the insertion of a transformer which promised to introduce distortion[7]. A suitable distortionless amplifier was first tried and subsequently distortion correcting networks were introduced between its stages. It was found that this permitted the use of more correcting elements than had been feasible in previous practice and thus indicated the possibility of attaining higher speeds than were usual at that time. The development of the shaping circuits employed was at first based on the principle of producing the various derivatives of the arriving current wave and adding them in proper phase relation to the arriving wave. This principle and methods of applying it had been developed mathematically by Mr. J. R. Carson of the American Telephone and Telegraph Company.[8] Mr. R. C. Mathes who conducted the experimental investigation beginning in 1916 simplified his work somewhat by recognizing that this principle was equivalent to a statement that the received signal would be satisfactory if the attenuation and phase distortion of the entire system of cable and amplifier for steady state alternating currents were corrected by the shaping networks over a range of frequencies from nearly zero to approximately the nominal signalling frequency. By the middle of August 1918 the employment of improved shaping methods made speeds of 22 cycles possible in simplex working on an artificial cable having a KR. of 2.7. The then standard cable apparatus would have permitted a speed of not more than 9 cycles, on a cable subject to interference of the magnitude usually encountered.

[6] B. W. Kendall, U. S. Patent No. 1,491,349, April 22, 1924.

[7] Expedients for avoiding distortion of this nature were suggested by Dr. H. W. Nichols of these Laboratories and by Mr. Lloyd Espenschied of the A. T. & T. Co. Their plans contemplated the modulation of an alternating current of relatively high frequency by the incoming signal, the amplification of the modulated current by suitable apparatus and its subsequent demodulation for obtaining the amplified low frequency signal (H. W. Nichols U. S. Patent No. 1,257,381, February 16, 1918; Lloyd Espenschied, U. S. Patent No. 1,428,156, September 5, 1922).

[8] See U. S. Patents No. 1,315,539, September 9, 1919, No. 1,450,969, April 10, 1923, No. 1,516,518, November 25, 1924 and No. 1,532,172, April 7, 1925. See also article by Dr. K. W. Wagner, *Electriche Nachrichten-Technik*, October 1924.

This apparatus [9] was then demonstrated to officials of the Western Union Telegraph Co. and with the cooperation of their engineers tests extending over a period of about a year were carried on at Rockaway Beach on several of the cables entering that station. It was shown in these experiments that while the vacuum tube amplifier together with suitable distortion-correcting networks would permit a considerable increase in the simplex speed (limited only by the interference present in the cable), the duplex speed was limited by the imperfect balance between the cable and the artificial line, and the increased sensitivity and signal shaping ability of the amplifier were of little value under the conditions then obtaining. Serious efforts were made to utilize the current limiting properties of vacuum tubes in conjunction with differentiating and integrating networks in reducing the effect of the unbalance on the signal and some successful results were attained.[10]

By 1920 the research leading to the development of the permalloy loaded cable had progressed to a point where it was evident that a new type of high speed cable amplifier would be required and the investigations were continued with this end in view. After the solution of numerous difficulties an amplifier was produced which was capable of correcting almost any variety of signal distortion which might be caused by a loaded cable. An amplifier of this type was tested on a trial length of 120 miles of loaded cable laid in a loop out of Devonshire Bay, Bermuda, and found to be generally satisfactory. The amplifier was then redesigned in a form suitable for commercial use and two amplifiers were built and installed at Rockaway Beach and Fayal in readiness for the New York-Azores loaded cable. They were put into successful operation and the predicted speed of 1,500 letters per minute was demonstrated within an hour after the cable had been released by the electricians of the ship which laid it.

Circuit Arrangements of Signal Shaping Vacuum Tube Amplifier

The electrical requirements of a cable signal shaping amplifier suitable for use on high speed loaded cables may be briefly stated as follows. It must take an input signal having components as low as one half millivolt and as high as possibly ten volts in amplitude, and correct the distortion by amplifying the weaker components much more than the stronger, at the same time making any necessary phase

[9] See U. S. Patents to R. C. Mathes No. 1,311,283, July 29, 1919, No. 1,426,755, August 22, 1922, No. 1,493,216, May 6, 1924 and No. 1,586,821, June 1, 1926; also Canadian Patent No. 207,231, January 4, 1921, granted to B. W. Kendall.

[10] D. K. Gannett and M. Kirkwood, U. S. Patent No. 1,483,172.

corrections. It must also be able to operate satisfactorily with signals in which the weaker components may be as strong as 100 millivolts. An output of about fifteen milliamperes at 15 volts should be available and this output must be adjustable by small steps. It must be capable of handling currents containing frequencies between a small fraction of a cycle and about 180 cycles, the particular part of this band of frequencies which is utilized depending on the nature of the cable and the speed at which it is operated.

These requirements are met in the present cable amplifier,[11] by circuits which are shown in the upper half of Figure 1. The amplifier proper consists of four stages of vacuum tubes, the first three being designed for high voltage-amplification and the last for large current output. An additional output stage is provided for the purpose of increasing the flexibility of the amplifier by permitting two separate classes of apparatus to be operated simultaneously. The coupling between stages is a combination of two types, the coupling for the very low frequencies being through a resistance capacity network while that for the higher frequencies is by means of an auto-transformer of special design or by highly damped resistance, inductance and capacity networks. The amplifier is connected to the cable through an input network and a shielded transformer. The input network assists in shaping the signal and prevents the first stage of the amplifier from being overloaded by the strong low frequency components of the signal arriving in the cable. The transformer permits earthing the filaments of the amplifier tubes and their associated batteries and avoids the short circuiting of the long balanced sea earth. The latter is used to reduce the effect of electrical disturbances on that part of the cable which lies in shallow water near the shore.[12] The requirements of this transformer are quite unusual as it must have a satisfactory voltage regulation from .2 to 200 cycles per second.

The ability to operate on voltages which may vary widely from time to time makes it necessary to provide a suitable range of adjustment of amplification. This is accomplished by providing that the secondary windings of the input transformer may be connected in series or in parallel and the plate coupling resistances of the tubes varied by a factor of four to one. A potentiometer placed between the second and third stages of the amplifier allows a variation of twenty to one in the voltage transmitted to the third stage, and with other adjustments as mentioned above, increases the total range

[11] A. M. Curtis, U. S. Patents Nos. 1,586,970 and 1,586,972, June 1, 1926, and 1,624,395 and 1,624,396, April 12, 1927.

[12] J. J. Gilbert, *Bell System Technical Journal*, July 1926; also British Patent No. 218,261, August 31, 1925, and Canadian Patent No. 265,944, Nov. 16, 1926.

of amplification adjustment to about 150 to 1. In addition a set of constant resistance potentiometers in the relay control panel associated with the amplifier allows the current through any of the relays to be varied in small steps without influencing the current in the other relays or changing the impedance of the amplifier output circuit.

The characteristic of the amplifier system may be measured by applying a certain input voltage at varied frequencies, and noting the corresponding output voltage. If the amplifier has been adjusted to give a satisfactory signal when connected to a cable, measurements will show that its amplification increases rapidly to a maximum which occurs at about 1.5 times the signaling frequency, and then falls to practically zero at about twice the signaling frequency. This elimination of the higher frequencies is effected by proper adjustment of the inter-stage shaping networks, and it results in suppressing that portion of the interfering currents received from the cable which lie above the band of frequencies required to form the signal.

The amplifier as described above is perfectly suitable for recorder operation and will permit communication at speeds up to at least ninety cycles per second which in cable code is equivalent to about 2,800 letters per minute, provided that a suitable recorder is employed. It is, however, not entirely suitable for multiplex printing telegraph operation under all conditions without the addition of apparatus to prevent "zero wander."

System for Correcting "Zero Wander"

In general, printing telegraph systems have been designed on the assumption that they were to work over land telegraph lines and they contain no provision for avoiding the effects of the "wandering zero" which is caused by the inability of a practical cable transmission system to transmit direct current. This inability to transmit direct currents is due to the fact that there is usually present in a submarine cable an earth current which is many times as strong as the signal, and it is necessary to block out this earth current by series condensers (as is usual in ordinary cable practice) or to keep it out of the amplifier by a transformer. The syphon recorder does not require that the zero of the signal be maintained closely but cable signal relays operate on a fixed value of current of either polarity and are incapable of determining whether or not part of this current is due to a displacement of the zero. It is therefore necessary to reconcile in some way the printing telegraph systems, which under some conditions require the reception of a direct current, with the cable system which cannot transmit a direct current. Several methods of doing this have been used with the mechanical amplifying systems on low speed cables; they usually supply directly to the relay a "zero correcting" current which depends upon the past history of the signal.[13]

When a vacuum tube amplifier is employed it is more convenient to apply the zero correction to the grid of the last stage vacuum tube as this results in the most economical utilization of the correcting battery and its circuits. The zero correcting apparatus is mounted in a cabinet adjacent to the amplifier and differs considerably in principle from that hitherto used with mechanical amplifiers.

The three element moving armature polarized relay, which had been designed for use in loaded cable operation generally, was changed in some details and adapted for use in the zero corrector. It is capable of operating at a high speed and also discriminates very accurately between currents of slightly differing values. When actuated by the normal signal its armature contacts vibrate between the fixed contacts, not touching either unless the zero of the signal deviates more than about three per cent from its proper position. When this deviation does occur the relay contacts close the circuit for an instant at the peak of a signal wave, and permit the battery to which they are connected to charge a condenser through a comparatively low resistance. The charge on this condenser then passes gradually to a second condenser through a high resistance and at the same time commences to be discharged from the second condenser by a shunt resistance. The voltage on this second condenser is applied to the grid of the last stage of the amplifier in such a sense that it produces a deflection of the amplifier zero in the direction opposite to the deflection which caused the relay contacts to come together. This correcting voltage is applied at a rate which is slow enough to prevent it from distorting the signal and the rate at which it is dissipated by the shunt if no further contacts take place is still slower. It should be noted that these rates of charge and discharge, while adjustable, need not bear any accurate relation to the shape of the signal itself. The correction is usually applied rapidly enough so that the zero is brought back to normal within the duration of about five signal pulses and the proper operation of the circuit prevents the zero from passing beyond limits about five per cent of the signal amplitude either side of the normal position. A somewhat simplified circuit of the relay control unit which includes the zero corrector is shown in the lower part of Fig. 1.

[13] A system in common use is due to S. G. Brown (British Patent No. 6,275, February 20, 1913). Other systems have been invented by D. K. Gannett and M. Kirkwood, U. S. Patent No. 1,548,878, Aug. 11, 1925, and R. C. Mathes, U. S. Patent No. 1,295,553, February 25, 1919.

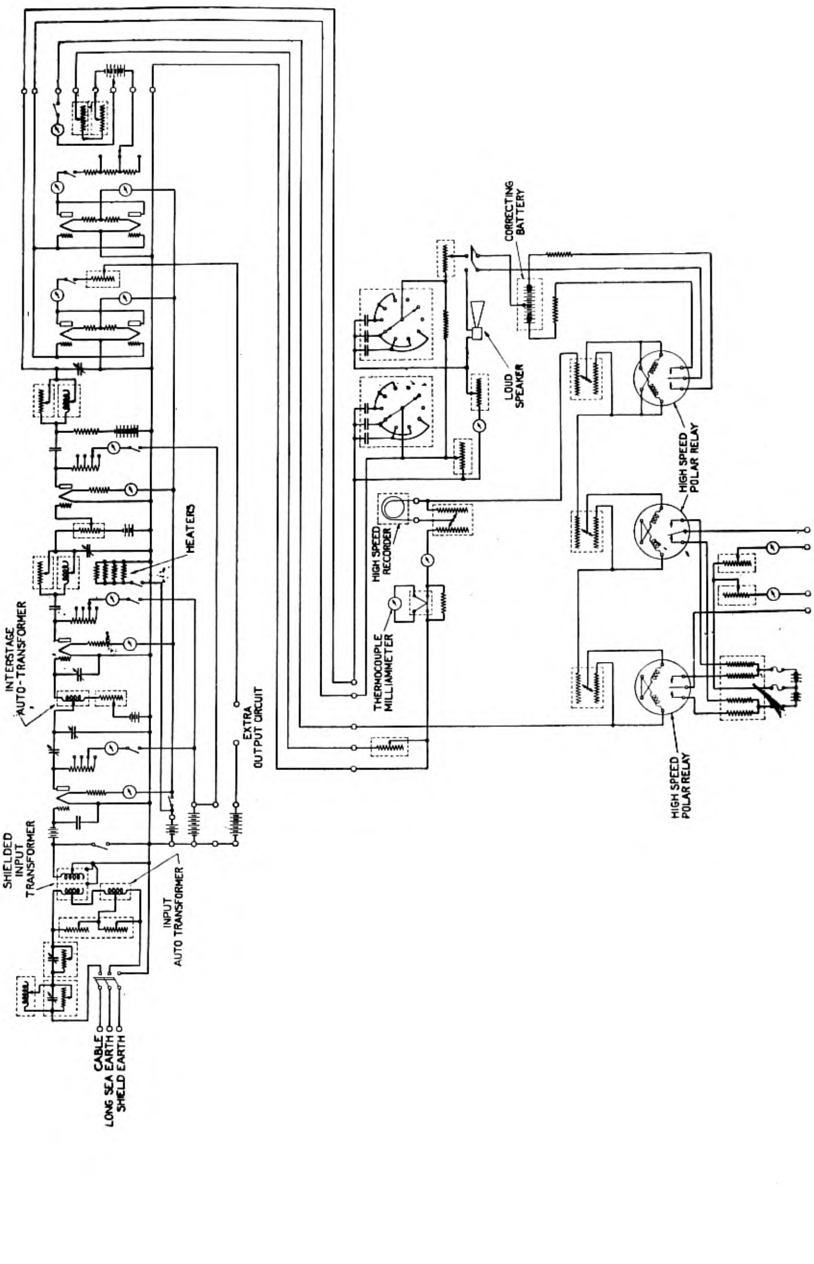

CABLE
LONG SEA EARTH
SHIELD EARTH
SHIELDED INPUT TRANSFORMER
INPUT AUTO TRANSFORMER
INTERSTAGE AUTO-TRANSFORMER
HEATERS
EXTRA OUTPUT CIRCUIT
THERMOCOUPLE MILLIAMMETER
HIGH SPEED RECORDER
LOUD SPEAKER
CORRECTING BATTERY
HIGH SPEED POLAR RELAY
HIGH SPEED POLAR RELAY
HIGH SPEED POLAR RELAY

FIG. 1.

Mechanical Details of Amplifier and Relay Control Desk

The construction of the amplifier and relay control desk is shown in the accompanying figures. Fig. 2 is a front and Fig. 3 a back view of the amplifier in its cabinet, Fig. 4 is a back view of the panel frame removed from the cabinet and Fig. 5 is a front view of the relay control panel associated with the amplifier. Mechanically the amplifier

Fig. 2.

consists of a frame of brass angles supporting on its front face four large hard rubber panels and sixteen small ones. The four large panels are mounted on the upper part of the frame, and hold the switches and the meters which it has been found desirable to provide in the filament and plate circuit of each vacuum tube. The adjustable elements of the amplifier interstage shaping networks are contained in the sixteen smaller panels mounted in the lower part of the framework. Each of these panels is a complete unit, comprising either a variable condenser, a variable resistance, or a switch for the adjustment of an inductance. As these panels are all of the same size, it is possible if necessary to change radically the arrangement of the interstage networks, or substitute entirely different ones without any difficult mechanical work on the amplifier. Each of the condensers and resistances is contained in an earthed metal box, into which it is sealed by a wax insulating compound. The frame in which the panels are

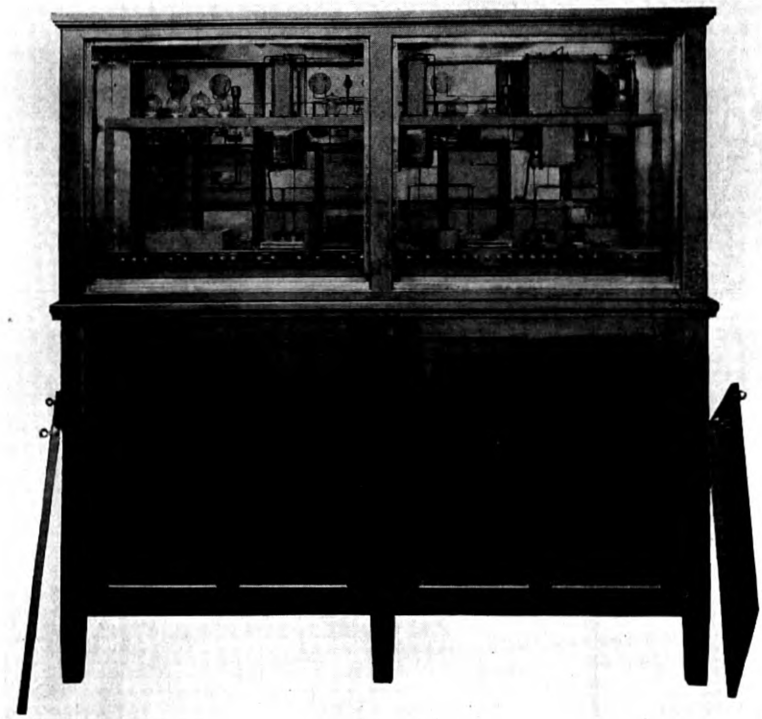


FIG. 3.

mounted is earthed and metallic fins are provided between the various panels in such a manner that any current which leaks through a film of moisture which might be deposited on the surface of the panels must pass to earth. The plan of mounting each individual piece of apparatus in a metal box, and sealing it in with insulating compound, has been adhered to throughout, the tubes, meters and switches of course being excepted. This is principally for protection from the serious humidity frequently found in cable stations. Additional protection is provided by electric heaters drawing current from the battery which operates the filaments of the vacuum tubes. Framework shelves provide space for mounting the tubes and heavy apparatus such as coils and coupling condensers. The tube of the first stage is held in a spring suspended socket, damped by an oil dashpot.[14] The socket of the second stage tube is sufficiently protected from vibration by a sponge rubber mounting, while the sockets of the third and fourth stages need no special protection. Dry cell grid batteries are mounted on the lower shelf of the framework. The numerous external connections are brought to the terminal strip which may be seen along the lower part of the back of the framework. The panel assembly is mounted in the upper part of a mahogany case, lined with copper. The lower part of the case contains the shaping

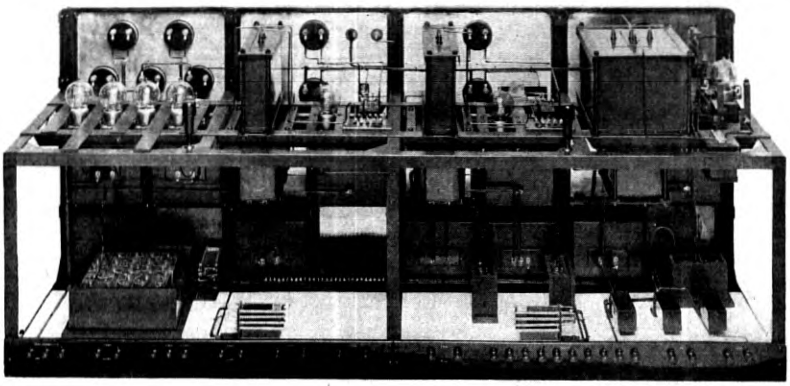


FIG. 4.

networks which are connected between the cable and the amplifier All adjustments are made either from the front of the panels or on apparatus contained in the lower part of the cabinet, and as all the elements of the amplifier are inherently stable, when the adjustments for shaping the signal at any given speed have once been determined they may be quickly duplicated at any time.

The relay control desk combines the apparatus for correcting the signal zero wander with means for adjusting the current through the relays used in the multiplex printing telegraph system and includes several switches used in the control of the latter apparatus.

In designing the amplifier the necessity of maintaining continuous operation and easily and quickly remedying any minor troubles was considered of the utmost importance, and this led to its being made large enough so that work may be done inside of it without having to remove

[14] W. A. Knoop, Patent U. S. No. 1,523,430, January 20, 1925.

it from the circuit and take it apart. All of the circuit elements may be reached from the back of the cabinet without disturbing anything else, and on several occasions this arrangement has proved of great value.

Power Supply for Amplifier

Three sets of storage batteries are required for an amplifier. The filaments are heated by a 6 V. 500 AH storage battery. The plate voltage for the first three stages is supplied by a 250 volt 1 AH storage

Fig. 5.

battery while the plate voltage for the last stage is supplied by a 275 V. 4 AH storage battery. These batteries are in the general battery room of each cable station and are handled by ordinary methods, the only special precautions necessary being the shielding of the leads from battery to amplifier and the avoiding of loose switch contacts.

Results Obtained in Service

The first two amplifiers built were put in operation on the New York-Azores cable in September 1924, and after a few weeks' testing a speed of 65 cycles per second or about 2,080 letters per minute in cable code was demonstrated. In the fall of 1926 three additional permalloy loaded cables were completed and equipped with vacuum tube amplifiers. They are laid between New York and Bay Roberts, Newfoundland, between Bay Roberts and Penzance, England, and between Fayal, Azores, and Emden, Germany. A speed of ninety cycles has been demonstrated on the New York-Bay Roberts cable, and the longer section from Bay Roberts to Penzance has worked at eighty cycles. The adjustment of amplification and the flexibility of the shaping networks is such that it has proved possible to remove an amplifier adjusted for operation at about 40 cycles from the long New York-Azores cable and readjust it for use on the short New York-Bay Roberts cable at 20 cycles in about fifteen minutes.

During the two and one half years of operation of amplifiers on the New York-Azores cable the maintenance required has been almost negligible and the rare cases of trouble have usually been found in the external connections. The longest delay to traffic caused by the amplifiers during this period was about two hours, and was due to the disarrangement of some temporary wiring. The reliability of the amplifiers is well attested by the fact that during the first two years there was only one available at each station, and there was no difficulty in keeping cable and amplifiers in continuous operation.

In connection with maintenance the vacuum tube amplifiers have a great advantage in that they do not require any delicate mechanical adjustments, while the electricians responsible for the operation of mechanical amplifiers must frequently spend hours at tasks requiring the skill and patience of a watchmaker.

It has been found possible to handle messages during thunderstorms which prevented operation of the non-loaded cables and their mechanical amplifiers for several hours. As an experiment the loaded cable and amplifier were worked continuously through an unusually severe thunderstorm during which stop watch observations of the intervals between lightning flashes and thunder claps showed that lightning had struck within a thousand feet of the cable terminal on three occasions. Although the induced surges were frequently several times as strong as the signals, the automatically limited output of the amplifier protected the recorders from damage, and the effect of each lightning discharge was limited to the possible mutilation of one or two letters.

The protection of these amplifiers from mechanical vibration has proved entirely satisfactory. During alterations to the Western Union Cable station building at Rockaway Beach a brick wall six feet from the amplifier was broken down with sledge hammers without interfering with the normal handling of messages.

Other Applications of Vacuum Tube Amplifiers in Cable Telegraphy

While as yet vacuum tube amplifiers have been utilized principally on high speed loaded cables they are not necessarily restricted to such use. It was mentioned in an early part of this paper that, since the non-loaded cables are ordinarily operated duplex at a speed which is set, not by the sensitivity of the receiver, but by the strength of the interference due to imperfect balance between cables and artificial line, no increase in speed might be expected to result from the substitution of vacuum tube amplifiers for the mechanical amplifiers now used. Nevertheless the superior ruggedness of the vacuum tube amplifier, combined with its ability to operate safely through thunderstorms which would ruin the mechanical amplifiers, might reduce appreciably the amount of lost time, particularly during the summer months, and thus improve the traffic capacity of these cables.

In addition to the use of vacuum tube amplifiers for operating terminal apparatus they have another important field as repeaters intermediate between two short sections of a long cable. As the speed at which any cable can be operated is roughly inversely proportional to the square of its length, it is customary to lay cables connecting distant centers of population in two or more sections, interrupted at some conveniently located but often inconveniently isolated island. This involves repeating the signals received from one section of the cable into the next section and for years this was done manually, that is, an operator received and translated the signals and passed them on to another operator for transmission on the other section. Within recent years through relay operation by means of repeaters has become general. At the intermediate station a device receives the signal from an amplifier and retransmits it to the next section usually correcting it completely to its original form. These repeaters, while very successful, are still quite complicated mechanically, and require skillful maintenance, particularly as they utilize delicate mechanical amplifiers and moving coil relays. It is possible to replace them by vacuum tube amplifiers having no moving parts, and thus requiring a minimum of maintenance. The vacuum tube amplifier is capable of reshaping the signal almost as completely as is done by the mechanical repeaters, and, in case of a cable worked simplex, the direction in which the amplifier at the intermediate station repeats the signal may be automatically controlled from either the sending or the receiving station.

Modulation in Vacuum Tubes Used as Amplifiers

By EUGENE PETERSON and HERBERT P. EVANS

SYNOPSIS: Recent developments in amplifier design tending toward more rigorous quality requirements have shown that the solutions of Van der Bijl and Carson are inadequate for certain purposes since they are based upon a convenient assumption which is not satisfied in fact. In particular, a detailed investigation of carrier current repeaters used for the simultaneous transmission of several channels, and upon which in consequence the modulation or crosstalk requirements are particularly severe, showed the modulation currents measured to be quite different from those specified by the theory, as was the law of variation of these currents with the circuit constants.

The cause of the discrepancy was found to reside in the neglect of the variation of the amplification factor (μ) with both plate and grid potentials. When the actual state of affairs was taken into account in the analysis by the application of a general method involving no assumptions, theory and experiment were found to be in good accord. The new expressions have been developed in terms of the amplification factor (μ), the internal output resistance of the tube (R_0), and their differential parameters, which are involved in the representation of the characteristic tube equation by a double power series. Expressions for the current components are developed in terms of the coefficients of the series, and modifications of Miller's method for greater convenience and precision in determinations of tube characteristics are described from which the series coefficients may be evaluated.

Conclusions are drawn from the solutions as to desirable tube characteristics by which, for example, a single tube may take the place of two tubes in push-pull connection. Finally, certain properties of different types of tubes under conditions of maximum output power are compared on the basis of μ constant and μ variable.

THE amount of modulation produced in vacuum tube amplifiers is in many cases a controlling factor in their application and it becomes of importance to determine how modulation products arise, so that the possibility of reducing them by tube and by circuit design may be studied.

In restricting discussion to amplifiers, and particularly to those used in communications, we are treating cases most amenable to analysis; in which, normally, the applied potential variation maintains the grid always negative so that conductive grid current does not flow, but in which, on the other hand, the greatest negative potential does not exceed the negative end of the plate current grid potential characteristic. In applying these two detailed restrictions we are incidentally insuring against prohibitive quantities of modulation; we know that, for example, the flow of conductive grid current may, under special conditions, produce an exceptionally efficient modulator of great service as such, but highly undesirable as an amplifier.

The necessity for suppressing modulation proceeds from the disturbing effects attendant on it, by which there may result reduction of quality in speech amplifiers, and crosstalk in the multi-channel amplifiers of carrier telephony, to take but two examples. The modulation level in the last case is restricted to much smaller values than are tolerated in the first; it is commonly required to reduce modulation products to the thousandth part, or even less, of the fundamentals which produce them. This last case is the one in which we are primarily interested; other cases of greater distortion referred to above may be treated by an extension of the methods used below in the case of grid current flow, and by Fourier series or expansions in terms of Bessel functions when the negative end of the tube characteristic is exceeded.

A thoroughgoing study of the amplifier problem would relate the static characteristics of a tube and the parameters of the circuit in which it works to its operating characteristics, and then would relate its static characteristics to the internal structure of the tube; it would in brief enable us to link the details of tube structure to the fundamental and modulation currents produced in the output wave of the amplifier. In the following, however, we shall treat only that part of the general problem which relates the operating characteristics and circuit parameters to the static characteristics.

A consideration of the usual plate current characteristics of a three electrode vacuum tube, as shown in Fig. 1, demonstrates the well-

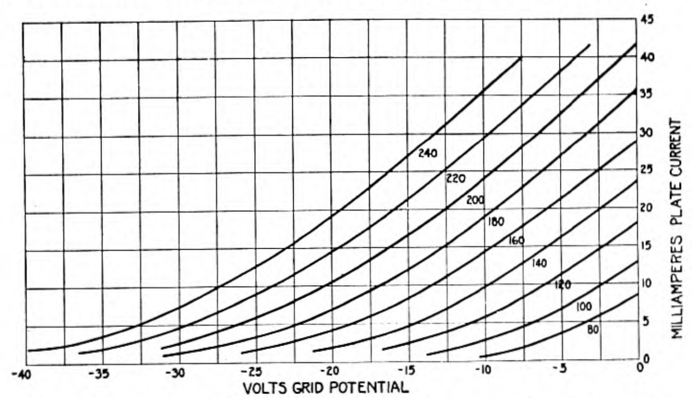


Fig. 1—Plate current as a function of grid potential with plate potential as a parameter. EL tube No. 109,150. $I_f = 1.1$ amperes

known dependence of the plate current upon the two variables, the grid and plate potentials. That is to say, the plate current varies with the grid potential when the plate potential is fixed, and it varies with the plate potential when the grid potential is fixed. It has been found of great convenience in the past to utilize an approximate relation between the grid and plate potentials as expressed in what is sometimes described as the fundamental theorem of the vacuum tube. The theorem states that a potential change in the grid circuit appears as a voltage generated in the plate circuit, the magnitude of which is equal to the grid potential change multiplied by the amplification factor μ. Solutions have been obtained for the output current components with the aid of this relation through the work of Van der Bijl and of Carson, which have been of great practical importance.

These solutions are approximate because of the simplifying assumption of the constancy of the amplification factor, which is certainly not accurate as the curves of Fig. 2 demonstrate. In this diagram

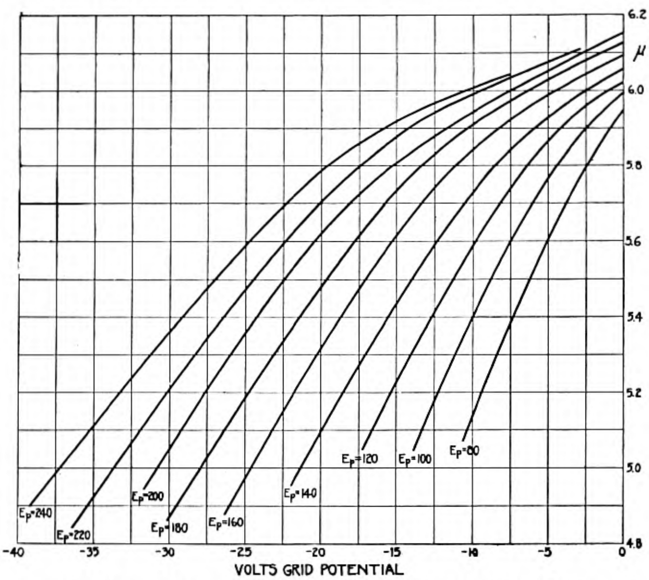


Fig. 2—μ as a function of the grid potential with plate potential as a parameter. EL tube No. 109,150. $I_f = 1.1$ amperes

the amplification factor for small applied potentials is plotted as ordinate with the grid potential as abscissa and the plate potential as parameter. The variation of μ is observed to be of the order of twenty per cent over the operating range. When we are interested in the distortion of the input wave, this variation cannot be ignored since to do so would in some cases yield results of another order of magnitude than those found experimentally. The treatment for our specific needs must therefore be modified to take account of the actual state of affairs.[1]

There are two ways open for a treatment involving the variation

[1] Early calculations to show the effect of a variable μ upon distortion were given by H. Nyquist of the American Telephone and Telegraph Company in an unpublished memorandum of April, 1921.

of μ; one is a modification of Carson's analysis while the other involves a reconsideration of the tube characteristic equation. The modification of Carson's treatment may be carried out with the aid of an expedient as follows: with a single input frequency the wave form of the generator voltage acting in the plate circuit is distorted by the variable amplification factor of the tube, and it is this distorted wave which acts in the plate circuit, instead of the pure sine wave operating with constant μ. The method of procedure is then evident; if we refer the actual distorted generator potential to the grid circuit—that is, if we divide it by the average value of μ—we have an input wave which, when treated by Carson's well-known method [2] in which μ is assumed constant, will yield correct results inasmuch as the μ variation has been taken into account—somewhat indirectly, it is true. When complex waves are applied to the grid circuit the effective grid potential is made up of numerous components and the treatment becomes very cumbersome.

Tube Characteristic Expressed by Double Power Series

A more direct method which has been used with some success consists in expressing the tube current-voltage relation by a double power series, without invoking any special relations regarding the connection between a grid potential change and the equivalent plate potential. If we use the symbols I_b, E_p, E_c to denote the plate current, plate potential, and grid potential, respectively, we may express the plate current as a double power series as follows:

$$\begin{aligned} I_b = f(E_p, E_c) &= \Sigma a_{mn} E_p{}^m E_c{}^n \\ &= a_{00} + a_{10}E_p + a_{01}E_c \\ &\quad + a_{20}E_p{}^2 + a_{11}E_pE_c + a_{02}E_c{}^2 \\ &\quad + a_{30}E_p{}^3 + a_{21}E_p{}^2E_c + a_{12}E_pE_c{}^2 + a_{03}E_c{}^3 \\ &\quad + \cdots \end{aligned} \tag{1}$$

where

$$a_{mn} = \frac{1}{m!\,n!}\,\frac{\partial^{m+n} f(0,\,0)}{\partial E_p{}^m \partial E_c{}^n}\,, \tag{2}$$

and in which it is understood that the development applies with the operating point on the characteristic. The derivatives, it will be noted, are evaluated at the point at which both E_p and E_c are zero. Some of the coefficients of Eq. (1) may of course be eliminated by reference to the evident properties of the tubes, but this need not concern us here since it is more convenient to formulate the tube equation in another way.

[2] *Proc. I. R. E.*, 1919.

Under normal conditions of amplifier operation E_p and E_c are fixed, and the alternating grid and plate potentials vary about these potentials. It then becomes convenient to express the coefficients as derivatives referred to the specific point E_{p_0}, E_{c_0}. If we indicate the variable components of the grid and plate potentials by e and v, respectively, the tube equation may be put in the form

$$\begin{aligned} I_b &= f(E_{p_0} + v,\ E_{c_0} + e) = \Sigma b_{mn} v^m e^n \\ &= f(E_{p_0}, E_{c_0}) + b_{10}v + b_{01}e \\ &\qquad + b_{20}v^2 + b_{11}ve + b_{02}e^2 \\ &\qquad + b_{30}v^2 + b_{21}v^2e + b_{12}ve^2 + b_{03}e^3 \cdots, \end{aligned} \tag{3}$$

where

$$b_{mn} = \frac{1}{m!\,n!}\,\frac{\partial^{m+n} f(E_{p_0}, E_{c_0})}{\partial^m E_p \partial^n E_c}.$$

The b coefficients are functions of the operating point—specified by E_{p_0}, E_{c_0}—and change with the operating point, in general. The a coefficients are definitely referred to the origin, however, so that each b coefficient may be expressed in terms of the a coefficients which correspond to it. It is possible, by determining this relation, to follow the variation of the b's in terms of E_{p_0} and E_{c_0}. To do this we substitute for E_c and E_p in Eq. (1) the expressions $E_{c_0} + e$, $E_{p_0} + e$, respectively, find the constant term, and obtain succeeding coefficients by differentiation. Thus

$$\begin{aligned} b_{00} &= a_{20}E_{p_0}^2 + a_{30}E_{p_0}^3 + a_{40}E_{p_0}^4 \\ &\quad + a_{21}E_{p_0}^2E_{c_0} + a_{31}E_{p_0}^3E_{c_0} + a_{22}E_{p_0}^2E_{c_0}^2, \end{aligned}$$

and further

$$b_{10} = \frac{\partial b_{00}}{\partial E_{p_0}}, \qquad b_{01} = \frac{\partial b_{00}}{\partial E_{c_0}}.$$

This concludes our consideration of the tube characteristics without reference to the circuit to which the tube may be connected. Eq. (3) rather than Eq. (1) will be used in the following.

It should be noted in terminating this part of the discussion that the treatment is capable of easy extension to characteristics depending upon a larger number of variables. Thus a four element (double grid) tube characteristic may be expressed by a triple power series, and so on. When the potential of the second grid is maintained constant it is evident that the tube characteristic is given by a double power series in which the coefficients depend in addition upon the potential of the second grid. To determine the dependence quantitatively, the triple series will serve.

Solutions for the Plate Circuit Components

We now pass on to a consideration of the operation of the tube working into a plate resistance. The more general case of a load impedance which is a function of frequency may be treated by application of the equations derived above, but it will serve our purpose here to deal with the case of a pure resistance load since the experimental work was done for that particular case which is of considerable practical importance.

If J is the alternating component of the plate current, we have from (3)

$$J = I_b - f(E_{p_0}, E_{c_0})$$

and J, it is seen, is a function of the two variables v and e. The quantity v depends on e of course, so that J may evidently be expressed as a function of e alone or

$$J = \sum_{k=1}^{k=\infty} C_k e^k. \tag{4}$$

A solution of the problem therefore consists in determining the C's in terms of the circuit and tube parameters.

The change in plate potential v may further be expressed as

$$v = -RJ = -R\sum_{k=1}^{k=\infty} C_k e^k. \tag{5}$$

The C's are then determined by putting (4) and (5) in (3) and identifying coefficients of similar powers of the variable. We have then

$$\begin{aligned} v &= -RC_1e - RC_2e^2 - RC_3e^3 \cdots, \\ v^2 &= R^2C_1^2e^2 + 2R^2C_1C_2e^3 \cdots, \\ v^3 &= -R^3C_1^3e^3 \cdots, \end{aligned}$$

in which powers higher than the third are neglected for this, the first approximation. Carrying through the substitutions we obtain the solutions

$$\begin{aligned} C_1 &= b_{01}/(1 + b_{10}R), \\ C_2 &= (b_{02} + b_{20}R^2C_1^2 - b_{11}RC_1)/(1 + b_{10}R), \\ C_3 &= \frac{b_{03} - RC_1b_{12} + R^2C_1^2b_{21} - R^3C_1^3b_{30} - RC_2b_{11} + 2R^2C_1C_2b_{20}}{1 + b_{10}R}. \end{aligned} \tag{6}$$

The first equation, which leads to the first approximation to the fundamental current, is identical with that obtained on the basis of μ constant, but the higher orders are distinctly changed. When e is a pure sine wave C_2 contributes to the constant term which represents the change in direct current, C_1 and C_3 contribute to the fundamental component of the plate current, C_2 gives rise to the second harmonic and C_3 gives rise to the third harmonic current. When e is a complex wave the subscripts indicate the order of modulation [3] to which each coefficient applies.

The b coefficients may be readily converted into quantities dependent on μ, R_0, and their derivatives; we have

$$\mu = \frac{\partial I_b/\partial E_{c_0}}{\partial I_b/\partial E_{p_0}} = \frac{b_{01}}{b_{10}}$$

and

$$1/R_0 = \partial I_b/\partial E_{p_0} = b_{10},$$

so that

$$\mu/R_0 = b_{01}.$$

Succeeding coefficients are obtained by differentiating with respect to E_p and to E_c. For example,

$$b_{20} = -\frac{1}{2R_0^2}\frac{\partial R_0}{\partial E_{p_0}},$$

$$b_{11} = \frac{1}{R_0}\frac{\partial \mu}{\partial E_{p_0}} - \frac{\mu}{R_0^2}\frac{\partial R_0}{\partial E_{p_0}} = -\frac{1}{R_0^2}\frac{\partial R_0}{\partial E_{c_0}},$$

$$b_{02} = \frac{1}{2R_0}\frac{\partial \mu}{\partial E_{c_0}} + \frac{\mu}{2R_0}\frac{\partial \mu}{\partial E_{p_0}} - \frac{\mu^2}{2R_0^2}\frac{\partial R_0}{\partial E_{p_0}}.$$

The b coefficients may be obtained directly from the family of characteristic curves either graphically or analytically, when the operating point is specified. If we obtain our coefficients from the μ and R_0 curves, however, derivatives of a lower order are required than we need in dealing directly with the static characteristics. A family of μ-curves is shown in Fig. 2, and a family of R_0 curves is shown in Fig. 2*a*.

Results applying to the four element tube which are obtained by methods analogous to the above may be stated briefly. If we express the change in plate current by

$$J = C_{10}\epsilon + C_{01}e + C_{20}\epsilon^2 + C_{11}\epsilon e + C_{02}e^2 \cdots,$$

where ϵ and e represent the alternating potentials on the two grids,

[3] The new frequencies produced by modulation are given by the expression

$$F = |mf_1 \pm nf_2 \pm \cdots|,$$

where f_1, f_2 are impressed frequencies and m, n are integers or zero; the order is simply the sum of m, n, $\cdots$.

we find

$$C_{10} = b_{010}/(1 + Rb_{100}),$$

$$C_{01} = b_{001}/(1 + Rb_{100}),$$

$$C_{11} = \frac{b_{011} + 2R^2C_{10}C_{01}b_{200} - RC_{01}b_{110} - RC_{10}b_{101}}{1 + Rb_{100}},$$

in which

$$b_{rst} = \frac{1}{r!s!t!}\frac{\partial^{r+s+t}f(E_{p_0}, E_{n_0}, E_{c_0})}{\partial E_p{}^r \partial E_n{}^s \partial E_c{}^t}$$

and E_{n_0} is the fixed potential of the second grid.

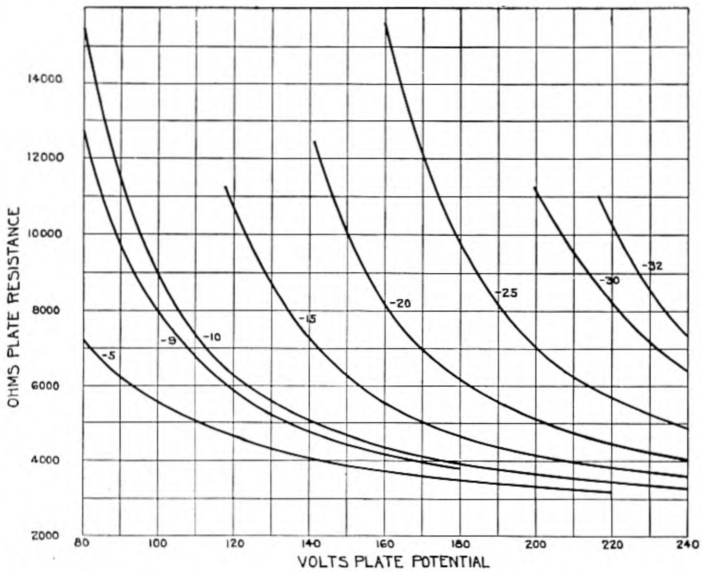


Fig. 2*a*—Plate resistance as a function of plate potential with grid potential as a parameter. EL tube No. 109,150. $I_f = 1.1$ amperes

We proceed to the methods used for the experimental determination of μ and R_0 for the three element tube.

Measurement of Tube Parameters

The amplification constant and plate impedance of a three element tube may be measured with precision by a well-known method due in principle to J. M. Miller [4] which requires no explanation here. The method as originally proposed is somewhat inconvenient in that the space current of the tube under test passes through a resistance common to the alternating measuring current, so that the operating

[4] *Proc. I. R. E.*, Vol. 6.

point of the tube is changed during manipulation for balance. Everitt's modification,[5] which consists in separating the direct and alternating current paths by a retard coil and condenser in the usual manner, is therefore preferable in this respect, but a complicating factor enters in the introduction of a reactive component due to the retard coil which cannot be balanced out by the variable resistances originally provided. This may be taken care of in a more or less obvious way by shunting a reactance around the grid resistance as shown in Fig. 3, the effect of which is to correct for the introduced

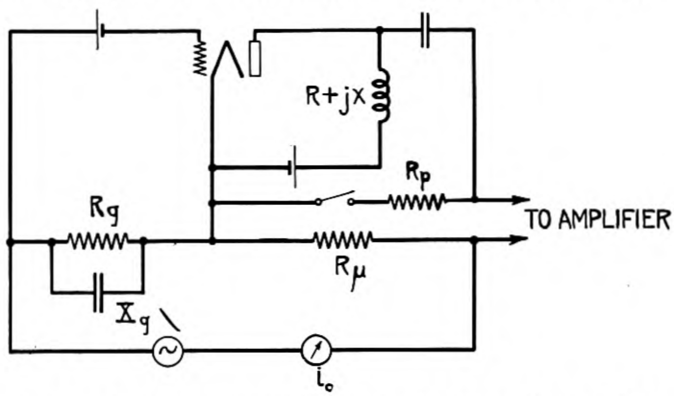


Fig. 3—Modification of Miller's method for determining μ and R_0

phase unbalance due to the retard coil and so lead to a precisely determinable null point instead of to a broad minimum, as is otherwise found.

The effect of the inserted reactance may be calculated by direct methods. Referring to the figure, there is no potential of fundamental frequency across the amplifier used to indicate balance at the null point, and if the grid-filament impedance of the tube is much greater than R_g (50 ohms), the total oscillator current passes through R_g and X_g in parallel, with R_μ in series. The potential impressed on the grid is then

$$e_g = jR_gX_gi_0/(R_g + jX_g),$$

which appears in the plate multiplied by the amplification factor of the tube and reversed in sign. The nomenclature is clearly indicated in the Figure. An alternating current flows in the plate circuit which is just balanced by the drop across R_μ so that we may write

$$\frac{j\mu R_gX_g}{R_g + jX_g} \cdot \frac{R + jX}{R_0 + R + jX} = R_\mu,$$

[5] See p. 201 of van der Bijl's "Thermionic Vacuum Tube."

which yields, after some reduction,

$$\mu = \frac{R_\mu}{R_g}\left(1 + \frac{R_0 R + R_0^2}{X^2 + R_0 R + R^2}\right).$$

As to the order of magnitude of the various quantities involved, R^2 is usually negligible before X^2, while R_0 and R may be of the same order of magnitude, so that we have

$$\mu = \frac{R_\mu}{R_g}\left(1 + \frac{2R^2}{X^2}\right).$$

In the specific case of a 101-D tube we had $X = 2.1 \times 10^5$, $R = 7 \times 10^3$ and

$$\mu = \frac{R_\mu}{R_g}(1 + 0.002).$$

The correction term, amounting to two parts in a thousand, drops out without the retard coil and we arrive at Miller's formula $\mu = R_\mu/R_g$. In measuring the output impedance of the tube after the settings for μ have been determined, R_g is doubled and R_p is connected in the plate circuit and varied until balance is again attained. It has been shown [6] by extension of the method used above that

$$R_0 = R_p\left[1 + \frac{3}{2}\left(\frac{R_p}{X}\right)^2\right]$$

and the correction term is of the same order of magnitude as that previously found for the amplification factor.

Balances may be obtained precise to one part in a thousand or better, but in much of our own work the observations are not ordinarily corrected for finite reactance. In order for the balancing action to take place the two reactances must be of opposite sign since amplification produces a 180° phase shift. If we balanced by a reactance shunted around R_μ instead of around R_g, the inserted reactance would, of course, be of the same sign as that of the plate retard coil, which was inductive at the frequency of 1,000 cycles at which the balances were made. The alternative scheme of shunting a variable condenser around R_g was adopted purely as a matter of convenience.

Applications of the Analysis

Second Order Modulation in Voltage Amplifiers

A striking illustration of the difference in the results of the two analyses, one based on the assumption of constant amplification

[6] By Mr. V. A. Schlenker.

factor and the other based on actual tube characteristics, is provided by a consideration of the second order modulation in the case of large external plate resistance.

The ratio of second harmonic to the fundamental, when μ is assumed invariable, comes out proportional to

$$(R + R_0)^{-2},$$

which shows that the ratio tends toward zero as R is made indefinitely great, a condition approximated in voltage amplifiers. According to this expression, the distortion would be eliminated by increasing the external plate resistance. That this is not really so is demonstrated by the analysis above which gives for the same ratio

$$\lim_{R \to \infty} \left(\frac{C_2}{C_1} \right) = (b_{02} + \mu^2 b_{20} - \mu b_{11})/b_{01}.$$

The ratio therefore approaches a constant value different from zero as R is indefinitely increased. The second harmonic level referred to the fundamental is about 40 T.U. down with a 101-D tube, which is prohibitively large distortion for certain classes of work such as multi-channel amplification used in carrier telephony; for a 104-D tube the level is about 32 T.U. down on the fundamental.

In order to bring out some important points involved in the theory, we shall discuss them in connection with experimental data on a standard type of tube (101-D) which are due to Mr. A. G. Landeen. The method used in measuring the current components is described in his paper on current analysis in the *Bell System Technical Journal* for April 1927.

Output Currents of a Representative Tube

Fig. 4 shows the calculated effect of varying the plate resistance on the fundamental, second, and third harmonic currents produced by a representative 101-D tube, which are indicated by circles, triangles, and crosses, respectively. In this drawing the values of the coefficients as calculated by Mr. J. G. Kreer are plotted as ordinates, and the external plate resistances are plotted as abscissæ. The agreement with the values obtained from experiment, and shown by the full lines, is seen to be rather close and within the limits of accuracy of the measurements except perhaps for the third harmonic at high load resistances. The coefficients used in the calculation of the quantities C_1, C_2, C_3 were obtained by graphical methods, which consisted in determining tangents to curves derived from μ and R_0 measurements. The precision obtained is sufficient for our present purposes, but for greater precision it may be desirable to use analytical methods for the determination of the b coefficients.

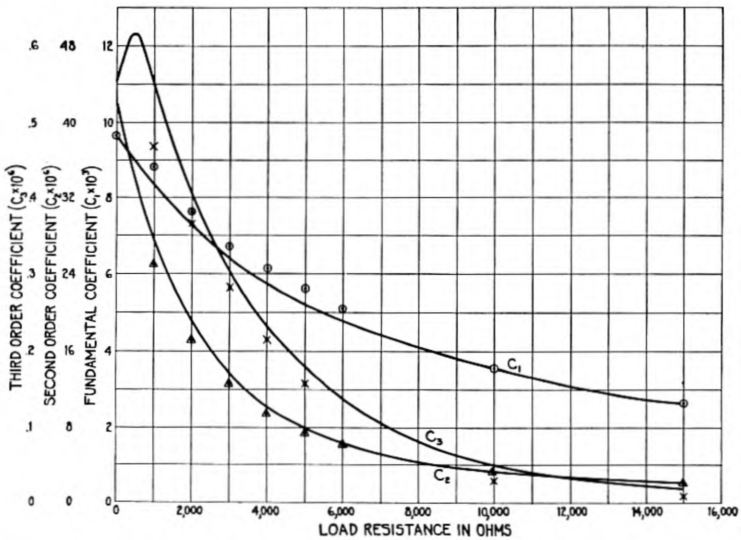


Fig. 4—Modulation coefficients for EL tube No. 109,150. $E_c = -9$, $E_p = 120$

Consideration of the expression for the second order coefficient, C_2, shows that the three terms of the numerator are all important, in general, except that the last term is negligible at very low resistances,

$$C_2 = \left[\frac{\partial \mu}{\partial E_{c_0}} - C_1(R - R_0) \frac{\partial \mu}{\partial E_{p_0}} - R_0 C_1^2 \frac{\partial R_0}{\partial E_{p_0}} \right] \frac{C_1}{2\mu}. \qquad (7)$$

Now in amplifiers, the condition for maximum power delivered to the load resistance at maximum gain demands equality of internal and external resistances, and this coincides with the requirement for minimum reflection coefficient [7] when the amplifier is connected to a line of definite characteristic impedance.

Under normal conditions, then, we have for the second order coefficient

$$C_2 = \left[\frac{\partial \mu}{\partial E_{c_0}} - \frac{\mu^2}{4R_0} \frac{\partial R_0}{\partial E_{p_0}} \right] \frac{1}{4R_0}, \qquad (8)$$

in which the variation of μ with respect to E_p does not enter, the only determining quantities being the variation of μ with respect to E_{c_0}, and the variation of R_0 with respect to E_{p_0}. The second order modula-

[7] The reflection coefficient is expressed as the quotient of the difference by the sum of the two connected impedances.

tion vanishes when

$$\frac{\partial\mu}{\partial E_{c_0}} \bigg/ \frac{\partial R_0}{\partial E_{p_0}} = \mu^2/4R_0, \tag{9}$$

which is a valuable property of amplifier tubes when the equality can be secured.

A more general relation for which the second order vanishes occurs when we set Eq. (7) equal to zero. As a matter of experience these conditions are not satisfied with the usual type of tube; they are found to hold in tubes of rather special construction. The null points are, of course, independent of the character of the applied grid potential, provided that the restrictions on the original development for the tube characteristic are not exceeded and that contributions of higher to lower order terms are negligible.

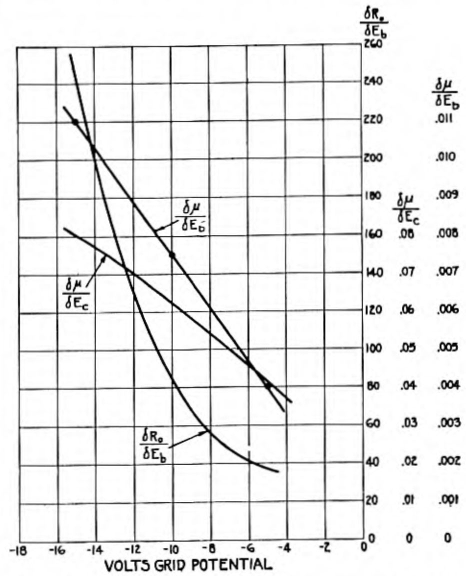


Fig. 5—Variation of tube parameters with grid potential. EL tube No. 109,150. $E_p = 120$

The expression for the third order coefficient contains six terms in the numerator, three of which are of opposite sign. If we consider the contribution of each of these terms as a function of the external plate resistance, we find that at very low resistances the single term b_{03} is predominant, while for resistances comparable to that of the internal plate resistance of the tube itself, no one of the six terms, of which three are negative and three are positive, may be neglected.

As a consequence of the subtraction of quantities of the same order of magnitude, the calculation for the third order coefficient is not capable of any great precision.

Fig. 5 shows how the fundamental coefficients $\partial\mu/\partial E_p$, $\partial\mu/\partial E_c$, $\partial R_0/\partial E_p$, which are involved in the second order term, vary as a function of the grid potential when the plate potential is maintained constant at 120 volts; $\partial R_0/\partial E_p$ is negative in sign.

Variation of C_1 and C_2 with Grid and Plate Potentials

To summarize our analysis up to this point, we have formulated an expression for the characteristic surface of a vacuum tube and have manipulated it to derive expressions for the fundamental and for the second and third order current coefficients. These theoretical

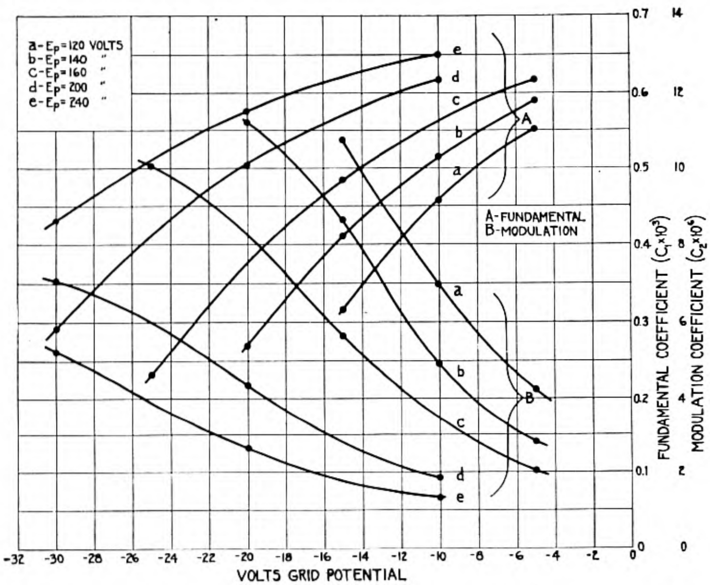


Fig. 6—EL tube No. 109,150. External resistance = 6,000 ohms. I_f = 1.1 amperes

relations have been compared with experimental determinations of the three quantities involved as a function of the external plate resistance, and a sufficiently good agreement has been obtained to indicate that the processes which we have treated are sufficient to account for experimental observations. We now present calculations of the coefficients C_1 and C_2 as a function of plate potentials and grid potentials for several values of the external plate resistance. It is seen from Figs. 6, 7, and 8 that these coefficients vary inversely with the plate and grid potentials.

As the plate resistance is increased all coefficients decrease, but C_2 decreases more rapidly than does C_1. The question then arises as to the conditions which would lead to the smallest amount of distortion while maintaining a definite fundamental power output at a definite

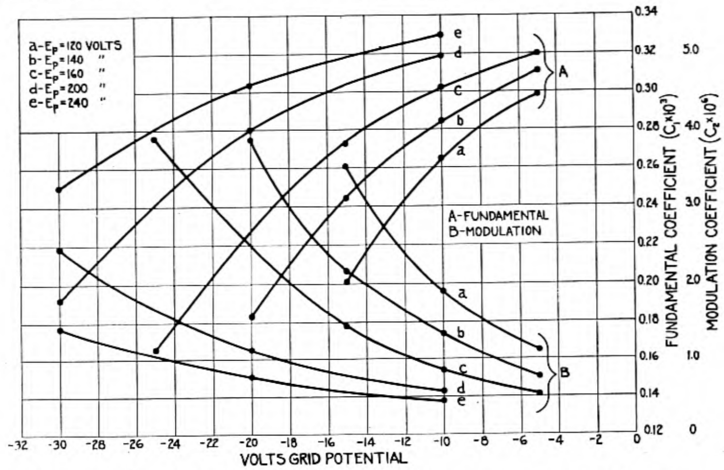


Fig. 7—EL tube No. 109,150. External resistance = 15,000 ohms.
I_f = 1.1 amperes

plate potential. This depends evidently upon the desired power. The results in an illustrative case are depicted by Fig. 9, which represents the level of second harmonic current referred to the fundamental current (Δ) plotted as a function of the external plate resistance. At the point of minimum distortion—in which the second harmonic

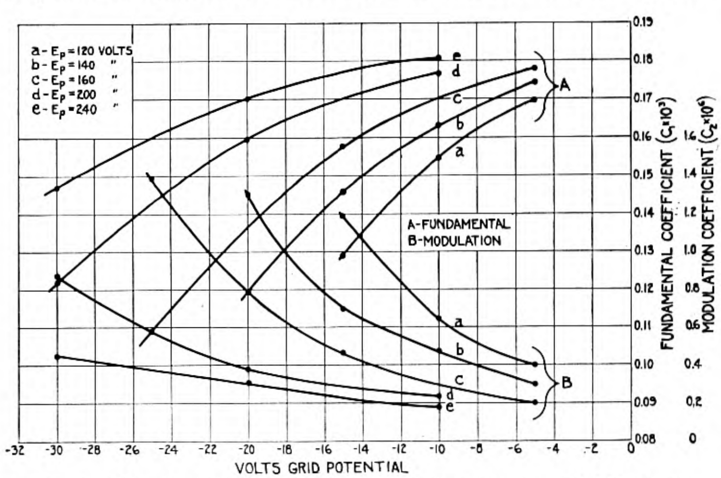


Fig. 8—EL tube No. 109,150. External resistance = 30,000 ohms.
I_f = 1.1 amperes

level is 27.5 T.U. down on the fundamental—the external resistance is twice the internal, the grid potential is -10, and the improvement in the relative reduction of J_2 over customary conditions (for which $R = R_0$ and $E_c = -9$) is about 3.5 T.U. A similar survey made

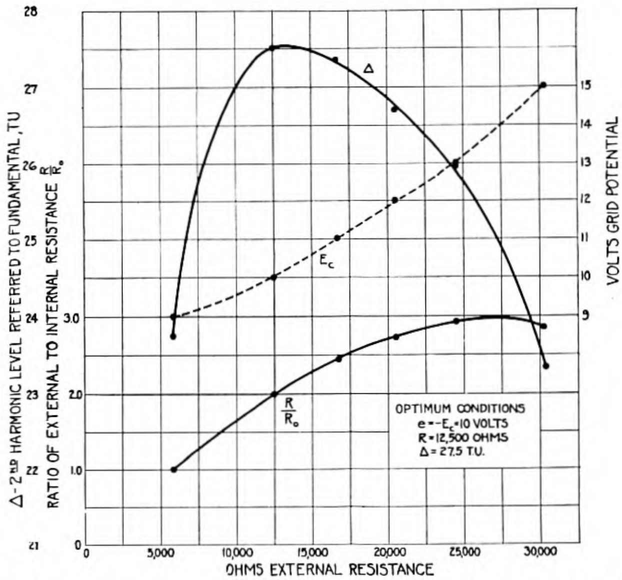


Fig. 9—Ratio of second harmonic to fundamental as a function of load resistance. EL tube No. 109,150. $E_p = 120$, $e = -E_c$ (variable). Power output constant $= .056$ watt.

with regard to the maximum fundamental power obtainable with constant plate potential, and with plate resistances matched, shows it to be had at -13.5 volts grid potential (Fig. 10), and to represent a gain of about 35 per cent in output power over that obtainable at the customary operating point. Other considerations such as stability with battery voltage variations operate in repeater practice to fix the grid potential at the customary values.

There has been considerable discussion recently as to the maximum undistorted power obtainable from a tube with fixed plate potential, and variable load resistance and grid potentials. The above analysis shows that in the strict sense of the word, distortion (C_2, C_3) always exists, while with some arbitrary criterion of distortion, results must depend upon the specific criterion adopted. Now as to the maximum fundamental power obtainable, it may be shown that with a parabolic tube characteristic the maximum is obtained for

$$R \doteq \frac{4}{3} R_0,$$

$$E_c \doteq -0.58 E_b / \mu.$$

These values are not very critical, however, since when we put $R = R_0$ and $E_c = - E_b/2\mu$ the power drops by about ten per cent. The last condition is that for maximum power at maximum gain,

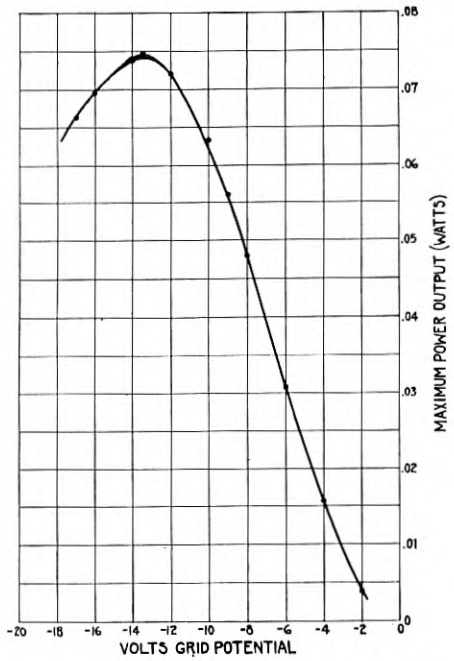


Fig. 10—Maximum power output as a function of the grid potential. EL tube No. 109,150. $E_p = 120$, $E = E_c$, $R = R_0$

in which the output power for a definite alternating grid potential is maximum. In view of the fact that it approximates optimum conditions and is much more convenient as a basis for calculation, we shall use the maximum power—maximum gain criterion of the performance of tubes.

Current and Power Relations.

The preceding analysis has shown that the maximum power obtainable at maximum gain without drawing grid current and without exceeding the negative end of the characteristic—at which grid potential variation produces substantially no plate current variation—is to be found at a grid voltage about midway between the two voltages specified by these conditions. It is again instructive to compare the predictions of the two theories as to the power dissipated in the tube, the a.c. power delivered to the load, the second order modulation, and the relations between these quantities at this operating point. It is a simple matter to calculate these quantities on the basis of Carson's and of van der Bijl's relations.

Thus the plate current is given by

$$i = \alpha(E_p + \mu E_c)^2,$$

the internal plate resistance to alternating currents is

$$R_0 = 1/2\alpha(E_p + \mu E_c),$$

and the internal plate resistance to direct current is

$$R_{dc} = E_b/\alpha(E_p + \mu E_c)^2.$$

At the operating point for maximum power we have $E_b = -2\mu E_c$, and when the alternating grid potential is equal in amplitude to the grid bias, the above expressions may be manipulated to give

$$\begin{aligned} &\text{1. The d.c. power dissipation} && P = \alpha E_p^3/4, \\ &\text{2. The a.c. power delivered} && W = \alpha E_p^3/32, \\ &\text{3. The 2d harmonic current} && J_2 = \alpha E_p^2/64, \\ &\text{4. The fundamental current} && J_1 = \alpha E_p^2/4. \end{aligned} \qquad (10)$$

From these we have for the ratio of d.c. to a.c. powers

$$P/W = 8,$$

or the efficiency of power conversion at the maximum power condition is $12\frac{1}{2}$ per cent. We find also

$$W/J_2 = 2E_p,$$

or the relation of the fundamental power to the second harmonic current depends upon just one parameter, the plate potential. Other relations of interest are the two following:

$$R_{dc}/R_0 = 4,$$
$$J_1/J_2 = 16.$$

These four relations are independent of tube structure (μ, R_0) and we know that they cannot be accurate in view of the assumptions made in deriving them. In view, however, of the importance of general relations of this type in the design of amplifiers, it is of interest to compare these relations with the ones existing, as calculated by the more accurate theory in which μ variation enters.

Accordingly there are tabulated below the maximum power conditions for a number of tubes of different structure with plate potentials from 120 to 350 volts, operated under the conditions for maximum power output.

CONDITIONS FOR MAXIMUM POWER

Tube	E_b	$-E_c$	μ	R_0	P/W	W/J_2	R/R_0	J_1/J_2
101-D	120	13.5	5.36	8,800	4.8	132	4.55	7.17
101-D	240	28.5	5.43	5,900	4.1	270	4.68	6.95
104-D	120	34.0	2.13	2,930	5.4	132	4.10	7.50
104-D	240	72.0	2.13	2,200	4.3	211	4.54	5.48
205-D	350	32.0	6.95	5,780	3.8	350	5.40	6.53
Special A	250	16.9	9.35	5,800	3.9	267	5.07	6.76
" B	240	9.0	18.1	7,400	3.9	224	4.45	5.50
" B	120	4.4	17.6	9,500	4.5	116	4.14	5.96
" C	130	77.0	1.13	1,430	3.6	131	4.96	6.05

The last four columns represent computations by the double series method which are to be compared with the approximate relations of Eq. (10). Thus J_1/J_2 of the last column is given as 16 when μ-variation is neglected whereas it actually varies between 5.48 and 7.50; R/R_0 on the approximate theory is put at 4, whereas it varies actually between 4.1 and 5.4; W/J_2 is given as $2E_p$ which actually comes out close to half that, and finally the P/W is given as 8, while it really varies between 3.8 and 5.4. On the other hand, the approximate independence of tube structure is shown by these four ratios as given in the last four columns of figures.

Application of the Theory of Probability to Telephone Trunking Problems

By EDWARD C. MOLINA

IF telephone plants were provided in such quantities that when a subscriber makes a call there would be immediately available such switching arrangements and such trunks or paths to the desired point as may be necessary to establish the connection instantly, it would require that paths and switching facilities be provided to meet the maximum demand occurring at any time, with the result that there would be a large amount of plant not in use most of the time.

Obviously, this would result in high costs, particularly in cases of long circuits or where the switching arrangements are complex.

Sound telephone engineering requires, therefore, that we approach this condition only in so far as it is practical and economical to do so, consistent with good telephone service to the subscriber. To take an extreme case—if enough toll lines were provided between New York and San Francisco so that no call would ever be delayed because of busy lines or busy switching arrangements, the rates it would be necessary to charge would be prohibitive, although the speed of service would be very good. Obviously, it is necessary to adopt a compromise between the number of circuits and amount of equipment and the time required to complete a call.

Handling traffic on any other than an instantaneous basis is generally spoken of as handling it on a "delayed basis," even though this delay may be, and generally is, inappreciable to the subscriber. While most of the traffic is handled practically on a no-delay basis, there are certain kinds that are handled on a "delayed basis," such as

1. Calls handled by toll lines, when all toll lines happen to be busy.
2. Calls served by senders or line finders in machine switching systems.
3. Calls handled by operators; the delays implied here being due to the time required by an operator to perform her functions apart from delays due to limitations of equipment.

For traffic handled in this manner it is desirable to have formulas or curves for determining the percentage of calls delayed and the average delay on calls delayed. The product of these two figures will give, of course, the average delay on all calls.

The object of this paper is to present for consideration the results of some theoretical studies made with reference to calls handled on a delay basis.

It is felt that these results may be applied with but slight modifications to many of those traffic problems in which calls are subjected to delay rather than loss when idle mechanisms for advancing them are not immediately available. Such items as service to the subscribers, wear on selecting mechanisms, reliability of circuit operation, etc., are often dependent on the magnitudes of the delays encountered in such cases. Their application to problems in manual traffic is probably less immediate and precise due to the human element which enters into the reckoning. Such factors as an operator's ability to speed up at times of heavy traffic and the facility with which she may reach distant signals appearing before other positions make the problems rather more involved than those dealing with mechanisms whose reactions under various circumstances are more possible to predict. Nevertheless in such cases as these, as well as in problems relating to the delays encountered in clearing trouble conditions, installing telephones, awaiting elevator service, and many other problems of interest to engineers in general, the results and methods discussed here, though probably not directly applicable, may prove highly suggestive in a qualitative way.

Obviously the average number of calls to be handled per unit of time, the average length of holding time per call, and the number of trunks assigned to handle the traffic are factors entering into the mathematical results. A knowledge of these three quantities is not sufficient, however, for the solution of the problem. Quite different results will be secured according to the assumption made as to how the holding times of individual calls vary about their given average, i.e., one set of results follows from the assumption that holding times are all of the same length—other sets of results if holding times vary, the precise set depending on the particular law of variation assumed.

The choice made of laws representing holding time variations must be governed by two considerations:

1. An assumed law must agree at least approximately with the points obtained if we plot the way holding times vary as found from observations.

2. The form of the law must lend itself to the mathematical solution of the delay problem.

Case No. 1

An assumption which permits of an easy and exact mathematical solution of the problem may be stated as follows: If a call is picked

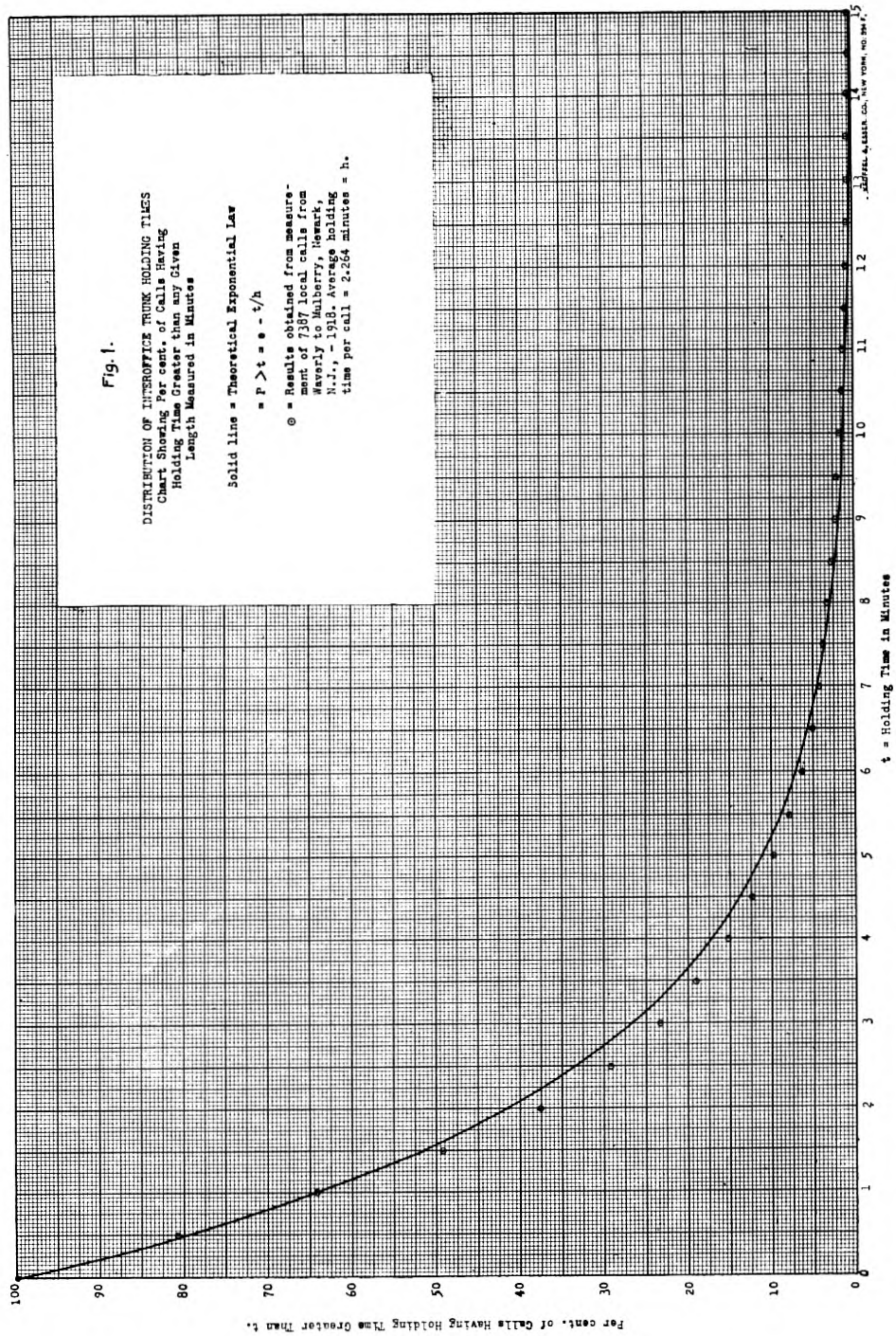


Fig. 1.
DISTRIBUTION OF INTEROFFICE TRUNK HOLDING TIMES
Chart Showing Per cent. of Calls Having
Holding Time Greater than any Given
Length Measured in Minutes
Solid line = Theoretical Exponential Law
= P >t = e - t/h
⊙ = Results obtained from measurement of 7387 local calls from Waverly to Mulberry, Newark, N.J., - 1918. Average holding time per call = 2.264 minutes = h.
t = Holding Time in Minutes
Per cent. of Calls Having Holding Time Greater Than t.
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
100
90
80
70
60
50
40
30
20
10
0

at random, the probability that its holding time is greater than an interval of time of length t is $e^{-t/h}$, where e is the base of natural logarithms and h is the average holding time of all calls. Fortunately there are cases in practice where the variations in length of holding times are closely represented by this exponential law. This is clearly shown by the following Fig. 1 entitled "Distribution of Interoffice-Trunk Holding Times." The points shown on the figure are the plots of actual holding times obtained from pen register records made on a group of trunks running from Waverly to Mulberry in Newark, New Jersey.

Case No. 2

Another assumption covered in this memorandum, because it checks closely with cases arising in practice, is that all calls have exactly the same holding time. The precise solution of the delay service problem becomes extremely difficult on this assumption of a constant holding time. An approximate solution is presented in this paper. Cases in practice where holding times are essentially constant are those of sender holding times of key indicator trunk groups and cordless B boards.

With reference to either Case No. 1 or Case No. 2 consider a group of a certain number of operators handling a certain number of calls having a certain average holding time. If we double the average holding time and halve the number of calls (so that the operators are busy for the same per cent of time on the average), the per cent of calls delayed will not change, but the average delay on calls delayed as measured in seconds will exactly double. Suppose, for example, we wish to obtain the same average speed of answer to line signals with two different groups or teams of operators, the first handling traffic which requires only a short operator holding time or work interval, and the second handling traffic which requires a longer work interval. If the teams are equal in number and ability, we must allow the second team a larger proportion of idle time than the first. If, on the other hand, the proportion of idle time is to be the same for both teams, the second team must be larger or more capable than the first. This general effect is well known, but it is hoped that the results herewith presented will supply more exact knowledge of the subject.

Before proceeding further it will be helpful to give here the notation used on the delay curves following this paper.

h = average holding time per call.

c = number of trunks in a straight multiple.

a = average number of calls originating per interval of time h.

$P\ (>t)$ = probability that a call is delayed for an interval of time greater than t. In other words, 100 times $P\ (>t)$ gives the per cent of calls which will, on the average, be delayed for an interval greater than t.

Charts

The two series of charts following this paper embody, for Case No. 1 and Case No. 2, respectively, curves giving for different values of c and the ratio a/c the probability that a call will be delayed to an extent which will exceed a given multiple of the average holding time. These curves may consequently be read to determine what proportion of the calls will be delayed on the average by an interval as measured in holding times. For example, consider the particular varying holding time chart which corresponds to $c = 10$; we see from the curve marked $a/c = 0.50$ that there is a probability of 0.001 that a call will be delayed for an interval of time which will exceed 0.72 of the average holding time; or, put another way, that 0.1 per cent of the calls will be, on the average, delayed this amount. Again there is a probability of only 0.000019 for a call being delayed at least 1.5 times the average holding time; or 0.0019 per cent of the calls will be delayed by this amount. If, on the same chart, we consider the curve marked $a/c = 0.70$, we find that 22 per cent of the calls will be delayed, 1 per cent will be delayed at least 1.04 times the average holding time, 0.01 per cent will be delayed 2.58 times the average holding time, or more.

The dotted line on each chart gives, at its points of intersection with the curves, the average delay on calls delayed as a multiple of the average holding time interval. For example, on the $c = 10$ varying holding chart we note that for $a/c = 0.70$ the *average delay on calls delayed* is $0.33h$. To obtain the average delay on all calls we multiply by the proportion of calls delayed, $P\ (> 0) = 0.22$, and obtain 0.073 times the average holding time.

A glance at the formulas, given in the Appendix, for the average delay on calls delayed shows that this delay does not reduce to zero when a/c becomes zero. It approaches a lower limit which has the value h/c in Case No. 1 and the value $h/(c+1)$ in Case No. 2. The latter limit may readily be anticipated from physial considerations as follows. Assume that the group consists of a single trunk; we have to show that the average delay when a call is delayed approaches the limit $h/(1+1) = h/2$ as the load approaches zero. Now when the load is very low, those cases where two or more calls have to wait for the trunk to become idle are quite negligible; we only have to consider the delay incurred by a single call originating while the trunk is busy. But as calls originate at random, the delayed call is just as likely to have fallen near the beginning as toward the end of the constant interval h during which the trunk is busy. In other words, on the average, the delayed call will have originated in the middle of the constant interval h and thus the average delay incurred will be $h/2$. This lower limit for the average delay on calls delayed is indicated in the lower left-hand corner of each sheet of curves by the point where the axis of abscissæ is intersected by a short vertical line.

It will be noted that the constant holding time delay curves change their direction at those points for which the abscissæ are exact multiples of the holding time interval h. No such discontinuities in slope

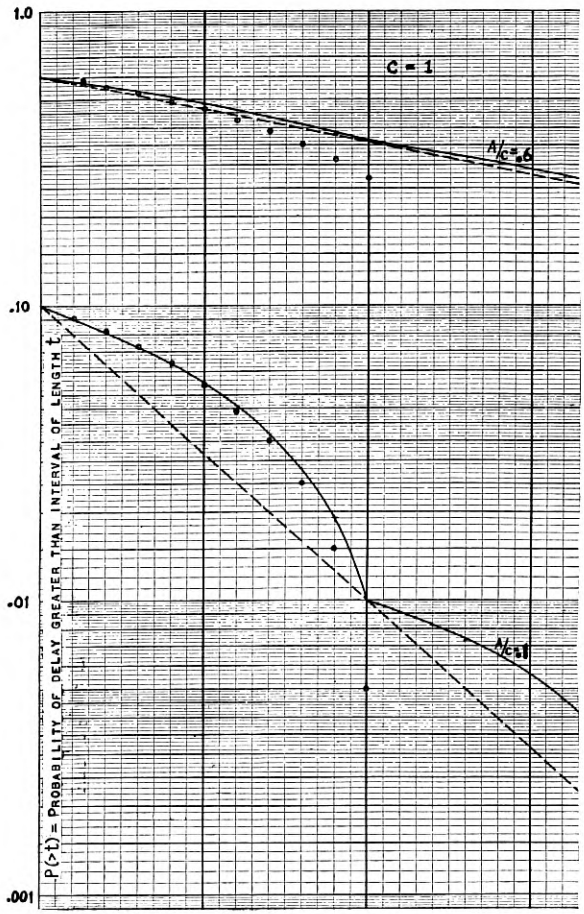


FIG. 4.

appear on the varying holding time curves. This difference in the two classes of curves should occasion no surprise. In the varying holding time case the quantity h has no physical significance; it is merely a numerical value obtained by an algebraic process called averaging. In the other case the quantity h represents a physical characteristic of each and every call.

As stated on page 464 the solution presented in this paper for the case where holding times are all of constant length is not exact. It is desirable therefore to have some idea of the degree of approximation attained.

Figure 4 shows a comparison between our tentative solution and true values which Erlang of Copenhagen, Denmark, succeeded in obtaining by a method which unfortunately becomes impracticable for values of c greater than 3. Our results are shown by the solid curve and Erlang's results by the small circles. We have also indicated on the $c = 1$, 2 and 3 constant holding time charts, Erlang's points for $a/c = 0.50$. For Erlang's work, reference may be had to the Elektrotechnische Zeitschrift, December 19, 1918, page 504.

It may be noted that for the higher values of the ratio a/c the curves are practically straight lines. They depart materially from straight lines for the lower values of the ratio a/c, particularly if c itself is not very large.

Assumptions Made in Mathematical Theory

The mathematical theory back of the curves accompanying this paper is based on the following assumptions:

1. Calls originating independently of each other, and at random with reference to time, have complete access to a single group of trunks.
2. The probability of a call originating during a particular infinitesimal interval, dt, is practically independent of the number of trunks busy or number of waiting calls at the beginning of said interval. This assumption implies that the total interval of time during which the calls fall at random is very large compared with the average holding time per call and that the total number of calls under consideration is very large compared with the number of calls originating per average holding time interval.
3. Calls are served in the order in which they originate. This restriction does not apply to the *average* delays obtained.

4*A*. The average holding time being h, the holding times of individual calls vary around this average in such a way that $e^{-t/h}$ is the probability that for a call taken at random the holding time is greater than t.

4*B*. The holding times of all calls are equal to a constant h.

5*B*. If, at any instant, s of the c trunks are busy, the distribution in time of the instants at which said s busy trunks were seized is identical with the distribution of s points picked individually at random on a straight line of length h.

Assumptions 1 and 2 together imply that the number of *sources* of calls is so large that any blocking of calls due to limitation of sources need not be considered. Assumptions 1, 2, 3 and 4*A* were made in deriving the formulas for Case No. 1. Assumptions 1, 2, 3, 4*B* and 5*B* were made in deriving the formulas for Case No. 2. Assumption 5*B* is not strictly compatible with the physics of the constant holding time case. The distribution in time of the s calls mentioned in 5*B* will, to a certain extent, depend on the history of previous calls. It is because this dependence is ignored that the solution for Case No. 2, presented in the Appendix, is only approximate.

APPENDIX

Mathematical Theory of Delay Formulas

The following mathematical analysis is based on the assumptions given above on pages 467 and 468.

Consider the state of affairs at the instant a particular call "X" originates. Suppose call "X" encounters x other calls; if x is less than c, call "X" will be served immediately but, if not, "X" will have to wait. Our problem is to determine the probability that the delay which "X" may suffer shall have a specified value.

We begin by determining the relative frequency with which the number of calls encountered by "X" has the value x. Let $f(x)$ be the relative frequency with which x calls are encountered by "X." At an instant of time t, x calls will be encountered if at the preceding instant $(t - dt)$ either x, $(x + 1)$ or $(x - 1)$ calls would have been encountered. We ignore as of too rare occurrence the cases where more than $(x + 1)$ or less than $(x - 1)$ calls would have been encountered at time $(t - dt)$ with x at time t. Now in passing from time $(t - dt)$ to time t the probability of an *increase* of one call is proportional to the difference in time, dt, and to the average number of calls, a, falling per holding time interval. Likewise the probability of a *decrease* of one call is proportional to the time difference, dt, and to the number of calls occupying trunks (a decrease must be due to a busy trunk becoming idle). Therefore

$$f(x) = f(x - 1)\frac{adt}{h} + f(x + 1)\frac{(x + 1)dt}{h} + f(x)\left[1 - \frac{adt}{h} - \frac{xdt}{h}\right]$$

when $x < c$, and

$$f(x) = f(x-1)\frac{adt}{h} + f(x+1)\frac{cdt}{h} + f(x)\left[1 - \frac{adt}{h} - \frac{cdt}{h}\right]$$

when $x \nless c$.

For these two equations we may substitute the simpler equations

$$xf(x)\frac{(dt)}{(h)} = af(x-1)\frac{(dt)}{(h)}$$

or

$$cf(x)\frac{(dt)}{(h)} = af(x-1)\frac{(dt)}{(h)}.$$

The solution of these equations gives

$$f(x) = f(0)\left[\frac{a^x}{c^{x-c}\underline{|c}}\right], \qquad x \nless c$$

and

$$f(x) = f(0)\frac{a^x}{\underline{|x}}, \qquad x \ngtr c,$$

where $f(0)$ is the arbitrary constant entering in the integration of the finite difference equations. But we must have, evidently,

$$\sum_{x=0}^{x=\infty} f(x) = 1.$$

Substituting in this equation the values for $f(x)$ given above, we obtain

$$1/f(0) = e^a\left[1 - P(c, a) + \frac{a^c e^{-a}}{\underline{|c}}\left(\frac{c}{c-a}\right)\right].$$

Since call "X" will be delayed whenever the number x of calls he encounters is equal or greater than c, we have

$$P(>0) = \sum_{x=c}^{x=\infty} f(x) = \frac{\left(\frac{a^c e^{-a}}{\underline{|c}}\right)\left(\frac{c}{c-a}\right)}{1 - P(c, a) + \left(\frac{a^c e^{-a}}{\underline{|c}}\right)\left(\frac{c}{c-a}\right)}.$$

The next question is to determine the probability, $P(>t)$, of a delay which is greater than an interval of length t.

We will get one answer to this question if we make use of assumption $4A$, and a different answer on the basis of assumptions $4B$ and $5B$. Therefore, from here on, it will be necessary to treat separately the varying and constant holding time cases.

Case No. 1—Holding Times Vary Exponentially

$P\,(> t)$ for this case was obtained by Erlang of Copenhagen. In 1917 he published the formula without its proof. The following deduction of his formula is therefore submitted.

The particular call "X" considered above will be delayed if the number of calls he encountered, x, is equal to or greater than the number, c, of trunks in the group. Suppose that the $x = c + (x - c)$ calls encountered by "X" are handled by the trunks in the manner indicated in the following Fig. 2, where $m_1 + m_2 + \cdots m_c = x - c$.

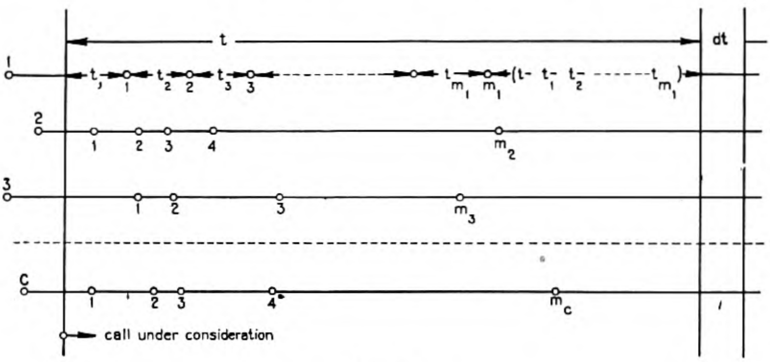


FIG. 2.

By assumption 4A and taking h as the unit of time, the probability that trunk No. 1 will be busy during an interval of time t is

$$(e^{-t_1}dt_1)(e^{-t_2}dt_2)\ \cdots\ (e^{-t_{m_1}}dt_{m_1})e^{-(t-t_1-t_2\cdots\ t_{m_1})}.$$

Giving $(t_1, t_2 \cdots t_{m_1})$ all positive values consistent with their sum $\not> t$, we obtain (see Todhunter's "Integral Calculus," sixth edition, art. 276)

$$e^{-t}\frac{(t^{m_1})}{(\underline{|m_1})}.$$

Therefore the compound probability that all c trunks are busy during the interval t with the x calls existing at the instant under consideration and that then one of the trunks becomes idle in the interval dt is

$$(e^{-t})^c\left[\sum\left(\frac{t^{m_1}\,t^{m_2}\,\cdots\,t^{m_c}}{\underline{|m_1}\ \underline{|m_2}\ \cdots\ \underline{|m_c}}\right)\right](cdt),$$

where $\sum$ means that we are to give $m_1, m_2 \cdots m_c$ all values such that

$$m_1 + m_2 + \cdots m_c = x - c.$$

By the multinomial theorem

$$\sum \left(\frac{(m_1 + m_2 + \cdots m_c)!}{m_1!\, m_2! \cdots m_c!} \right) (a^{m_1} b^{m_2} \cdots K^{m_c}) = (ct)^{x-c}$$

for $a = b = \cdots K = t$ and $(m_1 + m_2 + \cdots m_c) = x - c$.

The expression above reduces to

$$(e^{-ct}) \left[\frac{(ct)^{x-c}}{(x-c)!} \right] cdt.$$

Now all this is on the assumption that "X" encountered x other calls. Therefore we must multiply by $f(x)$ and sum for all values of x from c to ∞. We obtain

$$\sum_{x=c}^{\infty} f(x) e^{-ct} \frac{(ct)^{x-c}}{(x-c)!} cdt =$$

$$\sum_{x=c}^{\infty} f(0) \frac{a^x}{c^{x-c}\, c!} e^{-ct} \frac{(ct)^{x-c}}{(x-c)!} cdt =$$

$$f(0) \left(\frac{a^c}{c!} \right) ce^{-ct} dt \sum_{x=c}^{\infty} \frac{(at)^{x-c}}{(x-c)!} =$$

$$f(0) \left(\frac{a^c}{c!} \right) cdte^{-ct} e^{at} = f(0) \left(\frac{a^c}{c!} \right) ce^{-(c-a)t} dt.$$

This is the probability that "X" will be delayed for an interval of length t. To obtain the probability that the delay will be greater than t we must integrate with reference to t from t to ∞. But

$$\int_t^{\infty} e^{-(c-a)t} dt = \frac{e^{-(c-a)t}}{c-a}.$$

Thus, finally,

$$P(>t) = f(0) \frac{a^c}{c!} \left(\frac{c}{c-a} \right) e^{-(c-a)t}$$

$$= P(>0) e^{-(c-a)t}.$$

This formula for $P(>t)$ has been deduced by taking h as the unit of time. Evidently we would have obtained

$$P(>t) = P(>0) e^{-(c-a)t/h}$$

if h had not been taken as unit of time.

For the average delay on all calls we have

$$\bar{t} = \int_0^{\infty} t \frac{dP(>t)}{dt} dt = P(>0) \left(\frac{h}{c-a} \right),$$

and the average delay on calls delayed is

$$\left(\frac{h}{c - a}\right).$$

Case II—Holding Times Constant

Write x, the number of calls encountered by "X," in the form

$$x = nc + m - 1,$$

where n and m are positive integers such that $m \not> c$. In the Fig. 3 below these $nc + m - 1$ calls are shown in groups arranged according

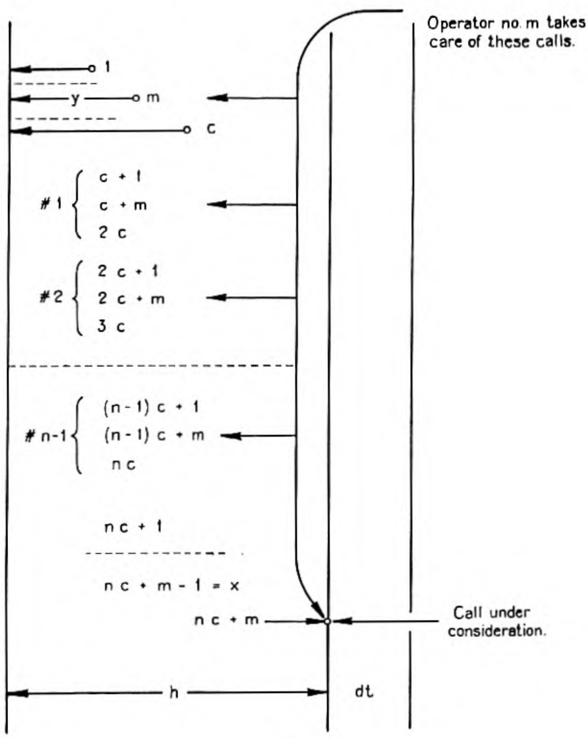


FIG. 3

to the order in which they originated. The first c calls are occupying the trunks. Then $n - 1$ groups, each consisting of c waiting calls, are shown and finally a remainder of $m - 1$ waiting calls; our particular call "X" is the $(nc + m)$th in order of time.

Now evidently, as indicated in the figure, trunk number m will serve call "X" after said trunk has served the call occupying it and also the mth call in each of the $n - 1$ waiting groups. Therefore our call will suffer a delay of length

$$(n - 1)h + y,$$

where y is the time which elapsed between the beginning of the interval h and the instant at which the mth trunk was seized by the call occupying it. The probability of this delay is a compound one made up of two factors.

1st—The probability that x calls are encountered. This probability is, as derived above,

$$\begin{aligned} f(x) &= f(0)\frac{a^x}{c^{x-c}\underline{|c}} \\ &= \frac{f(0)c^c(a/c)^{nc}}{\underline{|c}}\left(\frac{a}{c}\right)^{m-1}, \end{aligned}$$

since $x = nc + m - 1$.

2d—The probability that the distance from the beginning of the interval h to the instant at which the mth trunk was seized is y or, in more precise terms, lies between y and $y + dy$. This probability is, on the basis of assumption $5B$,

$$c\left(\frac{dy}{h}\right)\left[\binom{c-1}{m-1}\left(\frac{y}{h}\right)^{m-1}\left(1-\frac{y}{h}\right)^{(c-1)-(m-1)}\right].$$

The product of these two probabilities gives, writing $y = uh$, $(a/c) = R$,

$$\frac{f(0)c^{c+1}(R)^{nc}}{\underline{|c}}\left[\binom{c-1}{m-1}\left(\frac{uR}{1-u}\right)^{m-1}(1-u)^{c-1}\right]du.$$

But the subscribers' interest in a delay of magnitude $(n - 1)h + y$ is totally independent of what value m might have. Therefore the last probability expression must be summed for all permissible values of m, that is from $m = 1$ to $m = c$. We then obtain for the total probability of a delay of extent between $(n - 1)h + y$ and $(n - 1)h + y + dy$:

$$\frac{f(0)c^{c+1}R^{nc}(1-u)^{c-1}}{\underline{|c}}\left[\sum_{m=1}^{c}\binom{c-1}{m-1}\left(\frac{uR}{1-u}\right)^{m-1}\right]du =$$

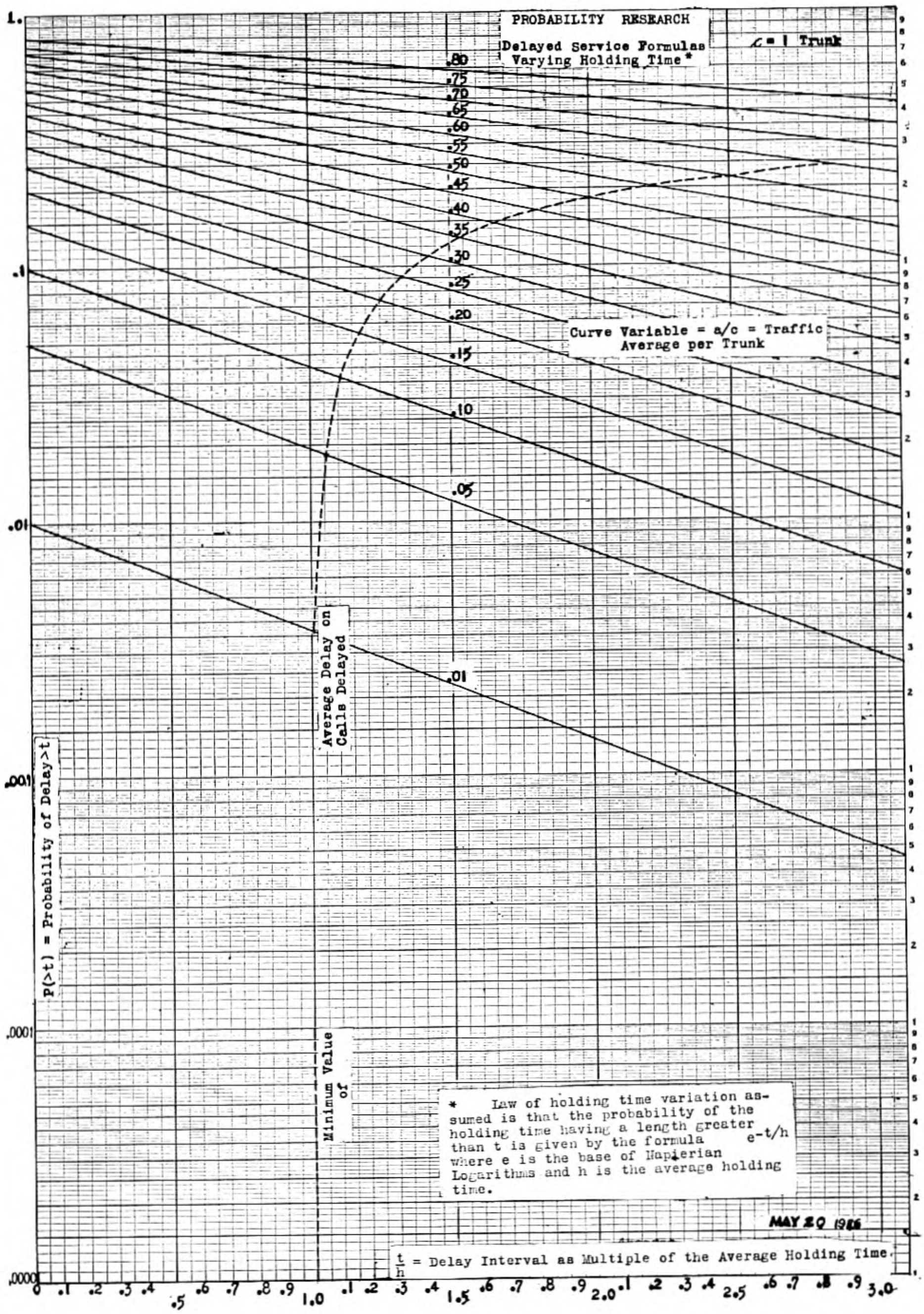
$$\frac{f(0)c^{c+1}R^{nc}(1-u)^{c-1}}{\underline{|c}}\left[1+\frac{uR}{1-u}\right]^{c-1}du =$$

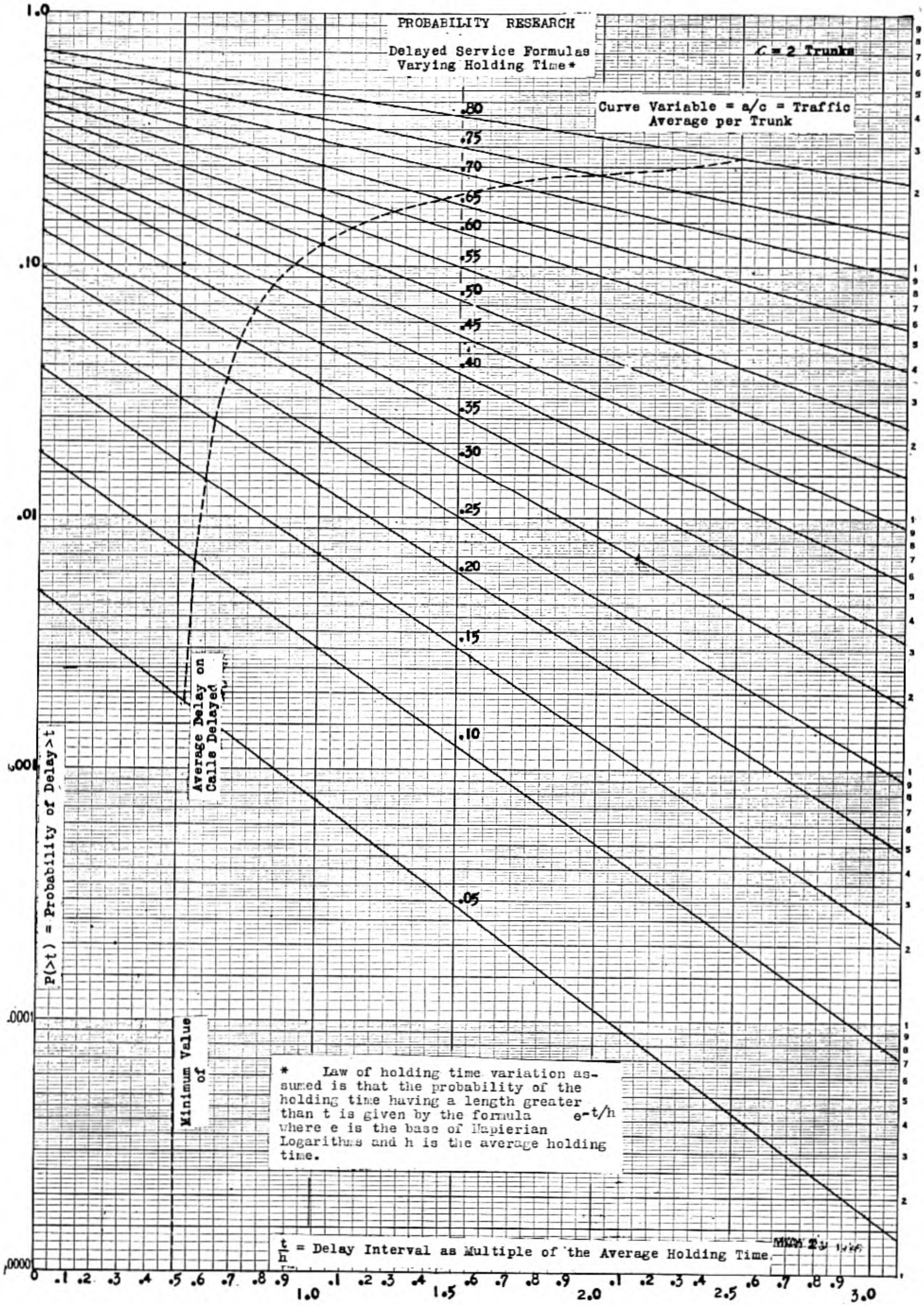
$$\frac{f(0)c^{c+1}R^{nc}}{\underline{|c}}[1-(1-R)u]^{c-1}du =$$

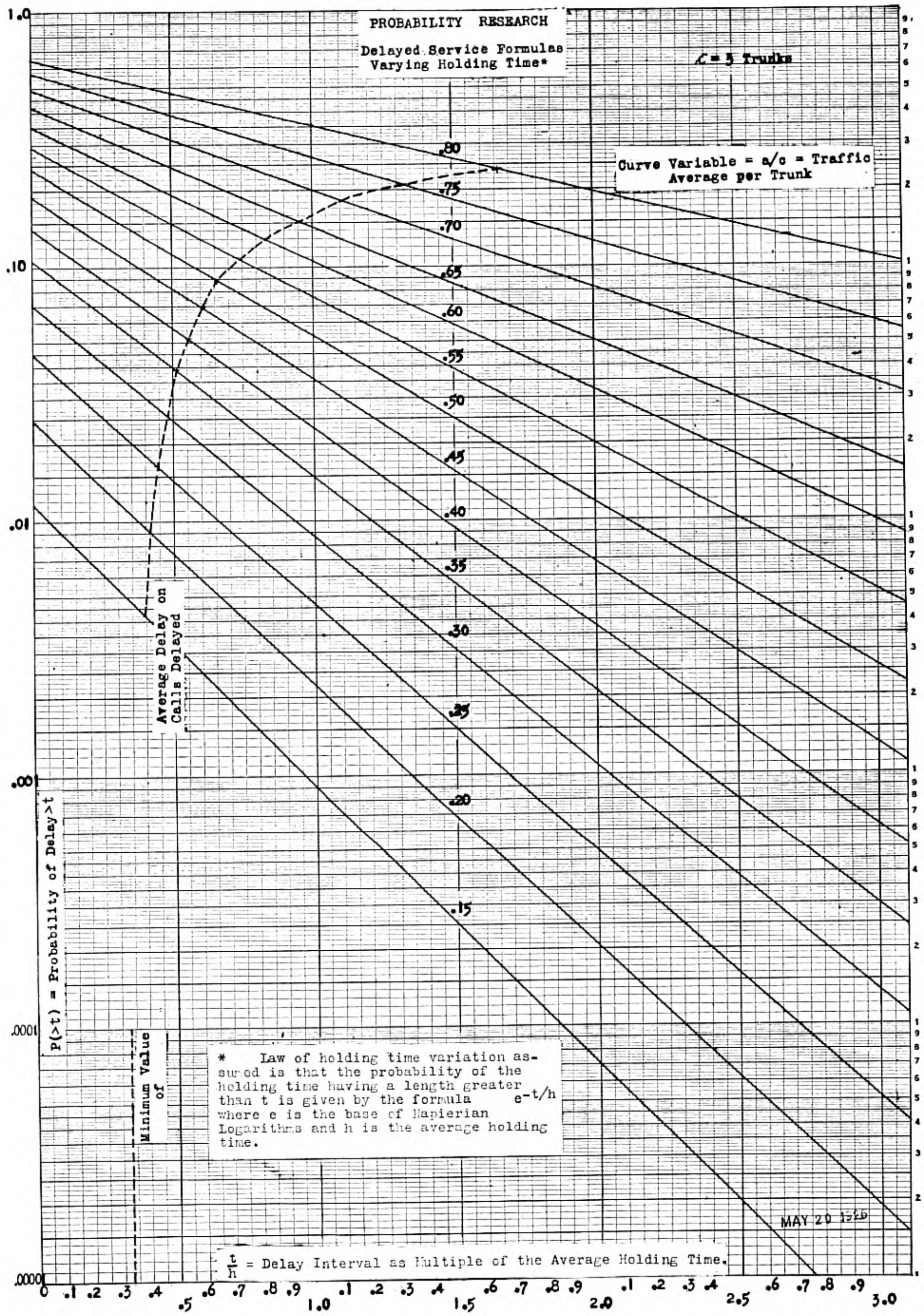
$$P(>0)(c-a)R^{(n-1)c}[1-(1-R)u]^{c-1}du.$$

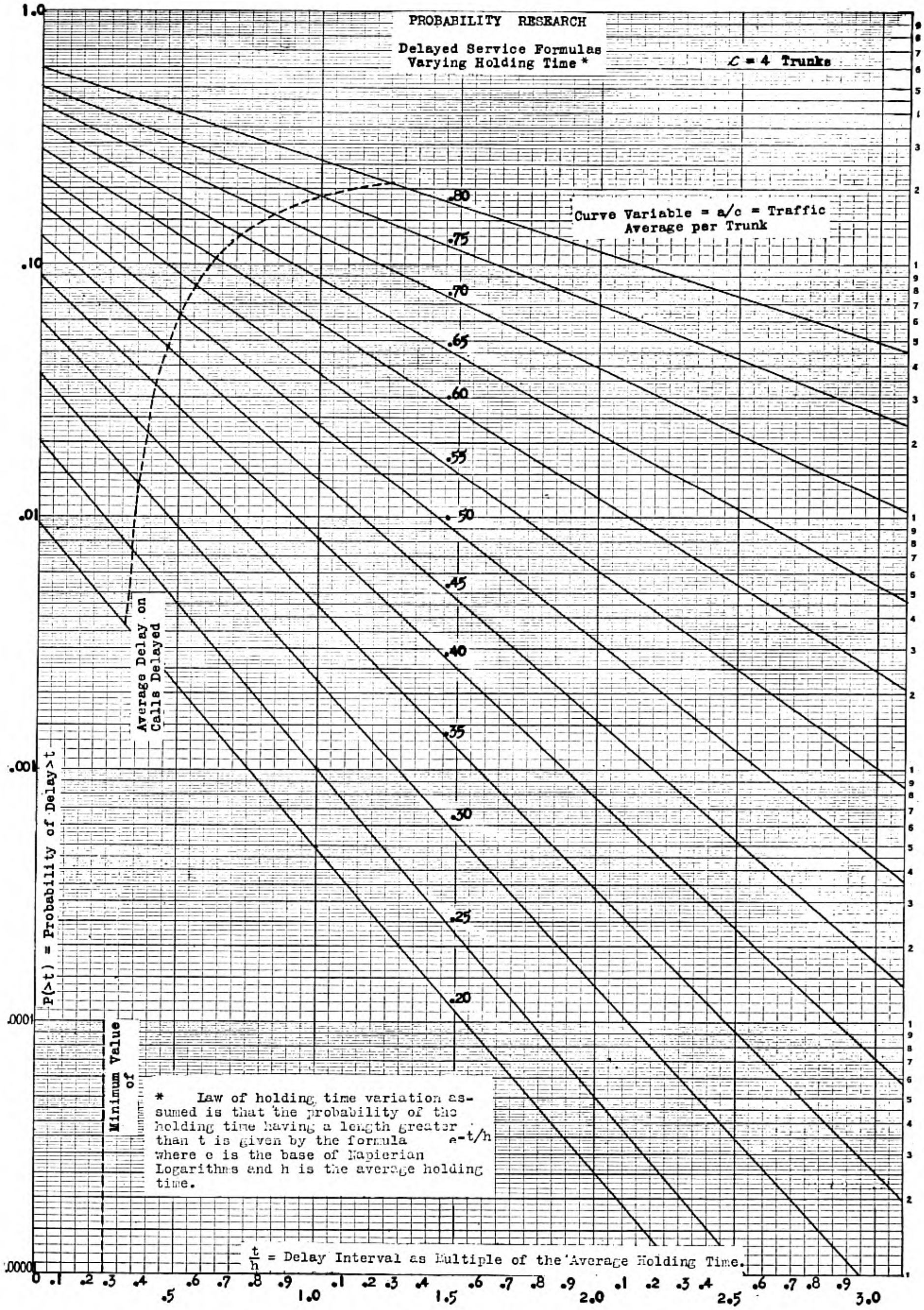
We now obtain by integration with reference to u, summation with reference to n, or both, the following results.

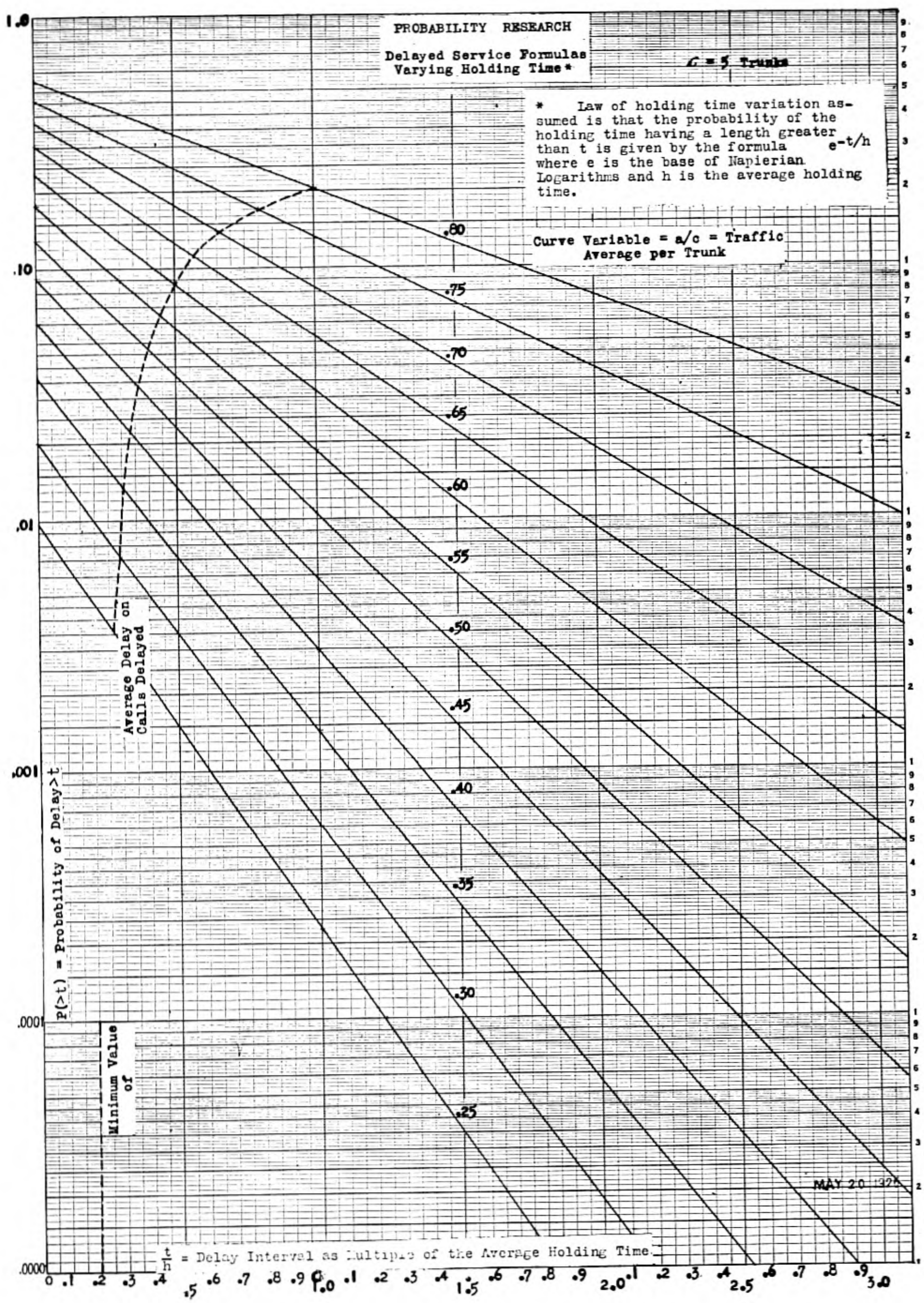
Character of Delay	Probability of Delay
From $(n-1+u)h$ to nh	$P(>0)R^{(n-1)c}([1-(1-R)u]^c - R^c)$
From $(n-1)h$ to nh	$P(>0)R^{(n-1)c}(1-R^c)$
Greater than $(n-1)h$	$P(>0)R^{(n-1)c}$
Greater than $(n-1+u)h$	$P(>0)R^{(n-1)c}[1-(1-R)u]^c$
Average delay on all calls	$P(>0)\left(\frac{h}{c-a}\right)\left(\frac{c}{c+1}\right)\left(\frac{1-R^{c+1}}{1-R^c}\right)$

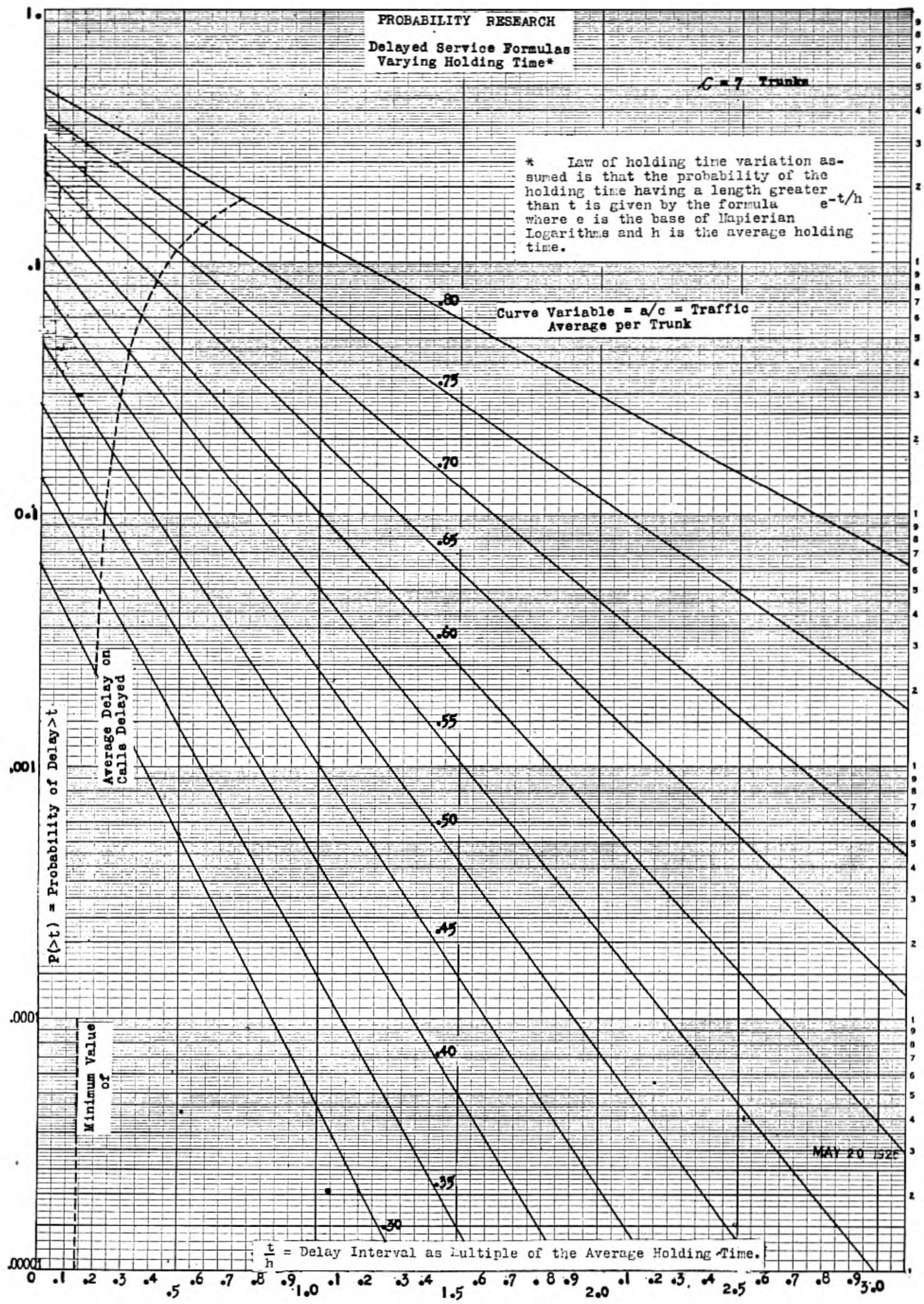

PROBABILITY RESEARCH
Delayed Service Formulas
Varying Holding Time *
c = 1 Trunk
.80
.75
.70
.65
.60
.55
.50
.45
.40
.35
.30
.25
.20
.15
.10
.05
.01
Curve Variable = a/c = Traffic Average per Trunk
Average Delay on Calls Delayed
Minimum Value of
P(>t) = Probability of Delay>t
1.
.1
.01
.001
.0001
.00001
* Law of holding time variation assumed is that the probability of the holding time having a length greater than t is given by the formula e-t/h where e is the base of Napierian Logarithms and h is the average holding time.
MAY 20 1926
t/h = Delay Interval as Multiple of the Average Holding Time.
0 .1 .2 .3 .4 .5 .6 .7 .8 .9 1.0 .1 .2 .3 .4 1.5 .6 .7 .8 .9 2.0 .1 .2 .3 .4 2.5 .6 .7 .8 .9 3.0

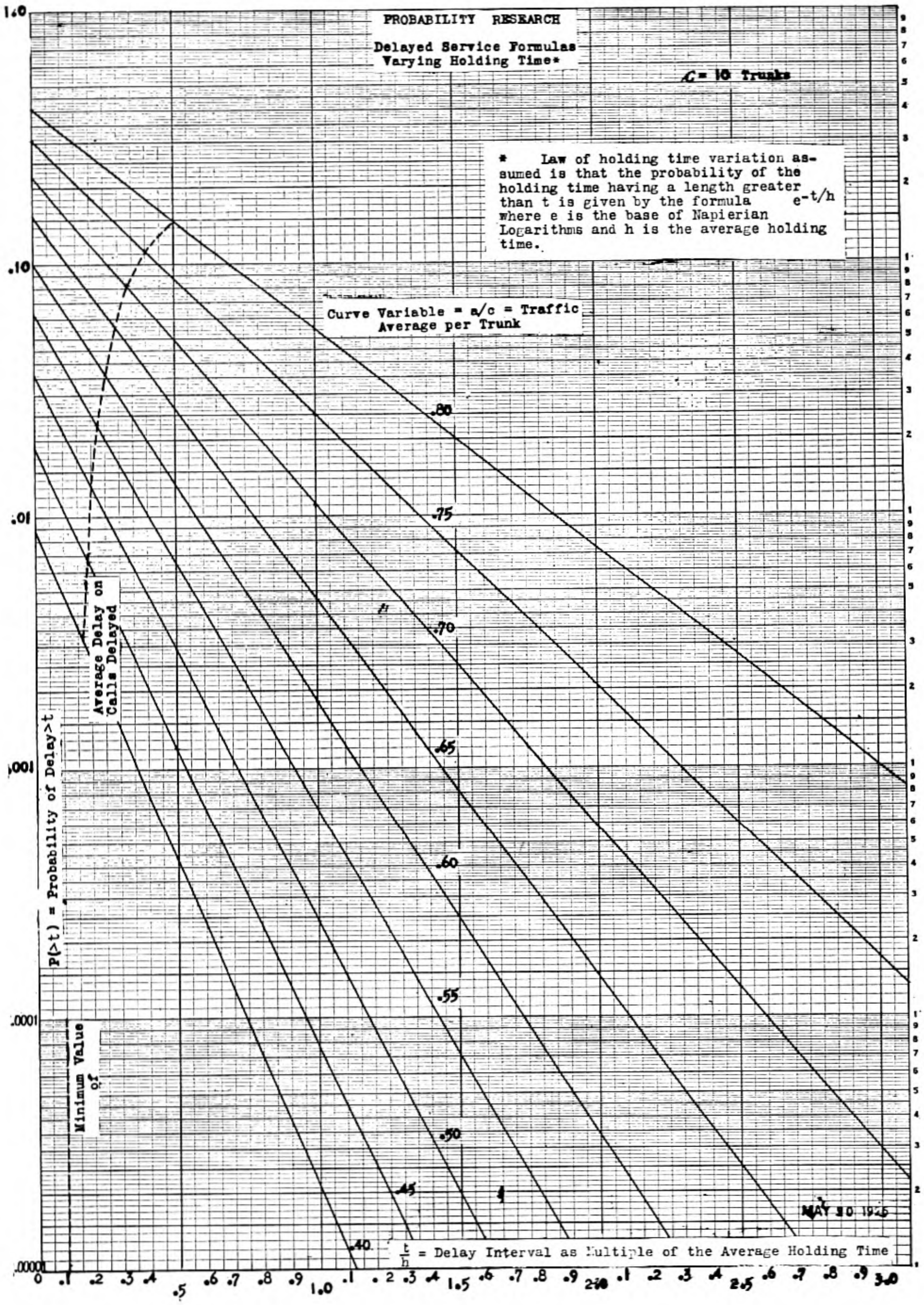

PROBABILITY RESEARCH
Delayed Service Formulas Varying Holding Time*
c = 2 Trunks
Curve Variable = a/c = Traffic Average per Trunk
.80
.75
.70
.65
.60
.55
.50
.45
.40
.35
.30
.25
.20
.15
.10
.05
Average Delay on Calls Delayed
Minimum Value of
P(>t) = Probability of Delay>t
1.0
.10
.01
.001
.0001
.00001
* Law of holding time variation assumed is that the probability of the holding time having a length greater than t is given by the formula e-t/h where e is the base of Napierian Logarithms and h is the average holding time.
t/h = Delay Interval as Multiple of the Average Holding Time
0
1.0
1.5
2.0
2.5
3.0

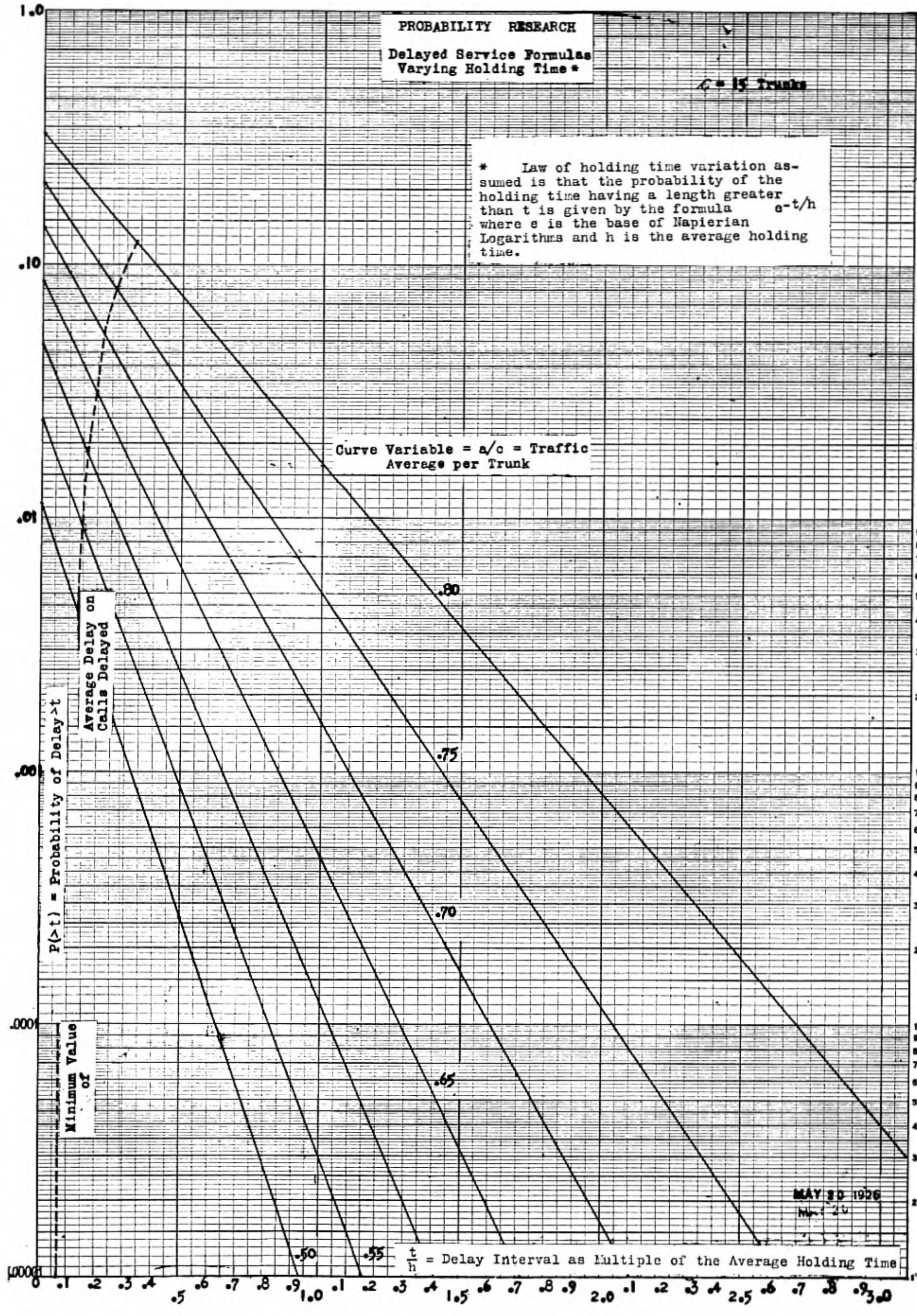

PROBABILITY RESEARCH
Delayed Service Formulas
Varying Holding Time*
c = 5 Trunks
Curve Variable = a/c = Traffic Average per Trunk
.80
.75
.70
.65
.60
.55
.50
.45
.40
.35
.30
.25
.20
.15
Average Delay on Calls Delayed
Minimum Value of
P(>t) = Probability of Delay>t
1.0
.10
.01
.001
.0001
.00001
* Law of holding time variation assumed is that the probability of the holding time having a length greater than t is given by the formula e-t/h where e is the base of Napierian Logarithms and h is the average holding time.
t/h = Delay Interval as Multiple of the Average Holding Time.
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
1.5
2.0
2.5
3.0
MAY 20 1926

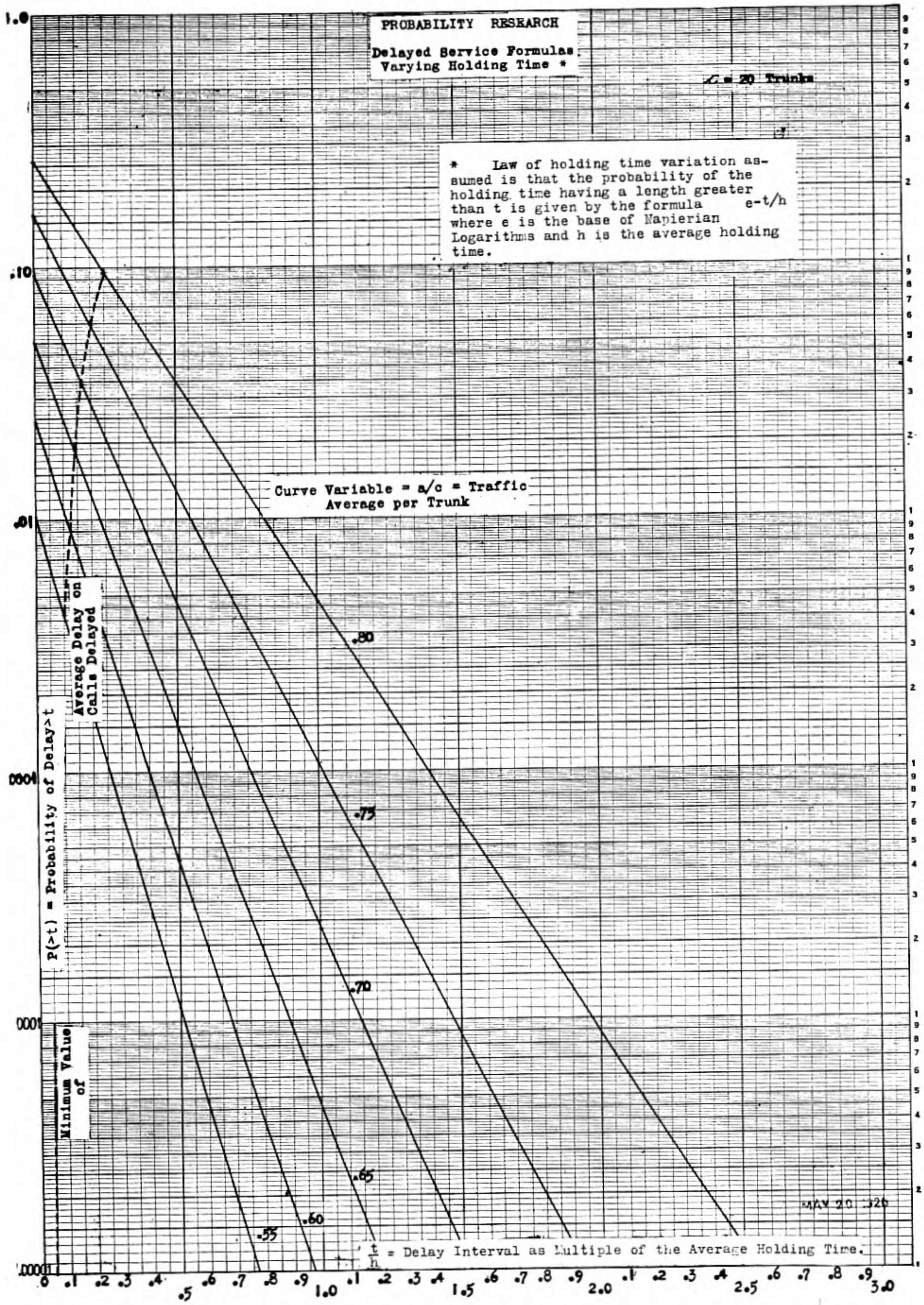

PROBABILITY RESEARCH
Delayed Service Formulas
Varying Holding Time *
c = 4 Trunks
Curve Variable = a/c = Traffic Average per Trunk
P(>t) = Probability of Delay>t
Average Delay on Calls Delayed
Minimum Value of
* Law of holding time variation assumed is that the probability of the holding time having a length greater than t is given by the formula $e^{-t/h}$ where e is the base of Napierian Logarithms and h is the average holding time.
$\frac{t}{h}$ = Delay Interval as Multiple of the Average Holding Time.
.80
.75
.70
.65
.60
.55
.50
.45
.40
.35
.30
.25
.20
1.0
.10
.01
.001
.0001
.00001
0
.5
1.0
1.5
2.0
2.5
3.0

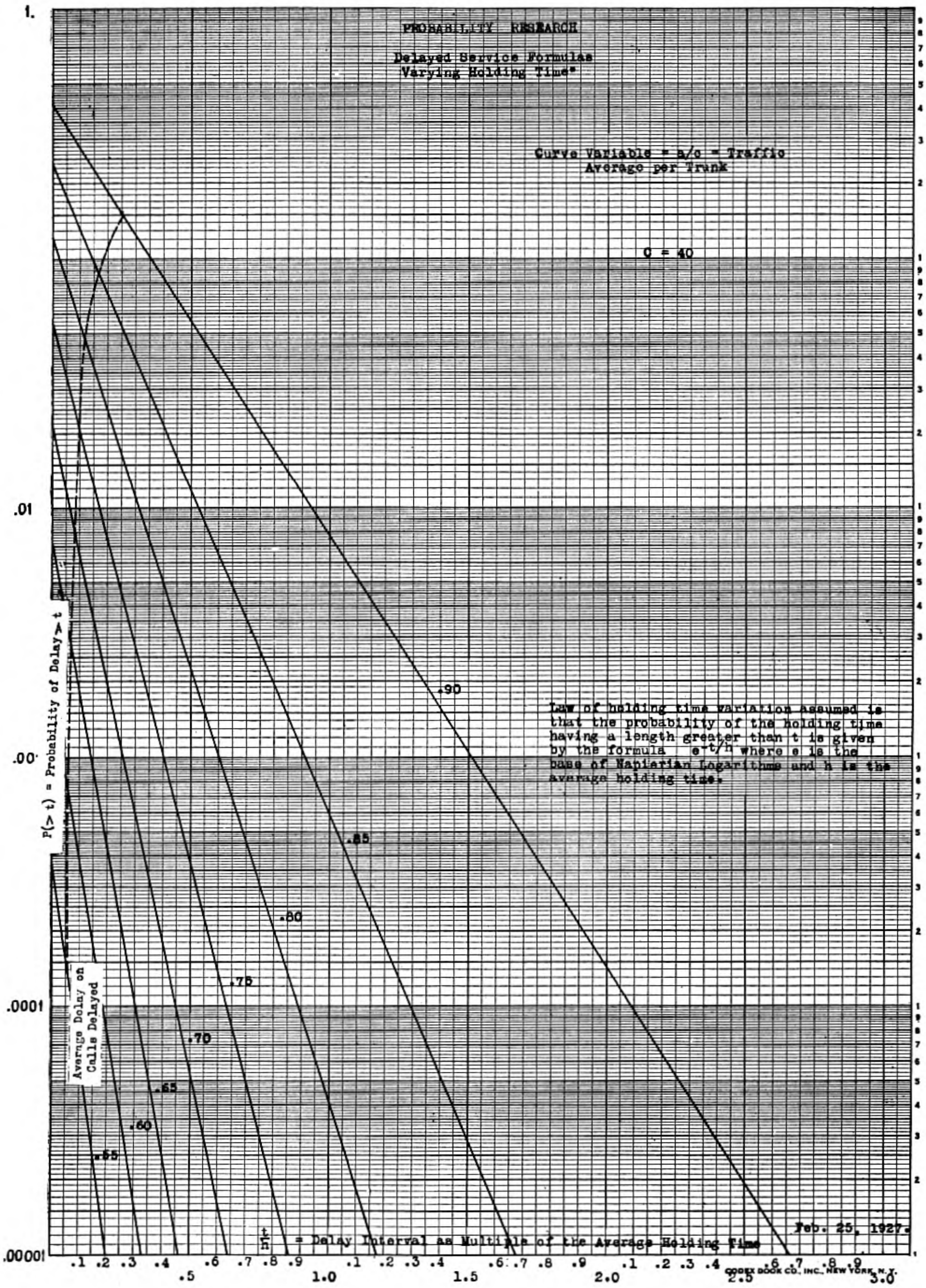

PROBABILITY RESEARCH
Delayed Service Formulas
Varying Holding Time*
C = 5 Trunks
* Law of holding time variation assumed is that the probability of the holding time having a length greater than t is given by the formula e-t/h where e is the base of Napierian Logarithms and h is the average holding time.
Curve Variable = a/c = Traffic Average per Trunk
.80
.75
.70
.65
.60
.55
.50
.45
.40
.35
.30
.25
Average Delay on Calls Delayed
P(>t) = Probability of Delay>t
Minimum Value of
t/h = Delay Interval as Multiple of the Average Holding Time
1.0
.10
.01
.001
.0001
.00001
MAY 20 1925

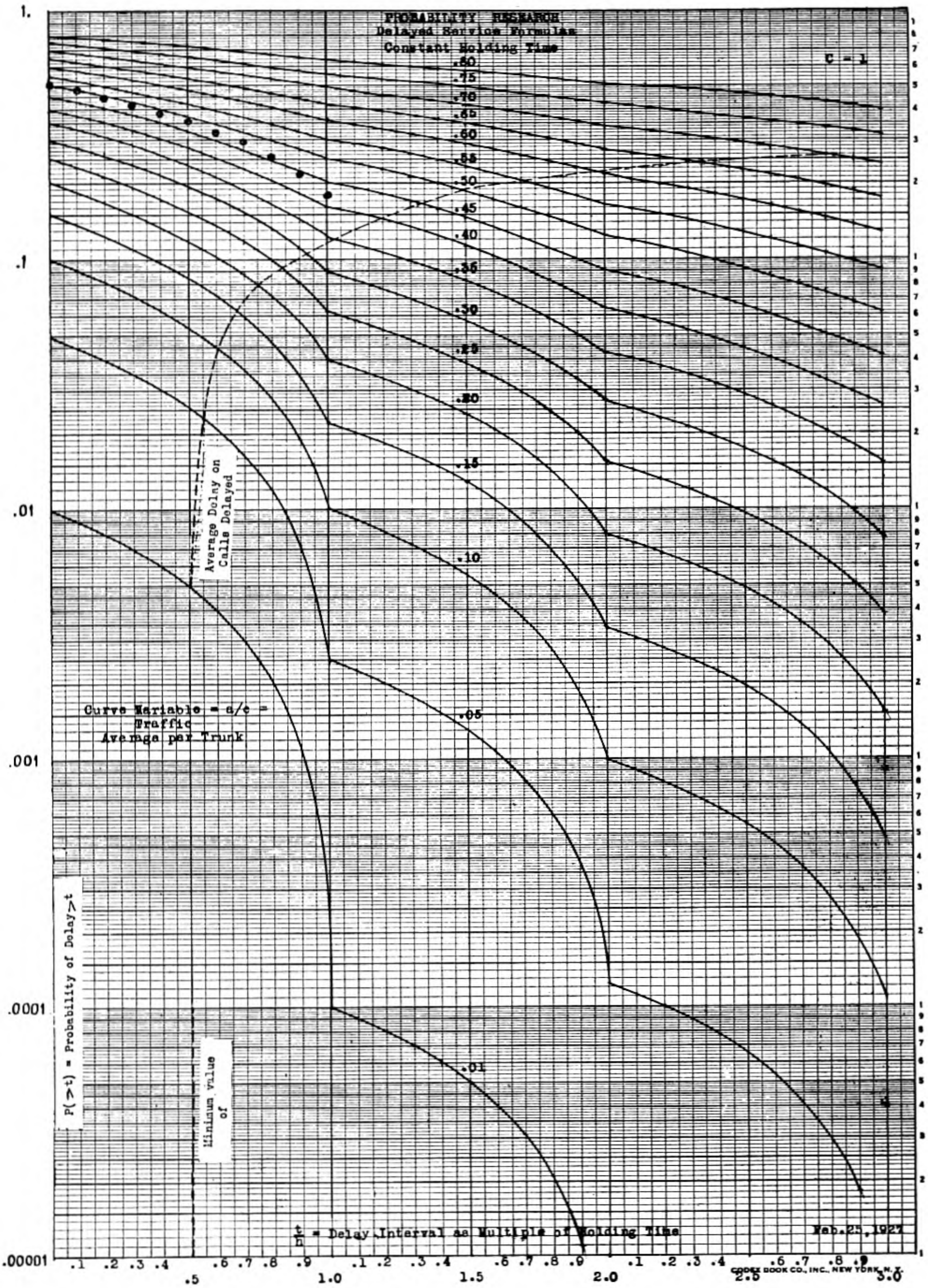

PROBABILITY RESEARCH
Delayed Service Formulas
Varying Holding Time*
C = 7 Trunks
* Law of holding time variation assumed is that the probability of the holding time having a length greater than t is given by the formula $e^{-t/h}$ where e is the base of Napierian Logarithms and h is the average holding time.
Curve Variable = a/c = Traffic Average per Trunk
.80
.75
.70
.65
.60
.55
.50
.45
.40
.35
.30
Average Delay on Calls Delayed
P(>t) = Probability of Delay>t
Minimum Value of
t/h = Delay Interval as Multiple of the Average Holding Time.
1.
.1
0.1
.001
.0001
.00001
0
.5
1.0
1.5
2.0
2.5
3.0
MAY 20 1925

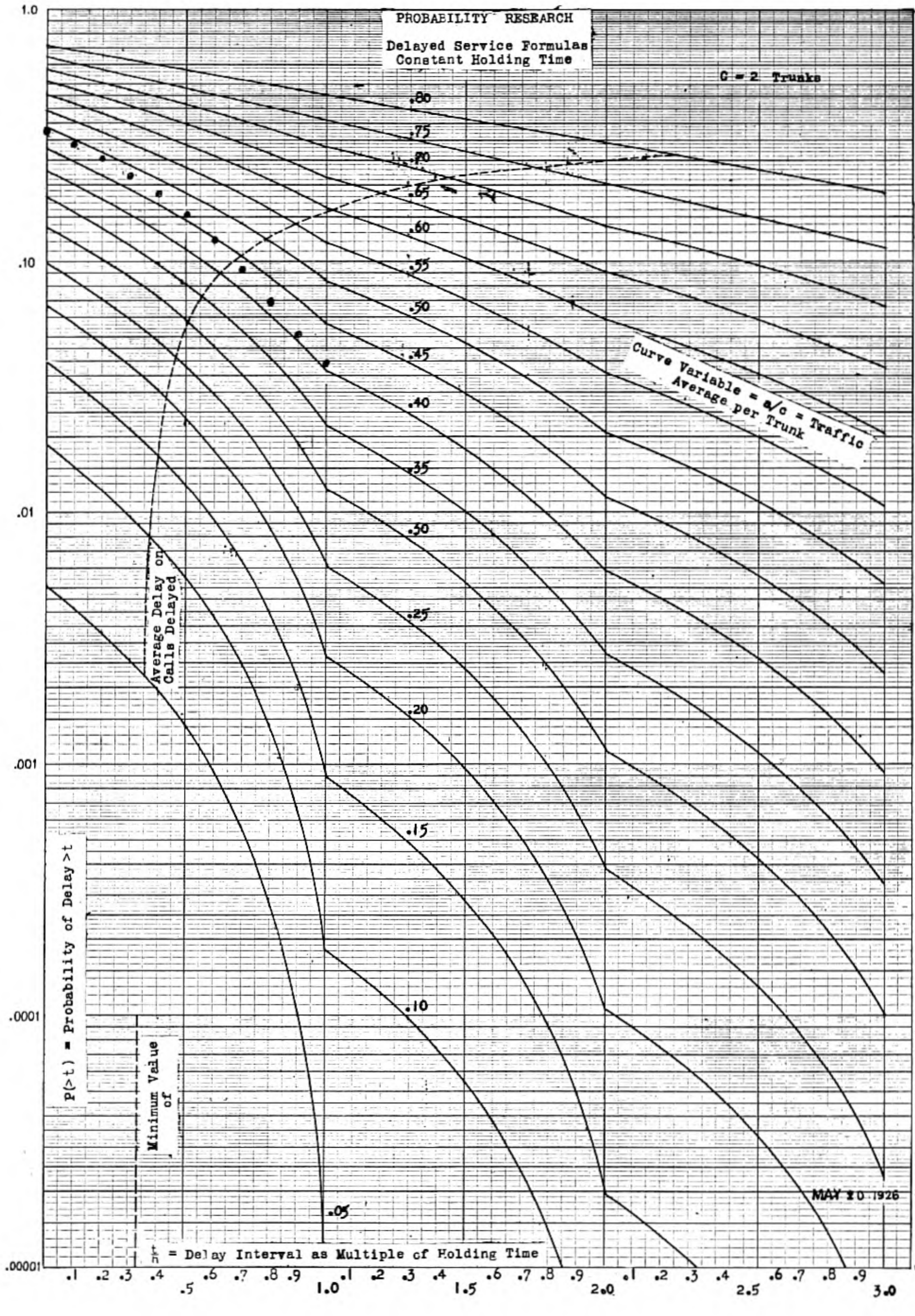

PROBABILITY RESEARCH
Delayed Service Formulas
Varying Holding Time*
C = 10 Trunks
* Law of holding time variation assumed is that the probability of the holding time having a length greater than t is given by the formula e^-t/h where e is the base of Napierian Logarithms and h is the average holding time.
Curve Variable = a/c = Traffic Average per Trunk
.80
.75
.70
.65
.60
.55
.50
.45
.40
Average Delay on Calls Delayed
Minimum Value of
P(>t) = Probability of Delay>t
t/h = Delay Interval as Multiple of the Average Holding Time
MAY 30 1925
1.0
.10
.01
.001
.0001
.00001

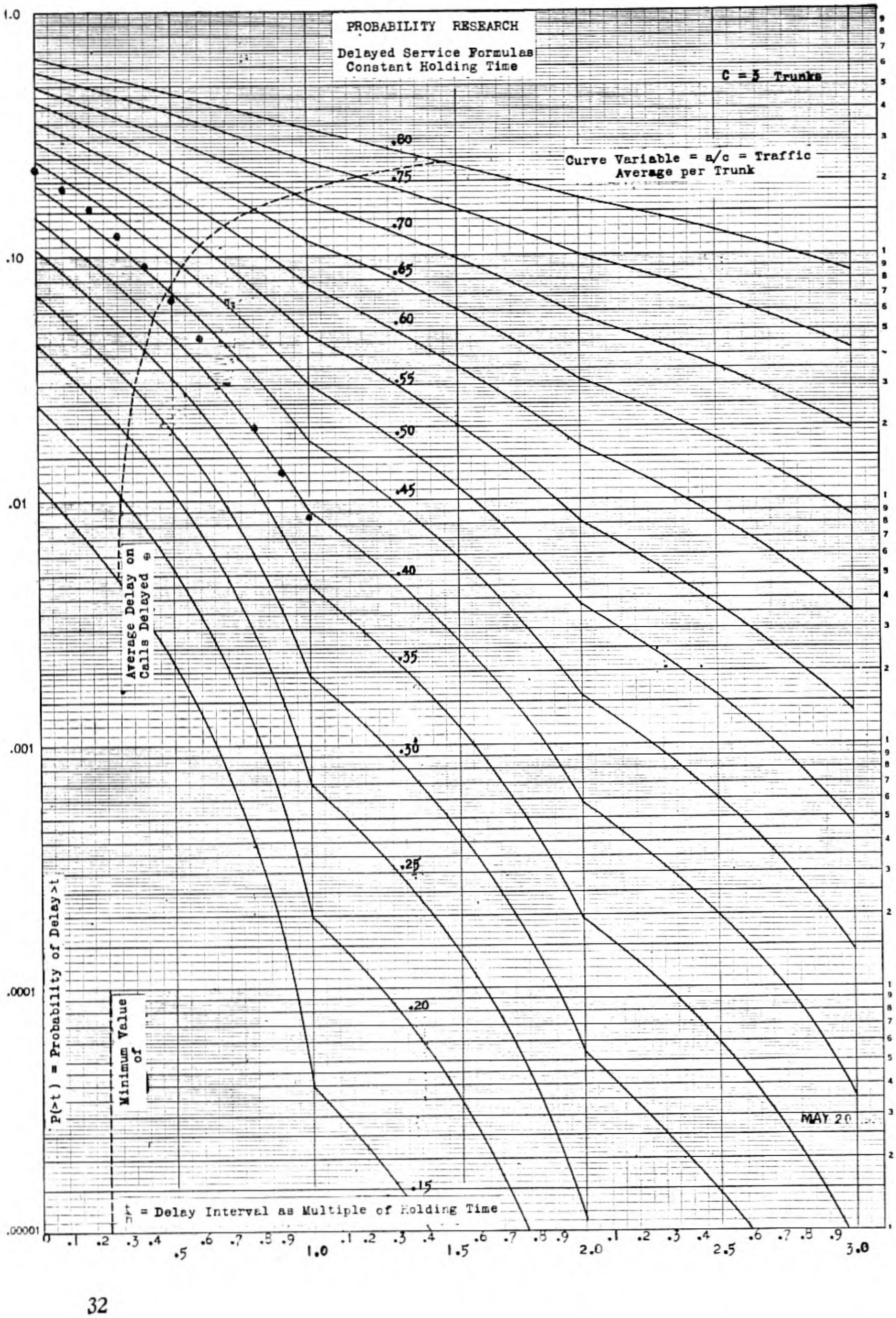

PROBABILITY RESEARCH
Delayed Service Formulas Varying Holding Time *
c = 15 Trunks
* Law of holding time variation assumed is that the probability of the holding time having a length greater than t is given by the formula e-t/h where e is the base of Napierian Logarithms and h is the average holding time.
Curve Variable = a/c = Traffic Average per Trunk
.80
.75
.70
.65
.50
.55
P(>t) = Probability of Delay >t
Average Delay on Calls Delayed
Minimum Value of
t/h = Delay Interval as Multiple of the Average Holding Time
1.0
.10
.01
.001
.0001
.00001
MAY 20 1926

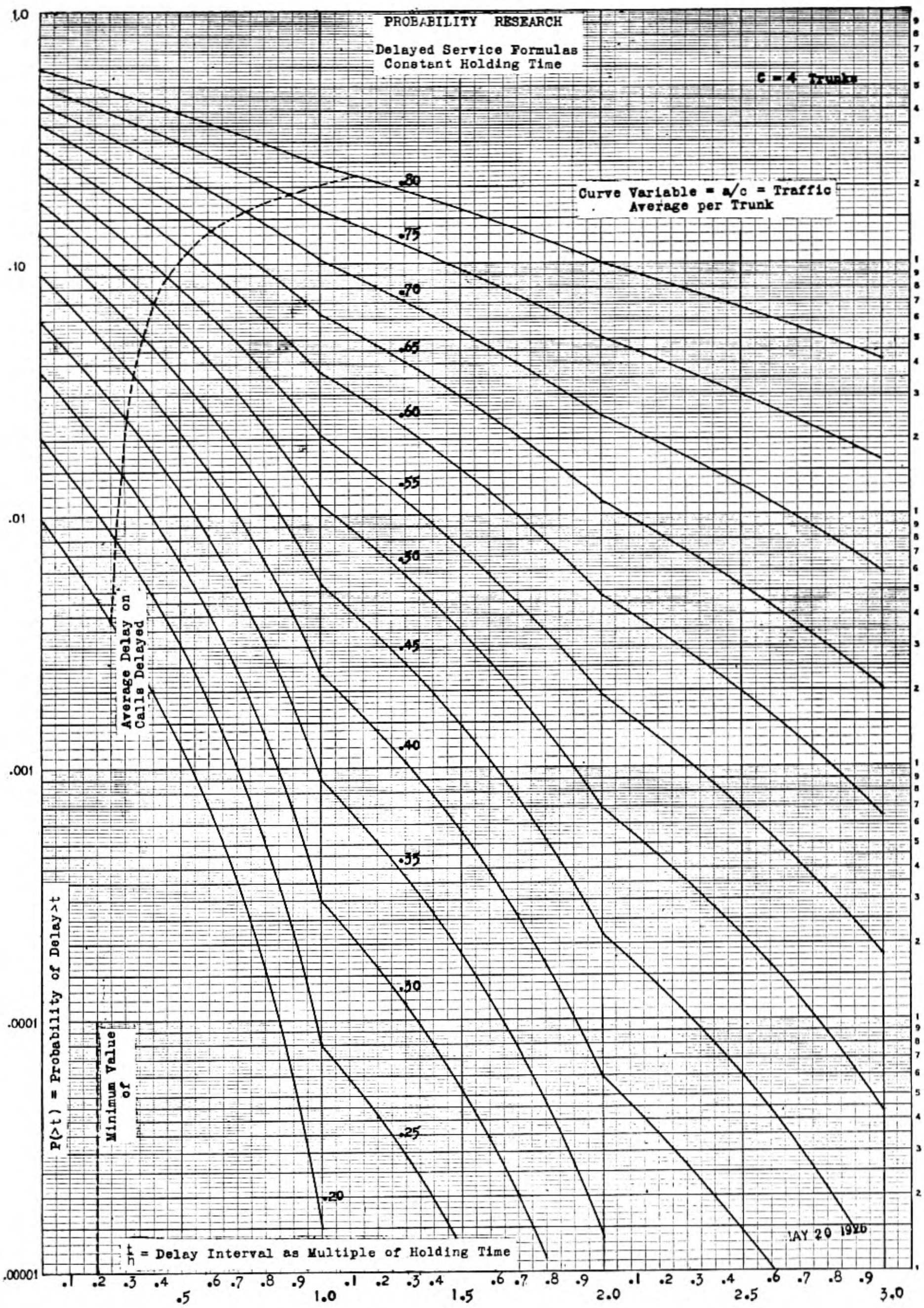

PROBABILITY RESEARCH
Delayed Service Formulas
Varying Holding Time *
c = 20 Trunks
* Law of holding time variation assumed is that the probability of the holding time having a length greater than t is given by the formula e-t/h where e is the base of Napierian Logarithms and h is the average holding time.
Curve Variable = a/c = Traffic Average per Trunk
P(>t) = Probability of Delay>t
Average Delay on Calls Delayed
Minimum Value of
.80
.75
.70
.65
.60
.55
t/h = Delay Interval as Multiple of the Average Holding Time.
1.0
.10
.01
.0001
.00001
MAY 20 1920

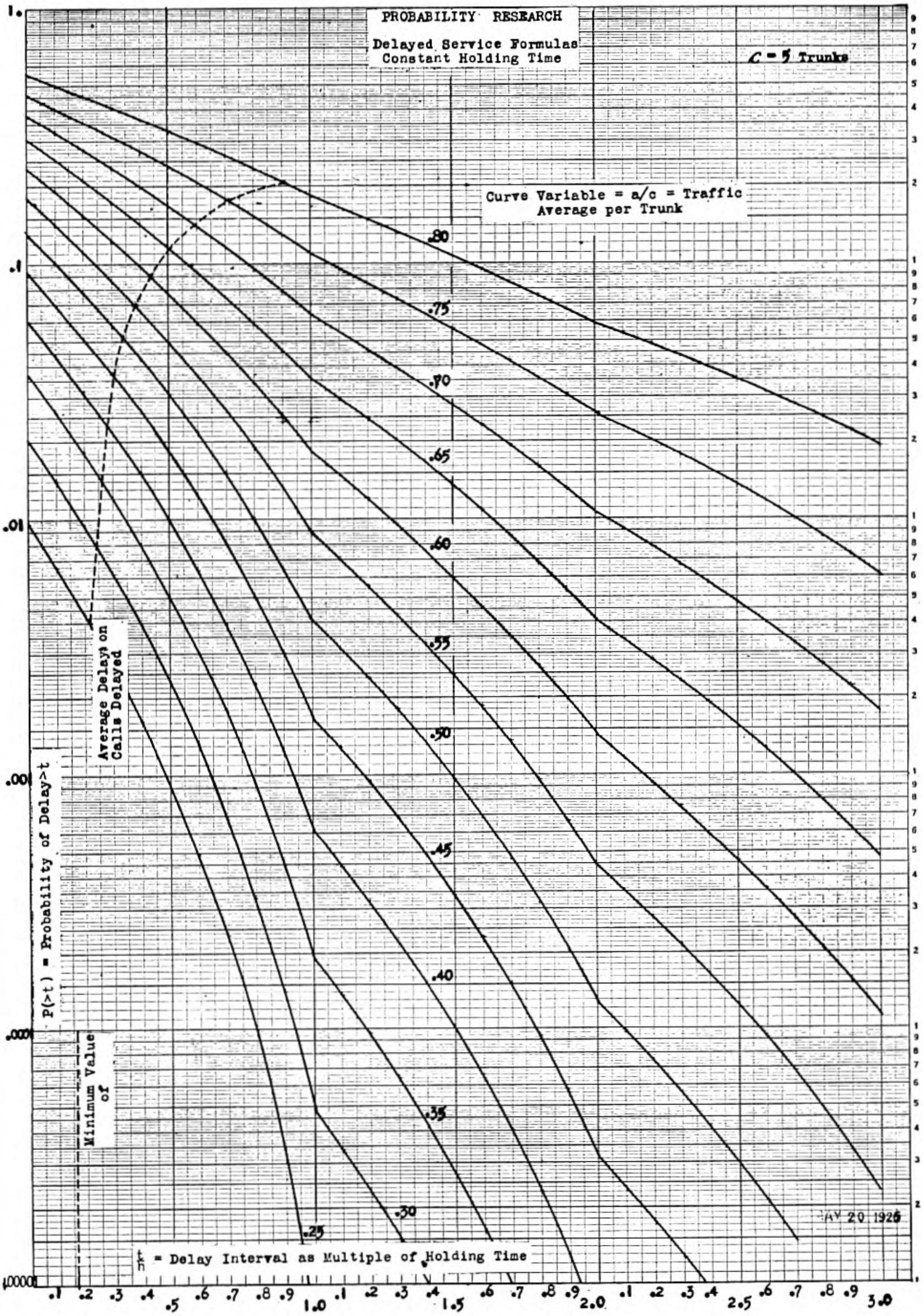

PROBABILITY RESEARCH
Delayed Service Formulas
Varying Holding Time"
Curve Variable = a/c = Traffic Average per Trunk
C = 40
Law of holding time variation assumed is that the probability of the holding time having a length greater than t is given by the formula e-t/h where e is the base of Napierian Logarithms and h is the average holding time.
.90
.85
.80
.75
.70
.65
.60
.55
P(> t) = Probability of Delay > t
Average Delay on Calls Delayed
t/h = Delay Interval as Multiple of the Average Holding Time
Feb. 25, 1927.
1.
.01
.001
.0001
.00001
.5
1.0
1.5
2.0
2.5
3.0
CODEX BOOK CO., INC., NEW YORK, N.Y.

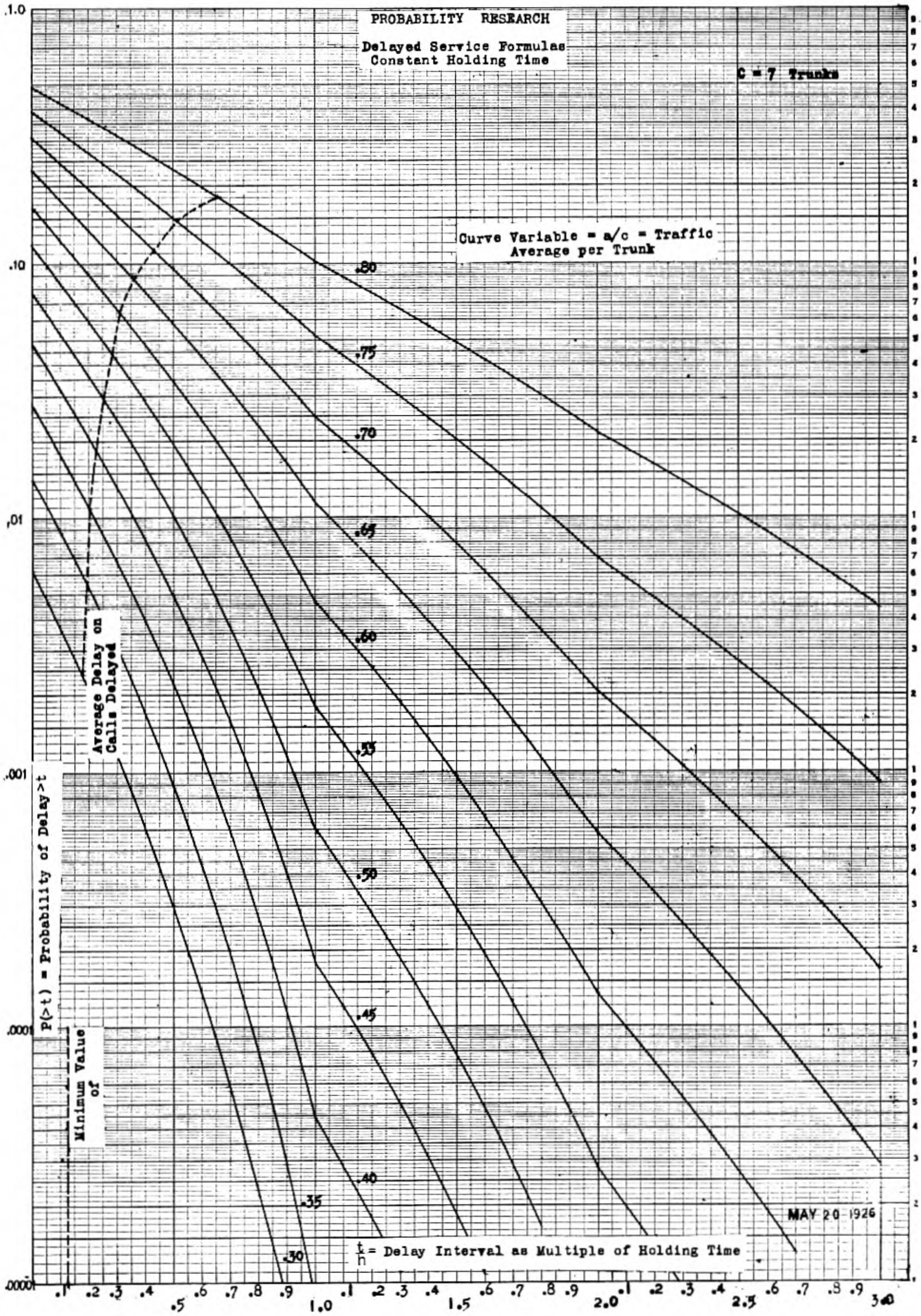

PROBABILITY RESEARCH
Delayed Service Formulas
Constant Holding Time
C = 1
Curve Variable = a/c = Traffic Average per Trunk
Average Delay on Calls Delayed
Minimum Value of
P(>t) = Probability of Delay>t
t/h = Delay Interval as Multiple of Holding Time
Feb.25,1927
CODEX BOOK CO., INC., NEW YORK, N. Y.


PROBABILITY RESEARCH
Delayed Service Formulas
Constant Holding Time
C = 2 Trunks
Curve Variable = a/c = Average per Trunk Traffic
Average Delay on Calls Delayed
Minimum Value of
P(>t) = Probability of Delay >t
t = Delay Interval as Multiple of Holding Time
MAY 20 1926
.80
.75
.70
.65
.60
.55
.50
.45
.40
.35
.30
.25
.20
.15
.10
.05
1.0
.10
.01
.001
.0001
.00001
.1 .2 .3 .4 .5 .6 .7 .8 .9 1.0 .1 .2 .3 .4 1.5 .6 .7 .8 .9 2.0 .1 .2 .3 .4 2.5 .6 .7 .8 .9 3.0

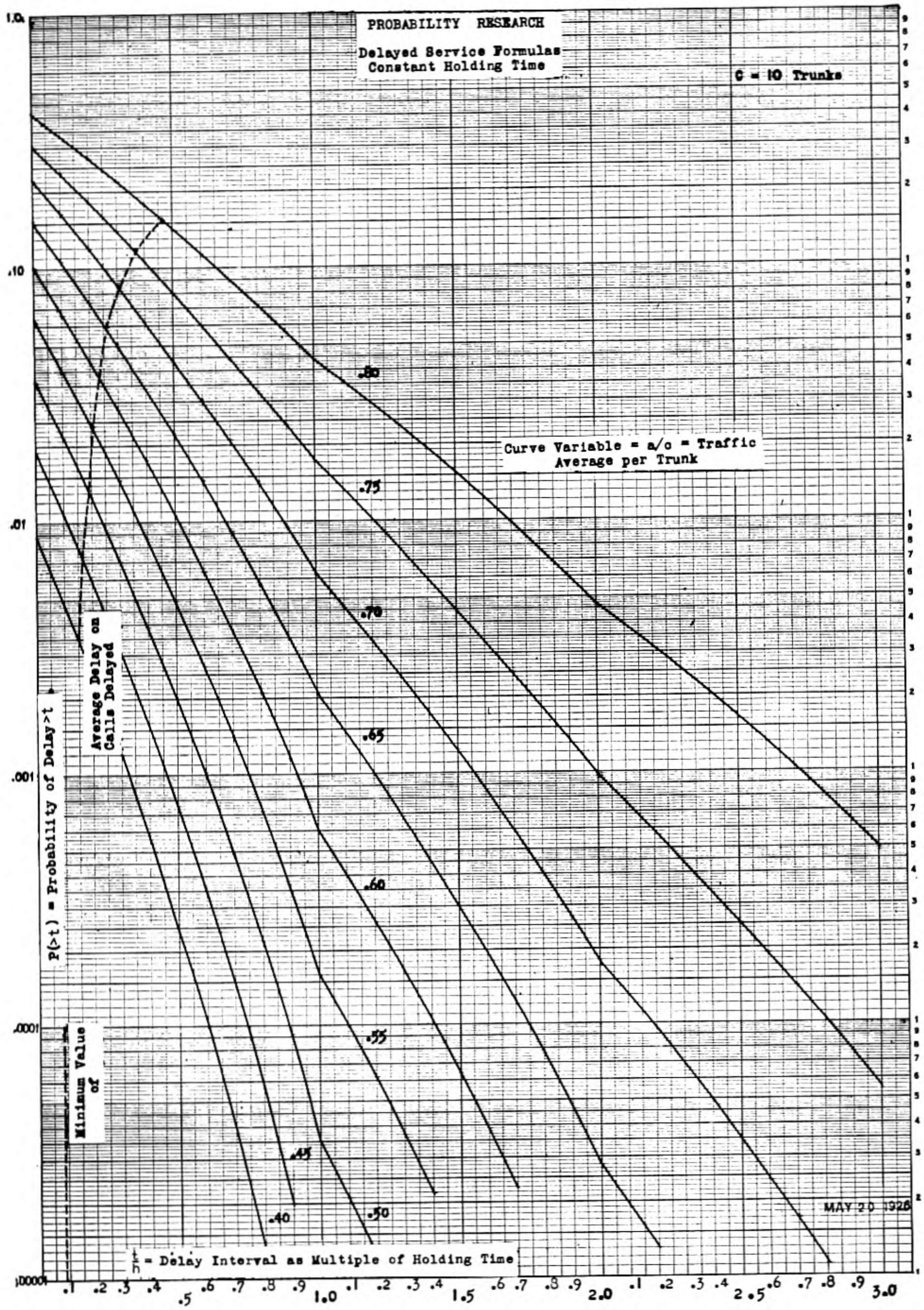

PROBABILITY RESEARCH
Delayed Service Formulas
Constant Holding Time
C = 3 Trunks
Curve Variable = a/c = Traffic Average per Trunk
Average Delay on Calls Delayed
Minimum Value of
P(>t) = Probability of Delay>t
t/h = Delay Interval as Multiple of Holding Time
.80
.75
.70
.65
.60
.55
.50
.45
.40
.35
.30
.25
.20
.15
1.0
.10
.01
.001
.0001
.00001
0
.5
1.0
1.5
2.0
2.5
3.0
MAY 20

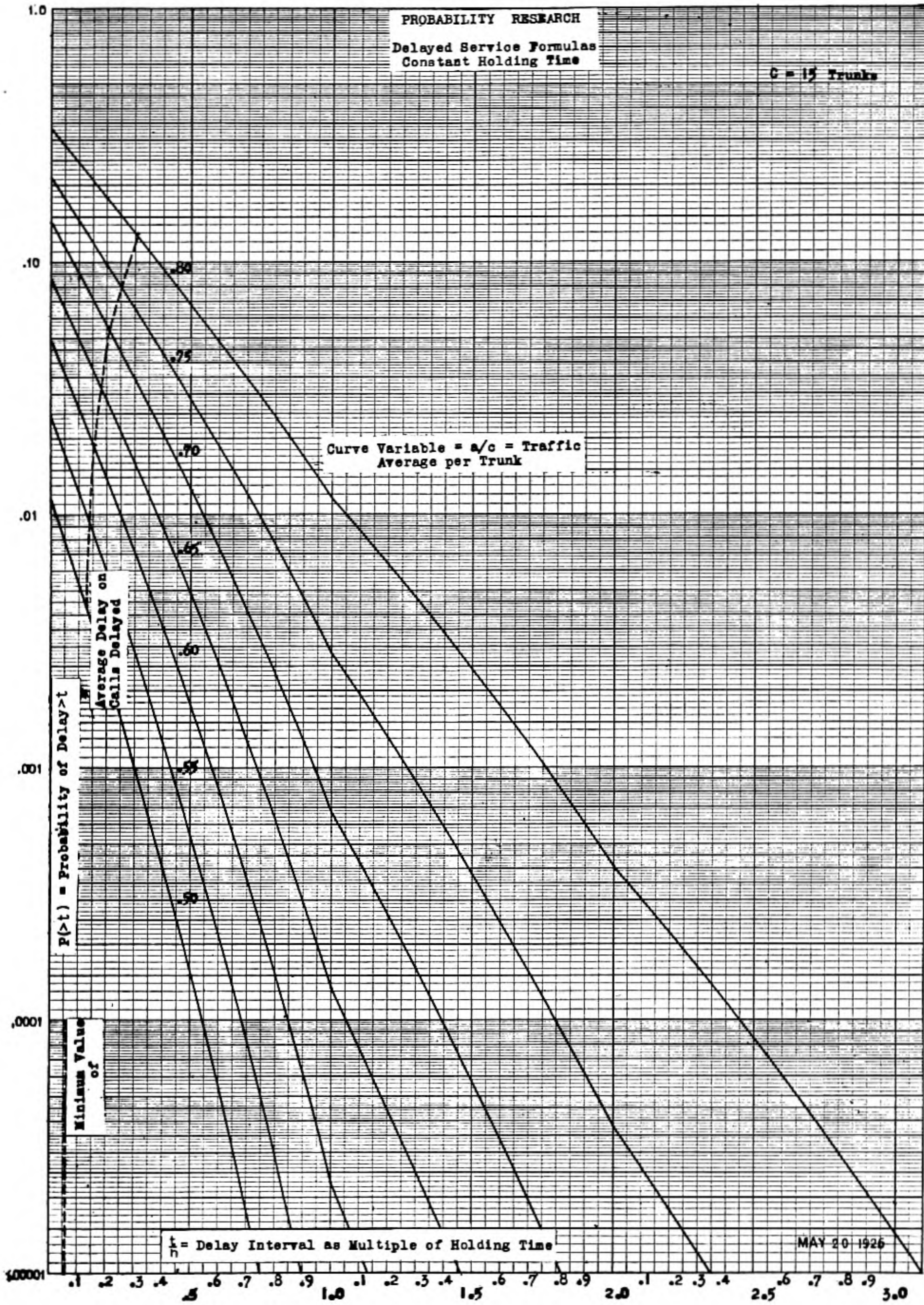

PROBABILITY RESEARCH
Delayed Service Formulas
Constant Holding Time
C = 4 Trunks
Curve Variable = a/c = Traffic Average per Trunk
Average Delay on Calls Delayed
Minimum Value of
P(>t) = Probability of Delay>t
t/h = Delay Interval as Multiple of Holding Time
.80
.75
.70
.65
.60
.55
.50
.45
.40
.35
.30
.25
.20
1.0
.10
.01
.001
.0001
.00001
.5
1.0
1.5
2.0
2.5
3.0
MAY 20 1926

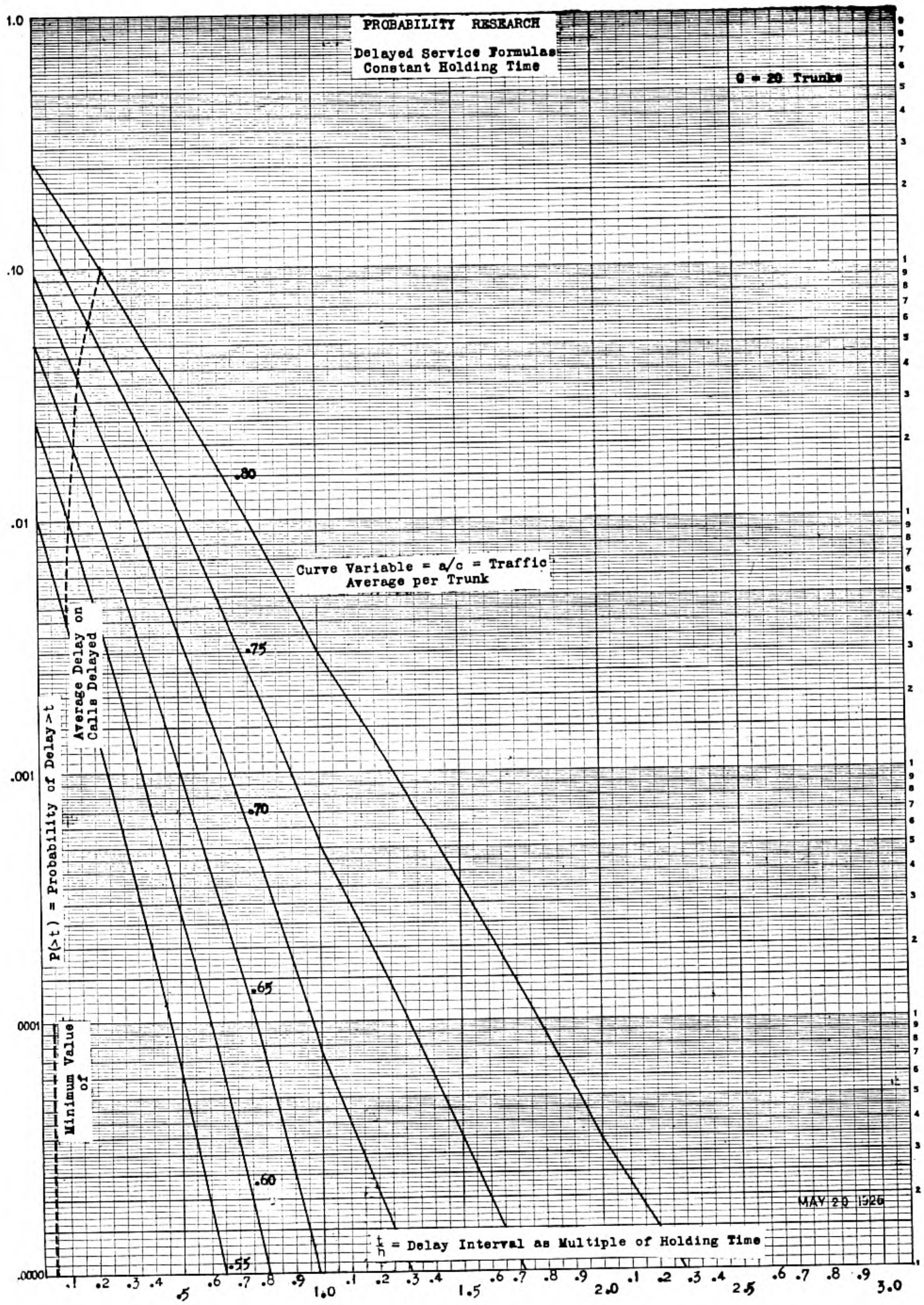

PROBABILITY RESEARCH
Delayed Service Formulas
Constant Holding Time
C = 5 Trunks
Curve Variable = a/c = Traffic Average per Trunk
.80
.75
.70
.65
.60
.55
.50
.45
.40
.35
.30
.25
Average Delays on Calls Delayed
Minimum Value of
P(>t) = Probability of Delay>t
t/h = Delay Interval as Multiple of Holding Time
1.
.1
.01
.001
.0001
.00001
.1 .2 .3 .4 .5 .6 .7 .8 .9 1.0 .1 .2 .3 .4 1.5 .6 .7 .8 .9 2.0 .1 .2 .3 .4 2.5 .6 .7 .8 .9 3.0

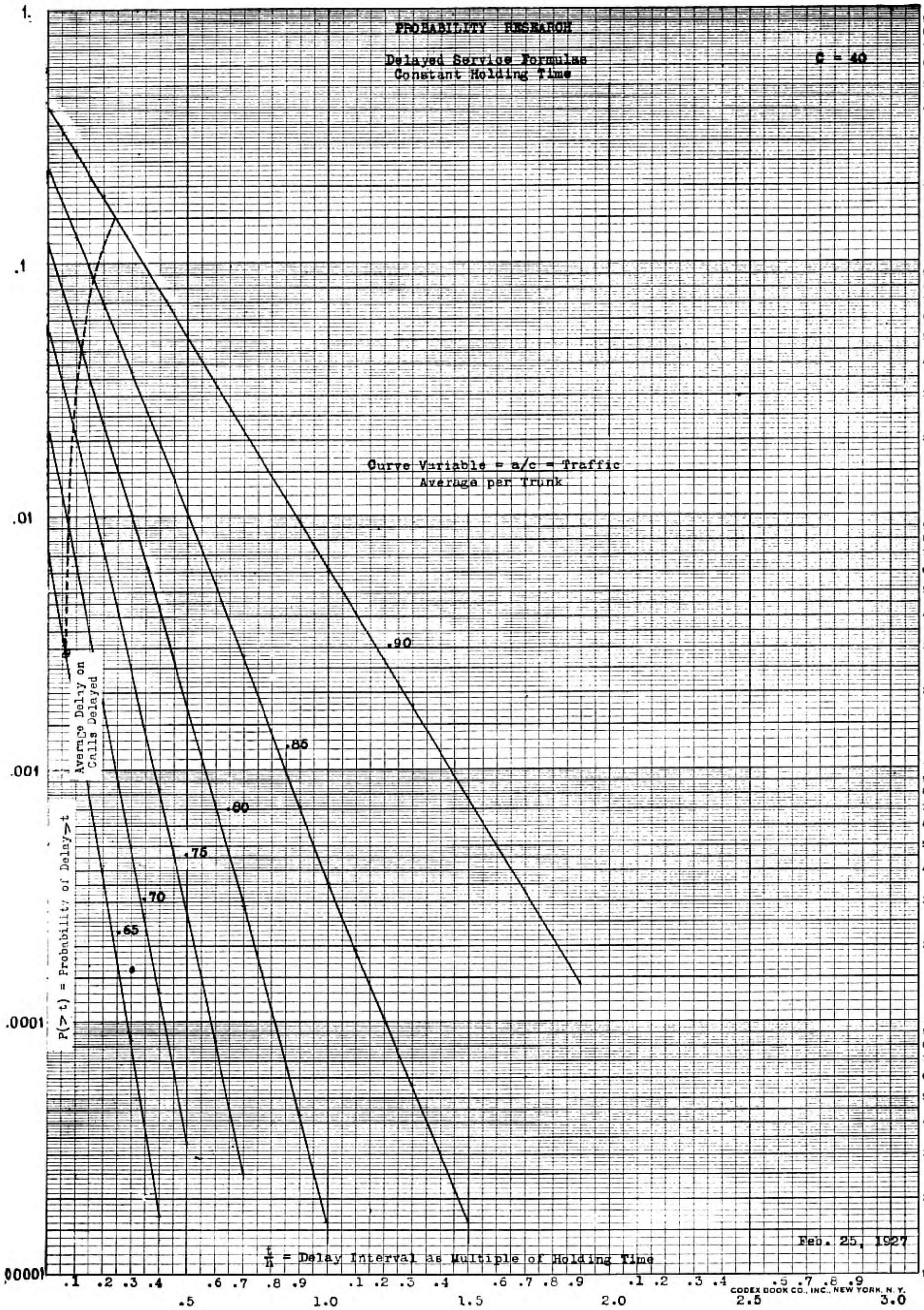


PROBABILITY RESEARCH
Delayed Service Formulas
Constant Holding Time
C = 7 Trunks
Curve Variable = a/c = Traffic Average per Trunk
Average Delay on Calls Delayed
P(>t) = Probability of Delay>t
Minimum Value of
t/h = Delay Interval as Multiple of Holding Time
.80
.75
.70
.65
.60
.55
.50
.45
.40
.35
.30
1.0
.10
.01
.001
.0001
.00001
.5
1.0
1.5
2.0
2.5
3.0
MAY 20 1926


PROBABILITY RESEARCH
Delayed Service Formulas
Constant Holding Time
C = 10 Trunks
Curve Variable = a/c = Traffic Average per Trunk
P(>t) = Probability of Delay>t
Average Delay on Calls Delayed
Minimum Value of
t/h = Delay Interval as Multiple of Holding Time
.80
.75
.70
.65
.60
.55
.50
.45
.40
1.0
.10
.01
.001
.0001
.00001
1.0
1.5
2.0
2.5
3.0
MAY 20 1926


PROBABILITY RESEARCH
Delayed Service Formulas
Constant Holding Time
C = 15 Trunks
Curve Variable = a/c = Traffic Average per Trunk
.80
.75
.70
.65
.60
.55
.50
Average Delay on Calls Delayed
P(>t) = Probability of Delay>t
Minimum Value of
t/h = Delay Interval as Multiple of Holding Time
MAY 20 1926
1.0
.10
.01
.001
.0001
.00001
.1 .2 .3 .4 .5 .6 .7 .8 .9 1.0 .1 .2 .3 .4 1.5 .6 .7 .8 .9 2.0 .1 .2 .3 .4 2.5 .6 .7 .8 .9 3.0


PROBABILITY RESEARCH
Delayed Service Formulas
Constant Holding Time
C = 20 Trunks
Curve Variable = a/c = Traffic Average per Trunk
P(>t) = Probability of Delay >t
Average Delay on Calls Delayed
Minimum Value of
t/h = Delay Interval as Multiple of Holding Time
.80
.75
.70
.65
.60
.55
1.0
.10
.01
.001
.0001
.00001
.5
1.0
1.5
2.0
2.5
3.0
MAY 20 1926


PROBABILITY RESEARCH
Delayed Service Formulas
Constant Holding Time
C = 40
Curve Variable = a/c = Traffic Average per Trunk
.90
.85
.80
.75
.70
.65
Average Delay on Calls Delayed
P(> t) = Probability of Delay > t
t/h = Delay Interval as Multiple of Holding Time
Feb. 25, 1927
1.
.1
.01
.001
.0001
.00001
.5
1.0
1.5
2.0
2.5
3.0
CODEX BOOK CO., INC., NEW YORK, N. Y.

Propagation of Periodic Currents over a System of Parallel Wires

By JOHN R. CARSON and RAY S. HOYT

SYNOPSIS: The first section of this paper is devoted to the formal mathematical theory of the propagation of periodic currents over a system of parallel wires energized at its physical terminals only. The theory developed is essentially a generalization of the classical theory of transmission over a single wire (with ground return) or over a balanced metallic circuit. The solution here given furnishes the fundamental formulas and a good deal of information regarding what takes place in a system of parallel wires; for actual calculations, however, the method of treatment is not so well adapted as that developed in the remaining sections of the paper.

The second section deals analytically with the problem of propagation over a line or a circuit exposed throughout its length to an arbitrary impressed field of force. The resulting solution is immediately applicable to problems of crosstalk and interference, and to the theory of the wave antenna.

The last two sections are devoted to the development and application of a more physical or synthetic method of treatment, based on the substitution of 'equivalent electromotive forces' for the arbitrary impressed field. This synthetic treatment, which permits of an intuitive or physical grasp of the various problems, has been found quite useful in dealing with crosstalk and interference, and also with the wave antenna. The method is illustrated (in the last section) by application to two representative problems of a diverse nature.

IN the modern telephone system, transmission takes place over a circuit which is usually in close juxtaposition to a number of parallel circuits, and which may be, and frequently is, exposed to interference from power circuits or other disturbing sources. The mathematical theory of wave propagation over such a circuit involves two problems: (1) propagation over a system of parallel wires, and (2) propagation over a wire or metallic circuit in an arbitrary impressed field of force.

The first Section of this paper is devoted to the formal mathematical theory of the propagation of periodic currents over a system of parallel wires, energized at its physical terminals only.[1] This problem is essentially a generalization of the problem of transmission over a line of uniformly distributed resistance, inductance, capacity and leakage; and involves the formulation and solution of a differential equation which may be termed the *generalized telegraph equation* in contradistinction to the well-known *telegraph equation* which characterizes transmission over a single wire (with ground return) or a balanced

[1] This is the assumption underlying ordinary transmission theory.

metallic circuit. The analysis of this problem, while furnishing the fundamental formulas and a good deal of information regarding what takes place in a system of parallel wires, is not well adapted for actual calculations, except for relatively simple systems; in particular it is not adapted to deal with the important problems of crosstalk and interference.

In Section II the problem of propagation over a circuit or line exposed throughout its length to an arbitrary impressed field of force is taken up. The resulting solution is immediately applicable to interference problems, where the field of the disturbing source is supposed known, and to the theory of the wave antenna. Moreover, as shown in Section II*a*, it is particularly well adapted to the problems of 'crosstalk,' or interference between circuits of the parallel system.

Sections I, II and II*a* furnish the formal analysis and the fundamental formulas. Sections III and IV, constituting the remainder of the paper, are devoted to the development and the application to representative problems of a more physical or synthetic treatment, in which the general theory and formulas are interpreted in terms of 'equivalent electromotive forces'; this concept permits of an intuitive or physical grasp of the various problems, and has been found quite useful in dealing with crosstalk and interference, and also with the wave antenna.

I

Propagation of Periodic Currents over a System of Parallel Wires, with Impressed Field Concentrated at Terminals

The physical system under consideration is supposed to consist of n parallel wires, numbered from 1 to n, which may either be a system of overhead wires parallel to the surface of the earth, or a multi-wire cable enclosed in a sheath. The formal analysis applies equally well to both cases; but the calculation of the circuit parameters is a matter of considerably greater difficulty in the case of the cable, due to the close juxtaposition of the wires. Even in this case, however, the circuit constants are rather easily calculable to a first approximation from the dimensions of the system; and they are, in any case, experimentally determinable.

Let $I_1, I_2 \cdots I_n$ be the currents in the n wires, which are taken as parallel to the x-axis, which is itself parallel to the surface of the earth or to the sheath (in the cable case). A steady state is assumed; that is to say, the currents are sinusoidal and involve the time t only through the common factor $\exp(i\omega t)$, where $\omega/2\pi$ is the frequency and i denotes $\sqrt{-1}$; consequently the differential operator d/dt is replaceable by $i\omega$ in accordance with the usual methods of alternating current theory.

The first equations of the problem are derived by applying the law

$$\text{curl } E = -\mu(dH/dt)$$

to a contour bounded by a length dx in the surface of the jth wire, a corresponding length dx in the surface of the earth, and two lines normal to the axis of the wire and joining the corresponding ends of the two line elements dx. This gives

$$z_{jj}I_j - E_{gj} = -\frac{dV_j}{dx} - \frac{d\phi_j}{dt}, \qquad (j = 1, 2 \cdots n). \tag{1}$$

In this set of equations, z_{jj} denotes the 'internal' impedance per unit length of the jth wire, that is, the ratio of the axial electric force at the surface of the wire to the current I_j. E_{gj} is the electric force, parallel to the axis of the wire, in the earth's surface. V_j is the line integral of the electric force from the wire to the surface of the earth, that is, the 'potential,' or 'voltage,' of the wire. Finally, ϕ_j is the magnetic flux,[2] per unit length, threading the contour.

Now, both ϕ_j and E_{gj} are linear functions of the n currents $I_1, \cdots I_n$; consequently (1) is reducible to the form [3]

$$z_{jj}I_j = -\frac{dV_j}{dx} - \sum_{k=1}^{n} Z_{jk}I_k, \qquad (j = 1, 2 \cdots n). \tag{2}$$

The calculations of the impedance functions Z_{jk} and, in particular, the effect of the finite conductivity of the earth are dealt with in detail in an earlier paper.[4] The internal impedance, z_{jj}, is of the general form $r_{jj} + i\omega l_{jj}$, where r_{jj} is the resistance of the jth wire and l_{jj} its 'internal inductance.' In the ideal non-dissipative system the mutual impedance Z_{jk} is a pure imaginary of the form $i\omega L_{jk}$, where L_{jk} is the mutual inductance between the jth and kth wires; actually, however, due to the finite conductivity of the earth and to 'proximity effect' between the wires, it is always complex and of the form $R_{jk} + i\omega L_{jk}$. A similar statement holds for the self impedance Z_{jj}. The 'proximity effect' [5] is, of course, the increased internal impedance of the wire due to the currents in the neighboring wires. It may be taken as negligible in open wire lines but is quite appreciable, at telephonic frequencies, in cable circuits.

[2] Expressed in 10^{-8} maxwells if the remaining quantities are in 'practical units.'
[3] It is to be noted that Z_{jj} does *not* include the internal impedance z_{jj} of wire j.
[4] 'Wave Propagation in Overhead Wires with Ground Return,' John R. Carson, *B. S. T. J.*, October, 1926.
[5] See 'Wave Propagation over Parallel Wires: The Proximity Effect,' John R. Carson, *Phil. Mag.*, April, 1921. Rigorously the term $z_{jj}I_j$ of equations (1) and (2) should be replaced by $\Sigma z_{jk}I_k$, the additional terms formulating the proximity effect. This effect will not be explicitly included in the following analysis and the term $z_{jk}I_k$ may be regarded as incorporated with $Z_{jk}I_k$.

Let $Q_1, \cdots Q_n$ denote the charges per unit length on the n wires; the potentials and charges are then related by the set of linear equations

$$Q_j = \sum_{k=1}^{n} q_{jk} V_k, \qquad (j = 1, 2 \cdots n), \tag{3}$$

$$V_j = \sum_{k=1}^{n} p_{jk} Q_k, \qquad (j = 1, 2 \cdots n), \tag{4}$$

in which the q and p coefficients are Maxwell's capacity and potential coefficients. They are calculable by the usual methods of electrostatics, on the assumption that all the conductors, including the earth, are of perfect conductivity.

To complete the specification of the system we have the further set of relations

$$i\omega Q_j + I_j' = -\frac{dI_j}{dx}, \qquad (j = 1, 2 \cdots n). \tag{5}$$

Here I_j' is the 'leakage' current from the jth wire; it is, in general, a linear function of the n potentials, that is,

$$I_j' = \sum_{k=1}^{n} g_{jk} V_k, \qquad (j = 1, 2 \cdots n), \tag{6}$$

where the coefficients g_{jk} depend on the geometry of the system and the conductivity of the dielectric medium. From (3), (5) and (6) we have

$$-\frac{dI_j}{dx} = \sum_{k=1}^{n} (i\omega q_{jk} + g_{jk}) V_k, \qquad (j = 1, 2 \cdots n). \tag{7}$$

This system of linear equations, when solved for the potentials, gives

$$V_j = -\frac{d}{dx} \sum_{k=1}^{n} w_{jk} I_k, \qquad (j = 1, 2 \cdots n). \tag{8}$$

For the special case where the dielectric medium surrounding the conductors is homogeneous and isotropic, the coefficient w_{jk}, which in general is obtained by solving (7), is given by

$$w_{jk} = p_{jk}/(i\omega + \delta), \tag{9}$$

where $\delta = 4\pi\sigma/\epsilon\xi$, σ and ϵ being the conductivity and specific inductive capacity of the dielectric, and ξ a constant whose value depends only on the units.[6] In many cases w_{jk} is calculable with sufficient accuracy from equation (9), so that the solution of (7) is then unnecessary.

[6] A derivation of the formula for δ is outlined shortly after equation (18) of Appendix I.

We are now prepared to write down the *generalized telegraph equation*, which is obtained by eliminating $V_1, \cdots V_n$ from equations (2) by aid of (8); it is:

$$z_{jj}I_j = \sum_{k=1}^{n}\left(w_{jk}\frac{d^2}{dx^2} - Z_{jk}\right)I_k, \qquad (j = 1, 2 \cdots n). \tag{10}$$

This system of n equations, which constitutes the *generalized telegraph equation*, will be written as:

$$\begin{aligned} m_{11}I_1 + m_{12}I_2 + \cdots + m_{1n}I_n &= 0, \\ m_{21}I_1 + m_{22}I_2 + \cdots + m_{2n}I_n &= 0, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad & \cdot \\ m_{n1}I_1 + m_{n2}I_2 + \cdots + m_{nn}I_n &= 0, \end{aligned} \tag{11}$$

where

$$m_{jk} = Z_{jk} - w_{jk}\frac{d^2}{dx^2}, \qquad (j \neq k), \tag{11.1}$$

$$m_{jj} = z_{jj} + Z_{jj} - w_{jj}\frac{d^2}{dx^2}. \tag{11.2}$$

Equations (11) are a system of n homogeneous equations; a finite solution for the currents $I_1, \cdots I_n$ therefore necessitates the vanishing of the determinant of the system, that is,

$$\begin{vmatrix} m_{11} & m_{12} & m_{13} & \cdots & m_{1n} \\ m_{21} & m_{22} & m_{23} & \cdots & m_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ m_{n1} & m_{n2} & m_{n3} & \cdots & m_{nn} \end{vmatrix} = 0. \tag{12}$$

In order to solve this equation the operator d^2/dx^2 is to be replaced by γ^2, which is equivalent to the assumption that the n currents $I_1, \cdots I_n$ involve the variable x only through the common factor exp (γx). With this substitution, equation (12) is of the nth order in γ^2 and its solution gives, in general, $2n$ values of γ, namely, $\gamma_1, \gamma_2, \cdots \gamma_n$ and $-\gamma_1, -\gamma_2, \cdots -\gamma_n$. The general solution of equations (11) is accordingly of the form

$$I_j = \sum_{k=1}^{n}(A_{jk}e^{-\gamma_k x} - B_{jk}e^{\gamma_k x}), \qquad (j = 1, 2 \cdots n). \tag{13}$$

The potentials are then determined from (8) and (13) in terms of the parameters of the system, the propagation constants, and the arbitrary constants of integration A_{jk}, B_{jk}.

By means of the relations (11) obtaining among the currents, it is easy to show that the number of independent arbitrary constants of integration is $2n$. These are determined by the $2n$ boundary conditions to be satisfied at the physical terminals of the system. In general these boundary conditions specify $2n$ relations among the impressed voltages, the terminal impedances, and the line currents and voltages. While the evaluation of the constants of integration from these $2n$ relations is formally straightforward, it is actually a matter of very considerable complexity if the system is composed of a large number of wires; furthermore, the evaluation of the propagation constants $\gamma_1, \cdots \gamma_n$ presents great difficulties in such cases.

The results of the foregoing formal analysis may be summarized as follows: In a system of n parallel wires there are in general n modes of propagation, corresponding to the n roots $\gamma_1, \cdots \gamma_n$ of the generalized telegraph equation; these may be termed the normal modes of propagation. Except when special boundary conditions obtain, the current in each and every wire is made up of component waves of all n modes of propagation, and the distribution of energy among the n modes is determined by the boundary conditions at the terminals of the wires. A characteristic and fundamental property of the normal modes of propagation is that a normal mode of propagation is the type which can exist alone. That is to say, if the boundary conditions have a particular set of values, the currents in all the wires may be made up of one mode only; unless, however, the particular conditions obtain, the currents involve components of all modes.

The existence of n modes of propagation in a system of n parallel wires follows from the fact that the determinant is of the nth order in γ^2 and therefore has n roots. In certain cases of practical importance, however, we may have multiple roots, so that the number of distinct modes of propagation is reduced. For example, in the ideal case of perfect conductors and perfect ground conductivity, $L_{jj}/p_{jj} = L_{jk}/p_{jk} = 1/c^2$ and therefore $\gamma = i\omega/c$, where c is the velocity of propagation in the medium. In this case, only one mode of propagation exists, namely, unattenuated transmission with the velocity of propagation of light in the dielectric medium; thus, for the direct wave,

$$I_j = A_j e^{-i\omega/c} \tag{14}$$

and the n constants $A_1, \cdots A_n$ are independent.

Another case of some interest is that in which the wires are all alike, so that $m_{11} = m_{22} = \cdots m_{nn} = m$ and furthermore $m_{jk} = m'$ (a condition which is partially realizable by a properly designed system of transpositions). In this case equation (12) becomes

$$\begin{vmatrix} m & m' & m' & \cdots & m' \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ m' & m' & m' & \cdots & m \end{vmatrix} = 0, \tag{15}$$

and there are only two modes, γ_1 and γ_2, corresponding to $m - m' = 0$ and $m + (n - 1)m' = 0$ respectively. The first mode obviously corresponds to metallic transmission, the second to ground return transmission. It is easily shown that the direct current waves are expressed by

$$I_j = A_j e^{-\gamma_1 x} + B e^{-\gamma_2 x}, \tag{16}$$

$$\sum A_j = 0, \tag{17}$$

with corresponding expressions for the reflected waves. The corresponding potentials are

$$V_j = \tfrac{1}{2} K_1 A_j e^{-\gamma_1 x} + n K_2 B e^{-\gamma_2 x}. \tag{18}$$

Here K_1 is the characteristic impedance of a metallic circuit composed of two wires, and K_2 is the characteristic impedance of the n wires in multiple, with the ground for return.

A case of greater practical importance is that of n balanced pairs, which is the ideal telephone transmission system. To consider this case let

$$\begin{aligned} m_{11} &= m_{22} = \cdots = m_{nn} = m_{2n,\,2n} = m, \\ m_{jk} &= m' \text{ between wires of the same pair}, \\ m_{jk} &= m'' \text{ between wires of different pairs.} \end{aligned} \tag{19}$$

In this case the determinant becomes

$$\begin{vmatrix} m & m' & m'' & m'' & \cdots & \cdots & m'' \\ m' & m & m'' & m'' & \cdots & \cdots & m'' \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ m'' & m'' & m'' & \cdots & \cdots & m & m' \\ m'' & m'' & m'' & \cdots & \cdots & m' & m \end{vmatrix} = 0. \tag{20}$$

There are therefore three modes of propagation γ_1, γ_2, γ_3 corresponding respectively to:

$$\begin{aligned} & m - m' = 0, \\ & m + m' - 2m'' = 0, \\ & m + m' + 2(n - 1)m'' = 0. \end{aligned} \tag{21}$$

The first mode of propagation corresponds to metallic transmission over a pair, the second to transmission over a four wire phantom circuit, and the third to ground return transmission. It may be easily shown that the solution for the direct waves is (writing I_j and I_j' for the currents in the two wires, j and j', of the jth pair, and V_j, V_j' for the corresponding potentials):

$$\begin{aligned} I_j &= \quad A_j e^{-\gamma_1 x} + B_j e^{-\gamma_2 x} + C e^{-\gamma_3 x}, \\ I_j' &= -A_j e^{-\gamma_1 x} + B_j e^{-\gamma_2 x} + C e^{-\gamma_3 x}, \end{aligned} \tag{22}$$

with the further condition $\sum B_j = 0$. The corresponding potentials are:

$$\begin{aligned} V_j &= \quad \tfrac{1}{2}K_1 A_j e^{-\gamma_1 x} + K_2 B_j e^{-\gamma_2 x} + 2nK_3 C e^{-\gamma_3 x}, \\ V_j' &= -\tfrac{1}{2}K_1 A_j e^{-\gamma_1 x} + K_2 B_j e^{-\gamma_2 x} + 2nK_3 C e^{-\gamma_3 x}. \end{aligned} \tag{23}$$

The characteristic impedances K_1, K_2, K_3 correspond to the three modes of propagation; from the foregoing and equation (8) they are found to have the following values:

$$\begin{aligned} K_1 &= 2(w - w')\gamma_1, \\ K_2 &= (w + w' - 2w'')\gamma_2, \\ K_3 &= \frac{1}{2n}(w + w' + 2[n - 1]w'')\gamma_3. \end{aligned} \tag{24}$$

The importance of the case just considered lies in the fact that the conditions of symmetry which obtain are those of the ideal multi-circuit telephone system. Two of the normal modes of propagation, physical and phantom circuit transmission, are those actually employed in telephone transmission and these modes can exist in any physical or phantom circuit without crosstalk or the induction of current in any other circuit. In fact the problem of crosstalk is the designing of the system, by means of transpositions, to approximate the ideal case, and the calculation of the effect of small departures from the ideal conditions of symmetry.

In investigating problems of crosstalk and of induction from foreign disturbing sources, the general formulas developed in the preceding pages do not lend themselves readily to the necessary calculations and interpretations. In the first place, calculation by means of the general formulas involves the location of the n roots of an nth order equation and thus presents the same difficulties as those encountered in the calculation of the transient oscillations of a network of n degrees of freedom; in fact the problems are mathematically the same, the space variable x of the present problem corresponding to the time variable t of the transient problem. In the second place the formulas, as they stand, are inapplicable to the important case where the circuit or line is exposed throughout its length to an arbitrary impressed disturbance. Furthermore, in the problem of crosstalk the departures from the conditions of balanced symmetry are necessarily very small, whereas the formulas are so general as to make it very difficult to introduce the essential simplifications which follow from the condition of small departures. For example, if the foregoing formulas are applied, as they stand, to transposed lines, it is necessary to set up new boundary conditions and evaluate a new set of integration constants at every transposition point, since a transposition point is a discontinuity. The difficulty of such a procedure is very great, aside from the fact that it requires as a preliminary the calculation of the n modes of propagation of the system in each transposition interval. In view of these difficulties a more powerful method of attack is required in the analytical investigation of the problems of crosstalk and of interference in general. Fortunately this is furnished by the solution of the problem dealt with in the next section: the propagation of periodic currents over wires in an arbitrary impressed field of force.

II

Propagation of Periodic Currents Over Wires in an Arbitrary Impressed Field of Force

We shall consider first the simplest case, namely, a single wire with ground return. The impressed or disturbing field is assumed to be periodic, of frequency $\omega/2\pi$, so that the problem is a steady state one and the time is involved only through the factor exp $(i\omega t)$. The resultant field is made up of two parts: first that due to the primary impressed field; and secondly that due to the current in and the charge on the wire, and the corresponding induced currents and charges in the ground. Let $f(x) = f$ denote the component of the electric force of the primary or impressed field parallel to the axis of the wire at its surface,[7] and $F(x) = F$ the line integral of the impressed or primary field from the surface of the wire to ground. It is then proved in Appendix I that the differential equation of the problem is

$$\frac{K}{\gamma}\left(\gamma^2 - \frac{d^2}{dx^2}\right) I = f, \tag{25}$$

the solution of which is [8]

$$\begin{aligned} I = e^{-\gamma x}&\left[A + \frac{1}{2K}\int_v^x dy\cdot f(y)e^{\gamma y}\right] \\ &- e^{\gamma x}\left[A' + \frac{1}{2K}\int_v^x dy\cdot f(y)e^{-\gamma y}\right]. \end{aligned} \tag{26}$$

[7] $f(x)$ is assumed to be sensibly constant over the cross-section of the wire.

[8] The lower limit, v, of integration is at our disposal. In case the line begins at $x = 0$, it may be convenient to take $v = 0$.

The corresponding potential of the wire is

$$V = Ke^{-\gamma x}\left[A + \frac{1}{2K}\int_v^x dy \cdot f(y)e^{\gamma y}\right] + Ke^{\gamma x}\left[A' + \frac{1}{2K}\int_v^x dy \cdot f(y)e^{-\gamma y}\right] + F(x). \tag{27}$$

In these equations, γ and K are the *propagation constant* and the *characteristic impedance* of the circuit [9] consisting of the wire with ground return, while A and A' are arbitrary or integration constants which are determined from the boundary conditions. It will be observed that if the arbitrary impressed field is removed ($f = F = 0$), the solution reduces to the usual form. If the terminal impedances are specified, it follows from (26) and (27) that the problem is completely solvable provided that f is specified along the wire and F at its physical terminals.

Two more general cases of practical importance will next be formulated:

(1) Balanced Pair of Wires

Let K_1, γ_1 be the characteristic impedance and propagation constant of transmission over the metallic circuit; and K_3, γ_3 the corresponding quantities for the case of the two wires in multiple, with ground return. Let f_1 and f_2 be the electric force of the primary or impressed field along the surfaces of the wires No. 1 and No. 2 respectively, and I_1 and I_2 the currents in the wires. The solution may then be written as

$$I_1 = a + c, \qquad I_2 = -a + c,$$

where

$$\begin{aligned} a &= e^{-\gamma_1 x}\left[A + \frac{1}{2K_1}\int_v^x \{f_1(y) - f_2(y)\}e^{\gamma_1 y}dy\right] \\ &\quad - e^{\gamma_1 x}\left[A' + \frac{1}{2K_1}\int_v^x \{f_1(y) - f_2(y)\}e^{-\gamma_1 y}dy\right], \\ c &= e^{-\gamma_3 x}\left[C + \frac{1}{4K_3}\int_v^x \tfrac{1}{2}\{f_1(y) + f_2(y)\}e^{\gamma_3 y}dy\right] \\ &\quad - e^{\gamma_3 x}\left[C' + \frac{1}{4K_3}\int_v^x \tfrac{1}{2}\{f_1(y) + f_2(y)\}e^{-\gamma_3 y}dy\right]. \end{aligned} \tag{28}$$

The component a corresponds to transmission over the metallic or physical circuit; while the component c corresponds to transmission over the two wires in multiple, with ground return. A and C are the

[9] It will be observed that in these equations the characteristics of the ground do not appear explicitly. They are, however, implicitly involved in K and γ of the ground return circuit.

integration constants of the direct wave, while A' and C' are those of the reflected wave. The first component, as regards the impressed force f along the wires, depends on the difference $f_1 - f_2$ at the surface of the two wires, while the second depends on the mean value $(f_1 + f_2)/2$. In the case of interference from external sources the latter is usually much the larger and consequently the induction mainly corresponds to the ground return mode of propagation, γ_3.

(2) System of n Balanced Pairs

We shall now write down the expressions for the currents in a system of n balanced pairs ($2n$ wires) when exposed to an arbitrary impressed field. The properties of this system were discussed briefly in the preceding section and formulated in equations (19), $\cdots$ (24). Let I_j and I_j' be the currents in the two wires j and j' respectively of the jth pair, and let f_j and f_j' be the corresponding impressed forces along the surfaces of the two wires, while F_j and F_j' are the corresponding line integrals of the impressed force to ground. By an extension of the previous formulas it is easy to show that the currents are made up of three components:

$$\begin{aligned} I_j &= a_j + b_j + c_j, \\ I_j' &= -a_j + b_j + c_j. \end{aligned} \tag{29}$$

If we write $\bar{f}_j = (f_j + f_j')/2$, the components a_j, b_j, c_j are given by:

$$\begin{aligned} a_j = e^{-\gamma_1 x}&\left[A_j + \frac{1}{2K_1}\int_0^x \{f_j(y) - f_j'(y)\} e^{\gamma_1 y} dy \right] \\ &- e^{\gamma_1 x}\left[A_j' + \frac{1}{2K_1}\int_0^x \{f_j(y) - f_j'(y)\} e^{-\gamma_1 y} dy \right], \end{aligned} \tag{30}$$

$$\begin{aligned} b_j = e^{-\gamma_2 x}&\left[B_j + \frac{1}{2K_2}\int_0^x \{\bar{f}_j(y) - \frac{1}{n}\sum \bar{f}_k(y)\} e^{\gamma_2 y} dy \right] \\ &- e^{\gamma_2 x}\left[B_j' + \frac{1}{2K_2}\int_0^x \{\bar{f}_j(y) - \frac{1}{n}\sum \bar{f}_k(y)\} e^{-\gamma_2 y} dy \right], \end{aligned} \tag{31}$$

$$\sum B_j = \sum B_j' = 0, \tag{32}$$

$$\begin{aligned} c_j = e^{-\gamma_3 x}&\left[C + \frac{1}{4nK_3}\int_0^x \frac{1}{n}\sum \bar{f}_k(y) e^{\gamma_3 y} dy \right] \\ &- e^{\gamma_3 x}\left[C' + \frac{1}{4nK_3}\int_0^x \frac{1}{n}\sum \bar{f}_k(y) e^{-\gamma_3 y} dy \right]. \end{aligned} \tag{33}$$

Here the a component corresponds to transmission over a pair, the b component to transmission over a phantom circuit, and the c component to ground return transmission. It will be observed that, as regards the impressed field, the a component depends on the difference $f - f'$ of the impressed force at the two sides of the circuit, while the c component depends on the mean impressed electric force averaged over the $2n$ conductors of the system; the b, or phantom component, involves the impressed field in a slightly more complicated way, depending on both the mean impressed force averaged for the two conductors of a pair and also averaged over all the $2n$ conductors of the system.

The extension of the preceding analysis to the general case of n parallel wires, in general dissimilar, is straightforward. The resulting formulas are, however, extremely complicated and for this reason, as well as their small practical utility, they will not be written down.

Formulas (25), $\cdots$ (33) are immediately applicable to the wave antenna and to interference problems in general where the impressed disturbance is supposed to be known. Their application to the problem of crosstalk, which will now be taken up, is not immediate in the same sense because here the primary disturbance which sets up crosstalk is itself a function of the unbalances among the wires composing the system. That is to say, the primary disturbance or impressed field causing the crosstalk is implicitly rather than explicitly given.

II*a*

In discussing the theory of crosstalk a representative problem will be dealt with rather than a formulation of the general problem. The types of problem encountered in practice are extremely varied, depending on whether we have to do with 'side-to-side,' 'side-to-phantom' or 'phantom-to-phantom' crosstalk, etc.; and each problem may call for special treatment. The representative problem, however, besides showing the underlying mathematical theory should serve to indicate the correct procedure in other specific problems.

Let us return to the general system of n parallel wires, dealt with in Section I, and let us suppose that two of them, say wires No. 1 and No. 2, constitute a metallic circuit which, for convenience, we shall suppose would be balanced with respect to ground if the other wires were removed. We now suppose that this metallic circuit is energized by an electromotive force impressed at $x = 0$, which in the absence of the other wires would produce a current I^0 in wire No. 1 and an equal and opposite current $-I^0$ in wire No. 2. Our problem is now to calculate the currents induced in the neighboring wires and the additional currents induced in wires No. 1 and No. 2 due to the reactions in the system.

It will be observed that, if the system were ideally balanced, no currents would be induced in the neighboring wires and we should simply have the current I^0 in the metallic circuit; in engineering language there would be no crosstalk. This is the ideal to which the correctly designed telephone system approximates by means of 'transpositions.' It is never, of course, completely realized but the approximation, as regards the neighboring metallic circuits, must be extremely close, since the allowable amount of crosstalk is very small.

Let us now return to the original system of equations for n parallel wires discussed in Section I and let us write $I_1 = I^0 + I_1'$, $I_2 = -I^0 + I_2'$ and replace $I_2, I_3 \cdots I_n$ by $I_2', I_3' \cdots I_n'$ respectively, the primes indicating that the currents are 'unbalance' currents. Similarly write $V_1 = V^0 + V_1'$, $V_2 = -V^0 + V_2'$; $Q_1 = Q^0 + Q_1'$, $Q_2 = -Q^0 + Q_2'$; and for the rest of the wires add primes to the symbols for potential and charge. Equations (2) may then be written as

$$\begin{aligned}
(z_{11} + Z_{11})I_1' + \frac{dV_1'}{dx} &= -Z_{12}I_2' - Z_{13}I_3' - Z_{14}I_4' - \cdots, \\
(z_{22} + Z_{22})I_2' + \frac{dV_2'}{dx} &= -Z_{21}I_1' - Z_{23}I_3' - Z_{24}I_4' - \cdots, \\
(z_{jj} + Z_{jj})I_j' + \frac{dV_j'}{dx} &= -(Z_{j1} - Z_{j2})I^0 - Z_{j1}I_1' \\
&\quad - Z_{j2}I_2' - Z_{j3}I_3' - \cdots, \\
&\qquad (j = 3, 4 \cdots n),
\end{aligned} \tag{34}$$

or, denoting the right hand sides of the equations by f_1', f_2', f_j', respectively,

$$\begin{aligned}
(z_{11} + Z_{11})I_1' + \frac{dV_1'}{dx} &= f_1', \\
(z_{22} + Z_{22})I_2' + \frac{dV_2'}{dx} &= f_2', \\
(z_{jj} + Z_{jj})I_j' + \frac{dV_j'}{dx} &= f_j', \\
&\qquad (j = 3, 4 \cdots n).
\end{aligned} \tag{35}$$

This set of equations in the unbalance currents $I_1', \cdots I_n'$ and unbalance potentials $V_1', \cdots V_n'$ admits of immediate interpretation. This is to the effect that the unbalance currents may be regarded as due to an impressed field characterized by an axial electric intensity

$f_j' - dF_j'/dx$ along the jth wire ($j = 1, 2, \cdots n$), and an impressed potential F_j' (line integral of impressed field from jth wire to ground), where

$$\begin{aligned} F_1' &= p_{12}Q_2' + p_{13}Q_3' + p_{14}Q_4' + \cdots, \\ F_2' &= p_{21}Q_1' + p_{23}Q_3' + p_{24}Q_4' + \cdots, \\ F_j' &= (p_{j1} - p_{j2})Q^0 + p_{j1}Q_1' + p_{j2}Q_2' + p_{j3}Q_3' + \cdots, \\ & \qquad (j = 3, 4 \cdots n). \end{aligned} \tag{36}$$

Consequently if f_j' and F_j' were known, equations (25), $\cdots$ (27) would be immediately applicable to the calculation of the unbalance currents. Inspection of equations (34), $\cdots$ (36) shows, however, that while I^0, V^0, Q^0 are supposed known, the expressions for f_j' and F_j' involve the unbalance currents and charges themselves. The solution of the equations calls therefore for a process of successive approximation, now to be discussed. While this method of solution is theoretically sound and applicable in all cases, its success in practical applications depends largely on the fact that the unbalance currents must be extremely small, compared with the primary current I^0, if the crosstalk is to be kept within tolerable limits.

Returning to equations (34), $\cdots$ (36), the first approximate solution is obtained by (1) ignoring the unbalance currents and charges in their effect on the current in the primary wires (No. 1 and No. 2), and (2) replacing $f_3', \cdots f_n'$ and $F_3', \cdots F_n'$ by

$$\begin{aligned} f_j' &= -(Z_{j1} - Z_{j2})I^0, \\ F_j' &= (p_{j1} - p_{j2})Q^0, \\ & \qquad (j = 3, 4 \cdots n). \end{aligned} \tag{37}$$

Consequently in the first approximate solution the primary current, charge and potential are I^0, Q^0, V^0, which are calculable in terms of the impressed e.m.f. and the terminal impedances for the circuit composed of the primary wires (No. 1 and No. 2) by ignoring the reaction of the other wires.[10] The unbalance currents, charges and potentials of the other wires are then calculable on the supposition that those wires are energized by the known impressed field f_j', F_j', as given by (37), which depends only on I^0 and Q^0.

The second approximate solution is obtainable by substituting the first approximate values of I_j' and Q_j' in the right hand side of equations (34) and (36) and then proceeding precisely as in the first approxi-

[10] It should be clearly understood that this particular procedure is not required and is not always followed in practice. For example, it is customary in calculating the crosstalk induced in a metallic or 'side' circuit to take into account, in the first approximate solution, the reaction between the wires making up the disturbed circuit.

mate solution. This process can, theoretically, be repeated indefinitely and successively closer approximations thereby obtained. Practically, however, even in a system of only a few wires, the process rapidly becomes prohibitively laborious and complicated, so that only the first and perhaps the second approximate solutions are practicable. Theoretically, however, the process is straight-forward and the successive approximate solutions form a convergent sequence. Fortunately, in engineering applications the allowable amount of crosstalk is so strictly limited that higher approximations than the second at most are not usually required.

It is an important and valuable property of the solution by successive approximations that the 'datum configuration' is not uniquely fixed, but is at our disposal, within limits. By 'datum configuration' is meant the assumed distribution from which the first approximate solution is derived. In the preceding the datum configuration for the primary wires is taken as

$$\begin{aligned} I_1 &= I^0 = -\ I_2, \\ Q_1 &= Q^0 = -\ Q_2, \end{aligned} \tag{38}$$

while in calculating any I_j' $(j = 3, \cdots n)$ it is assumed that the unbalance currents and charges of the other disturbed wires are zero. From the form of the equations this is certainly the natural configuration with which to start. It does not at all follow, however, that this datum configuration results in the optimum first approximate solution.

Another datum configuration which may be taken and which appears to possess practical advantages in certain cases is the following: [11]

$$\begin{aligned} I_1 &= I^0 = -\ I_2, \\ Q_1 &= Q^0 = -\ Q_2, \end{aligned} \tag{39}$$

for the primary wires, while in calculating any I_j' $(j = 3, \cdots n)$ it is assumed that the unbalance currents and *potentials* (instead of *charges*) of the other disturbed wires are zero.[12] Higher successive approximate solutions then follow the same scheme of procedure as in the first case.

The foregoing completes the formal analytical theory. The remaining sections of the paper will be devoted to the interpretation of the fundamental mathematical theory and its formulation along more physical and engineering lines, together with applications to representative problems.

[11] This is essentially the basis of the crosstalk formulas developed, in terms of a different mathematical treatment, by Dr. G. A. Campbell of the American Telephone and Telegraph Co., in his early and fundamental work on crosstalk and transposition theory.

[12] See, however, the preceding footnote as to possible modification of the datum configuration.

III

Representation of Impressed Field by Equivalent Electromotive Forces

In the present section we shall start anew with the problem dealt with in Section II, and attack it by a synthetic method, as distinguished from the analytical method employed there. While the results so derived are all deducible from the analytical theory and formulas of Section II, the synthetic or physical mode of attack has important advantages in engineering applications, in giving a physical picture of the phenomena and an intuitive grasp of the problem. In many cases it enables us to deduce results very simply, when the physical picture is well in mind, whereas the purely analytical solution may be laborious.

The essence of this synthetic method consists in replacing the known electric field impressed on the physical system by a set of equivalent electromotive forces; the current at any point in the system can then be calculated when the transfer admittances between that point and the points where the electromotive forces are situated are known or calculable (as is often the case in practical applications). For, considering any linear system containing any number m of electromotive forces inserted at any points $1, \cdots m$, it is known, from the principle of superposition, that the current I_h at any point h is a linear function of all the electromotive forces, that is,

$$I_h = \sum_{k=1}^{m} A_{hk} E_k. \tag{40}$$

The coefficient A_{hk} is called the 'transfer admittance' from k to h, because A_{hk} is equal to the ratio of I_h to E_k when all of the electromotive forces except E_k are zero. If the system contains any unilateral element (such as a one-way amplifier, for instance), A_{hk} is not in general equal to A_{kh}.

Fundamental Set of Equivalent Electromotive Forces: General Formulation

Consider any system of parallel wires situated in an arbitrary impressed field, with any number of localized admittance bridges between wires or between wires and ground. (Evidently, distributed bridged admittance can be analyzed into infinitesimal elements, and these can be regarded as localized.) The cross-sectional dimensions of the wires are assumed to be small enough so that the axial (longitudinal) impressed electric force is sensibly constant over each cross-section.

The electric constituent of the impressed field is assumed to be specified at every point along the wires by the impressed axial electric force and the impressed potential. At any point x in any wire, h, the impressed axial electric force will be denoted by $f_h(x)$ and the impressed potential by $F_h(x)$; these are to be regarded as arbitrary functions of x, and may even be discontinuous.

The following set of electromotive forces is easily seen to be equivalent to the above-specified arbitrary impressed field, in the sense of producing the same currents and charges. This set will be termed the 'fundamental' set of equivalent electromotive forces; for, from the physical viewpoint of this paper, it is in fact the fundamental set.[13]

(*A*) In each wire a distributed axial electromotive force whose value, per unit length, at each point is equal to the impressed axial electric force there; thus, at any point x in wire h, an electromotive force $f_h(x)dx$ in the differential length dx.

(*B*) At each point where the impressed potential is discontinuous, an axial electromotive force equal to the decrement in the impressed potential there; thus, at any point of discontinuity $x = u$ in any wire h, an electromotive force equal to $-\Delta F_h(u) = F_h(u-) - F_h(u+)$.

(*C*) In each bridge an electromotive force equal to the impressed voltage in that bridge; thus, in a bridge at any point $x = b$, from wire h to any other wire k (or to ground), an electromotive force equal to $F_h(b) - F_k(b)$.

(*D*) In case a point $x = b$ where a bridge is situated coincides with a point $x = u$ where the potential $F_h(x)$ impressed on wire h is discontinuous, the corresponding electromotive forces are as follows: Axial electromotive forces equal to $F_h(b-)$ and $-F_h(b+)$ at points $b-$ and $b+$ respectively in wire h; no electromotive force in the bridge itself, which is connected to the point b situated between $b-$ and $b+$ in wire h.[13a]

[13] For a one-wire line and for a balanced two-wire line, five other sets of equivalent electromotive forces are formulated in a later subsection.

[13a] By supposing points b and u to be not quite coincident, say $b = u-$ or $b = u+$, item (*D*) can be derived by first applying items (*B*) and (*C*) and then applying the 'branch-point theorem' formulated in the second paragraph following equation (75).

A further application of the 'branch-point theorem' yields for item (*D*) the following alternative set of electromotive forces: Axial electromotive forces each equal to $[F_h(b-) - F_h(b+)]/2$ at $b-$ and $b+$ in wire h; an electromotive force equal to $[F_h(b-) + F_h(b+)]/2$ in the bridge at b. Clearly, this set reduces to (*C*) when $F_h(x)$ is continuous at $x = b$, and it reduces to (*B*) when there is no bridge at $x = u$.

A physical verification of the correctness of the foregoing set of equivalent electromotive forces can be obtained by starting with the given system, situated in the specified arbitrary impressed field (but not otherwise energized), and then inserting in the wires and bridges a set of electromotive forces, which will be termed the 'annulling electromotive forces,' such as to annul all currents in the wires and bridges. The resultant axial electric force in the wires will then be zero, and furthermore the wires will be uncharged; hence the inserted axial electric force must be equal and opposite to the impressed axial electric force. Since the wires are uncharged their potentials will be those of the impressed field; hence, since no current flows in the bridges, the electromotive forces inserted in the bridges must be equal and opposite to the voltages of the impressed field at the bridges. Evidently the negatives of the annulling electromotive forces constitute a set of electromotive forces equivalent to the impressed field; for, insertion of the negatives of the annulling electromotive forces restores the system to its original state, in which it is acted on by only the original impressed field.

From the nature of this demonstration it is seen that the 'fundamental set' of equivalent electromotive forces is not limited to a system of parallel horizontal wires. In the general case, where the wires are neither straight nor parallel nor horizontal, x (and hence u and b) is to be interpreted as being the 'intrinsic coordinate' of a point in the particular wire contemplated, that is, the distance measured along that wire from any arbitrary fixed point therein. Thus, for wires h and k respectively, x becomes x_h and x_k, which in general are independent of each other.

For the case of a one-wire line, an analytical derivation of this set of equivalent electromotive forces is given in a later subsection by interpretation of the fundamental differential equations of the line.

A One-Wire Line in an Arbitrary Impressed Field

As indicated by Fig. 1, the line extends from $x = 0$ to $x = s$, and is terminated in impedances Z_0 and Z_s respectively. γ denotes the propagation constant per unit length, and K the characteristic impedance.[14] The direct leakage admittance from the wire to ground, per unit length, is denoted by Y'; this is the generalization of a mere leakage conductance.[31]

The impressed field is specified by the functions $f(x)$ and $F(x)$; $f(x)$ denoting the impressed axial electric force and $F(x)$ the impressed

[14] Given by formulas (12) and (11) of Appendix I.

potential, at any point x of the wire. For generality, $F(x)$ is assumed to be discontinuous at any point $x = u$ by the increment

$$\Delta F(u) = F(u+) - F(u-).$$

The problem is to calculate the current $I(x)$ produced at any point x by the impressed field.

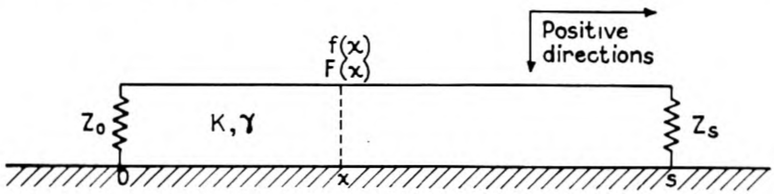


Fig. 1.

Case 1: General Case

The current $I(x)$ is the sum of two constituents: $j(x)$ due to the impressed axial electric force, and $J(x)$ due to the impressed potential. Formulas for these constituents will now be written down by aid of the fundamental set of equivalent electromotive forces formulated in the preceding subsection. Thus [15]

$$j(x) = \int_0^s A(x, y) f(y) dy, \tag{41}$$

$A(x, y)$ denoting the transfer admittance between points x and y. $J(x)$ itself consists of four constituents: $J_0(x)$ and $J_s(x)$, originating in the terminal impedances Z_0 and Z_s respectively, $J_{0s}(x)$ originating in the direct leakage admittance of the whole line 0-s, and $J_u(x)$ originating at the point $x = u$ where $F(x)$ is discontinuous. Thus

$$J_0(x) = - A(x, 0) F(0), \tag{42}$$

$$J_s(x) = A(x, s) F(s), \tag{43}$$

$$J_{0s}(x) = \int_0^s Y' F(y) B(x, y) dy, \tag{44}$$

$$J_u(x) = - A(x, u) \Delta F(u), \tag{45}$$

$B(x, y)$ denoting a current transfer factor representing that fraction

[15] With regard to the analytical evaluation of the integrals, attention should perhaps be called to the fact that the integrand may be discontinuous or may change its functional form at one or more points within the range of integration; whence the integral must be broken up into a sum of integrals.

of the current contribution originating in the direct leakage admittance element $Y'dy$ at y, which reaches point x.

Case 2: Terminal Impedances Equal to Characteristic Impedances

Here we have

$$Z_0 = Z_s = K, \tag{46}$$

$$A(x, y) = \frac{1}{2K} e^{-\gamma|x-y|}, \tag{47}$$

$$B(x, y) = \mp \tfrac{1}{2} e^{-\gamma|x-y|}, \qquad y \lesseqgtr x, \tag{48}$$

whence

$$j(x) = \frac{1}{2K} \int_0^s e^{-\gamma|x-y|} f(y) dy \tag{49}$$

$$= \frac{1}{2K} \int_0^x e^{-\gamma(x-y)} f(y) dy + \frac{1}{2K} \int_x^s e^{-\gamma(y-x)} f(y) dy, \tag{50}$$

$$J_0(x) = -\frac{F(0)}{2K} e^{-\gamma x}, \tag{51}$$

$$J_s(x) = \frac{F(s)}{2K} e^{-\gamma(s-x)}, \tag{52}$$

$$J_{0s}(x) = -\frac{Y'}{2} \int_0^x F(y) e^{-\gamma(x-y)} dy + \frac{Y'}{2} \int_x^s F(y) e^{-\gamma(y-x)} dy, \tag{53}$$

$$J_u(x) = -\frac{\Delta F(u)}{2K} e^{-\gamma|x-u|}. \tag{54}$$

A Balanced Two-Wire Line in an Arbitrary Impressed Field

Because the metallic circuit here contemplated is balanced, its treatment can be made formally the same as the treatment of the one-wire line in the preceding subsection.

This fact is immediately evident in two special cases of the impressed electric field (potential and axial electric force): (1) the case where the impressed electric field has equal values at the two wires, and (2) the case where it has equal but opposite values at the two wires.

The general case where the impressed field at the two wires has any values can be treated as a superposition of the two special cases just

ERRATA: Propagation of Periodic Currents Over a System of Parallel Wires—J. R. Carson and R. S. Hoyt

Bell System Technical Journal, July, 1927

Page 514, Equations (48), (49) and (54) *should read:*

$$B(x, y) = \mp \tfrac{1}{2} e^{-\gamma |x-y|}, \qquad y \lesseqgtr x, \tag{48}$$

whence

$$j(x) = \frac{1}{2K} \int_0^s e^{-\gamma |x-y|} f(y) dy \tag{49}$$

$$J_u(x) = -\frac{\Delta F(u)}{2K} e^{-\gamma |x-u|}. \tag{54}$$

Page 519, Line 1: *read* "Set 2 (Fig. 7)" *for* "Set 4 (Fig. 8)."

Page 530, Equation (91) *should read:*

$$V_{k\pi} = -\sum_{h=1}^{n} W_{kh} \left(\frac{dI_h}{dx} + \sum_{r=1}^{n} X_{hr} F_r \right), \qquad (k = 1, \cdots n), \tag{91}$$

mentioned, by the simple device of resolving the impressed field at each wire into two constituents one of which has equal values at the two wires while the other has equal but opposite values at the two wires. This resolution is always possible, for it is merely in accordance with the following pair of algebraic identities:

$$\eta_1 = \tfrac{1}{2}(\eta_1 + \eta_2) + \tfrac{1}{2}(\eta_1 - \eta_2), \tag{55}$$

$$\eta_2 = \tfrac{1}{2}(\eta_1 + \eta_2) - \tfrac{1}{2}(\eta_1 - \eta_2). \tag{56}$$

Although η_1 and η_2 may in general denote any two quantities whatever, in the present application they refer to the impressed electric field at the two wires No. 1 and No. 2 of the contemplated two-wire line. It is convenient to introduce the symbols η_c and η_a defined by the equations

$$\eta_c = \tfrac{1}{2}(\eta_1 + \eta_2), \tag{57}$$

$$\eta_a = \eta_1 - \eta_2, \tag{58}$$

so that the resolutions (55) and (56) of η_1 and η_2 can be written in the more compact forms

$$\eta_1 = \eta_c + \tfrac{1}{2}\eta_a, \tag{59}$$

$$\eta_2 = \eta_c - \tfrac{1}{2}\eta_a. \tag{60}$$

η_c and η_a will be termed respectively the mode-*c* and mode-*a* constituents of the impressed field, because they give rise to mode-*c* and mode-*a* effects respectively; mode-*c* effects being defined as those which are equal in the two wires, mode-*a* effects as those which are equal but opposite in the two wires—as discussed in connection with equations (28). From (57) and (58) respectively it will be noted that the mode-*c* effects and the mode-*a* effects depend respectively on the average and on the difference of the impressed fields at the two wires.

As in treating the one-wire line (in the preceding subsection), so also in treating the balanced two-wire line (in the present subsection) it is usually advantageous to deal separately with the axial electric force and the potential of the impressed field. Furthermore, in the case of the two-wire line each of these constituents of the impressed field is to be resolved into two modes, *c* and *a*, in the manner represented by equations (59) and (60) together with (57) and (58).

Owing to the balance (bilateral symmetry) of the assumed two-wire line, the mode-*c* constituent η_c of the impressed field will produce only mode-*c* effects, and the mode-*a* constituent only mode-*a* effects. Thus, η_c will produce equal currents I_c and I_c in the two wires, while η_a will produce equal but opposite currents I_a and $-I_a$ in the two wires. The total mode-c current $2I_c$ along the two wires in parallel is calculable from η_c through the mode-c parameters (γ_c, K_c, and terminal impedances) of the system; while the mode-a or loop current (I_a and $-I_a$ in the two wires respectively) is calculable from η_a through the mode-a parameters (γ_a, K_a, and terminal impedances). The connection of each current constituent with the corresponding field constituent, through the corresponding parameters, is formally the same as for the one-wire line (treated in the preceding subsection).

Finally, it may be remarked that the assumption of balance (bilateral symmetry) for the two-wire line is essential to the above simplicity; for otherwise each mode of the impressed field would produce components of both modes of effects, instead of only the appropriate single mode of effects.

Illustrative Special Case

For illustration it will suffice to choose the simple case of a balanced two-wire line terminated at each end in its mode-c and mode-a characteristic impedances simultaneously. That is, the line consisting of the two wires in parallel, with ground return, is terminated at each end in the mode-c characteristic impedance K_c; while the loop circuit is terminated at each end in the mode-a characteristic impedance K_a. (Evidently these two modes of terminating can be simultaneously accomplished by means either of a balanced T-network or of a balanced Π-network at each end.)

Let $f_1(x)$ and $f_2(x)$ denote the axial impressed electric forces at any point x in wires No. 1 and No. 2 respectively; and let them be resolved into mode-c and mode-a constituents $f_c(x)$ and $f_a(x)$, respectively, such that

$$f_c(x) = \tfrac{1}{2}[f_1(x) + f_2(x)], \tag{61}$$

$$f_a(x) = f_1(x) - f_2(x), \tag{62}$$

in accordance with equations (57) and (58). Similarly, let $F_1(x)$ and $F_2(x)$ denote the impressed potentials at point x; and let them be resolved likewise, so that

$$F_c(x) = \tfrac{1}{2}[F_1(x) + F_2(x)], \tag{63}$$

$$F_a(x) = F_1(x) - F_2(x). \tag{64}$$

Thus, formulas (49), $\cdots$ (54) of the one-wire line are seen to be formally applicable to the balanced two-wire line, for calculating separately the two modes of currents. This is with the understanding that they give the sum of the mode-c currents in the two wires, hence twice the mode-c current in each wire; and that they give the loop current (the current circulating in the metallic loop), which is equal to the mode-a current in one of the wires and hence to the negative of the mode-a current in the other wire.

Six Different Sets of Equivalent Electromotive Forces for a One-wire[16] *Line in an Arbitrary Impressed Field*

The physical system here contemplated (Fig. 2 below) is a one-wire transmission line consisting of a uniform horizontal straight wire situated in an arbitrary impressed field and terminated in any arbitrary impedances to ground. The wire extends from $x = 0$ to $x = s$; the arbitrary terminal impedances[17] are denoted by Z_0 and Z_s. The arbitrary impressed field is specified, at each point of the wire, by the impressed axial electric force $f(x)$ and the impressed potential $F(x)$, as previously.

Six different sets of 'equivalent electromotive forces' are formulated in the early part of this subsection; while their derivations are briefly outlined toward the latter part. Set 1 will be recognized as a particular case of the 'fundamental' set already formulated in the early part of Section III. The five remaining sets are derived from Set 1. In the actual formulations of these various sets of equivalent electromotive forces, the impressed potential $F(x)$ is assumed to be a continuous function of x; the extension to the case where $F(x)$ is discontinuous is a simple matter and is formulated in connection with equations (65) and (66).

In the following diagrams (Figs. 2, $\cdots$ 11) it is found convenient to represent any localized electromotive force by the conventional battery-symbol. This symbol is intrinsically directional; the longer of the two plates is to be regarded as at the higher potential, so that there is an internal rise of potential in passing through the symbol from the shorter to the longer plate.

In some of the figures the actual line is represented as replaced by the corresponding artificial line composed of differential elements, each of length dx. (For clearness, the line is represented as composed of only a small number of such elements.)

The letters Z, Y, Y', Y^0 denote certain line parameters per unit length, as follows: Z and Y respectively denote the 'complete series impedance' and the 'complete shunt admittance' or, briefly, the 'series impedance' and the 'shunt admittance.' These may be regarded as defined by the equations

$$Z = \gamma K, \qquad Y = \gamma / K,$$

[16] The case of a balanced two-wire line is outlined in the next subsection.
[17] See also the remarks under the subheading following shortly after equation (67).

γ denoting the propagation constant of the line per unit length, and K the characteristic impedance. Or they may be regarded as defined by the differential equations

$$- dV/dx = ZI, \qquad - dI/dx = YV,$$

characterizing the line when there is no impressed field present. Y' denotes the 'direct leakage admittance' and Y^0 the 'basic shunt admittance,' the latter defined as being the value of Y when $Y' = 0$, whence $Y = Y^0 + Y'$. On referring to equations (12), (11), (8), (1) of Appendix I, and also to equations (2) and (7) in Section I, it is seen that [18]

$$Z = z + i\omega L, \qquad Y = G + i\omega C,$$
$$G = G^0 + Y', \qquad Y^0 = G^0 + i\omega C.$$

The various sets of equivalent electromotive forces remain valid even when the line parameters are functions $Z(x)$, $Y(x)$, etc., of position x along the system. For the 'fundamental' set this fact can be readily seen by reference to the formulation and verification of the fundamental set, in the early part of Section III.

As indicated by the arrows, the positive axial (longitudinal) direction is the direction of increasing x, and the positive vertical direction is downward.

Six Different Sets of Equivalent Electromotive Forces

Set 1 (Fig. 4)

(A) In the wire, a distributed electromotive force, $f(x)dx$ in each differential length dx.

(B) In the distributed direct leakage admittance, a distributed electromotive force, $F(x)$ in each differential element $Y' \cdot dx$ of direct leakage admittance.

(C) In the termial impedances Z_0 and Z_s, electromotive forces $F(0)$ and $F(s)$ respectively.

From the physical viewpoint of the present paper, Set 1 is the fundamental set of equivalent electromotive forces.

This set is particularly simple when there is no direct leakage admittance ($Y' = 0$), for then it reduces to merely the axial constituents (A) and the terminal constituents (C).

[18] Thus Z, unsubscripted, includes the internal impedance $z = z_w + z_g$ of the circuit, and hence is to be sharply distinguished from the double-subscripted Z occurring frequently in this paper; for, as remarked in connection with equation (2), Z_{jj} does not include the internal impedance z_{jj} of wire j, whence it is seen that $Z = Z_{jj} + z_{jj}$ for wire j.

Set 4 (Fig. 8)

(A) In the wire, a distributed electromotive force, $[f(x) + (Y'/Y) \times dF(x)/dx]dx$ in each differential length dx.

(B) In the terminal impedances Z_0 and Z_s, electromotive forces $(1 - Y'/Y)F(0)$ and $(1 - Y'/Y)F(s)$ respectively.

Set 2 is distinguished by containing no electromotive forces in the shunt admittance even when the direct leakage admittance Y' is not negligible. Thus Set 2 with the direct leakage admittance not negligible is *formally* as simple as Set 1 with the direct leakage admittance negligible. However, in Set 2 the element of axial electromotive force is a much more complicated function than in Set 1.

Set 3 (Fig. 9)

(A) In the wire, a distributed electromotive force, $[f(x) + dF(x)/dx]dx$ in each differential length dx.

(B) In the distributed basic shunt admittance, a distributed electromotive force, $- F(x)$ in each differential element Y^0dx of the distributed basic shunt admittance.

In Set 3 it should be noted that the electromotive force $- F(x)$ is in the basic shunt admittance element Y^0dx, not in the complete shunt admittance element Ydx.

It will be observed that this set contains no electromotive forces in the terminal impedances.

When the ground is a perfect conductor, so that $f_g(x) = 0$, the differential element of axial electromotive force in this set reduces to merely $- [d\Phi(x)/dt]dx$, as is shown by equation (2) of Appendix I, $\Phi(x)$ denoting the impressed magnetic flux.

Set 4 (Fig. 8)

(A) In the wire, a distributed electromotive force, $[f(x) + (1 + Y'/Y)dF(x)/dx]dx$ in each differential length dx.

(B) In the distributed complete shunt admittance, a distributed electromotive force, $- F(x)$ in each differential element Ydx of the complete shunt admittance.

(C) In the terminal impedances Z_0 and Z_s, electromotive forces, $- (Y'/Y)F(0)$ and $- (Y'/Y)F(s)$ respectively.

In Set 4 it should be noted that the electromotive force $- F(x)$ is in the complete shunt admittance element Ydx.

The differential element of axial electromotive force in this set does not reduce to $- [d\Phi(x)/dt]dx$ when the ground is a perfect conductor $(f_g(x) = 0)$ unless also the direct leakage admittance is zero $(Y' = 0)$.

Set 5 (Fig. 6)

(A) In the wire, a distributed electromotive force, $f(x)dx$ in each differential length dx.

(B) In the distributed complete shunt admittance, a distributed electromotive force, $(Y'/Y)F(x)$ in each differential element Ydx of the complete shunt admittance.

(C) In the terminal impedances Z_0 and Z_s, electromotive forces $F(0)$ and $F(s)$ respectively.

In Set 5 it should be noted that the electromotive force $(Y'/Y)F(x)$ is in the complete shunt admittance element Ydx.

It will be observed that Set 5 (Fig. 6) is the same as Set 1 (Fig. 4) as regards the axial and the terminal electromotive forces.

Set 6 (Fig. 11)

(A) At any arbitrary fixed point $x = a$ in the wire, an axial electromotive force G_a,

$$G_a = \int_0^s \left[f(x) + \frac{dF(x)}{dx} \right] dx.$$

(B) In each differential element Ydx of the distributed complete shunt admittance, a distributed electromotive force E_x,

$$E_x = \int_0^x \left[f(x) + \frac{Y'}{Y}\frac{dF(x)}{dx} \right] dx, \qquad x < a,$$

$$E_x = -\int_x^s \left[f(x) + \frac{Y'}{Y}\frac{dF(x)}{dx} \right] dx, \qquad x > a.$$

(C) In the terminal impedances Z_0 and Z_s, electromotive forces $(1 - Y'/Y)F(0)$ and $(1 - Y'/Y)F(s)$ respectively.

Set 6 is perhaps mainly of academic interest.

Two limiting cases of Set 6 may be noted, corresponding to $a = 0$ and $a = s$ respectively, each characterized by containing no *internal* axial electromotive force: for when $a = 0$ the axial electromotive force G_a can be combined with the terminal electromotive force $(1 - Y'/Y)F(0)$ in the terminal impedance Z_0, and when $a = s$ it can be combined with $(1 - Y'/Y)F(s)$ in Z_s.

Extension to the Case where the Impressed Potential is Discontinuous

In the foregoing formulations of Sets 1, $\cdots$ 6 of equivalent electromotive forces it has been assumed that the impressed potential $F(x)$ is a continuous function of x throughout the length of the line.

Suppose now, for greater generality, that the impressed potential $F(x)$ is discontinuous at any point $x = u$ by the increment

$$\Delta F(u) = F(u+) - F(u-).$$

Then (as shown in the next paragraph), for the particular differential element which contains the point u, the quantities $f(x)dx$ and $[dF(x)/dx]dx$ must be replaced by $-\Delta F(u)$ and $\Delta F(u)$ respectively; that is,

$$f(u)du = -\Delta F(u), \tag{65}$$

$$\frac{dF(u)}{du}du = \Delta F(u). \tag{66}$$

Equations (65) and (66) can be obtained, by a limiting process, from equation (2) of Appendix I, which for the present purpose will be written in the form

$$f(x)dx = -\frac{dF(x)}{dx}dx - \frac{d\Phi(x)}{dt}dx + f_0(x)dx. \tag{67}$$

It will be recalled, from Appendix I, that this equation was derived by applying the second curl law to a differential rectangle extending from x to $x + dx$; but x may equally well be a point within the differential segment dx, and for the present purpose it will be so regarded. The limiting process now consists in letting $dF(x)/dx$ approach infinity while dx approaches zero, but in such a way that the product $[dF(x)/dx]dx$ approaches a preassigned finite value, denoted by $\Delta F(x)$. Then, in the limit, the last two terms on the right side of (67) vanish so that (67) reduces to

$$f(x)dx = -\Delta F(x).$$

Thus we obtain equations (65) and (66), where u denotes, for distinction, the particular value of x at which $F(x)$ is discontinuous.

Remarks on the Terminal Impedances and the Equivalent Electromotive Forces in Them

The arbitrary terminal impedances Z_0 and Z_s (Fig. 2) need not actually be localized. They may, for instance, be the impedances offered by other lines to which the given line 0-s may be connected, and these other lines may themselves be situated in arbitrary impressed fields; in particular, 0-s may be merely a segment, of any length, forming part of a given line in an arbitrary impressed field.

From this broad view, any 'equivalent electromotive forces' situated in the terminal impedances Z_0 and Z_s may advantageously be regarded as being situated in the ends of the line itself (that is, in the end-points $x = 0$ and $x = s$), these electromotive forces being then regarded as pertaining primarily to the line-segment 0-s rather than to the terminal

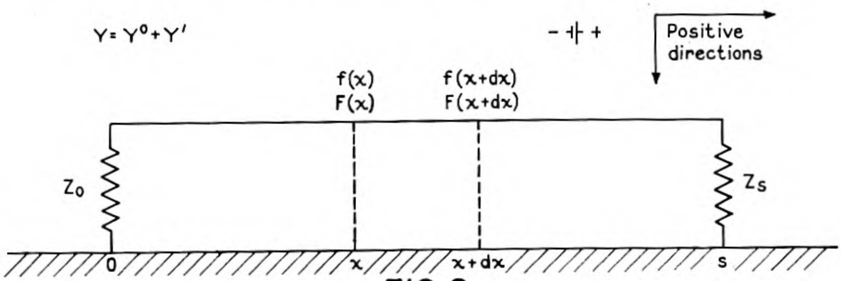


FIG. 2

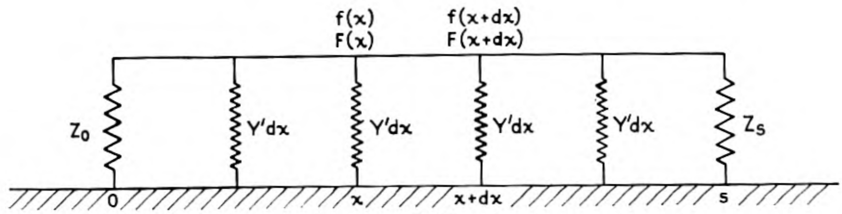


FIG. 3

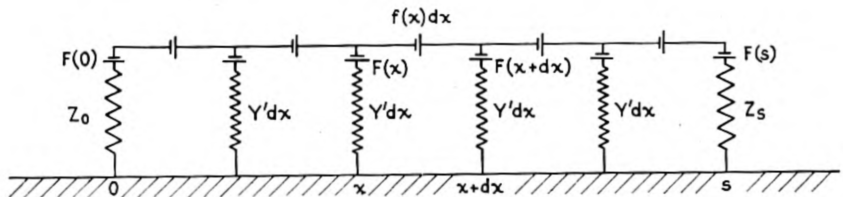


FIG. 4 (SET 1)

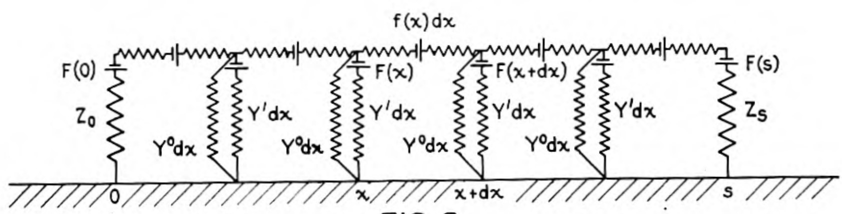


FIG. 5

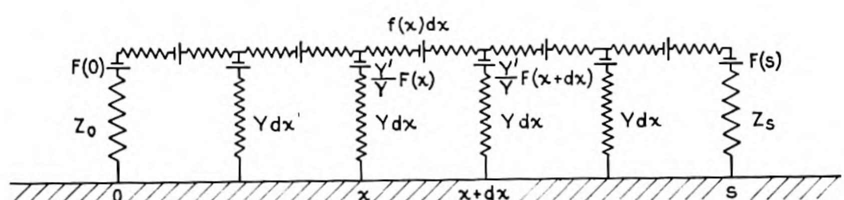


FIG. 6 (SET 5)

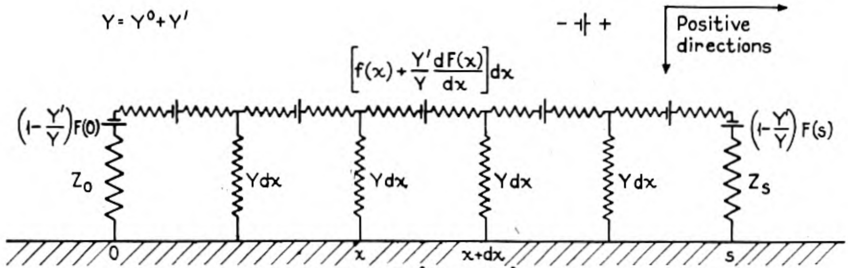


FIG. 7 (SET 2)

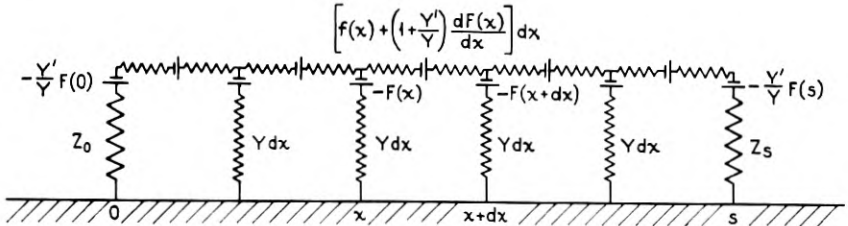


FIG. 8 (SET 4)

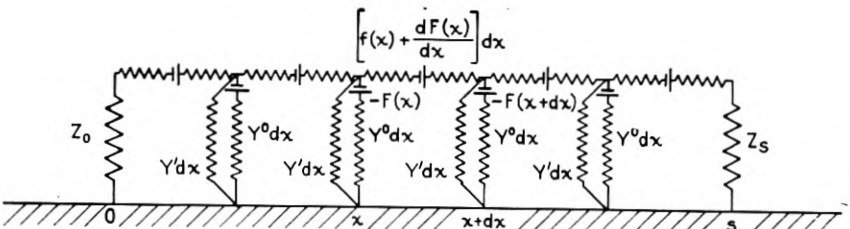


FIG. 9 (SET 3)

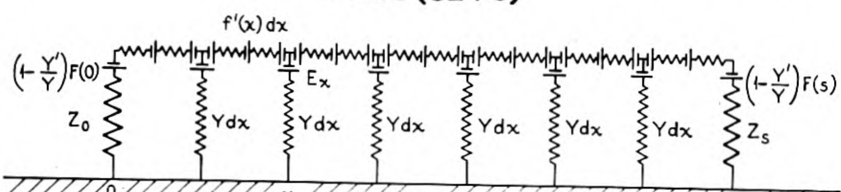


FIG. 10

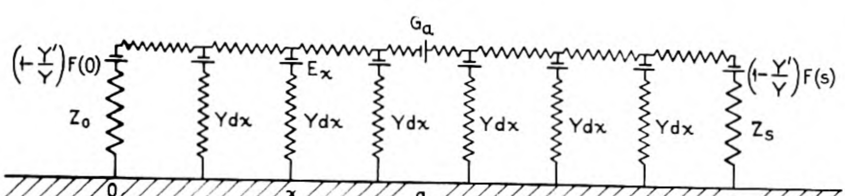


FIG. 11 (SET 6)

impedances. Thus, for instance, in the formulations of Set 1 and Set 5, item (C) would read: '(C) In the ends $x = 0$ and $x = s$ of the line, axial electromotive forces $- F(0)$ and $F(s)$ respectively.' (Observe, here, the negative sign before $F(0)$, in contrast to the positive sign in the original formulation.)

In this way it is readily seen that at a point $x = u$ where the impressed potential $F(x)$ is discontinuous, the equivalent electromotive force is an axial electromotive force equal to the decrement of the impressed potential, that is, equal to $F(u -) - F(u +)$; this agrees with equation (65), and with item (B) in the fundamental set of equivalent electromotive forces formulated in the early part of Section III.

Derivations of Set 1

A synthetic derivation of Set 1 has already been furnished in the early part of Section III. An analytical derivation will now be outlined; it is based on an interpretation of equations (68), (71), (73) below; these equations, in turn, are based on certain equations of Appendix I, as follows:

Combining equations (1) and (2) of Appendix I gives

$$zI + \frac{dV'}{dx} + \frac{d\phi'}{dt} = f, \tag{68}$$

where f denotes f_w; and V' is that part of the potential of the wire due to its charges (and the corresponding opposite charges on the surface of the ground), while ϕ' is that part of the magnetic flux due to the current in the wire (and the corresponding return current in the ground); that is,

$$V' = V - F = Q/C, \tag{69}$$

$$\phi' = \phi - \Phi = LI, \tag{70}$$

so that V' and ϕ' do not include the impressed potential and impressed magnetic flux F and Φ respectively.

By (5) and (7) of Appendix I the equation of current continuity can be written

$$-\frac{dI}{dx} = \frac{dQ}{dt} + \frac{G}{C}Q + Y'F. \tag{71}$$

The actual potential V of the line is of course the resultant of V' and F; that is,

$$V \equiv V(x) = V'(x) + F(x), \tag{72}$$

whence, in particular, at the ends $x = 0$ and $x = s$,

$$\begin{aligned} V(0) &= V'(0) + F(0), \\ V(s) &= V'(s) + F(s). \end{aligned} \tag{73}$$

Returning, now, to a consideration of equations (68), (71), (73), it is seen that they are identically the same as the equations for the same line without any impressed field but containing the set of electromotive forces formulated above under the heading 'Set 1 (Fig. 4)'; for an interpretation of equations (68), (71), (73) yields respectively (A), (B), (C) of Set 1.

It may be noted that equations (68) and (71) can be written in the following more compact forms:

$$- dV'/dx = ZI - f, \tag{74}$$

$$- dI/dx = YV' + Y'F, \tag{75}$$

whose interpretation yields immediately items (A) and (B) of Set 1.

Outline of Derivations of Sets 2, 3, 4, 5, 6

Synthetic derivations of Sets 2, 3, 4, 5, 6 from Set 1 will now be briefly outlined by aid of the diagrams in Figs. 2, $\cdots$ 11. The physical systems represented by these diagrams are all equivalent in the sense that the currents at corresponding points in all of them are equal.

In the derivation-work extensive use is made of an artifice which, for convenience, will be formulated in what may be termed the 'branch-point theorem,' as follows: *In any network of any number of branches the currents will not be affected by inserting at any branch-point a set of equal electromotive forces, one in each branch, directed either all toward or all from the branch-point.*

Fig. 2 represents the given one-wire line in an arbitrary impressed field, as already specified. For generality the line is assumed to have uniformly distributed leakage admittance of amount Y' per unit length.

Fig. 3 is derived from Fig. 2 by lumping the distributed direct leakage admittance into localized admittances each of amount $Y' \cdot dx$ at intervals of length dx.

Fig. 4 is derived from Fig. 3 by replacing the arbitrary impressed field by Set 1 of equivalent electromotive forces.

Fig. 5 is derived from Fig. 4 by replacing the line, exclusive of the direct leakage admittance Y', by its equivalent artificial line having 'complete series impedance' Z and 'basic shunt admittance' Y^0 per unit length. This replacement of the actual line by the corresponding artificial line is permissible now that the impressed field has been replaced by a set of equivalent electromotive forces (Fig. 4).

Fig. 6 is derived from Fig. 5 by replacing the compound shunt element, consisting of Y^0dx in parallel with $Y'dx$ containing the electromotive force $F(x)$, by the equivalent simple shunt element Ydx containing the electromotive force $(Y'/Y)F(x)$.

Figs. 7, 8, 9 are derived from Figs. 6, 7, 5 respectively by applying the 'branch point theorem.'

Fig. 11 is derived from Fig. 7 by applying the 'branch point theorem' in the manner indicated by Fig. 10, where $f'(x)dx$ denotes, for brevity, the original axial equivalent electromotive force situated between x and $x + dx$ of Set 2 as indicated by Fig. 7, so that

$$f'(x) \equiv f(x) + \frac{Y'}{Y}\frac{dF(x)}{dx}, \tag{76}$$

and a is the coordinate of the contemplated arbitrary point. The E's, of which E_x is typical, are sets of electromotive forces inserted at the branch-points. At first these electromotive forces are arbitrary, except that each set of three accords with the branch-point theorem, so as not to alter the original currents in the system. Next, starting at the ends, it is found that these electromotive forces can be so determined as to annul the original axial electromotive forces $f'(x)dx$ in all of the differential elements dx except in the one containing the point a; the requisite value of E_x and the resulting value of G_a are found to be as formulated in Set 6.

Finally it may be remarked that each of the Sets 2, 3, 4, 5, 6 can be verified against Set 1 by formulating the total current produced at any point x by each of the Sets 2, 3, 4, 5, 6 and then comparing the resulting formula with the sum of formulas (41), (42), (43), (44). Evidently it suffices to do this for the relatively simple case where the terminal impedances are equal to the characteristic impedance of the line; for this case, formulas (41), (42), (43), (44) reduce to (50), (51), (52), (53) respectively.

Sets of Equivalent Electromotive Forces for a Balanced Two-Wire Line in an Arbitrary Impressed Field

The foregoing six sets of equivalent electromotive forces for a one-wire line can be readily extended to a two-wire line after resolving the impressed field into mode-a and mode-c constituents, which are then dealt with separately. For Set 1 this procedure has been fully outlined above in the subsection entitled 'A Balanced Two-Wire Line in an Arbitrary Impressed Field,' and it has found a natural application in the 'Crosstalk Problem' treated below in Section IV.

It is clear that all of the sets of equivalent electromotive forces are immediately applicable to dealing with the mode-c constituent of the impressed field, since this constituent acts on the circuit consisting of the two wires in parallel with each other, with ground return, which is formally the same as a one-wire line with ground return.

All of the sets of equivalent electromotive forces become applicable to dealing with the mode-a constituent of the impressed field by an appropriate interpretation of the diagrams (Fig. 2, $\cdots$ 11), namely, the following interpretation:

1. In each diagram regard the wire-symbol as representing the outgoing wire of the actual two-wire line, and regard the ground-symbol as representing not the ground but the return wire of the two-wire line. (The presence of the earth is then to be regarded as implied, its effects appearing implicitly in the values of the line parameters.)
2. Hence regard Z_0 and Z_s as denoting the mode-a terminal impedances functioning as though connected directly across the two-wire line at its ends $x = 0$ and $x = s$ respectively.
3. Regard Y', Y^0, Y, Z as denoting the mode-a line constants (including implicitly the effects of the earth).
4. Regard $f(x)$, $F(x)$, and $\Phi(x)$ as denoting the mode-a constituents of the impressed field—that is, as denoting the difference of the actual values impressed at the two wires. (In order to maintain the balanced condition of the two-wire line, $f(x)$ is to be regarded as constituted of $f(x)/2$ in the outgoing wire and $-f(x)/2$ in the return wire; and similarly for $F(x)$ and $\Phi(x)$.)

The Electric Field Due to a System of n Parallel Wires in an Arbitrary Impressed Field

Thus far in the present section of this paper the field impressed on the given physical system has been supposed known and the problem has been to calculate the resulting currents. Actually, however, the impressed field is not usually known but has to be calculated—from a knowledge of the currents and charges producing it.

The present subsection deals with the problem of calculating the electric field impressed on a secondary system consisting of a single horizontal wire j by a primary system π consisting of n wires which are parallel to each other and to j. For generality, the primary and secondary systems are supposed to be in an arbitrary impressed field.[19]

Consider at first any parallel geometrical line i, not necessarily in any of the wires; and let $V_i = V_i(x)$ and $E_i = E_i(x)$ denote the

[19] Of course the field produced by any given system is directly due only to the currents and charges of the system, and does not depend directly on any field that may be impressed on the system; but, assuming the system to be energized only by the impressed field, the currents—and thence the charges—are directly due to the impressed field and can (theoretically, at least) be expressed in terms of it.

potential and the axial electric force at any point x in i. Then V_i is analyzable into three parts ($V_{i\pi}$, V_{ij}, F_i) and E_i into three parts ($E_{i\pi}$, E_{ij}, f_i) due respectively to the primary system π, to the secondary system j, and to the arbitrary impressed field; that is,[20]

$$V_i = V_{i\pi} + V_{ij} + F_i, \tag{77}$$

$$E_i = E_{i\pi} + E_{ij} + f_i. \tag{78}$$

In particular, at the secondary wire j the potential V_j and the axial electric force E_j are analyzable in accordance with the equations

$$V_j = V_{j\pi} + V_{jj} + F_j, \tag{79}$$

$$E_j = E_{j\pi} + E_{jj} + f_j. \tag{80}$$

In the present subsection, the problem to be dealt with is the calculation of $V_{j\pi}$ and $E_{j\pi}$, namely the potential and the axial electric force at any point x in the secondary j due directly to the currents and charges of the primary system π.

The n wires of the primary system π will be numbered $1, 2, 3, \cdots n$. The letters h, k, r will be employed generically: each may denote any one of the designation numbers $1, \cdots n$—as h in equation (81); or each may run through the whole set $1, \cdots n$—as in equation (91). (It is hardly necessary to remark that j is *not* a member of the set $1, \cdots n$, in the notation of this Section (III), where j *always* designates the secondary wire.)

The current at any point x in any wire h of the primary will be denoted by $I_h = I_h(x)$; and the charge on wire h, per unit length, by $Q_h = Q_h(x)$.

The potential V_h and the axial electric force E_h at any primary wire h are analyzable in the manner expressed by the equations

$$V_h = V_{h\pi} + V_{hj} + F_h, \tag{81}$$

$$E_h = E_{h\pi} + E_{hj} + f_h, \tag{82}$$

in accordance with the general equations (77) and (78) respectively. It will be found convenient to call $V_{h\pi}$ the 'systemic potential' and $E_{h\pi}$ the 'systemic axial electric force' at wire h, since $V_{h\pi}$ and $E_{h\pi}$ are due only to the system π of which h is a member, and do not include

[20] Regarding the use here of double subscripts, it will be noted that the first subscript designates the line or the wire where the effect occurs, and the second the wire or the system of wires which produce the effect. Thus, $V_{i\pi}$ is the potential produced in line i by the whole primary system π; the contribution of any one wire h would be denoted by V_{ih}.

any contributions from the secondary system or from the impressed field. (More fully, $V_{h\pi}$ may be termed the 'primary systemic potential' at h and $E_{h\pi}$ the 'primary systemic axial electric force' at h.)

For explicit use below, we may here note the formulas for the systemic potential $V_{k\pi}$ and the systemic axial electric force $E_{k\pi}$ at any wire k of the primary system π:

$$V_{k\pi} = \sum_{h=1}^{n} p_{kh}Q_h, \qquad (k = 1, \cdots n), \tag{83}$$

$$E_{k\pi} = -\sum_{h=1}^{n}\left(Z_{kh}I_h + p_{kh}\frac{dQ_h}{dx}\right), \qquad (k = 1, \cdots n), \tag{84}$$

p_{kh} and Z_{kh} being respectively the mutual potential coefficient and the mutual impedance[3] between wires h and k, per unit length. Equation (84) is obtainable by applying the second curl law to a differential rectangle substantially as in deriving equations (1) and (2); see also Appendix I.

As already stated in connection with equations (79) and (80) the problem to be considered in the present subsection is the calculation of the potential $V_{j\pi}$ and the axial electric force $E_{j\pi}$ produced at the secondary wire j by the primary system π. The fundamental formulas for $V_{j\pi}$ and $E_{j\pi}$ are:

$$V_{j\pi} = \sum_{h=1}^{n} p_{jh}Q_h = \sum_{h=1}^{n} V_{jh}, \tag{85}$$

$$E_{j\pi} = -\sum_{h=1}^{n}\left(Z_{jh}I_h + \frac{dV_{jh}}{dx}\right) = \sum_{h=1}^{n} E_{jh} \tag{86}$$

$$= -\sum_{h=1}^{n}\left(Z_{jh}I_h + p_{jh}\frac{dQ_h}{dx}\right), \tag{87}$$

where V_{jh} and E_{jh} are the contributions of wire h to $V_{j\pi}$ and $E_{j\pi}$ respectively.

With regard to applications of the equations (85) and (87) for the potential $V_{j\pi}$ and the axial electric force $E_{j\pi}$ impressed on the secondary j by the primary π, it will be supposed that all the primary currents $I_1, \cdots I_n$ are known. But the primary charges Q_h and their axial gradients dQ_h/dx (where $h = 1, \cdots n$) are usually not known; and therefore ways will now be indicated for expressing them in terms of quantities which may be known. For that purpose, the presence of the secondary will be entirely ignored, in all respects. (This procedure may be regarded as the first-approximation step in a solution by successive approximations.)

The charges Q_h can be expressed in terms of the systemic potentials $V_{k\pi}$ by solving the set of n equations (83). Thus

$$Q_h = \sum_{k=1}^{n} q_{hk} V_{k\pi}, \qquad (h = 1, \cdots n), \tag{88}$$

where q_{hk} is the Maxwell capacity coefficient between wires h and k; in terms of the potential coefficients, its value is

$$q_{hk} = D_{kh}(p)/D(p), \tag{89}$$

$D(p)$ being the determinant of all the potential coefficients (the p's) in the set of n equations (83) and $D_{kh}(p)$ the cofactor of p_{kh} in $D(p)$.

The systemic potentials $V_{k\pi}$, occurring in (88), can be obtained by solving the equations of current continuity, namely the set of n equations [21]

$$-\frac{dI_h}{dx} = \sum_{k=1}^{n} (Y_{hk} V_{k\pi} + X_{hk} F_k), \qquad (h = 1, \cdots n), \tag{90}$$

Y_{hk} and X_{hk} being of the nature of admittances (per unit length), and F_k the impressed potential at wire k; it is thus found that

$$V_{k\pi} = -\sum^{n} W_{kh}\left(\frac{dI_h}{dx} + \sum_{r=1}^{n} X_{hr} F_r\right), \qquad (k = 1, \cdots n), \tag{91}$$

where the coefficient W_{kh} is the same function of the Y's that q_{kh} is of the p's, that is,

$$W_{kh} = D_{hk}(Y)/D(Y). \tag{92}$$

It is seen that W_{kh} is of the nature of an impedance (per unit length), though it is not a simple impedance.

The charges can now be expressed in terms of the impressed potentials F_r and the axial gradients of the currents by substituting (91) in (88).

The axial gradients of the charges can be expressed in various ways. They can be immediately expressed in terms of the axial gradients of the systemic potentials V_k by merely differentiating (88) with respect to x. Also, they can be expressed in terms of the currents I_r and the systemic axial electric forces $E_{k\pi}$ at the wires, by solving the set of n equations (84); thus

$$\frac{dQ_h}{dx} = -\sum_{k=1}^{n} q_{hk}(E_{k\pi} + \sum_{r=1}^{n} Z_{kr} I_r), \qquad (h = 1, \cdots n), \tag{93}$$

[21] Derived in the latter part of Appendix I.

q_{hk} being the Maxwell capacity coefficient given by (89). Furthermore, the systemic axial electric force $E_{k\pi}$ occurring in (93) is expressible in terms of the current I_k in wire k and the axial electric force f_k impressed on wire k, by the simple relation

$$E_{k\pi} = z_k I_k - f_k, \tag{94}$$

z_k denoting the internal impedance of wire k, per unit length; for, the resultant axial electric force at wire k must be equal to $z_k I_k$ and must also be equal to $E_{k\pi} + f_k$. Thus the axial gradients of the charges can be expressed explicitly in terms of the currents and the impressed axial electric forces at the wires, by substituting (94) in (93). The axial gradients of the charges can be expressed still otherwise by differentiating (88) with respect to x after substituting (91).

Substituting into (85) and (87), the various foregoing expressions for the charge Q_h and its axial gradient dQ_h/dx, and in some cases transforming and rearranging the results, gives the following formulas for the potential $V_{j\pi}$ and the axial electric force $E_{j\pi}$ produced at any point x in the secondary wire j by the primary system π, when the presence of the secondary j is entirely ignored in calculating the currents, charges, and potentials of the primary (in accordance with the statement of the paragraph following equation (87)):

$$V_{j\pi} = \sum_{h=1}^{n} p_{jh} Q_h \tag{95}$$

$$= \sum_{h=1}^{n} T_{jh} V_{h\pi} \tag{96}$$

$$= -\sum_{h=1}^{n} T_{jh} \left[\sum_{k=1}^{n} W_{hk} \left(\frac{dI_k}{dx} + \sum_{r=1}^{n} X_{kr} F_r \right) \right], \tag{97}$$

where T_{jh}, which may be termed a 'potential transfer factor' or 'voltage transfer factor,' has the value

$$T_{jh} = \sum_{k=1}^{n} p_{jk} q_{kh}. \tag{98}$$

$$E_{j\pi} = -\sum_{h=1}^{n} Z_{jh} I_h - \frac{dV_{j\pi}}{dx} \tag{99}$$

$$= -\sum_{h=1}^{n} \left(Z_{jh} I_h + p_{jh} \frac{dQ_h}{dx} \right) \tag{100}$$

$$= -\sum_{h=1}^{n} \left(Z_{jh} I_h + T_{jh} \frac{dV_{h\pi}}{dx} \right) \tag{101}$$

$$= -\sum_{h=1}^{n}\left(Z_{jh}I_h - \sum_{k=1}^{n} T_{jh}W_{hk}\left[\frac{d^2I_k}{dx^2} + \sum_{r=1}^{n} X_{kr}\frac{dF_r}{dx}\right]\right) \quad (102)$$

$$= -\sum_{h=1}^{n}\left(Z_{jh}I_h - T_{jh}E_{h\pi} - T_{jh}\sum_{r=1}^{n} Z_{hr}I_r\right) \quad (103)$$

$$= -\sum_{h=1}^{n}\left(Z_{jh}I_h - T_{jh}[z_hI_h - f_h] - T_{jh}\sum_{r=1}^{n} Z_{hr}I_r\right). \quad (104)$$

Formula for $E_{j\pi}$ when the Earth is a Perfect Conductor

When the earth is a perfect conductor, all of the external mutual and self impedances (Z_{jk}, Z_{hk}, Z_{kk}, etc.) are pure reactances and are proportional to the corresponding potential coefficients (p_{jk}, p_{hk}, p_{kk}, etc.), the proportionality factor being merely $i\omega/\tau$, where τ is an absolute constant whose value depends only on the units employed. Thence it can be shown that (103) and (104) respectively reduce to the very simple formulas

$$E_{j\pi} = \sum_{h=1}^{n} T_{jh}E_{h\pi}, \quad (105)$$

$$E_{j\pi} = \sum_{h=1}^{n} T_{jh}(z_hI_h - f_h), \quad (106)$$

with T_{jh} given by (98). It is seen that (105) corresponds exactly to (96).

As at least of some academic interest, it may be remarked that equations (105) and (106) hold even when the earth is imperfect, provided

$$\frac{Z_{jh}}{p_{jh}} = \frac{Z_{kh}}{p_{kh}}, \qquad (h = 1, \cdots n;\ k = 1, \cdots n).$$

IV

Practical Applications

For illustrative purposes, the methods presented in the foregoing sections will now be applied to two practical problems of a rather diverse nature. The first application will be to the wave antenna employed in certain important cases of long-distance radio reception,[22] the second, to a problem in crosstalk.

The Wave Antenna

The wave antenna, in its usual form, may be described as a transmission line with ground return, utilized for the reception of radio waves.

[22] Notably in transoceanic radio telephony.

In its simplest form, as here contemplated, the wave antenna consists of a long straight horizontal wire terminated at each end in its characteristic impedance K, as represented by Fig. 12a, which gives

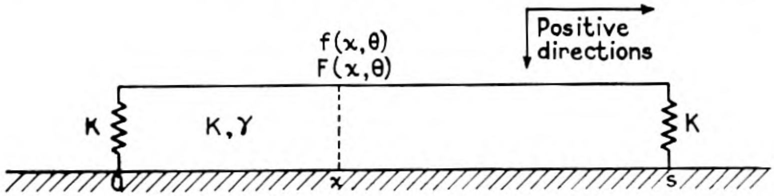


Fig. 12a

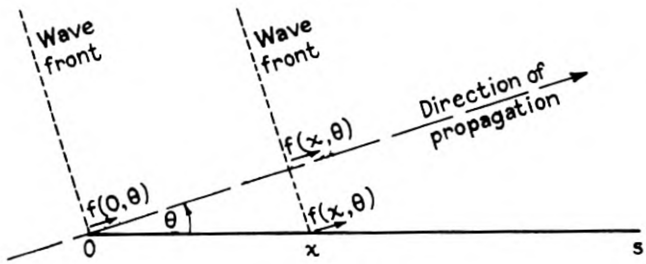


Fig. 12b

an elevation view. This is seen to be the same as Fig. 1 when $Z_0 = Z_s = K$; and γ is now the propagation constant, per unit length, of the wave antenna regarded as a transmission line. Hence formulas (46), $\cdots$ (54), pertaining to Fig. 1, are immediately applicable for calculating the current at any point x in the wave antenna of Fig. 12a, after the appropriate formulation of the functions $f(x)$ and $F(x)$—namely the impressed axial electric force and the impressed potential, respectively, at any point x in the wave antenna.

These functions can be evaluated by aid of Fig. 12b, which gives a plan view representing a train of plane radio waves (whose magnetic component is horizontal) incident on the wave antenna at an arbitrary angle θ measured horizontally from the wave antenna to the direction of propagation of the wave train along the earth's surface. By the 'direction of propagation' is here meant the horizontally specified direction of a vertical plane which is normal to the plane of the wave front. $f(x, \theta)$ denotes the horizontal component of electric force in the impressed waves at any point x of the wave antenna, and $F(x, \theta)$ the potential of the impressed waves there.[23] Then the axial electric

[23] The presence of θ in the functional symbols is of course to allow for a possible

force $f(x)$ and the potential $F(x)$ impressed at point x on the wave antenna by the radio waves are given by the equations

$$f(x) = f(x, \theta) \cos \theta, \tag{107}$$

$$F(x) = F(x, \theta). \tag{108}$$

It is convenient to take one end, say $x = 0$, as a fixed reference point, and then to express $f(x, \theta)$ and $F(x, \theta)$ in terms of their values $f(0, \theta)$ and $F(0, \theta)$ at $x = 0$. For this purpose, it will be assumed that the radio waves are propagated in a simple exponential manner, so that

$$\frac{f(x, \theta)}{f(0, \theta)} = \frac{F(x, \theta)}{F(0, \theta)} = e^{-\Gamma x \cos \theta}, \tag{109}$$

Γ denoting the propagation constant of the radio waves, per unit length measured horizontally along the direction of their propagation. Then the equations (107) and (108) become

$$f(x) = f(0, \theta) \cos \theta \, e^{-\Gamma x \cos \theta}, \tag{110}$$

$$F(x) = F(0, \theta) e^{-\Gamma x \cos \theta}, \tag{111}$$

wherein $f(0, \theta)$ and $F(0, \theta)$ may be supposed known. In this connection it should be remarked that $f(0, \theta)$ and $F(0, \theta)$—and, more generally, $f(x, \theta)$ and $F(x, \theta)$—are not in phase.[24]

On substitution of (110) and (111), equations (49), $\cdots$ (53) now become applicable for calculating the current $I(x)$ at any point x of the wave antenna; this current will be written $I(x, \theta)$ because it depends on the incidence-angle θ, even when $f(x, \theta)$ and $F(x, \theta)$ are independent of θ. In the engineering of wave antennæ, we are usually concerned merely with the current $I(s, \theta)$ received at the end $x = s$. In general there will be four constituents of $I(s, \theta)$, corresponding to equations (50), (51), (52), (53) when $x = s$. From the discussion of the corresponding more general equations (41), (42), (43), (44), it will be recalled that the current-constituent $j(s, \theta)$ is due to the impressed axial electric force acting throughout the length of the wave antenna, $J_0(s, \theta)$ is due to the impressed voltage $F(0, \theta)$ acting at the end $x = 0$, $J_s(s, \theta)$ is due to the corresponding impressed voltage $F(s, \theta) = F(0, \theta)e^{-\Gamma s \cos \theta}$ acting at the end $x = s$, and $J_{0s}(s, \theta)$

dependence on θ. It may be noted that, in the calculation of the ordinary polar diagram representing the directional selectivity of a wave antenna, the functions $f(x, \theta)$ and $F(x, \theta)$ are regarded as independent of θ, in accordance with the very definition of the directional selectivity.

[24] The ratio of the horizontal electric force $f(x, \theta)$ to the vertical electric force $F(x, \theta)/H$—where H here denotes the height of the wave antenna above the earth's surface—is a complex number whose value depends on the conductivity, dielectric constant, and permeability of the ground, and on the frequency.

is due to the distributed impressed voltage acting in the leakage admittance from the wave antenna to ground (this leakage admittance being regarded as uniformly distributed). By substituting the values $f(y)$ and $F(y)$ given by (110) and (111) when x is replaced by y, then carrying out the indicated integrations, and finally transforming the results somewhat, the constituents corresponding to (50), (51), and (52) are found to have the following formulas:

$$j(s,\,\theta) = \frac{sf(0,\,\theta)\cos\theta}{2K}\,\frac{\sinh\left[(\gamma - \Gamma\cos\theta)s/2\right]}{(\gamma - \Gamma\cos\theta)s/2}\,e^{-(\gamma+\Gamma\cos\theta)s/2}, \tag{112}$$

$$J_0(s,\,\theta) = -\frac{F(0,\,\theta)}{2K}\,e^{-\gamma s}, \tag{113}$$

$$J_s(s,\,\theta) = \frac{F(0,\,\theta)}{2K}\,e^{-\Gamma s\cos\theta}. \tag{114}$$

The fourth constituent, $J_{0s}(s,\,\theta)$, corresponding to (53), will be omitted, because it is relatively unimportant and also because its formula is found to be somewhat lengthy.

The valuable directional selectivity of a wave antenna resides mainly in the directional properties of the admittance $j(s,\,\theta)/sf(0,\,\theta')$ whose value is found by dividing equation (112) through by $sf(0,\,\theta')$, where θ' denotes some fixed value of θ (usually $\theta' = 0$). This ratio may properly be termed a 'directional admittance.' The corresponding admittances obtained by dividing (113) and (114) through by $sf(0,\,\theta')$ are not usefully directional, the former being entirely non-directional, and the latter only directional as regards its phase angle—not as regards its absolute value. By suitable choice of the length of the wave antenna, the constituent represented by (112) can be made to have high directional selectivity, while the constituents corresponding to (113) and (114) become relatively unimportant (except over a few narrow ranges of the incidence angle θ).[25]

A Crosstalk Problem

This problem is concerned with the derivation of formulas for the first-order crosstalk between two simple open-wire telephone circuits of which one is non-transposed and the other is once-transposed, as represented in plan view by Fig. 13.

The once-transposed circuit is taken as the primary, and the non-transposed as the secondary. Each extends from $x = 0$ to $x = s$; and the primary is transposed at its mid-point $x = s/2$.

[25] For a detailed study of the wave antenna, the reader is referred to the well-known paper by Beverage, Rice, and Kellog entitled 'The Wave Antenna' in *J. A. I. E. E.* beginning with March, 1923.

The primary is energized by an alternating electromotive force E_0 inserted at $x = 0$. Stated precisely, the problem here contemplated is the derivation of formulas for the currents produced in the two ends, $x = 0$ and $x = s$, of the secondary circuit by the primary circuit, when all reactions of the secondary on the primary are neglected.

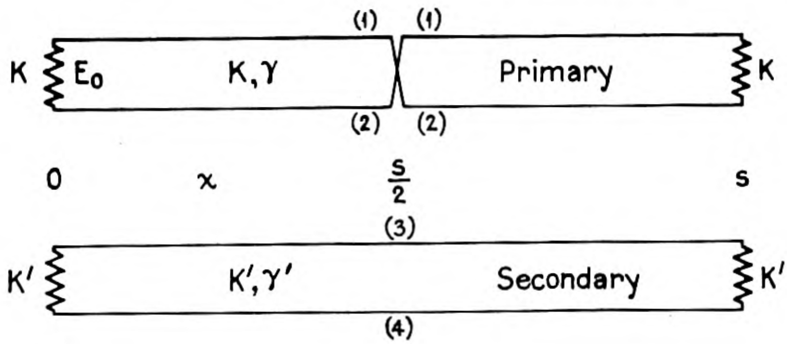


Fig. 13

The wires of the secondary circuit are numbered 3 and 4. The primary wires are numbered 1 and 2, in the sense that 1 and 2 designate the *positions* the wires would occupy if non-transposed; in this sense, wires 1 and 2 are each discontinuous at $x = s/2$, the transposition cross thus being regarded as extraneous to the wires.[26]

Each circuit is terminated at each end in its mode-a characteristic impedance—K for the primary, K' for the secondary. The mode-a propagation constants of the primary and secondary, per unit length, are denoted by γ and γ' respectively.

The earth is assumed to be a perfect conductor. This assumption is effectively a good approximation because the contemplated circuits are such that the distance between the two wires of each circuit is small compared with their height above ground; at the same time the assumption greatly facilitates and simplifies the solution.

Evidently the first step is to formulate the primary current and the primary systemic potential at any point x. The second step is to formulate the electric field impressed on the secondary by the primary. The third and final step is to formulate the currents produced in the secondary by the impressed field of the primary; this third step will be carried through by means of the synthetic method, employing the set of equivalent electromotive forces formulated in the early part of Section III.

[26] The transposition cross may conveniently be regarded as merely a particular kind of transducer (four-terminal network) inserted in the primary, namely a reversing transducer.

Let $I_r(x) = I_r$ and $V_r(x) = V_r$ denote the current and the potential, respectively, at any point x of wire r, where $r = 1, 2, 3, 4$. Then, evidently, for the primary currents and potentials we have:

$$I_1 = -I_2, \qquad V_1 = -V_2, \tag{115}$$

$$V_1 - V_2 = KI_1. \tag{116}$$

For $x \lesseqgtr s/2$:

$$I_1 = \pm \frac{E_0}{2K} e^{-\gamma x}, \tag{117}$$

$$V_1 = \pm \frac{E_0}{4} e^{-\gamma x}. \tag{118}$$

Thus the primary currents, $I_1(x)$ and $I_2(x)$, and the corresponding primary potentials, $V_1(x)$ and $V_2(x)$, are each discontinuous at $x = s/2$ (by reason of the transposition there).

The electric field impressed on the individual wires of the secondary circuit by the primary circuit can be formulated by means of equations (106) and (96). Thus

$$E_3 = (T_{31} - T_{32})zI_1, \tag{119}$$

$$E_4 = (T_{41} - T_{42})zI_1, \tag{120}$$

$$V_3 = (T_{31} - T_{32})V_1, \tag{121}$$

$$V_4 = (T_{41} - T_{42})V_1, \tag{122}$$

where $z = z_1 = z_2$ is the internal impedance of each wire of the primary, per unit length, and the T's are 'voltage transfer factors' given by (98).

Evidently the secondary circuit constitutes a balanced two-wire line in an arbitrary impressed field (the field due to the primary), and hence is amenable to the treatment already fully described and formulated in the subsection following equation (54). Thus the current at any point x in the secondary consists of two modes, a and c. However, as already indicated, we shall ultimately be concerned only with the currents in the ends $x = 0$ and $x = s$ of the secondary; evidently these are mode-a currents, for at each end the mode-c currents must be zero, since the circuit is insulated from ground at each end.

As we shall be concerned only with the mode-a currents produced in the secondary, the next step is to formulate the mode-a constituents of the electric field impressed by the primary. If E' and V' denote the mode-a constituents of the axial electric force and of the potential impressed by the primary, then

$$E' = E_3 - E_4 = 4TzI_1, \tag{123}$$

$$V' = V_3 - V_4 = 4TV_1, \tag{124}$$

where

$$T = (T_{31} - T_{32} - T_{41} + T_{42})/4. \tag{125}$$

Remembering that I_1 and V_1 are discontinuous at $x = s/2$, in accordance with equations (117) and (118), it is seen that E' and V' are discontinuous at $x = s/2$ in accordance with the following equations.[27] For $x \lesseqgtr s/2$:

$$E' = \pm E_0 T \frac{2z}{K} e^{-\gamma x}, \qquad V' = \pm E_0 T e^{-\gamma x}. \tag{126}$$

We are now prepared to formulate the mode-a currents produced in the secondary (3, 4) by the field arising from the primary (1, 2). Since the secondary constitutes a balanced two-wire line in an arbitrary impressed field, it is amenable to the treatment formulated in the subsection following equation (54); thence equations (41), $\cdots$ (45) and equations (50), $\cdots$ (54) are formally applicable.

If $I_3(x)$ and $I_4(x)$ denote the mode-a currents at any point x in the secondary wires 3 and 4 respectively, then

$$I_3(x) = -I_4(x) = I(x), \text{ say.} \tag{127}$$

On referring to the subsection containing equations (41), $\cdots$ (45) and applying it to the mode-a effects in the present problem, it will be seen that $I(x)$ is the sum of the five mode-a constituents $j(x)$, $J_0(x)$, $J_s(x)$, $J_{0s}(x)$, $J_u(x)$, corresponding to equations (41), (42), (43), (44), (45) respectively. From the discussion in connection with those equations and from the analysis of the impressed field into two modes, a and c, as described and formulated in the subsection following equation (54), it will be seen that $j(x)$ is due to the mode-a axial electric force $E_3(y) - E_4(y)$ acting at all points y of the secondary, $J_0(x)$ is due to the mode-a impressed voltage $V_3(0) - V_4(0)$ acting at the end $y = 0$, $J_s(x)$ is due to the mode-a impressed voltage $V_3(s) - V_4(s)$ acting at the end $y = s$, $J_{0s}(x)$ is due to the mode-a impressed voltage $V_3(y) - V_4(y)$ acting at all points y in the leakage admittance[28] between the secondary wires, and $J_u(x)$ is due to the discontinuity $V'(u -)$

[27] From mere physical considerations, it is evident that the whole primary field is reversed at $x = s/2$.

[28] That is, the 'mutual' leakage admittance (equal to the 'direct' leakage admittance between wires plus one half of the 'direct' leakage admittance from each wire to ground).

$- V'(u+)$ in the mode-a impressed voltage $V'(y) \equiv V_3(y) - V_4(y)$ at $y = u = s/2$. (In what follows, the constituent $J_{0s}(x)$ will be omitted because it is relatively unimportant.)

Thus, for the formulation of the mode-a effects the functions $f(y)$ and $F(y)$, representing the electirc forces and potentials in equations (41), $\cdots$ (45) and (50), $\cdots$ (54), have the following mode-a values:

$$f(y) = E_3(y) - E_4(y) = E'(y), \tag{128}$$

$$F(y) = V_3(y) - V_4(y) = V'(y), \tag{129}$$

whence, in particular,

$$F(0) = V'(0), \tag{130}$$

$$F(s) = V'(s), \tag{131}$$

$$F(u-) - F(u+) = V'(u-) - V'(u+). \tag{132}$$

Substituting these values into equations (50), (51), (52), (54), and carrying out the indicated integrations [15] when $x = 0$ and when $x = s$, and finally dividing each equation by the value of the primary current $I_1(0) = E_0/2K$ at $x = 0$, we obtain the following formulas (134), $\cdots$ (137) for the four mode-a current ratios at $x = 0$, and the formulas (141), $\cdots$ (144) for those at $x = s$. Also, there are included formulas for $J(x)/I_1(0)$ at $x = 0$ and at $x = s$, $J(x)$ denoting the sum of the mode-a current constituents due to the impressed potential, that is,

$$J(x) = J_0(x) + J_s(x) + J_u(x) \tag{133}$$

since $J_{0s}(x)$ is neglected.

At $x = 0$ the formulas for the four current-ratios are

$$\frac{j(0)}{I_1(0)} = T\frac{K}{K'}\frac{sz}{K}\frac{[1 - e^{-(\gamma+\gamma')s/2}]^2}{(\gamma + \gamma')s/2}, \tag{134}$$

$$\frac{J_0(0)}{I_1(0)} = -T\frac{K}{K'}, \tag{135}$$

$$\frac{J_s(0)}{I_1(0)} = -T\frac{K}{K'}e^{-(\gamma+\gamma')s}, \tag{136}$$

$$\frac{J_u(0)}{I_1(0)} = 2T\frac{K}{K'}e^{-(\gamma+\gamma')s/2}. \tag{137}$$

The sum of the last three is

$$\frac{J(0)}{I_1(0)} = -T\frac{K}{K'}[1 - e^{-(\gamma+\gamma')s/2}]^2. \tag{138}$$

On dividing (134) by (138) the ratio of $j(0)$ to $J(0)$ is found to have the simple value

$$\frac{j(0)}{J(0)} = -\frac{z}{K(\gamma + \gamma')/2}. \tag{139}$$

In particular, when the two circuits have equal propagation constants ($\gamma' = \gamma$),

$$\frac{j(0)}{J(0)} = -\frac{z}{Z}, \tag{140}$$

where $Z = \gamma K$ is the mode-a 'complete series impedance'[18] of the primary, per unit length; it will be recalled that z is the 'internal impedance' of each primary wire, per unit length. The ratio z/Z is ordinarily a very small fraction.

At $x = s$ the formulas for the four current-ratios are

$$\frac{j(s)}{I_1(0)} = T\frac{K}{K'}\frac{sz}{K}\frac{[1 - e^{-(\gamma-\gamma')s/2}]^2}{(\gamma - \gamma')s/2}e^{-\gamma's}, \tag{141}$$

$$\frac{J_0(s)}{I_1(0)} = -T\frac{K}{K'}e^{-\gamma's}, \tag{142}$$

$$\frac{J_s(s)}{I_1(0)} = -T\frac{K}{K'}e^{-\gamma s}, \tag{143}$$

$$\frac{J_u(s)}{I_1(0)} = 2T\frac{K}{K'}e^{-(\gamma+\gamma')s/2}. \tag{144}$$

The sum of the last three is

$$\frac{J(s)}{I_1(0)} = -T\frac{K}{K'}[1 - e^{-(\gamma-\gamma')s/2}]^2e^{-\gamma's}. \tag{145}$$

On dividing (141) by (145) the ratio of $j(s)$ to $J(s)$ is found to have the simple value

$$\frac{j(s)}{J(s)} = -\frac{z}{K(\gamma - \gamma')/2}. \tag{146}$$

When the absolute value of $(\gamma - \gamma')s/2$ is small compared to unity, equations (141) and (145) become approximately

$$\frac{j(s)}{I_1(0)} = T\frac{K}{K'}\frac{sz}{K}\frac{(\gamma - \gamma')s}{2}e^{-\gamma's}, \tag{147}$$

$$\frac{J(s)}{I_1(0)} = -T\frac{K}{K'}\left[\frac{(\gamma - \gamma')s}{2}\right]^2e^{-\gamma's}. \tag{148}$$

Thus:

$$\text{When} \quad \gamma' = \gamma: \ j(s) = 0, \qquad J(s) = 0.$$

For some cases, particularly those where the attenuation is neglected, it is advantageous to express the square-bracketed factors in equations (134), (138), (141), (145) partially in terms of hyperbolic sines.

Appendix I

Derivations of Equations (25) and (90)

Equation (25)

Let the primary or impressed field of force be specified by an electric intensity f_a parallel to the axis of the wire (and to the surface of the earth), and an electric intensity f_n normal to the surface of the earth and measured downward. We denote by f_w the value of f_a at the axis of the wire,[29] and by f_g its value at the surface of the ground in the plane which is normal to the ground and which includes the axis of the wire. The impressed or primary potential F of the wire, due to the impressed field, is then

$$F = \int_0^h f_n dy,$$

where h is the height of the wire above ground and the integral is taken along the vertical from the wire ($y = 0$) to ground ($y = h$).

Due to the impressed field, specified above, a current I flows in the wire and a corresponding *superposed* current distribution is induced in the ground. The *resultant* axial electric intensity at the surface of the wire is then $z_w I$ (where z_w is the internal impedance of the wire, per unit length); correspondingly the *resultant* electric intensity along the surface of the ground is $f_g - z_g I$. Application of the second curl law to a contour composed of two verticals from the wire to ground and the line segments dx in the surfaces of the wire and ground gives

$$(z_w + z_g)I - f_g + \frac{dV}{dx} = -\frac{d\phi}{dt},$$

which is preferably written as

$$zI - f_g + \frac{dV}{dx} = -\frac{d\phi}{dt}, \tag{1}$$

where $z = z_w + z_g$ is the internal impedance of the circuit, per unit length; V is the *resultant* potential of the wire; and ϕ is the *resultant* magnetic flux threading the contour, per unit length. But we have

[29] f_w is assumed to be sensibly constant over the cross-section of the wire.

also

$$f_w - f_g + \frac{dF}{dx} = -\frac{d\Phi}{dt}, \tag{2}$$

where Φ denotes the impressed magnetic flux threading the contour, per unit length. Subtracting (2) from (1) and observing that

$$\begin{aligned} V - F &= \frac{1}{C}Q, \\ \phi - \Phi &= LI, \end{aligned} \tag{3}$$

we get

$$zI + L\frac{dI}{dt} + \frac{1}{C}\frac{dQ}{dx} = f_w, \tag{4}$$

where, of course, Q is the charge, C the capacity to ground and L the external inductance, per unit length of the wire.

To eliminate Q from (4) we make use of the equation of current continuity, namely

$$-\frac{dI}{dx} = \frac{dQ}{dt} + I', \tag{5}$$

where I' is the leakage current per unit length of the wire. If the wire is embedded in a homogeneous leaky medium, then

$$I' = \frac{G^0}{C}Q = G^0(V - F), \tag{6}$$

where G^0 is proportional to the conductivity of the medium.[30] If, furthermore, there is direct leakage admittance from the wire to ground (as at poles and insulators) of amount Y' per unit length,[31] when regarded as uniformly distributed, then

$$I' = \frac{G^0}{C}Q + Y'V = \frac{G}{C}Q + Y'F, \tag{7}$$

where

$$G = G^0 + Y'. \tag{8}$$

On substituting the last value of I' into (5), setting $d/dt = i\omega$, then differentiating with respect to x, and finally substituting the resulting value of dQ/dx into (4), we get

$$(z + i\omega L)I - \frac{1}{G + i\omega C}\frac{d^2I}{dx^2} = f_w + \frac{Y'}{G + i\omega C}\frac{dF}{dx}, \tag{9}$$

[30] A formula for G^0 is equation (18) derived below.

[31] While G^0 is merely a pure conductance, Y' is in general an admittance (leakage admittance), because the insulators and poles have capacity as well as conductance. Hence G, defined by (8), is an admittance.

which can be written

$$\frac{K}{\gamma}\left(\gamma^2 - \frac{d^2}{dx^2}\right) I = f_w + \frac{Y'K}{\gamma}\frac{dF}{dx}, \tag{10}$$

where

$$K = \sqrt{\frac{z + i\omega L}{G + i\omega C}}, \tag{11}$$

and

$$\gamma = \sqrt{(z + i\omega L)(G + i\omega C)}. \tag{12}$$

Thus K is the characteristic impedance and γ the propagation constant of the transmission system composed of the overhead wire with ground return; it is to be noted that $G = G^0 + Y'$, in accordance with (8), and hence that G is in general an admittance—not a pure conductance.

If we define f' by the equation

$$f' = f_w + \frac{Y'K}{\gamma}\frac{dF}{dx}, \tag{13}$$

then (10) becomes

$$\frac{K}{\gamma}\left(\gamma^2 - \frac{d^2}{dx^2}\right) I = f', \tag{14}$$

which is formally the same as equation (25) of the text. There, however, it is assumed that the term $(Y'K/\gamma)dF/dx$ is negligible; probably this is usually the case but circumstances may arise where it is not negligible. Its inclusion, however, introduces no formal modification of the analysis.

The foregoing derivation has been given in detail because prior derivations known to the writers have not been entirely satisfactory. Their chief defect has been that no explicit consideration was given to the finite conductivity of the ground (except that it produces a tangential component f_a). In the derivation given above, the effect of ground conductivity is expressly recognized and in the final equation appears implicitly in the values of K and γ. These parameters, it will be observed, are experimentally determinable, and are the only parameters besides Y' appearing in the final differential equation.

A formula for the quantity G^0 occurring in equations (6) and (7) can be derived by application of Gauss' theorem, as follows: Let E_r denote the radial component of the total electric force at the surface of the wire, dS a differential element of the surface of the wire, and σ and ϵ the conductivity and specific inductive capacity of the medium, which is homogeneous and isotropic by assumption. Then the leakage current I' flowing outward, per unit length of the wire, is given by

$$I' = \int \sigma E_r dS = \frac{\sigma}{\epsilon} \int \epsilon E_r dS, \tag{15}$$

the surface integral being taken over the unit length of the wire. But, by Gauss' theorem (ξ being a constant whose value depends only on the units),

$$\int \epsilon E_r dS = 4\pi Q/\xi, \tag{16}$$

the resultant axial electric flux from the ends of the element being negligible compared with the radial electric flux from the lateral surface. Thus

$$I' = \frac{4\pi\sigma}{\epsilon\xi} Q, \tag{17}$$

and comparison of this equation with (6) gives the result

$$G^0/C = 4\pi\sigma/\epsilon\xi. \tag{18}$$

In this connection it may be noted that the leakage current represented by (15) does not directly depend on the impressed field, but only on the field produced by the wire itself. This is because the assumed medium is homogeneous and isotropic; hence σ in (15) can be taken outside the sign of integration, and then the conclusion follows from (15) by noting that one of the constituents of E_r is the impressed radial electric force f_r, and that

$$\int f_r dS = \int \operatorname{div} f \cdot dv = 0,$$

since the divergence of the impressed electric force must be zero. The conclusion would not follow, in general, if the medium were either heterogeneous or æolotropic. It may be noted that a homogeneous isotropic medium surrounding a wire and containing direct leakage admittance paths from the wire to ground may be regarded as a heterogeneous æolotropic medium.

The value given for δ in equation (9) of the text, namely $\delta = 4\pi\sigma/\epsilon\xi$, is readily derivable by combining equations (4), (10), (5) of the text with (17) of this Appendix.

Equations (90)

If there were no impressed potential at the primary wires ($F_k = 0$), the equations of continuity would be merely

$$-\frac{dI_h}{dx} = \sum_{k=1}^{n} Y_{hk} V_{k\pi}, \qquad (h = 1, \cdots n), \tag{19}$$

where, in accordance with equations (7) in Section I,

$$Y_{hk} = g_{hk} + i\omega q_{hk}. \tag{20}$$

It should here be remarked that Y_{hk} depends not only on the geometry of the system and on the conductivity of the medium but also on any direct leakage admittance existing between the wires themselves and also on any between the wires and ground. The direct leakage admittance between wires h and k, per unit length, will be denoted by Y'_{hk} and that between wire h and ground, by Y'_{hh}; these are regarded as being uniformly distributed along the system.

When there is present an impressed potential, the existence of the direct leakage admittances gives rise to the following supplementary terms for the right side of equation (19):

$$F_h Y'_{hh} + \sum_{\substack{k=1\\k\neq h}}^{n} (F_h - F_k) Y'_{hk} = \sum_{k=1}^{n} X_{hk} F_k,$$

where

$$\begin{aligned} X_{hk} &= -Y'_{hk} \quad \text{for} \quad k \neq h,\\ X_{hh} &= \sum_{k=1}^{n} Y'_{hk}. \end{aligned} \tag{21}$$

It is seen that Y_{hk} and X_{hk} are of the nature of admittances (per unit length), although they are not 'direct admittances.' Their precise meanings are readily deducible from equations (90).

When the medium itself is of zero conductivity, g_{hk} reduces to X_{hk}.

Abstracts of Bell System Technical Papers Not Appearing in this Journal

Direct Determination of Hydrocarbon in Raw Rubber, Gutta-Percha, and Related Substances.[1] A. R. KEMP. Iodochloride in glacial acetic acid is shown to be a suitable reagent to determine the unsaturation of the hydrocarbon in rubber or gutta-percha. The influence of time, temperature, sunlight, and reagent concentration upon the reaction is shown.

Comparisons are made between iodochloride, iodobromide and bromine relative to their reactions with raw rubber and some of the terpenes.

Results of analyses of several rubber and gutta-percha samples are given.

The effects of mastication and heat upon the unsaturation of raw rubber are shown.

Microtomic Preparation of Soft Metals for Microscopic Examination.[2] F. F. LUCAS. This paper outlines the apparent limitations of polishing methods for preparing specimens of soft metals for metallographic examination. A microtome method has been developed and its successful application to the study of lead cable sheath alloys illustrated. Much time and labor are saved and results have been obtained which were impossible by polishing methods.

In lead-antimony cable sheath alloys a widened grain boundary phenomenon was disclosed by the new method and the probable nature of the structural changes determined. Changes in structure due to aging are shown and those which accompany thermal or mechanical treatment of the metal may be followed clearly.

Distribution of Energy in Worked Metals.[3] LYALL ZICKRICK and R. S. DEAN. The purpose of this paper is to give experimental results that seem to be in accord with the theory, that in the deformation of a crystal the energy supplied is distributed to the atoms of the lattice, probably by the forced formation of molecules.

[1] *Journ. Ind. and Engr. Chem.*, Vol. 19, p. 531, 1927.
[2] Institute of Metals Division, A. I. M. E., February 15, 1927.
[3] *Wire*, Vol. 2, p. 161, 1927.

Modern Developments in Inspection Methods.[4] E. D. HALL. This extensively illustrated article describes inspection methods as carried out at the Hawthorne plant of the Western Electric Company. The number of individual piece parts manufactured is more than 100,000 and the number of inspection gauges employed totals more than 25,000. Machine testing and gauging for certain parts is described and cost savings resulting therefrom are given. One of the machines described tests a porcelain protector block containing a carbon insert. The machine is adjusted to accept blocks from which the recess distance of the carbon lies between 0.0024 and 0.0032 inch and to reject, when the distance is 0.0023 inch or less or 0.0033 inch or more. The machine performs its operation at the rate of 2,800 blocks per hour. It is used to test about 4,500,000 blocks per year and represents an annual saving of approximately $2,500.

A Direct Comparison of the Loudness of Pure Tones.[5] B. A. KINGSBURY. The loudness of eleven pure tones was studied by adjusting the voltage applied to a telephone receiver to make these tones as loud as certain fixed levels of a 700-cycle tone. The average results of 22 observers, 11 men and 11 women, were arranged as contour lines of equal loudness through the normal auditory sensation area in terms of r.m.s. pressure in car canal as a function of frequency. Frequencies from 60 to 4,000 cycles were used and intensities from threshold of audibility to 90 T. U. above the 700-cycle threshold. It was found that if the amplitudes of pure tones are increased in equal ratios the loudness of low frequency tones increases much more more rapidly than that of high frequency tones. For frequencies above 700 cycles the rate is nearly uniform.

As a loudness unit the least perceptible increment of loudness of a 1,000-cycle tone was employed. In absolute magnitude this varies from level to level, but in the ordinary range of loudness it becomes constant. This unit takes into account the subjective character of loudness.

The variability of the data from which the averages were computed was separated into a factor expressing dissimilarity of ears and another expressing errors of observers' judgment. There was no level at which the variances were a minimum. Dissimilarity of ears causes more variation than errors of observers' judgment. The variances showed no significant sex difference.

[4] *Mechanical Engineering*, Vol. 48, p. 1435, 1926.
[5] *Physical Review*, Vol. 29, p. 588, 1927.

The Scattering of Electrons by a Single Crystal of Nickel.[6] C. DAVISSON and L. H. GERMER. Preliminary announcement is made in this note of the discovery that a beam of swiftly moving electrons in its reaction with a single crystal of nickel behaves in some respects as if it were a beam of wave radiation such as light or x-rays. As the speed of the electron beam is increased a series of critical speeds is found at which sharply defined beams of scattered electrons issue from the crystal. This is similar to what is observed when a beam of monochromatic x-rays is sent into a crystal—as the wave-length of the x-rays is decreased a series of critical wave-lengths is found at which sharply defined beams of scattered x-rays issue from the crystal. This x-ray phenomenon is quantitatively accounted for as due to the interference of waves scattered by the regularly arranged atoms of the crystal. In fact, it was this phenomenon discovered by Laue, Friedrick and Knipping in 1913 that established the wave nature of x-radiation, and it is from measurements based on this phenomenon that the lengths of x-ray waves are determined.

The analogous electron phenomenon is less simple, and yet it is simple enough and of such a nature as to leave little doubt that a beam of swiftly moving electrons is in some sense equivalent to a beam of wave radiation. The wave-length of the equivalent radiation can be measured, and is found to be in satisfactory agreement with requirements of the new theory of wave or undulatory mechanics: namely, that the wave-length of the equivalent radiation shall be equal to h/mv, where h represents Planck's universal constant of action, and mv the momentum of an individual electron.

Structure of a Protective Coating of Iron Oxides.[7] RICHARD M. BOZORTH. It is shown that the Bower-Barff protective coating, produced by the action of steam on iron at about 700° with subsequent cooling in air, is built up of layers of ferrous oxide, magnetite and ferric oxide, arranged in this order (the order of oxidation) upon the iron base. The thicknesses of these layers are estimated to be of the order of 10^{-2}, 2×10^{-4} and 2×10^{-5} cm., respectively. The data on which the above conclusions are based are the positions and intensities of lines on powder photographs taken with molybdenum, iron and copper $K\alpha$ X-rays. The iron and copper $K\alpha$ X-rays penetrate the coating to different depths and give information about different parts of its structure because their wave-lengths are, respectively, a little greater and a little less than the critical-absorption wave-length of the iron which forms the greater part of the coating.

[6] *Nature, 119*, 558 (1927).
[7] *J. Amer. Chem. Soc.*, Vol. 49, pp. 969–976 (1927).

Contributors to this Issue

J. G. FERGUSON, B.S., University of California, 1915; M.S., 1916; research assistant in physics, 1915–16; Engineering Department of the Western Electric Company and Bell Telephone Laboratories, 1917–. Mr. Ferguson's work has been in connection with the development of methods of electrical measurement.

J. J. GILBERT, A.B., University of Pennsylvania, 1909; Harvard, 1910–11; Chicago, 1911–12; E.E., Armour Institute, 1915; instructor of electrical engineering, Armour, 1912–17; Captain Signal Corps, 1917–19; Engineering Department, Western Electric Company, 1919; Bell Telephone Laboratories, Inc., 1925–. Mr. Gilbert has worked primarily on submarine cable problems.

A. A. CLOKEY was employed before the war in the Engineering Department of the Western Union Telegraph Co. During the war, he served as captain in the Signal Corps working on experimental investigations to improve cable communication. In 1919, he joined Bell Telephone Laboratories and has since been in charge of the development of terminal equipment for permalloy loaded cables.

A. M. CURTIS came to the Engineering Department of the Western Electric Company in 1913 after having spent several years as radio engineer for the Brazilian Government. During the war he was commissioned and sent to France to serve in the Division of Research and Inspection of the Signal Corps. In 1919, he returned to Bell Telephone Laboratories and has since been engaged with the applications of vacuum tube amplifiers to submarine cables.

E. PETERSON, Cornell University, 1911–14; Brooklyn Polytechnic, E.E., 1917; Columbia, A.M., 1923; Ph.D., 1926; Electrical Testing Laboratories, 1915–17; Signal Corps, U. S. Army, 1917–19; Engineering Dept., Western Electric Co., 1919–24; Bell Tel. Labs., 1924–.

H. P. EVANS, B.S. in electrical engineering, University of Wisconsin, 1923; M.S., 1927. Mr. Evans was with Bell Telephone Laboratories from 1923 to 1925, returning to the University of Wisconsin as instructor in electrical engineering and then as research assistant in physics.

Mr. Evans' contributions to the present paper were made while still a member of the Laboratories staff.

EDWARD C. MOLINA, Engineering Department of the American Telephone and Telegraph Company, 1901–19, as engineering assistant; transferred to the Circuits Design Department to work on machine switching systems, 1905; Department of Development and Research, 1919–. Mr. Molina has been closely associated with the application of the mathematical theory of probabilities to trunking problems and has taken out several important patents relating to machine switching.

JOHN R. CARSON, B.S., Princeton, 1907; E.E., 1909; M.S., 1912; Research Department, Westinghouse Electric and Manufacturing Company, 1910–12; instructor of physics and electrical engineering, Princeton, 1912–14; American Telephone and Telegraph Company, Engineering Department, 1914–15; Patent Department, 1916–17; Engineering Department, 1918; Department of Development and Research, 1919–. Mr. Carson's work has been along theoretical lines and he has published several papers on theory of electric circuits and electric wave propagation.

RAY S. HOYT, B.S. in electrical engineering, University of Wisconsin, 1905; Massachusetts Institute of Technology, 1906; M.S., Princeton, 1910; American Telephone and Telegraph Company, Engineering Department, 1906–07; Western Electric Company, Engineering Department, 1907–11; American Telephone and Telegraph Company, Engineering Department, 1911–19; Department of Development and Research, 1919–. Mr. Hoyt has made contributions to the theory of transmission lines and associated apparatus, and more recently to the theory of crosstalk and other interference.