The Bell System Technical Journal

January, 1931

The Detection of Two Modulated Waves Which Differ Slightly in Carrier Frequency *

By CHARLES B. AIKEN

The present paper contains an analysis of the detection of two waves modulated with the same, or with different, audio frequencies and differing in carrier frequency by several cycles or more. Both parabolic and straight line detectors are treated and there are derived the expressions for all of the important audio frequencies present in the output of these detectors when such waves are impressed. There are discussed the types of interference which result when one station is considerably weaker than the other and simple attenuation formulæ are employed in estimating the character and extent of the interference areas around the two transmitters. Beyond the use of such formulæ no attention is given to phenomena which may occur in the space medium such as fading, diurnal variations in field intensity, etc.

WHENEVER one of two stations operating on the same wave-length assignment wanders from its proper frequency, waves are likely to be received which differ in carrier frequency by several cycles or more. Under such conditions the two signals may be thought of as made up of entirely distinct frequencies and phase relations between analogous components of the two waves need not be considered. In the important case in which the carriers are of identical frequency this is no longer true and phase and its dependence on position and transmission phenomena must be taken into account. This case will be reserved for future study, the present work being limited to a consideration of the phenomena connected with the detection of distinct frequencies.

The most important undesired frequency which is present in the output of the detector is the beat note between the two carriers. It is sometimes carelessly assumed that if the frequency of this beat note is reduced below the audible range the only remaining interference will be due to the speech from the undesired station. Such is not the case and it will be shown later on that when the beat frequency is reduced below the audible range, but not to zero, there remains a group of spurious frequencies which will introduce an interfering background. When the undesired carrier is of relatively small intensity this background is a great deal stronger than the interfering speech. It is therefore desirable to obtain quantitative data on the interfering spec-

* *Proc., I. R. E.*, Jan., 1931.

trum which occurs in the receiver output, in terms of the intensities and degrees of modulation of the input signals.

It is to be expected that the results obtained will depend, to some extent at least, on the type of detector which is used. The square law characteristic is a fair approximation to that of any detecting device which is worked over only a small range and hence an analysis of this characteristic may be expected to serve as an excellent guide to general detector performance. When large signals are impressed on the detector the functioning of the device may approximate more closely to that of the ideal straight line detector. It has been felt that a study of these two types would furnish data from which the performance of any intermediate type of detector could be inferred without great error. As the problem of the square law detector is very much the simpler it will be considered first.

Mathematical Analysis

There will be assumed two broadcasting stations transmitting on frequencies which differ by a relatively small amount, the beat frequency being restricted to the audible range or less. Each of the carriers will be assumed to be modulated by a single audio frequency, the modulating frequencies at the two stations being, in general, different. The total signal impressed on the receiving detector will then be of the form

$$v = E(1 + M \cos pt) \cos \omega_1 t + e(1 + m \cos qt) \cos \omega_2 t, \qquad (1)$$

in which

v is the total alternating voltage impressed on the detector.
E is the amplitude of the desired carrier.
e is the amplitude of the undesired carrier.
M is the degree of modulation of the desired signal.
m is the degree of modulation of the undesired signal.
$\omega_1/2\pi$ is the frequency of the desired carrier.
$\omega_2/2\pi$ is the frequency of the undesired carrier.
$p/2\pi$ is the frequency of the desired modulation.
$q/2\pi$ is the frequency of the undesired modulation.

Square Law Detector

We shall first suppose this signal to be impressed on a detector which will be assumed to have a characteristic in the neighborhood of the operating point, of the form

$$i = A_0 + A_1 v + A_2 v^2. \qquad (2)$$

An expression of this type will accurately represent a small portion of any continuous characteristic. The present analysis requires that the impressed e.m.f. shall be of small amplitude in order that the limits of the portion of the characteristic thus represented may not be exceeded. This restriction is necessary in treating square law detectors.

The audio frequency output of the detector will be due entirely to the second order term in (2). Hence it will be sufficient, for our purposes, to square the expression for v. We are interested primarily in the ratios of the amplitudes of the various undesired audio frequencies produced to the amplitude of the desired signal of frequency $p/2\pi$. Such a ratio will be designated as a relative amplitude. Neglecting circuit constants, etc., which will apply equally in all the expressions for the various frequencies, the amplitude of the desired component of the audio frequency output is readily shown to be E^2M. The expression for v^2 is reduced to first power sinusoids and the amplitude of each frequency converted to a relative amplitude by dividing by E^2M. The case in hand yields twelve undesired audio frequencies, the relative amplitudes of which are listed in table I. Before commenting on these results we shall consider the straight line detector.

TABLE I

Angular Velocity	Ratio to E^2M	Angular Velocity	Ratio to E^2M
$2p$	$\frac{M}{4}$	$p \pm u$	$\frac{e}{2E}$
q	$\frac{e^2m}{E^2M}$	$q \pm u$	$\frac{em}{2EM}$
$2q$	$\frac{e^2m^2}{4E^2M}$	$p \pm q \pm u$	$\frac{em}{4E}$
u	$\frac{e}{EM}$		

in which $u = \omega_1 - \omega_2$.

The Straight Line Detector

In making analyses of rectification by a straight line detector it is customary to reduce the sum of the various impressed radio frequencies to a single radio frequency, the amplitude and phase angle of which are slow functions of time. The most common example of this type of treatment is a combination of the carrier and two side bands of single frequency modulation into the familiar expression for a modulated wave in which the amplitude of the radio frequency is an audio frequency function. In this case the radio frequency phase angle is constant. In the case of a single frequency modulation with one side-band eliminated there are impressed on the detector input only two frequencies. These may be combined in a well known manner.[1] Thus, if the impressed voltages are of the form acosx and bcosy, then the amplitude is given by

$$\sqrt{a^2 + b^2 + 2ab \cos (x - y)}. \tag{3}$$

The expression for the phase angle will not be given here as it can be shown that if a and b are unequal and the difference between the frequencies $x/2\pi$ and $y/2\pi$ is small compared with either frequency, then the variation of the phase angle with time may be neglected in computing the audio frequency components. In the present case we have two radio frequency waves the amplitudes of which are not constants but are slow functions of time and these may be substituted for a and b in (3). Thus the effective amplitude of the total input signal may be taken to be

$$S = \sqrt{A^2 + B^2 + 2AB \cos ut}, \tag{4}$$

in which

$$A = E(1 + M \cos pt),$$
$$B = e(1 + m \cos qt),$$

and

$$u = \omega_1 - \omega_2.$$

The problem then resolves itself into an analysis of the detection, by a straight line detector, of a single radio frequency component. The results of such an analysis are well known and it can be readily shown that the audio frequency output may be obtained, except for a factor of proportionality, by resolving the amplitude into its audio frequency components. In the present case the amplitude to be resolved is given by (4) which may be written

$$S = \sqrt{(A + B)^2 - 2AB(1 - \cos ut)}.$$

The interfering signal B will be taken to be always less than the desired signal A, and hence $A^2 + B^2 > 2AB$, from which it follows that $(A + B)^2 > 2AB(1 - \cos ut.)$ Hence the radical may be expanded by the binominal theorem, giving

$$S = A + B - \frac{AB(1 - \cos ut)}{A + B} - \frac{A^2B^2(1 - \cos ut)^2}{2(A + B)^3} - \frac{A^3B^3(1 - \cos ut)^3}{2(A + B)^5} \cdots \tag{5}$$

[1] Vide: Lord Rayleigh, "Theory of Sound," page 23, sec. ed.

It is to be observed that each of the terms of this series, except the first, contains time in the denominator and hence further expansions are necessary. The denominators of the various terms can be expanded by the binominal theorem in such a way as to put all the expressions containing time in the numerators, the expansions being in powers of

$$(ME \cos pt + me \cos qt)/(E + e).$$

By the proper trigonometric transformations it is possible to reduce the final expression for S to frequencies in p, q, u and the sums and differences of the various multiples of these quantities. An additional discussion of this analysis is given in an appendix. In order that the various series involved may converge with a manageable degree of rapidity it is necessary to limit the relative amplitudes of the interfering carriers and the degrees of modulation as well. Consequently the solutions are restricted to intensities of the interfering carrier of 0.1, or less, of the desired carrier and to degrees of modulation of either signal ranging from 0.1 to .5. These limits are suitable also because we are interested chiefly in interference by a relatively weak signal, the interference caused by a signal, the carrier amplitude of which is greater than 0.1 of that of the desired carrier amplitude being near the tolerable limit in the majority of cases. The upper value for the modulation of 0.5 is approximately equal to the average degree of modulation of a station employing as deep modulation as is practical, only the peaks running up to nearly unity. The value of 0.1 for the lower limit is of course transgressed by soft passages in speech or music. However, the range here specified is sufficiently large to give an excellent idea of what may be expected from various degrees of modulation of desired and interfering signals and the results of more extreme cases may be inferred from the data here developed. Under these limits it is found that the only audio frequencies of any importance which appear in the output are:

$$\begin{aligned} S = &\left(ME - eg\left(a_0 M - a_1 + a_2 \frac{M}{2} \right) + \frac{m^2 e^2 M g^2}{2E} \right) \cos pt \\ &+ \left(me - eg\left(a_0 m - \frac{a_1 M m}{2} - \frac{meg}{E} \right) - \frac{3e^2 g^3 b_0 m}{2E} \right) \cos qt \\ &+ \left(\frac{m^2 e^2 g^2}{2E} - \frac{b_0 e^2 g^3 m^2}{4E} \right) \cos 2qt \\ &+ \left(eg\left(a_0 - \frac{a_1 M}{2} - \frac{m^2 eg}{2E} \right) + \frac{b_0 e^2 g^3}{2E}(2 + m^2) \right) \cos ut \\ &- \frac{b_0 e^2 g^3}{4E} \cos 2ut \end{aligned}$$

$$+ eg\left(\frac{a_0M}{2} - \frac{a_1}{2} + \frac{a_2M}{4} - \frac{m^2Meg}{4E}\right)\cos\,(p \pm u)t$$
$$+ \left(eg\left(\frac{a_0m}{2} - \frac{a_1Mm}{4} - \frac{meg}{2E}\right) + \frac{b_0e^2g^3m}{E}\right)\cos\,(q \pm u)t. \tag{7}$$

In which

$$\left.\begin{aligned} a_0 &= 1 + \frac{M^2g^2}{2} + \frac{3M^4g^4}{8}, \quad a_1 = Mg + \frac{3M^3g^3}{4} + \frac{5M^5g^5}{8},\\ a_2 &= \frac{M^2g^2}{2} + \frac{M^4g^4}{2}, \qquad\quad b_0 = 1 + 3M^2g^2,\\ g &= \frac{E}{E+e}. \end{aligned}\right\} \tag{7a}$$

Comparison Between Detectors

It is now possible to make a comparison between the performance of the straight line and the square law detectors. In Figs. 1 to 4 are shown the relative amplitudes of the interfering frequencies in the two cases for various degrees of modulation. The data for the square law case are indicated by dashed lines and for the straight line case by solid lines, and where the two coincide this is noted on the figures. It is to be noted that the expression for the amplitude of the desired frequency $p/2\pi$ is a complicated function. However, computation shows that over the range in which we are interested, the value of this expression does not differ from ME by more than 1 per cent and, therefore, this value has been assumed in computing the relative amplitudes of the other frequencies.

Probably the most striking feature to be noted in comparing the two cases is the similarity of the results. This is particularly evidenced by the carrier beat note of frequency $u/2\pi$ the amplitude of which differs in the two cases by an inappreciable amount. The spurious frequencies $(q \pm u)/2\pi$ also are practically identical for both detectors. There are, however, several important differences as follows:

The group of spurious frequencies of angular velocity $p \pm q \pm u$, which is of appreciable importance in the square law case, is entirely absent from the range of magnitude considered when a straight line detector is employed. The frequencies $(p \pm u)/2\pi$ are greater in the square law case over the range which we have considered, but the curve which represents them has a smaller slope than in the straight line case and for larger values of the interfering signal the intensities of these frequencies would be relatively less with the square law detector. The intensity of the undesired speech q is definitely less in the straight line case than in the square law case but the slope of the q curves is about the same for both except for $M = m = 0.5$. It is of interest to observe that the interfering speech received on the straight line detector is very much less in intensity than would be the case if the strong desired signal were absent, and that the variation of the amplitude of this frequency with intensity of the undesired carrier is greater when the desired frequency is present. We have here an analytical description of the familiar masking effect which occurs when a strong unmodulated carrier is received simultaneously with a weak modulated signal. For example, when $e/E = 0.1$ it can be seen from Fig. 1 that the relative amplitude of the component of frequency $q/2\pi$ is 0.0063 for the case of the straight line detector. If this component were unaffected by the presence of the strong signal it would have an amplitude proportional to em and a relative amplitude of em/EM which for the values here considered is 0.1. Hence the "masking" effect is here responsible for a reduction of 24 db.

Lastly, it may be mentioned that there are in the case of the straight line detector certain frequencies of small amplitude which are entirely absent from the square law case. However, no frequency is shown the relative amplitude of which is less than 0.01 for all four pairs of values of M and m, as such frequencies are unimportant. An exception is made with regard to $p \pm u$. This is always less than 0.01 over the range considered but is included for the sake of comparison with the square law results.

Further Consideration of Detector Output

The second harmonic of the desired signal is of importance only in the square law case. It is of the nature of a distortion which is independent of the interference and may be omitted from the consideration of the undesired audio frequencies which are a result of the interference. From Figs. 1 to 4 it is evident that the most important interfering frequencies are those of angular velocity, u, $q \pm u$, $p \pm u$ and $p \pm q \pm u$, the last being of importance only in the case of the square law detector. It is with these frequencies, together with that of the interfering speech $q/2\pi$, that we shall be chiefly concerned.

When the relative magnitudes of the interfering frequencies, which are tabulated on page 3, are multiplied by E^2M, the resulting quantities are proportional to the absolute magnitudes of these frequencies. It is to be noted that the frequencies of greatest interest have absolute magnitudes which are linear functions of M or m except $(p \pm q \pm u)/2\pi$ which is proportional to mM, and $u/2\pi$ which is independent of both M and m and will, therefore, be unaffected by the type of modulation employed at either station. In case there are several frequencies present in the modulation of each station the radio frequency waves will be of the form $E(1 + M_1 \cos p_1 t + M_2 \cos p_2 t + \cdots) \cos \omega_1 t$ and $e(1 + m_1 \cos q_1 t + m_2 \cos q_2 t + \cdots) \cos \omega_2 t$. For every frequency of the former case which contained M as a factor of its amplitude we shall now have several frequencies respectively proportional to M_1, M_2 etc. while an analogous new group will correspond to the former frequencies

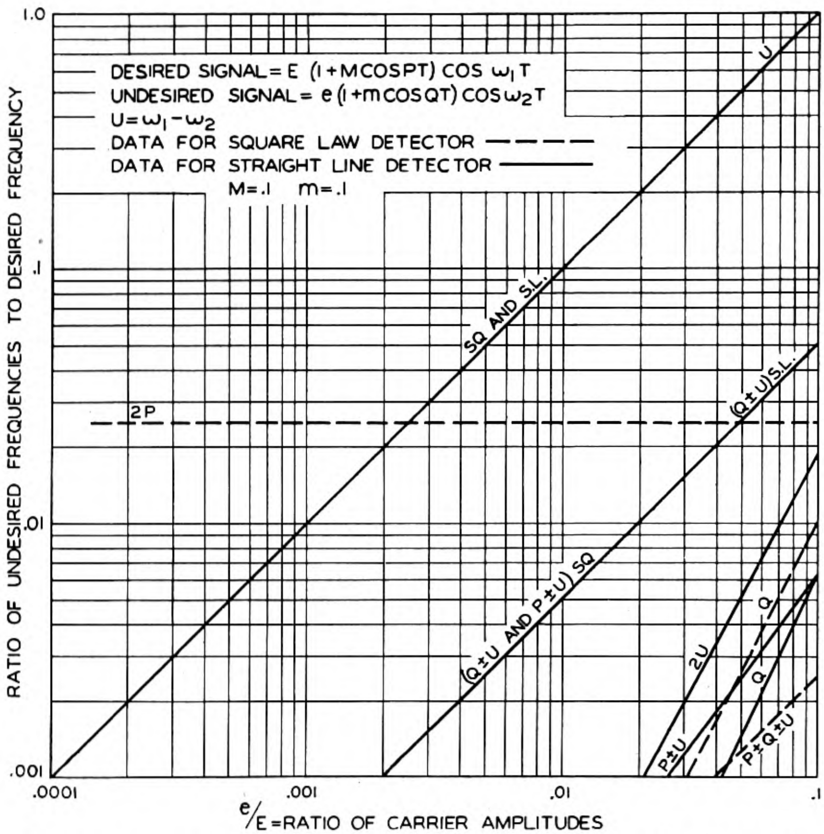


Fig. 1—Relative amplitudes of undesired frequencies as a function of the ratio of the amplitudes of the desired and the interfering carriers. Modulation of both stations small and equal.

containing m. Hence we shall have two frequency spectra derived from the desired speech spectrum containing the p's, but one of the spectra will be shifted upward in frequency by an amount $u/2\pi$ and the other downward by the same amount. Two additional spectra will be derived in a similar manner from the undesired speech spectrum containing the q's. The frequencies of the type $(p \pm q \pm u)/2\pi$ will be numerous as there will be a product of the M's with each of the m's. However, these are of even moderate importance only when the modulations of both stations are high, and a square law detector is employed at the receiver.

Hence we may picture the interference as made up chiefly of displaced frequency spectra of the type mentioned above, of a carrier

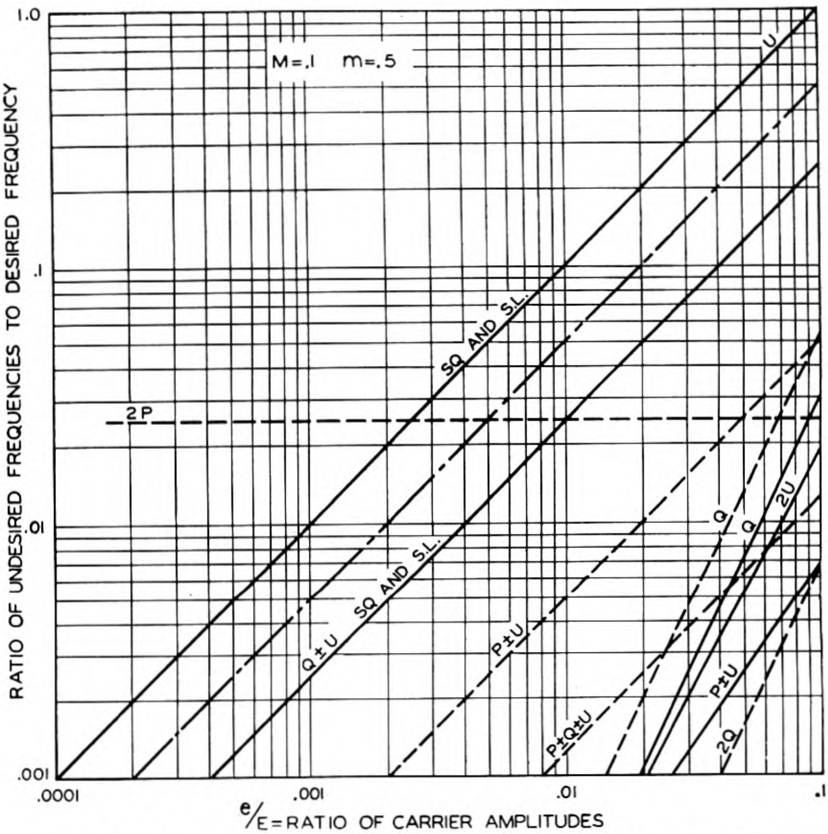


Fig. 2—Relative amplitudes of undesired frequencies as a function of the ratio of the amplitudes of the desired and the interfering carriers. Modulation of desired station small and of interfering station large.

beat and of the interfering speech, which is weak but important because of its intelligibility. The results in the case of a straight line detector would not be very greatly different. The frequencies of the type $(p \pm q \pm u)/2\pi$ would be negligible, the two spectra derived from $p \pm u$ would be much less important and certain new, but rather small cross product frequencies would appear.

In estimating the interference the carrier beat can be considered by itself and from the data at hand there can be derived the areas around each of two stations having approximately the same carrier frequency, inside of which the amplitude of the beat note will be down a given number of db from that of the desired speech. The same is true of the interfering speech when it is different from the desired speech. The

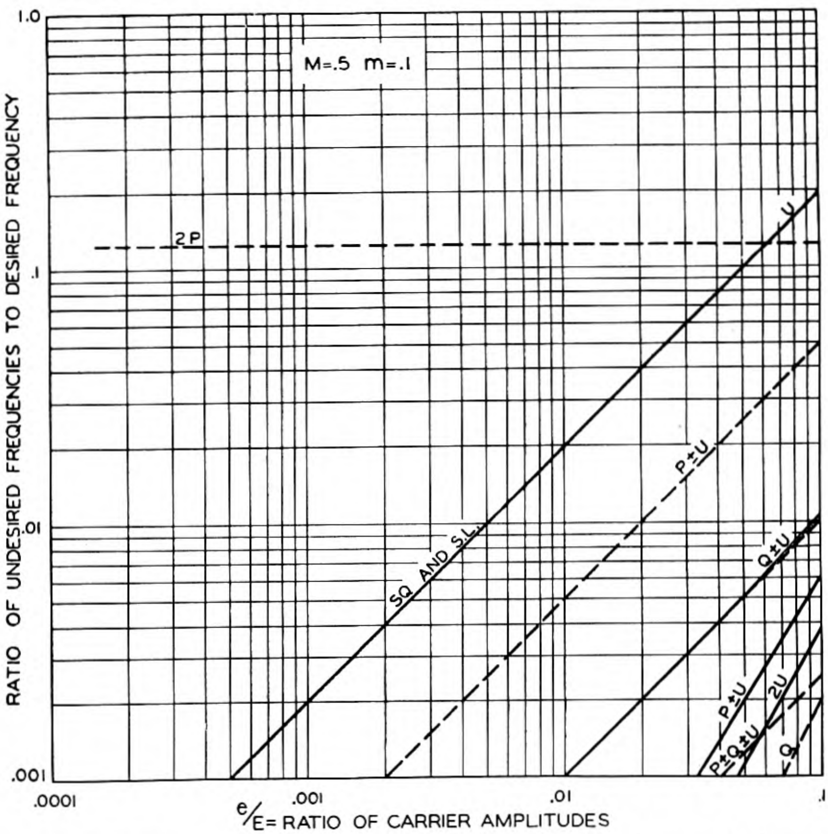


Fig. 3—Relative amplitudes of undesired frequencies as a function of the ratio of the amplitudes of the desired and the interfering carriers. Modulation of desired station large and of interfering station small.

frequencies $(p \pm u)/2\pi$, $(q \pm u)/2\pi$, $(p \pm q \pm u)/2\pi$, etc., will combine to form a disturbing background which we shall designate as "displaced side band interference." This may be taken to include all of the interfering frequencies except those of the undesired speech and its entirely unimportant harmonics. (The frequency $2p/2\pi$ is not here classed as an interfering frequency.)

From Figs. 1 and 4 it is to be noted that when $m = M$ the frequencies $(q \pm u)/2\pi$ are the largest components of the displaced side band interference if a straight line detector is used and have the same amplitude as the $(p \pm u)/2\pi$ components if a square law detector is used. When $m > M$ the $q \pm u$ group is much more important than the $p \pm u$ group as is evident from Fig. 2. When $M > m$ the $q \pm u$

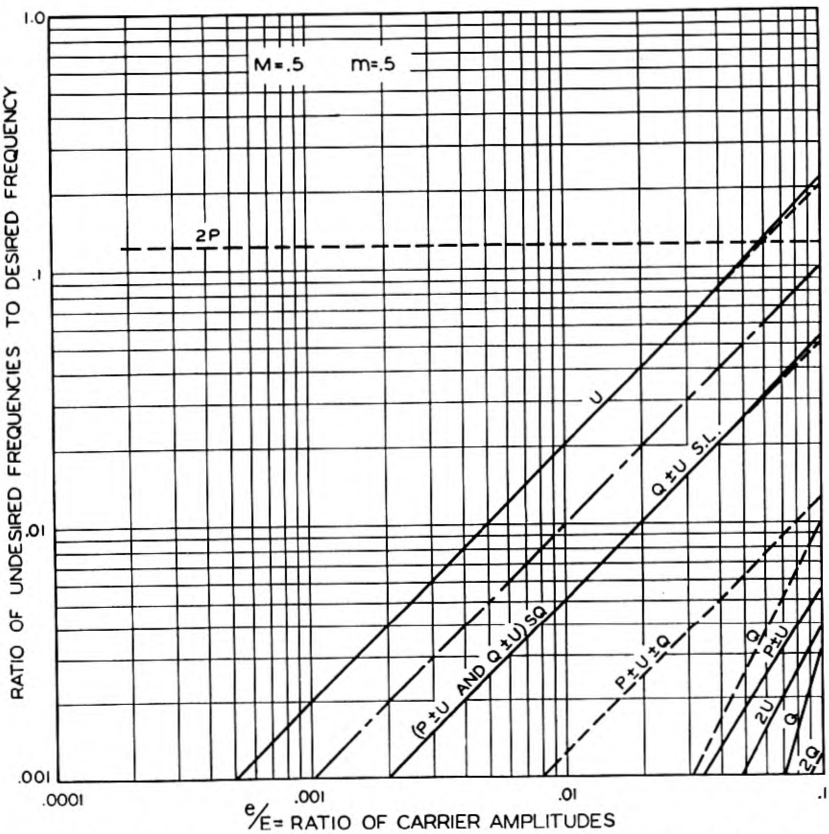


Fig. 4—Relative amplitudes of undesired frequencies as a function of the ratio of the amplitudes of the desired and the interfering carriers. Modulation of both stations large and equal.

group is less important but this case is of no great interest for if the stations are transmitting identical programs, with similar degrees of modulation, it cannot occur and if the programs are different then the interference is determined primarily by what happens when $m > M$. Consequently we may consider that the $q \pm u$ group constitutes the most important part of the displaced side band interference except when a square law detector is used and the programs are identical. In such a case we shall assume that both stations employ the same degree of modulation and that therefore the $q \pm u$ and $p \pm u$ groups are of the same importance.

Interference Areas of Stations

We have distinguished between three types of interference, namely, carrier beat, unwanted speech and displaced side band. We shall now compute, for several values of attenuation, percentage modulation etc., the areas around a transmitting station inside of which each of these types of interference, due to a second station, will have a relative importance which is not greater than a certain specified amount.

In estimating these areas we must deal with two possible cases which may arise in practice: (1) The two stations transmit different programs. (2) The programs are the same. The carriers are assumed to differ in frequency in both cases.

Case 1

The importance of the various types of interference which are present, will be determined by their ratios to the intensity of the desired speech. In the present case in which the two stations transmit different programs, the amount of interference which may be tolerable will be determined by what occurs when the modulation of the desired station is low, while that of the interfering station is high. Hence, in studying this case we shall make use of Fig. 2, which gives data computed on the basis of a modulation of 0.1 for the desired station and 0.5 for the interfering station.

Taking up first the consideration of the carrier beat note, we shall determine the curve along which the intensity of the beat is down a given number of db from the desired speech. The position of this curve will depend on the degree of modulation of the desired signal, since the lower the modulation the more noticeable will be a beat note of a given intensity. When we have specified the db difference which must exist between these two components of the receiver output the carrier ratio can be picked off from the u line of Fig. 2.

In order to determine the curve along which this carrier ratio exists we shall proceed as follows:

The desired station will be considered to be at the origin of a system of rectangular coordinates and the undesired station will be at the point (D, O). We shall assume that the powers of the desired and undesired stations are P_1 and P_2, respectively, and that their distances from a point in the coordinate plane are d_1 and d_2; then if we denote the ratio of the carriers by $K = e/E$ the equation of the curve along which the value of K is constant is given by:

$$\frac{K\sqrt{P_1}}{d_1}\epsilon^{-gd_1} = \frac{\sqrt{P_2}}{d_2}\epsilon^{-gd_2}. \tag{8}$$

This equation is based upon a convenient form of the Austin-Cohen [2] formula for the intensity of the field radiated from a radio transmitter. This formula is:

$$E = A\,\epsilon^{-101.5\alpha d/\lambda^{0.6}}, \tag{9}$$

in which λ is the wave-length in meters, d is the distance from the transmitter in miles and α is an attenuation constant which may range from zero up to 0.01 or even more. In writing down equation (8) we have used the abbreviation:

$$g = \frac{101.5\alpha d}{\lambda^{0.6}}, \tag{10}$$

From (8) there have been computed curves for the case in which $P_1 = P_2$ and for various values of K and α. λ has been taken as 300 meters and D, the distance between the stations, as 1,000 miles.

In Fig. 5 are shown several curves for $\alpha = 0.001$. For small values of K, the curves are practically circular and are of small area. As K increases, the curves become oval shaped and it can be readily shown that for values of K greater than a certain critical amount, the curves will not close but will be of a shape which is roughly hyperbolic.

In Fig. 6 are shown curves corresponding to a value for α of 0.002. It is to be noted that an increase in α enormously increases the area inside of which the ratio of the carriers is less than a certain value. The effect of α will of course be dependent upon the magnitude of the distance between the stations and will be more pronounced the larger this distance. For the present case in which $D = 1{,}000$ miles, there is not much point in considering values of α larger than 0.002, since the attenuation would be so great as to make the effect of one station on the service area of the other of very little consequence.

If we specify that the carrier beat must be at least 40 db down from the speech output due to a 10 per cent modulated signal, then curve 1 of Figs. 5 and 6 will represent the areas inside of which this requirement will be met, while if we call for an interval of 20 db between these two components, curve 5 of Figs. 5 and 6 will represent the areas in which the condition is satisfied. It is evident that if a rigid restriction is placed on the permissible beat note interference which may be allowed, and if the attenuation is of a small value then the area in which the beat

[2] L. W. Austin, *Proc., I. R. E.*, Vol. 14, p. 377.

note may be neglected is extremely small. On the other hand this area increases very rapidly as the attenuation increases.

We may use the same sets of curves in considering the displaced side band interference. From Fig. 2 it is evident that by far the most important components of this interference are those represented by the $(q \pm u)$ group. In order to estimate this interference we must follow some rule for combining the $q + u$ component with the $q - u$ component. In order to do this in a strictly correct manner we should have to take into account the frequencies and sensation levels of the components. However, it has been shown [3] that over a considerable portion of the audio frequency range, and for sensation levels of approximately the magnitude in which we are interested, the interfering effect of these frequencies may be taken to be approximately equal to that due to a single frequency of twice the amplitude of either component. We shall therefore take our data from the dash-dot curve of Fig. 2. From this curve it appears that if the displaced side band interference is to be 40 db down from the desired speech, we must have a carrier ratio of 0.002, while if it is to be 20 db down from the desired speech the corresponding carrier ratio is 0.02. The curves corresponding to these values are shown by 2 and 6, respectively, on Figs. 5 and 6.

From this it appears that the area in which the side band noise is not objectionable may be a great deal larger than that in which the carrier beat is of a tolerable intensity. If the frequency of the carrier beat is reduced below the useful audible range then the former area may be considered to be entirely free from interference of any kind. Consequently, it is highly desirable to limit the maximum possible differences in the carrier frequencies to a value which is definitely below the audio frequency pass band of commercial radio receivers and loud speakers.

Turning now to the undesired speech, we note that it is of very little importance compared with the displaced side band interference. Thus, if this speech is to be 40 db down from the desired speech, the value of the carrier ratio is 0.044 for the case of a square law detector, while for a difference in level of 20 db, the carrier ratio is 0.14. A curve for the case of a 40 db difference is indicated by 7 of Fig. 5.

The comparison between curves 7 and 6 emphasizes the fact that we may have considerable areas of intolerable displaced side band interference in which the intelligible speech from the undesired station is not noticeable. Of course, this interference is often classed as distorted speech but the distinction is convenient in the present discussion.

[3] J. C. Steinberg, "The Relation Between the Loudness of a Sound and its Physical Stimulus," *Phys. Rev.*, Sec. Ser., Vol. 26, pp. 507–523.

Case 2

In this case the programs are identical and consequently the speech from the two stations will undergo simultaneous fluctuations of intensity. We shall here assume that the two stations have the same degree of modulation at any instant. We may then take our data from the curves for which $M = m$. However, this does not apply to the carrier beat note, since its intensity is independent of the degree of

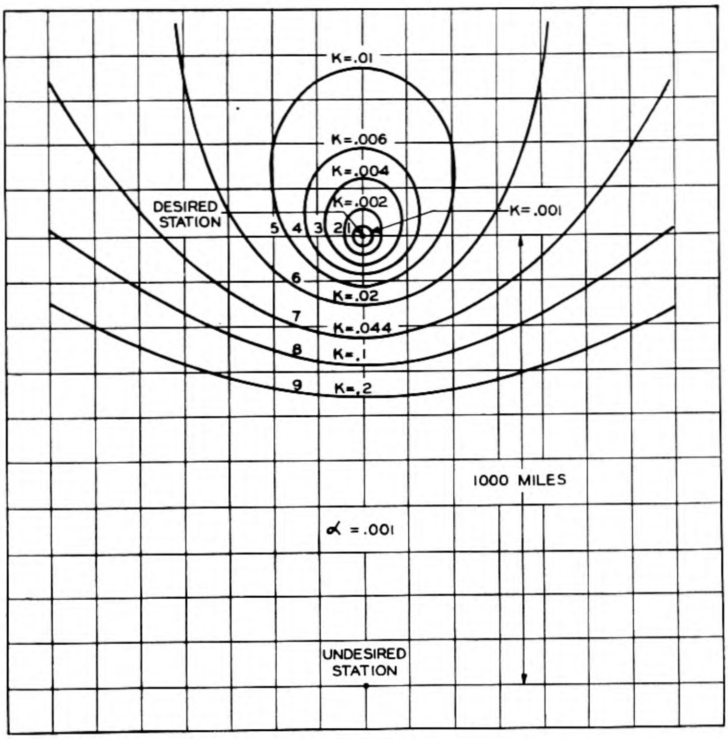


Fig. 5—Curves along which the ratio of the carrier amplitudes received from two stations has a constant value K, as indicated. Attenuation small.

modulation of either station and its interfering effect will be determined by conditions which exist when the desired station has a low degree of modulation. Hence the discussion of this component of the interference will be exactly the same as in the preceding case.

Referring to Figs. 1 and 4, it is evident that by far the greatest portion of the displaced side band interference is due to the $q \pm u$ components, in the case of the straight line detector, and the $q \pm u$ and $p \pm u$ components in the case of the square law detector. The identity of the curves for these components in the two figures shows that the degree of modulation has practically no effect on the relative importance of the interference which occurs when the same programs are transmitted.

If we again assume that the total interference may be represented by a fictitious component of twice the amplitude of the $q + u$ component,

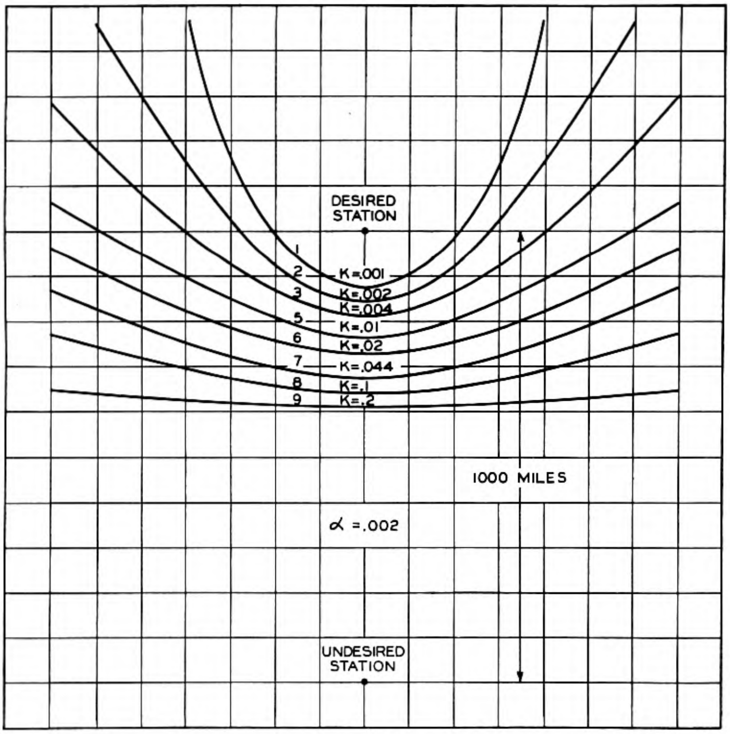


Fig. 6—Relative amplitudes of undesired frequencies as a function of the ratio of the amplitudes of the desired and the interfering carriers. Attenuation constant α twice that of Fig. 5.

we may take our data from the dash-dot line of Fig. 4. This should represent the case fairly well for the straight line detector but when a square law detector is used, greater interference should result due to the importance of the $p \pm u$ terms. However, we shall consider only the $q \pm u$ group and the phenomena associated with the square law case may be readily inferred. In order that the displaced side band interference may be 40 db down from the desired speech the carrier ratio must have a value of 0.01, while if it is to be 20 db down, this value must be 0.1. The first value corresponds to curves 5 of Figs.

5 and 6, while the second value corresponds to curves 8. We observe that there is a tremendous difference between the areas which may be considered to be free from displaced side band interference and those which will be free from carrier beat interference, in case the beat frequency is allowed to wander into the audible range. The comparison between the two areas is given by curves 1 and 5 for the 40 db interval and by curves 5 and 8 for the 20 db interval.

The speech from the interfering station will now be the same as the desired speech and can have effect only in so far as it adds to or subtracts from the desired speech. It will be noted from Figs. 1 and 5 that for carrier ratios of less than 0.1 this component is always down more than 40 db and may be safely neglected.

The foregoing discussion serves to illustrate the types of interference which may be expected when two stations are operated on approximately the same frequency. The data discussed have involved low values of attenuation. This is of particular interest when the distance between stations is large since with high values of attenuation either station will have very little effect on the service area of the other. Of course at night time we may have signal strengths which will be of the order of magnitude of that given by the simple inverse distance law involving zero attenuation. This possibility probably presents a serious limitation on night time common frequency broadcasting but should be of little consequence during the daylight hours. Conditions will be somewhat different for stations that are placed nearer together and specific results can be readily computed for any given spacing. The equations which have been discussed can be applied to any such case and the areas corresponding to those in Figs. 5 and 6 determined.

One point which is emphasized by the results which have been obtained is, that with a carrier frequency difference of several cycles satisfactory reception cannot be expected in the regions which lie midway between two transmitters. The field strength of one station must be at all times predominately higher than that of the other and consequently the use of pseudocommon frequency broadcasting should be restricted to stations of wide geographic separation. It should then be possible to furnish high grade service to relatively small densely populated areas in the immediate vicinity of either transmitter, reception at a considerable distance from both stations being admittedly unsatisfactory. However, if the carriers are strictly isochronous much larger service areas should be feasible.

Appendix

Equation (5) is

$$S = \overset{\text{I}}{A + B} - \overset{\text{II}}{\frac{AB(1 - \cos ut)}{A + B}} - \overset{\text{III}}{\frac{A^2B^2(1 - \cos ut)^2}{2(A + B)^3}} - \overset{\text{IV}}{\frac{A^3B^3(1 - \cos ut)^3}{2(A + B)^5}}. \tag{5}$$

To expand these terms we write

$$\begin{aligned}\frac{1}{(A+B)^n} &= \frac{1}{(E+e+ME\cos pt+me\cos qt)^n}\\ &= \frac{1}{(E+e)^n}\Bigg(1 - \frac{n(ME\cos pt+me\cos qt)}{E+e}\\ &\qquad + \frac{n(n+1)(ME\cos pt+me\cos qt)^2}{2(E+e)^2}\cdots\\ &\qquad +(-1)^r\frac{n(n+1)(n+2)\cdots(n+r-1)(ME\cos pt+me\cos qt)^r}{\underline{|r}(E+e)^r}\Bigg).\end{aligned} \tag{5a}$$

It is evident there are present in S an infinite number of frequencies and it is necessary to select those which are of appreciable magnitude relative to that of the desired frequency of amplitude EM. Fortunately these are not very numerous.

In deciding whether or not a given term should be retained there are two points to be considered: (1) whether all the terms of a given frequency total to a value sufficiently large to call for the presence of this term in the final result; (2) what per cent accuracy should be required in the frequencies which are retained. Thus if it is desired to retain all frequencies the relative amplitude of which is greater than 0.01 we cannot arbitrarily retain all individual terms which make a contribution of 0.01 or greater and neglect those of relative importance of less than 0.01. Thus if a term of a given frequency has a relative amplitude of 0.01 and another term of the same frequency a relative amplitude of 0.009 the second term should be retained. Otherwise we should have a large percentage error in the value of the amplitude of this frequency. On the other hand it is not desirable to maintain the same degree of accuracy for the case of retained frequencies of slight relative importance as for those of large importance. As a compromise all individual terms have been retained which, after division by EM, are of a magnitude greater than 0.005 for any values of M, m and e/E which are here dealt with. An exception is made in the case of a term in $\cos pt$ derived from term III of (5). This term is slightly larger than the above limit when $M = 0.5$ and $e/E = 0.1$ but as it decreases rapidly with a decrease in e/E it has been omitted for the sake of simplicity.

Having chosen this limit of 0.005 for the relative magnitude of individual terms it can be shown to be permissible to neglect term IV and all subsequent terms of (5). Furthermore, only a few of the large number of terms yielded by III need be retained.

After applying these rules there appear several frequencies that are never as large as 0.01 in relative magnitude and these have been omitted from consideration. As has been stated in the body of the paper, an exception is made in the case of the frequencies $(p \pm q \pm u)/2\pi$. If a given frequency exceeds 0.01 for any one of the four pairs of values of M and m, it has been shown on the figures for all of the pairs.

After the formula (5a) has been applied to S and the expressions for A and B inserted there remains the necessity of reducing products and powers of various sinusoidal terms to sums of simple first order sinusoids. This is a tedious procedure but is a matter of simple trigonometry and will not be set forth in detail.

From (5a) it can be seen that if M or m is near unity the series will converge very slowly. Furthermore, since to obtain relative magnitudes we divide by M, it is impossible to obtain satisfactory convergence due to small values of M in the denominator. Hence it is necessary to limit M and m to 0.5 or less and in addition M must be no smaller than 0.1. It would be permissible to allow m to become less than 0.1 but as little would be gained by this m has been restricted to the same range as M.

A Magnetic Curve Tracer

By F. E. HAWORTH

An apparatus for photographically recording hysteresis loops and initial magnetization curves is described. It employs a rotating drum and a fluxmeter, the restoring torque of the latter being completely counter-balanced by a photoelectric cell arrangement. With this apparatus curves may be taken so slowly that eddy currents are negligible. The accuracy of the instrument is intrinsically as great as that of a ballistic galvanometer. An analysis of sources of error is included.

FOR accurate determinations of hysteresis loops and initial magnetization curves of magnetic specimens, a laborious routine involving the use of a ballistic galvanometer is usually necessary. This article describes an apparatus by means of which these curves may be obtained photographically with quantitative accuracy. Attempts to devise such a scheme have previously been made. Ewing[1] describes one which was used with short, thick specimens in a magnetic yolk. Fleming[2] invented a device, the Campograph, which made use of a magnetometer and had the advantage of making possible the use of long, thin, specimens, thus reducing eddy current and demagnetization effects. J. B. Johnson[3] describes the most recently published design, embodying a vacuum tube amplifier and a Braun tube oscillograph. This hysteresigraph is used with frequencies of the order of five cycles per second, or higher, and consequently introduces an eddy current loss, a disadvantage in a great many measurements.

The greatest difficulty has always been to devise an instrument which would accurately record the total change in magnetic flux in the specimen. The ideal instrument would be a fluxmeter with no restoring force and no friction. Fluxmeters are on the market in which the restoring force is negligible only over short periods of time or in which there is no restoring force but where the friction is appreciable; but if it is required that the magnetic cycle have a period of more than a few seconds, such fluxmeters are out of the question. In addition they require that the search coil be of such low resistance that it must have too few turns for use with long thin specimens, in which the flux is small. These difficulties have been overcome in the apparatus described below, in which the principal feature is the use of a

[1] J. A. Ewing, "Magnetic Induction in Iron and Other Metals," 3d ed., p. 118.

[2] J. A. Fleming, *Proc. Phys. Soc. Lon.*, 27, 316–27 (1915).

[3] J. B. Johnson, *Bell System Tech. Jour.*, 8, 286–308 (1929).

fluxmeter in which the suspended coil has its restoring torque counterbalanced for all deflections within a range sufficient for accurate delineation of magnetic curves.

Description of the Apparatus

The operation of the apparatus is as follows: a long, sensitive, photoelectric cell is fitted with a V-shaped slit, as shown in Fig. 1; a beam of

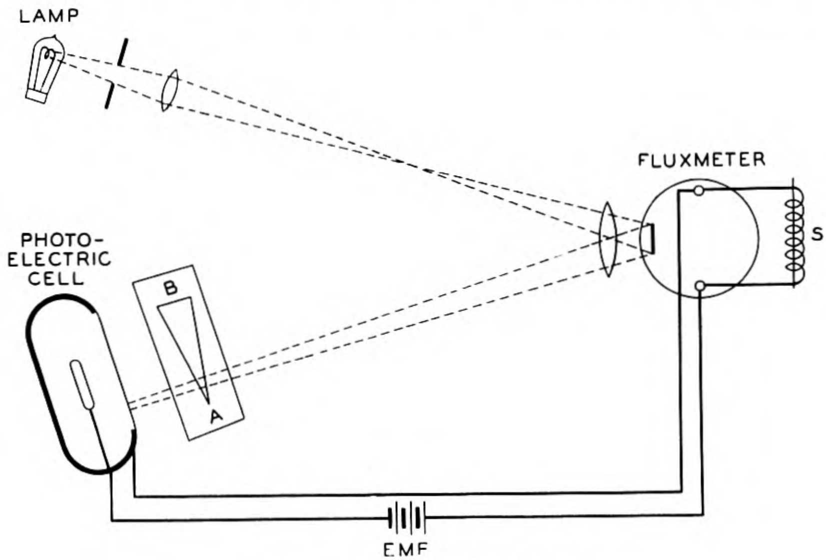


Fig. 1—The photoelectric cell circuit.

light is reflected from the mirror of the fluxmeter and focused on the slit of the photo-electric cell, which is connected, in series with a source of e.m.f., across the terminals of the fluxmeter; the e.m.f. is adjusted once for all to such a value that, if the beam is at rest when at the narrow end of the slit, at any other position the current controlled by the cell will develop a torque in the fluxmeter coil which just balances the restoring torque of the suspension. The fluxmeter deflection will then be proportional to the change of flux which has occurred within the search coil *S*. It may be found necessary to shape the slit empirically to correspond to the unequal sensitivities of the photo-electric cell at different spots. The fluxmeter used is a Leeds and Northrup type 2290 HS galvanometer. It has a critical damping resistance of about 100,000 ohms, and when used with about one hundred ohms in the external circuit it is much over damped.

The apparatus for registering the deflections photographically, and for changing the magnetic field in the specimen, is shown in Fig. 2. A drum D, carrying photographic paper, is placed in a light-tight box provided with a long, narrow slit parallel to the axis of rotation of the drum. A beam of light from a second lamp is reflected by the fluxmeter mirror and focused on the slit. This beam is reflected by the same mirror which reflects the beam onto the photo-electric cell, the two

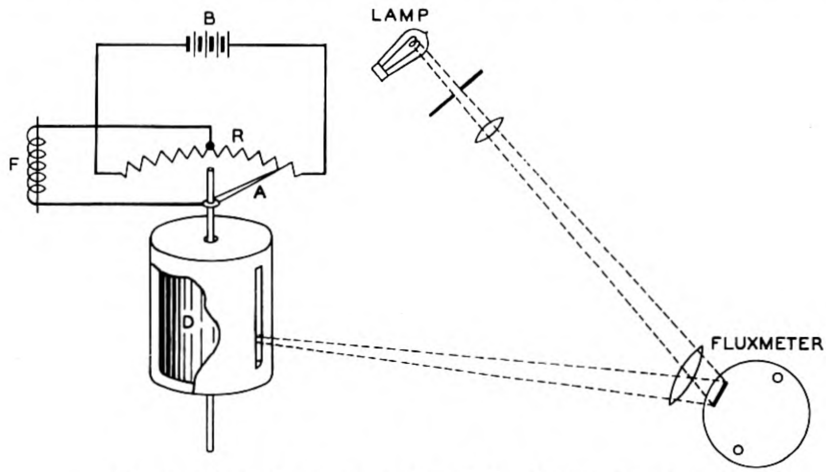


Fig. 2—The field current circuit and the photographic drum.

beams being incident at different angles. Attached to the shaft of the drum is an arm A, which slides along the rheostat R. A battery B is connected across R, and a center tap soldered to it. Between the arm A and the center tap a varying e.m.f. is produced which is applied to the field coil F. This e.m.f. reverses its sign every time the arm A slides past the center of the rheostat, and the latter is curved in a manner calculated so that the field current will be proportional to the angle of rotation of the drum from the position for zero current. The search coil S of Fig. 1 is placed within F, and consequently when D is rotated it moves the photographic paper past the slit so that the distance moved is proportional to the change in field current, while at the same time the fluxmeter deflects the beam of light along the slit so that the deflection is proportional to the time integral of the changes of flux within S. As the drum is turned from one position to another, a curve with rectangular axes is thus registered, the scales of which may be calibrated in terms of B and H. Figs. 4 to 7 are some examples of curves taken with the apparatus.

In Fig. 3 the electrical circuits are shown in detail. R_7 is the rheostat controlling the field current, and A is the arm which rotates with the drum. The battery B_2 supplies the field current, and B_3 furnishes the e.m.f. for the photo-electric cell, the value of the potential applied to the latter being regulated by R_1. The potential divider R_3, and dry cell B_1, in series with the 10 megohm resistance R_2, are used to balance out thermo-electric potentials and current from the photo-electric cell

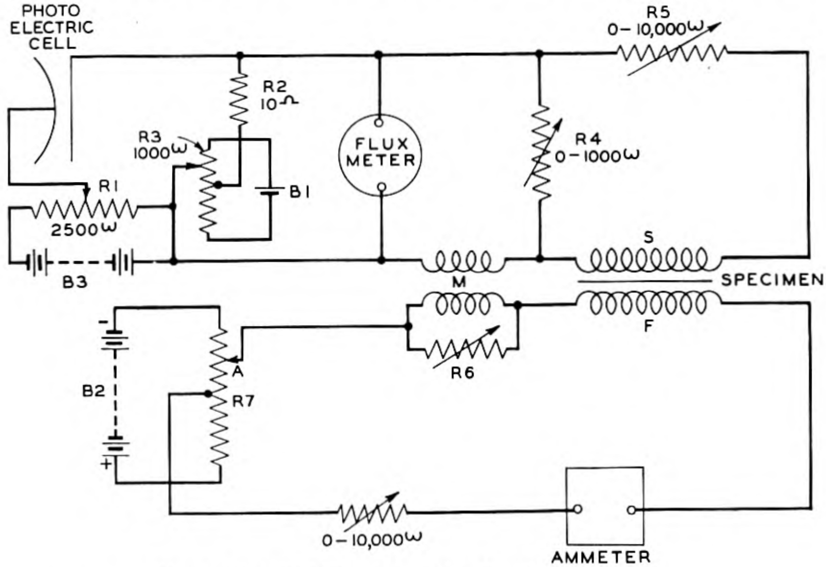


Fig. 3—Detailed diagram of the electrical connections.

due to stray light. R_4 and R_5 are adjusted according to the amount of flux in the specimen, in order to keep the maximum deflection within the desired limits. The mutual inductance M is used to balance out the potentials produced in S when no specimen is within it, so that the

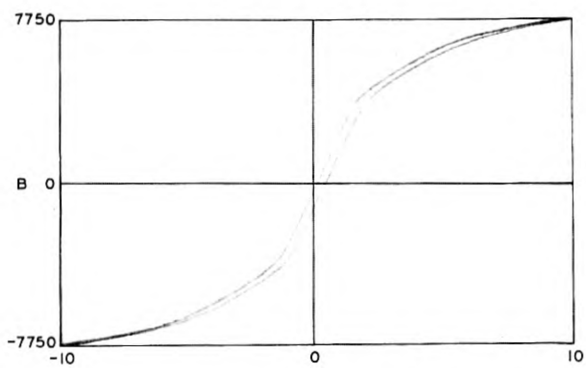


Fig. 4—Hysteresis loop of annealed iron.

fluxmeter deflection is proportional to the change in $B - H$. The drum is conveniently rotated by an electric motor, connected by gears so that the drum makes about one revolution in two minutes, and it is desirable to have this rate variable. The motor may be reversed, so that complete hysteresis loops may be recorded.

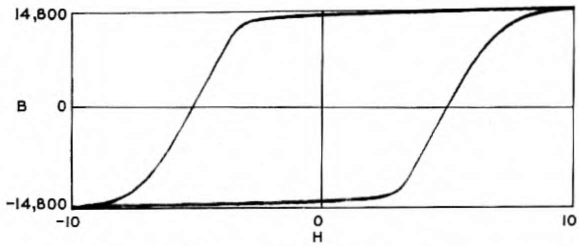


Fig. 5—Hysteresis loop of hard iron.

In setting up the apparatus the photo-electric cell may be conveniently placed above or below the drum, and one lamp above the other. The lamp used to illuminate the photo-electric cell should furnish a brilliant beam, and it was found that a 250 watt Mazda projection lamp was quite satisfactory.

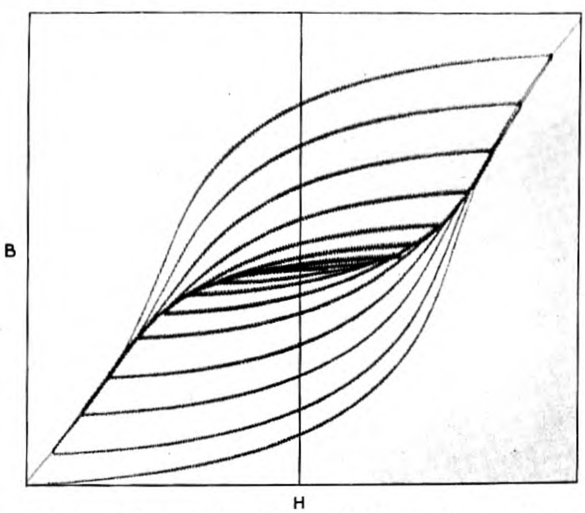


Fig. 6—Hysteresis loops of hard iron, with increasing maximum fields.

Calibration of the Circuit

The circuit is calibrated by passing a known current through the primary of a known mutual inductance, the secondary of which is connected in series with the search coil S. By measurement of the deflection produced the relation between the quantity of electricity passing through the fluxmeter and its deflection can be determined. From this relation and other known constants the change in induction of a magnetic specimen producing a given deflection may be calculated.

This calibration may be done in the following manner: Let the

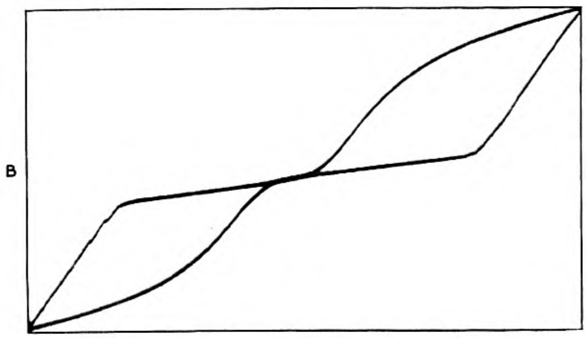


Fig. 7—Hysteresis loop of perminvar, showing the "wasp waisted" loop.

magnetic specimen be removed, R_4 and R_6 be set on infinite resistance, the magnetizing coil F be shorted, and a change in the field current made which will give a convenient deflection on the drum, as shown in Fig. 8.

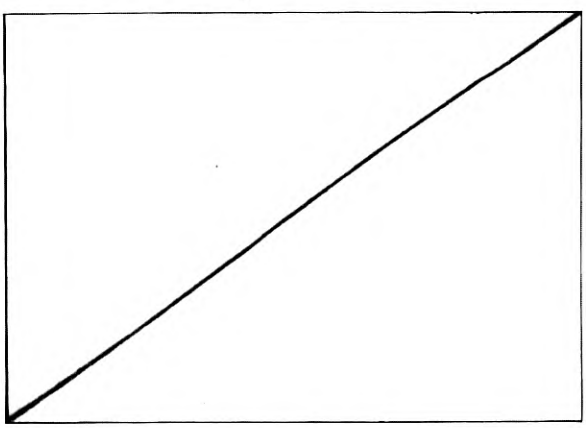


Fig. 8—Line taken for calibrating the apparatus.

Let i_M = instantaneous primary current,
i_2 = instantaneous secondary current,
r_2 = resistance of secondary circuit,
M = mutual inductance of M,
L_2 = self inductance of secondary circuit.

Then:

$$\frac{L_2}{r_2}\frac{di_2}{dt} + \frac{M}{r_2}\frac{di_M}{dt} + i_2 = 0.$$

Integrating from time $t = 0$ to $t = t_0$, the time at any later instant,

$$\frac{L_2}{r_2}\int_0^{t_0} di_2 + \frac{M}{r_2}\int_0^{t_0} di_M = -\int_0^{t_0} i_2 dt.$$

Now if i_M is changed slowly enough

$$\frac{L_2}{r_2}\int_0^{t_0} di_2$$

is negligible and we have:

$$\frac{M}{r_2}\int_0^{t_0} di_M = -\int_0^{t_0} i_2 dt,$$

or

$$\frac{M}{r_2} i_M = -Q_M,$$

where Q_M is the quantity of electricity that has passed through the fluxmeter in time t_0. Now let $Q_M = -K\delta_M$, where δ_M is the deflection produced when Q_M flows. Then:

$$\frac{M}{r_2} i_M = K\delta_M,$$

and

$$K = \frac{Mi_M}{r_2\delta_M},$$

and the quantity of electricity which has passed through the fluxmeter for any other deflection is

$$Q = -\frac{Mi_M}{r_2\delta_M}\delta. \tag{1}$$

This equation makes it possible to determine $B - H$, calculated from Q as described below, by observing the deflection δ. Relation (1) may be determined once for all as it is a constant of the fluxmeter only. The parts of Q passing through R_2 and the photo-electric cell will be negligible on account of their high resistances.

Now suppose a magnetic curve recorded with R_6 adjusted until the deflection is due solely to the magnetization of the specimen. Let the resistance of the fluxmeter plus that of the secondary of M be denoted by R_g, and that of S plus R_5 be denoted by R_s. Then if the field current i_H is varied slowly enough, the time lag in the secondary circuit will be negligible and we shall have for the instantaneous current in the fluxmeter:

$$i_g = \frac{e}{R_s + R_g + \dfrac{R_s R_g}{R_4}} \cdot$$

Now the e.m.f. in the search coil is

$$e = -AN\frac{d(B - H)}{dt},$$

where A is the cross sectional area of the specimen and N is the number of turns in the search coil. Then

$$i_g = \frac{-AN\dfrac{d(B-H)}{dt}}{R_s + R_g\left(1 + \dfrac{R_s}{R_4}\right)}$$

and therefore

$$Q = \int_0^{t_0} i_g dt = \frac{-AN}{R_s + R_g\left(1 + \dfrac{R_s}{R_4}\right)}\int_0^{t_0} d(B - H).$$

But by Eq. (1)

$$Q = -K\delta$$

therefore

$$\Delta(B - H) = \frac{K\delta\left[R_s + R_g\left(1 + \dfrac{R_s}{R_4}\right)\right]}{AN}, \qquad (2)$$

where

$$K = \frac{M\Delta i_m}{r_2 d_M},$$

r_2 being the total secondary resistance when K was determined. This equation, then, gives $B - H$ for any given deflection δ, in terms of known constants. For any fluxmeter, K is determined once for all by passing the current i_M through a mutual inductance and measuring the deflection δ_M on a photographic record. The other constants are changed in a calculable way when the number of turns in the search coil, the resistance settings, and the cross-sectional area of the sample are changed.

Sources of Error

Since it is the voltage applied to the magnetizing coil F which is proportional to the angle through which the drum has rotated, there is a lag in the field current behind the field registered on the drum, due to the self inductance of the coils. Added to this there is a lag in the secondary due to its self inductance, and another lag due to the time required for the fluxmeter to act. The effect of these is to widen the loop. In Fig. 9 is shown a curve traced with no magnetic sample in the field coil, and with dH/dt so great that the lag is appreciable.

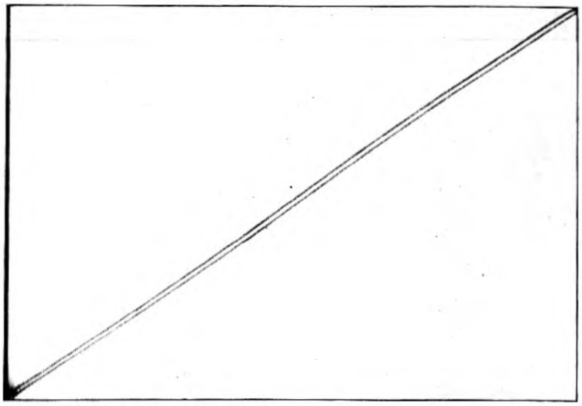


Fig. 9—Loop made with an air core mutual inductance at a very high dH/dt.

Fig. 10 shows two loops, the outer one representing a loop as taken on the apparatus, and the inner one the true loop corresponding thereto. Let B be some induction near zero, on the traced loop. B will be incorrect for the indicated value of H by an amount $B_0 - B$, such that if the field were held constant at that point while the drum continued to rotate the curve would approach B_0 as an asymptote, as indicated by the dotted curve. If dH/dt is not zero, B may be regarded as momentarily approaching B_0 as an asymptote. The equation for B at any instant is:

$$\lambda_1 \frac{dB}{dt} + B = B_0, \tag{3}$$

where λ_1 is the time constant of the circuit and B_0 is not a constant but a function of H and t. If we assume that dB/dH is constant for a small region in the neighborhood of $B = 0$, we have, putting ΔH_c equal to the error in coercive force H_c,

$$B_0 - B = \frac{dB}{dH} \Delta H_c.$$

Combining this with Eq. (3), we have

$$\Delta H_c = \lambda_1 \frac{dH}{dt} \cdot \tag{4}$$

Data taken with no magnetic specimen inserted show that this linear relation actually exists. Added to this there is an increase in H_c due to

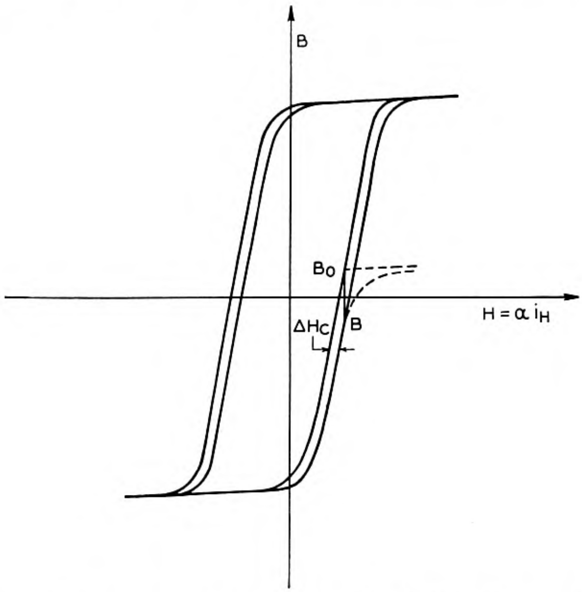


Fig. 10—A diagram to illustrate the widening of a loop due to inductance.

eddy current lag. Johnson [3] has derived an equation for this, and with a slight modification to make it applicable to cylindrical specimens, it is:

$$\Delta H_c = \frac{\Pi}{2} \frac{10^{-9}}{\rho} r^2 \frac{dB}{dH} \frac{dH}{dt}, \tag{5}$$

where ρ is the resistivity of the specimen, and r its radius. This gives us for the total error,

$$\begin{aligned} \Delta H_c &= \left(\lambda_1 + \frac{\Pi}{2} \cdot \frac{10^{-9}}{s} r^2 \frac{dB}{dH} \right) \frac{dH}{dt} \\ &= (\lambda_1 + \lambda_2) \frac{dH}{dt} \cdot \end{aligned}$$

This equation was tested experimentally by taking a series of loops with varying dH/dt. The specimen used was a cylinder of 81 per cent Ni permalloy, 60 cm. long and 0.1 cm. in diameter, and was placed in a magnetic yolk. Its hysteresis loop, as shown in Fig. 11, has an

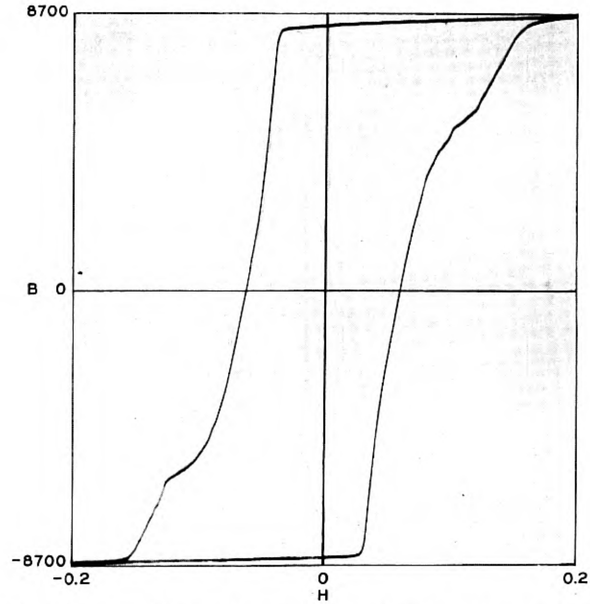


Fig. 11—A hysteresis loop of permalloy containing 81 per cent nickel.

unusual slope, 225,000 at $B = 0$. This gives $\lambda_2 = .055$ sec. From this series of curves the straight line shown in Fig. 12 was obtained, for which $\lambda_1 + \lambda_2 = 0.314$ sec. By another set of loops in which the

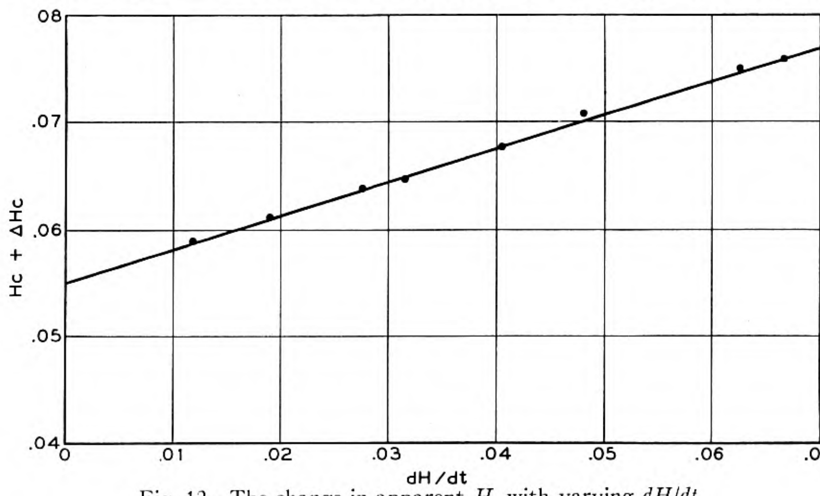


Fig. 12—The change in apparent H_c with varying dH/dt.

deflection is produced by an air core mutual inductance, λ_1 is found to be 0.134 sec. This determines λ_2 as 0.180 sec., in disagreement with the value 0.055 sec. calculated from Eq. 5. Johnson assumes in his derivation that dB/dH is constant and hence that the shape of the curve before H_c is reached has no effect on ΔH_c. It is probable that if the equation were changed to allow for dB/dH being a function H,

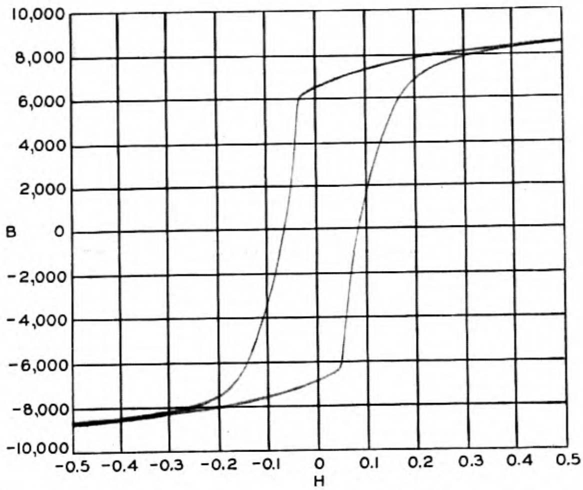


Fig. 13—A hysteresis loop of permalloy containing 78.1 per cent nickel.

that the difference could be accounted for. At any rate, this error is negligible for all but specimens with exceptionally high dB/dH or great thickness, and the true coercive force can always be found by taking two loops with different values of dH/dt and extrapolating to $dH/dt = 0$.

Another possible source of error is the passage of a large fraction of the photo-electric cell current through the search coil, the field being thereby altered. The maximum photo-electric cell current used is on the order of $5(10)^{-7}$ amperes. Since the search coil is unlikely to have more than about 400 turns per centimeter, this would make the maximum error in H about $2.5(10)^{-4}$ gauss, which is negligible for most measurements.

As a test of the accuracy of the instrument, a comparison was made with curves made by ballistic galvanometer measurements. Fig. 13 shows a loop taken of the specimen which Bozorth used in some previous measurements.[4] Both the coercive force and the maximum induction taken by the two methods agreed to within less than one

[4] R. M. Bozorth, *Phys. Rev.*, 32, 124–132 (1928).

per cent. Fig. 14 shows an initial magnetization curve which gives a value of the initial permeability agreeing accurately with the value determined ballistically.

A fluxmeter with no restoring torque is also useful in certain types of current measurements. If the average value of a current which fluctuates too much to be read on a slowly moving meter is desired, it

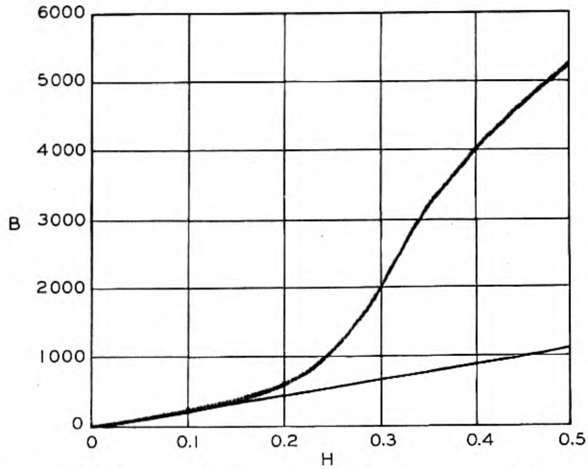


Fig. 14—An initial magnetization curve of the specimen of 78.1 per cent nickel permalloy.

can be integrated on the fluxmeter, and the average value obtained by dividing the total quantity of electricity which has passed through by the time during which the measurement was made. Also if a current is too small to be read directly on a galvanometer it may be possible to maintain it for a sufficient length of time to give a readable deflection on the fluxmeter, and again the current will be obtained by dividing by the time.

In conclusion I wish to thank Dr. R. M. Bozorth for suggestions given during the development of the apparatus, and Mr. A. W. Metz for his assistance in taking the curves.

A Multi-Channel Television Apparatus *

By HERBERT E. IVES

A bar to the attainment of television images having a large amount of detail is set by the practical difficulty of generating and transmitting wide frequency bands. An alternative to a single wide frequency band is to divide it among several narrow bands, separately transmitted. A three-channel apparatus has been constructed in which prisms placed over the holes in a scanning disc direct the incident light into three photoelectric cells. The three sets of signals are transmitted over three channels to a triple electrode neon lamp placed behind a viewing disc also provided with prisms over its apertures so that each electrode is visible only through every third aperture. An image of 13,000 elements is thus produced. For the successful operation of the multi-channel system, it is imperative to have very accurate matching of the characteristics in the several channels.

IF, in a received television image, the individual image elements are, as they should be, of such a size as to be just indistinguishable, or unresolved, at a given observing distance, the number of image elements increases directly with the area of the image. The number of such indistinguishable elements in everyday scenes, in the news photograph, or in the frame of an ordinary motion picture is astonishingly large. An electrically transmitted photograph 5 inches by 7 inches in size, having 100 scanning strips per inch, has a field of view and a degree of definition of detail, which, experience shows, are adequate (although with little margin) for the majority of news event pictures. It is undoubtedly a picture of this sort that the television enthusiast has in the back of his mind when he predicts carrying the stage and the motion picture screen into the home over electrical communication channels. In this picture, the number of image elements is 350,000. At a repetition speed of 20 per second (24 per second has now become standard with sound films) this means the transmission of television signals at the rate of 7,000,000 per second,—a frequency band of 3½ million cycles on a single sideband basis. This may be compared to the 5,000 cycles in each sideband of the sound radio program, or it may be evaluated economically as the equivalent of a thousand telephone channels.

When we examine what has been achieved thus far in television, we find that the type of image successfully transmitted falls very far short of the finely detailed picture just considered. Probably the most satisfactory example of television thus far demonstrated is the

* *Jour. Optical Soc.*, Jan., 1931.

72-line picture used in the two-way television-telephone installation of the American Telephone and Telegraph Company in New York.[1] Here the object to be transmitted is definitely restricted to the human face, which fills the whole field of view, and is adequately rendered by the 4,500 image elements used.

The gap between the 4,000 elements of this image and the 350,000 considered above is enormous, not only in figures, but in terms of technical possibility of bridging. Even if we are forced to content ourselves with relatively simple types of scenes for television transmission, still the fact must be squarely faced that a very much larger number of image elements must be transmitted than has thus far been found possible; and a far wider frequency band utilized than has ever been used in any communication problem. Now the situation is, simply stated, that *all parts of the television system are already having serious difficulty in handling the 4,500-element image.* Consequently, a *major problem in television progress is to develop means to extend the practical frequency range.*

It will be worth while to survey briefly the points in a television system where difficulty is now encountered when the attempt is made to increase the amount of image detail and the accompanying band of transmitted frequencies. Consider in turn the scanning discs at sending and receiving ends, the photoelectric cells, the amplifying systems, the transmission channels, the receiving lamps.

In the scanning disc at the sending end, which we shall assume arranged for direct scanning, increased detail means either loss of light or increase in the size of the disc. In either case, the factor of change involved is large. For instance, if the number of scanning holes is doubled in a disc of given size, providing four times the number of image elements, the holes must be spaced at half the angular distance apart, and twice the number of holes, imagined placed end to end, must be included in this half diameter scanning field. The holes will therefore be of one-quarter the diameter or 1/16 the area. The light falling on the photoelectric cell at any instant is the light transmitted by one hole; in this case, 1/16 the amount with the disc of half the number of holes. In general, the light transmitted by the disc to the cell decreases as the square of the number of image elements. If the disc is enlarged so as to hold the transmitted light unchanged, its diameter increases directly as the number of image elements. It is obvious that any considerable increase in the number of image elements—such as ten times—demands either enormously increased sensitiveness in our photo-responsive devices, or quite fabulous sizes of

[1] *Bell System Technical Journal*, July 1930, p. 448.

discs. Perhaps the most pertinent conclusion from this survey is that the disc, while quite the simplest means for scanning images of few elements, is entirely impractical when really large numbers of image elements are in question. As yet, however, no practical substitute for the disc of essentially different character has appeared.

Turning now to the photoelectric cell. The question of adequate sensitiveness to handle a large number of image elements is intimately connected with the method of scanning, as has just been brought out, so that no simple answer is possible. It is, however, probable that a very considerable increase in sensitiveness over anything now available must be anticipated, whatever scanning device is adopted. In the matter of frequency range there is definite information.[2] In cells depending on gas amplification (such as argon or neon) a characteristic behavior is a falling off of output with frequency, greater the higher the voltage used, which, becoming noticeable at about 20,000 cycles, may at 100,000 cycles be so considerable as to constitute a practical block to transmission. Vacuum cells are free from this failing, but are much less sensitive. Systematic experiment and development of photoelectric cells with particular reference to extending their range of frequency response is indicated as a necessary step in the attainment of a many-element image.

Taking up next the circuits associated with the photoelectric cell, we find, in general, that the higher frequencies progressively suffer from the electrical capacity of cells and associated wiring and amplifier tubes. This in turn calls for auxiliary equalizing circuits, with attendant problems of phase adjustment, and for increased amplification. Amplifiers capable of handling frequency bands extending from low frequencies up to 100,000 cycles or over offer serious problems.

Communication channels, either wire or radio, are characterized by increasing difficulty of transmission as the frequency band is widened. In radio, fading, different at different frequencies, and various forms of interference stand in the way of securing a wide frequency channel of uniform efficiency. In wire, progressive attenuation at higher frequencies, shift of phase, and cross-induction between circuits offer serious obstacles. Transformers and intermediate amplifiers or repeaters capable of handling the wide frequency bands here in question also present serious problems.

At the receiving end of the television system, conditions are similar to the sending end. The neon glow lamp, commonly used for reception, is already failing to follow the television signals well below 40,000 cycles, and, in the case of the 4,500-element image above

[2] Loc. cit., p. 456.

referred to, the neon must be assisted by a frequently renewed admixture of hydrogen, which again cannot be expected to increase the frequency range indefinitely. In the scanning disc, as at the sending end, increasing the number of image elements rapidly reduces the amount of light in the image. With a plate glow lamp of given brightness, the apparent brightness of the image is inversely as the number of image elements.

From this rapid survey, it is clear that at practically every stage in the television system, we encounter serious difficulties when a large increase in image elements is contemplated. It is not claimed that these difficulties are insuperable. One of the chief uses of a tabulation of difficulties is to aid in marshalling the attack upon them. But the existing situation is that if a many-element television image is called for today, it is not available, and *one of the chief obstacles is the difficulty of generating, transmitting, and recovering signals extending over wide frequency bands.*

One alternative, which prompted the experimental work to be described below, is the *use of multiple scanning, and multiple-channel transmission.* The general idea, which is obvious from the name given to the method, is to divide the image into groups of elements, the various groups to be simultaneously scanned, and to transmit the signals from the several groups through separate transmission channels. In place of apparatus to generate and transmit a frequency band of n cycles, we arrange m scanning processes each to provide frequency bands of n/m cycles width; n/m being chosen as within the limits set by the available practical elements of a television system. It will appear that the method which has been developed does provide an image of manyfold more image elements than heretofore, and may make easier the problem of transmission over practical transmission lines.

Description of a Three-Channel Apparatus

The multi-scanning apparatus which has been constructed and given experimental test uses scanning discs over whose holes are placed prisms of several different angles. At the sending end, the beams of light from successive holes are thereby diverted to different photoelectric cells. At the receiving end, the prisms similarly take beams of light from several lamps and divert them to a common direction. The mode of action of the prisms is illustrated in Fig. 1a, where a three-channel arrangement is shown, which is that actually used in the experimental apparatus. In the figure, the disc holes are shown disposed in a spiral, at such angular distances apart that three holes are always included in the frame f. Over the first hole of a set of three is placed a prism P_1 which diverts the normally incident light upward; the second hole is left clear; the third is covered by a prism P_2 turned to divert the light downward. If we imagine the prisms removed and a single channel used instead of the three that are proposed, it is clear that the holes would have to be spaced three times as far apart so that no more than one would be included in the frame f at one time. The diameters of the holes, and the radial separation of the first and last in the spiral would be unchanged. Quite apart, therefore, from the smaller frequency bands which are sufficient to carry each of the three sets of signals, which is the principal objective sought, there is realized in this arrangement a reduced size of apparatus for the same size of disc holes.

Studying more closely the division of the light into three sets of beams, it is important to note that the signals transmitted by any one of the three sets of holes are continuous—as one hole of a given prism series passes out of the frame the next of the same series comes in. The signals generated in each photoelectric cell are accordingly exactly like those of a single-channel system.

Before describing the details of the apparatus, the general relationship between the number of image elements, band width, number of channels, and shape of picture may be developed. For this purpose, let the following symbols be used.

B = frequency band available in one channel, in cycles per second.
F = repetition frequency of images, per second.
C = number of communication channels.
n = total number of scanning holes.
a/b = ratio of tangential to radial dimensions of frame.
α = angular opening of frame.

We shall assume that the picture elements into which the frame is imagined divided are symmetrical in shape, i.e. either circles or squares. We then have that

the number of picture elements in the radial direction = number of holes = n;

the number of picture elements in the tangential direction = $(a/b)\cdot n$.

Now the number of signal cycles corresponding to this number of elements is $(1/2)\cdot(a/b)n$.

The number of cycles per second in one transit along the frame $= (1/2)\cdot(a/b)\cdot n\cdot F$;

over the whole picture it is $(1/2)\cdot(a/b)\cdot n\cdot F\cdot n = (1/2)(a/b)Fn^2$;

and the number of cycles per second for each channel $= (1/c) \cdot (1/2)(a/b)Fn^2 = B$.

The angular opening of the frame $\alpha = 360/n \times C$.

The number of picture elements $= n^2 \cdot (a/b)$.

These formulæ may be utilized upon assuming values for any of the variables, to fix the values of the other. In the present case, it was decided for reasons of simplicity to restrict the number of channels to 3.

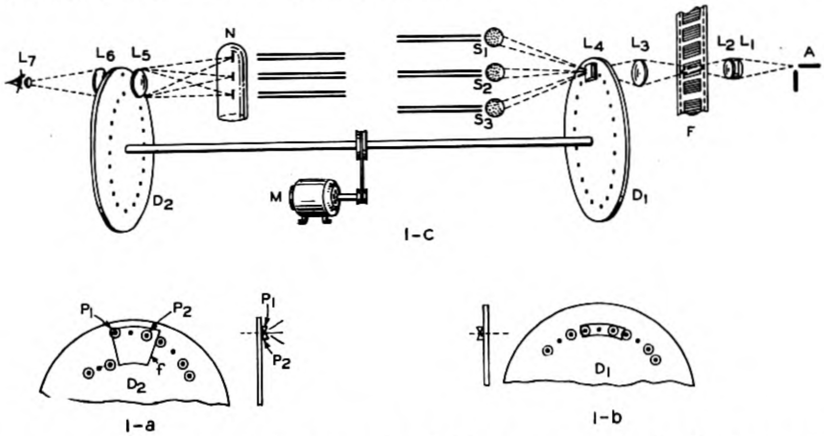


Fig. 1—Schematic of three-channel television apparatus. (*a*) Receiving end disc with spiral of holes provided with prisms. (*b*) Sending end disc with circle of holes provided with prisms. (*c*) General arrangement of apparatus.

The band width was chosen as that found feasible in the two-way television system, namely 40,000 cycles. The picture shape chosen was that of the sound motion picture, for which $a/b = 7/6$. The repetition frequency assumed was 18 per second, again following closely that of existing experimental synchronizing apparatus. Substituting these values in the formula rearranged to give n, we get for the number of holes,

$$n = \sqrt{\frac{2Bbc}{aF}} = 108$$

and for α,

$$\frac{360}{108} \times 3 = 10 \text{ degrees,}$$

for the number of picture elements,

$$n = (108)^2 \times \frac{7}{6} = 13{,}608.$$

In utilizing the prism disc principle at the sending end, direct scanning, in which the object is imaged on the disc, was chosen, since beam scanning would introduce the problem of separating the light reflected from the object from the several spots simultaneously projected from the disc. Since the light going through the disc must be separated into several beams to be directed into separate photoelectric cells, the full aperture of the image forming lens must be divided by C, the number of channels, with a consequent proportional loss of light to each cell. (This loss counterbalances the decreased size of disc above noted.) It therefore becomes necessary to insure a very high illumination of the object. In the present case, it was decided to use motion picture film to provide the sending end image, since this can have a large amount of light concentrated through it by an appropriate lens system.

The use of motion picture film permitted a simplification of the transmitting disc, which is illustrated in Fig. 1*b*. This consists in arranging the scanning holes in a circle instead of a spiral, and producing the longitudinal scanning of the film by giving it a continuous uniform motion at right angles to the motion of the scanning holes. The continuous motion of the film also avoids the loss of transmission time that an intermittent motion demands for the shutter interval.

At the receiving end, a spiral of holes is used as shown in Fig. 1*a*. There again, because of the division of the light into three beams, the angle which can be subtended by the light source (neon lamp) is much restricted. In consequence, the neon lamp cathodes are of small area, and a lens system has been used to focus their images on the pupil of the observer's eye. Other methods of receiving, which promise to be less restricted as to position of observation, are possible, however, as discussed below.

With this survey of certain of the more important features of the system, we may proceed to a more detailed account of the apparatus as constructed. The general arrangement of parts is shown in Fig. 1*c* and in the photographs, Figs. 2, 3, 4 and 5 in all of which the symbols are uniform. Both sending and receiving discs were, for simplicity of operation, mounted on the same axis, driven by the motor M. This means that no question of synchronization entered. Synchronization is in fact a separate problem, having nothing to do with multi-channel operation and has been very completely solved in connection with other television projects.[1] If it should be decided to transmit the multi-channel image to a distant point, the apparatus could be cut in two and each end, after separation to the desired distance, operated by synchronous motors controlled in approved fashion. Similarly, no long transmission lines were included.

Starting at the extreme right end of the schematic drawing Fig. 1*c*, we have an arc lamp A, a cylindrical lens L_1, a condensing lens L_2, the two lenses together concentrating a line of light on the film F. Between the film and the disc is a lens L_3 which images the film on the disc. Directly behind the disc D_1, with its circle of prism covered holes, is a second cylindrical lens L_4 which concentrates the transmitted

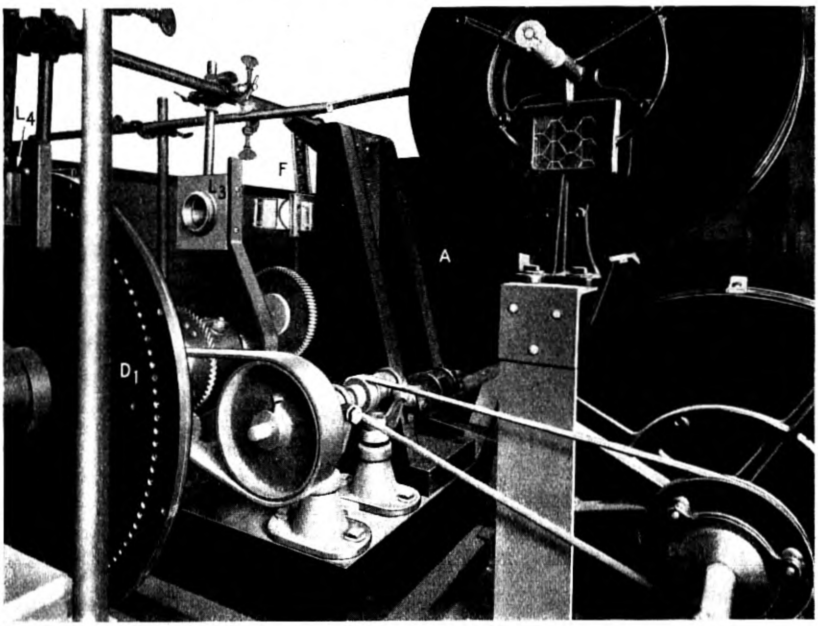


Fig. 2—Sending end of three-channel television apparatus, showing film driving arrangements.

light laterally, upon the three photoelectric cells S_1, S_2, S_3. By virtue of this lens arrangement, the light falls upon the cells in three small practically stationary spots. Additional apparatus not shown in the diagram but visible in the photographs are gears by means of which the film is driven from the disc axle through a differential, which permits the film to be framed up and down. The light beam is directed through the film at right angles to the axis of the discs by means of two prisms, whereby certain conveniences in driving and handling the film are attained.

The photoelectric cells are similar to ones previously described. The amplifier system was substantially identical with that used in the two-way television system, and need not be described again. Similarly, the amplifiers at the receiving end were the actual set used in the three-color television apparatus previously described.[3]

At the receiving end, the three sets of signals were supplied to the three electrodes of a special neon lamp N, shown in Fig. 5, which is provided with a hydrogen valve to enable it to respond to the higher frequencies. Condensing lenses L_5 and L_6 image the three electrodes

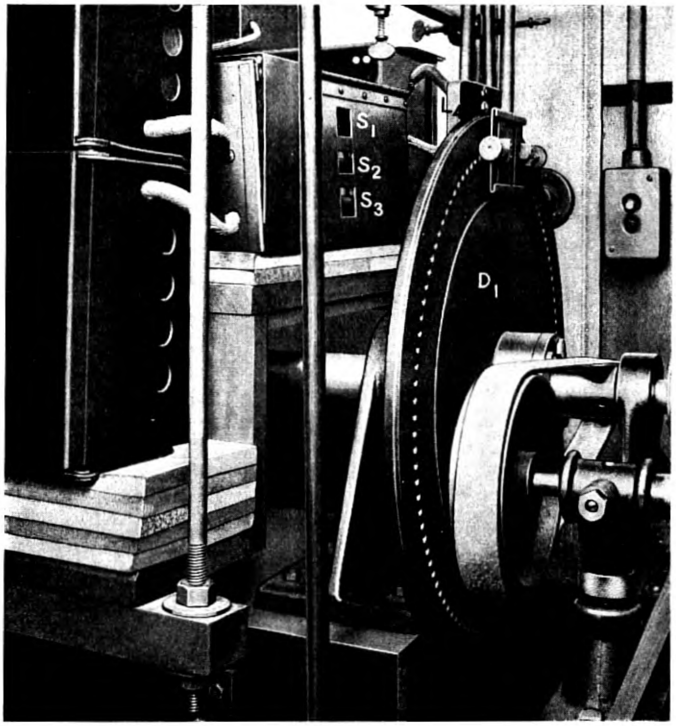


Fig. 3—Sending end of three-channel television apparatus, showing sending prism disc and photoelectric cells.

at the eye, where another lens L_7 is placed at the eye to focus the face of the disc D_2. By this system, nine electrode images are formed, of which three are superposed at the eye, and successive scanning holes are seen illuminated by each electrode in turn. This viewing arrangement, by which the image is visible to only a single eye, is adequate for an experimental investigation of the multi-channel method, but some other scheme would of course be needed if the method were developed into a practical form. Of several schemes, mention will be made here only

[3] *Journal of the Optical Society*, February, 1930, p. 11.

of the possible use of a triple grid of neon tubes, using a triple distributor of the type used in displaying images to a large audience in our initial work in 1927.[4]

Discussion of Results

The three-channel apparatus, when all parts are properly functioning, yields results strictly in agreement with the theory underlying its construction. The 13,500-element image, in resolving power and

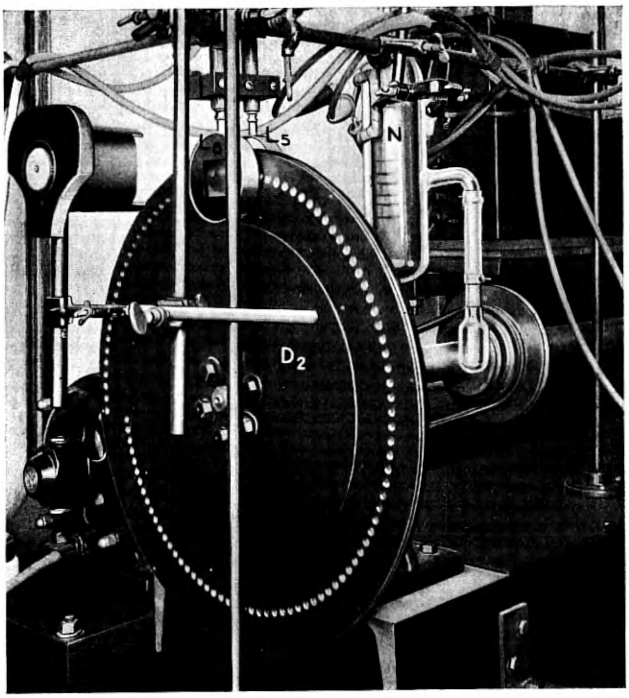


Fig. 4—Receiving end of three-channel television apparatus.

amount of detail handled, is a marked advance over the single-channel 4,500-element image. Even so, the experience of running through a collection of motion picture films of all types is disappointing, in that the number of subjects rendered adequately by even this number of image elements is small. "Close-ups" and scenes showing a great deal of action, are reproduced with considerable satisfaction, but scenes containing a number of full length figures, where the nature of the story is such that facial expressions should be watched, are very

[4] *Bell System Technical Journal*, October, 1927, pp. 551–652.

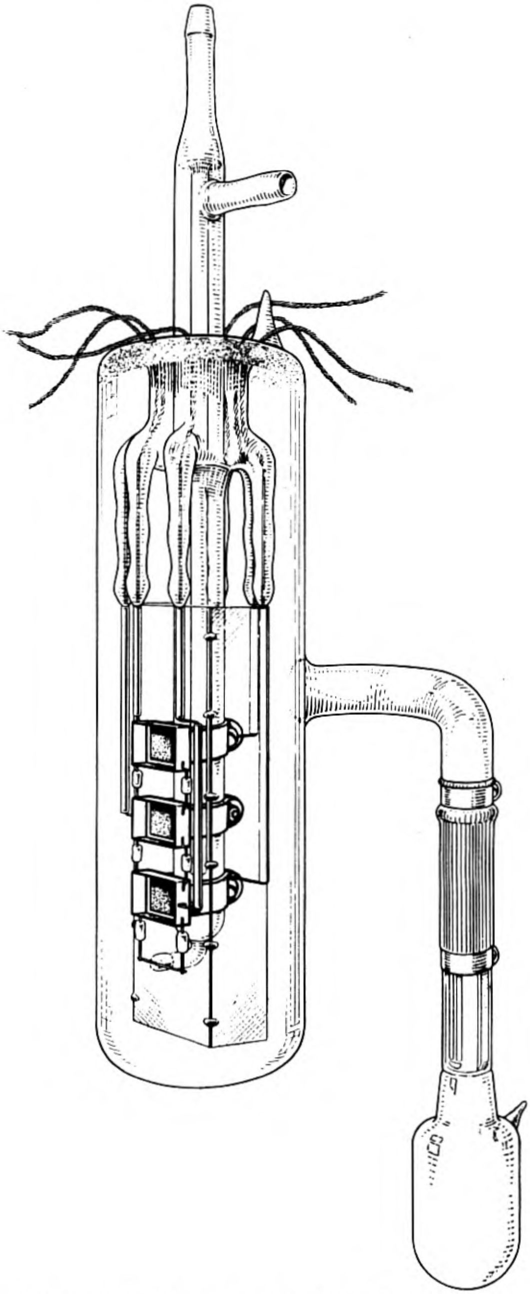


Fig. 5—Three-electrode neon lamp used for three-channel television reception.

far from satisfactory. On the whole, the general opinion expressed in an earlier paragraph is borne out, that an enormously greater number of elements is required for a television image for general news or entertainment purposes. This, however, was anticipated, and the real question is whether the results of this experiment indicate that the finer grain image is best attained by resort to multi-channel means.

This leads to a discussion of what has proved to be a serious practical difficulty with the multi-channel apparatus. This is *the problem of keeping the several channels properly related to each other in signal strength.* In the experimental apparatus, the direct current components (introduced at the receiving end) and the alternating current signals, are separately controlled, manually, by potentiometers. These have fine enough steps so that with care, with a non-changing image, a uniform picture may be obtained. If, however, for any reason the signals on one of the channels becomes too strong or too weak, the picture exhibits at once a strongly lined appearance. The eye is quite sensitive to irregularity of this sort, and the transition from a smooth grainless image to one showing a periodicity of 1/3 the number of constituent lines largely offsets the higher resolving power afforded by the actual number of scanning lines used. A characteristic practical defect of the system as set up is that any marked change in the general character of the signal, such as is produced by a shift from close-up to a wide angle view may throw out the existing signal balance sufficiently to show objectionable grain in the picture.

Differences of this sort in the three signals are of course caused in general by differences in the characteristics of the three circuits. Such differences can arise from overloading of amplifier tubes, whereby one or more may be working on a non-linear portion; by rectifying action of different amounts in the tubes immediately associated with the neon lamps, or in the neon lamp electrodes themselves. A remedy is the careful design and test of all parts of the system to insure the greatest possible uniformity of performance. When this is carefully done, the behavior of the three signals is reasonably satisfactory.

Conclusion

We are, as a consequence of this work, in a position to make a general comparison of the two chief theoretical means for achieving a television image of extreme fineness of grain, which are (1) extension of the frequency band, and (2) the use of several relatively narrow frequency bands. Both, because of the diminished amount of light which finer image structure entails, demand enhanced sensitiveness of the photo-sensitive elements at the sending end, and increased efficiency fo the light sources at the receiving end. The multi-channel scheme described has some advantage in compactness over the equivalent single-channel apparatus, but since it is restricted to narrow angles of illumination of the discs the overall efficiency of light utilization is not essentially different. Comparing now the demands made upon the electrical systems the differences between the two methods are clear cut. Method (1) demands an extension of the frequency range of all parts of the apparatus, the attainment of which depends upon physical properties and technical devices whose mastery lies in the indefinite future. Method (2) demands a multiplication of apparatus parts, and careful design and construction of these parts so as to insure accurately similar operation of a considerable number of electrical circuits and terminal elements. The attainment of the necessary uniformity of performance of the several electrical circuits and terminal elements, while involving no fundamental problems, must present increasing difficulty with the number of channels used.

Condenser and Carbon Microphones—Their Construction and Use *

By W. C. JONES

Of the numerous microphones which have been developed since Bell's original work on the telephone, only two are used extensively in sound recording for motion pictures, namely, the condenser microphone and the carbon microphone.

The condenser microphone was first proposed in 1881 but owing to its low sensitivity was limited in its field of usefulness until the development of suitable amplifiers. In 1917, E. C. Wente published an account of the work which he had done on a condenser microphone having a stretched diaphragm and a back plate so designed as to introduce an appreciable amount of air damping. The major portion of the condenser microphones used today in sound recording embody the essential features of the Wente microphone. Marked progress has, however, been made in the design and construction of these instruments with the result that they are not only more sensitive but also more stable. The factors which contribute to this improvement are described in detail in this paper. Recently a number of articles have appeared in the technical press calling attention to certain discrepancies between the conditions under which the thermophone calibration of the condenser microphone is made and those which exist in the studio. The nature of these discrepancies and their bearing on the use of the microphone are discussed.

Microphones in which the sound pressure on the diaphragm produces changes in the electrical resistance of a mass of carbon granules interposed between two electrode surfaces have been used commercially since the early days of the telephone. In recent years the faithfulness of the reproduction obtained with the carbon microphone has been materially improved by the introduction of an air damped, stretched diaphragm and a push-pull arrangement of two carbon elements. This instrument is finding extensive use in sound recording and reproduction fields where carbon noise is not an important factor. The outstanding design features of the push-pull carbon microphone are described in this paper and suggestions made as to the precautions to be taken in its use if the best quality, maximum life, etc. are to be obtained.

OF the numerous microphones which have been developed since Bell's original work on the telephone, only two are used extensively in sound recording for motion pictures, namely, the condenser microphone and the carbon microphone. It has therefore been suggested that it would be fitting to review at this time the construction of these instruments and consider some of their transmission characteristics and the precautions which should be exercised in their use.

Condenser Microphone

In 1881, A. E. Dolbear [1] proposed a telephone instrument which could be used either as an electrostatic microphone or receiver. This

* Presented at Soc. of Motion Picture Engineers' Convention, Oct. 20, 1930; Journal, Soc. of Motion Picture Engineers, Jan., 1931.

[1] "A New System of Telephony," A. E. Dolbear, *Scientific American*, June 18, 1881, p. 388.

instrument consisted of two plates insulated from one another and clamped together at the periphery. The back plate was held in a fixed position whereas the front was free to vibrate and served as a diaphragm. It is obvious that, if the diaphragm were set in vibration by sound pressure, the electrical capacitance between the two plates would be changed in response to the sound waves, and if a source of electrical potential were connected in series with the instrument a charging current would flow which would be a fairly faithful copy of the pressure due to the sound wave. Apparently Dolbear realized that the current developed in this way would be minute, for in the telephone system which he proposed as a substitute for the one using Bell's magnetic instruments he employed the electrostatic instrument only as a receiver and adopted the loose contact type of microphone. At approximately the same time an article appeared in the French press [2] calling attention to the use of a condenser as a microphone and commenting on the fact that this type of microphone had been found to be less sensitive than the loose contact type.

Owing to the low sensitivity of the condenser microphone, the field of usefulness of this instrument was extremely limited for a number of years and it did not assume a position of importance among the instruments used in acoustic measurements and sound reproduction until suitable amplifiers had been developed. The development of the vacuum tube amplifier, however, filled this need. In 1917 E. C. Wente [3] published an account of the work which he had done on an improved condenser microphone having a stretched diaphragm and a back plate so located relative to the diaphragm that in addition to serving as one plate of the condenser it added sufficient air damping to reduce the effect of diaphragm resonance to a minimum.[4] The response of this instrument was sufficiently uniform over a wide range of frequencies to make it not only useful in high quality sound reproduction but a valuable tool in acoustic measurements in general.

The major portion of the condenser microphones used today in sound recording embody the essential features of the Wente microphone. Marked progress has, however, been made in the design and construction of these instruments since the initial disclosure and it will no doubt be of interest to many to consider briefly the nature of this advance.

[2] "La Lumiere Electrique," 1881, p. 286.

[3] "A Condenser Transmitter as a Uniformly Sensitive Instrument for the Absolute Measurement of Sound Intensity," E. C. Wente, *Physical Review*, July 1917, pp. 39–63. "Electrostatic Transmitter," E. C. Wente, *Physical Review*, May 1922, pp. 498–503.

[4] A discussion of the theory of air damping is given in "Theory of Vibrating Systems and Sound," I. B. Crandall, pp. 28–39.

In the early microphones employing air damping the diaphragm was composed of a thin sheet of steel which was stretched to give it a relatively high stiffness. When assembled in the microphone the stiffness was further increased by that of the air film between diaphragm and the damping plate with the result that the resonant frequency was well above the frequencies which it was desired to transmit and the diaphragm vibrated in its normal mode over a wide frequency range. In such a structure the mechanical impedance for frequencies below resonance is due almost entirely to stiffness reactance. Hence a constant sound pressure produces substantially the same displacement of the diaphragm at all frequencies within this range and uniform response results except at the very low frequencies where an appreciable reduction in the stiffness of the air film occurs. The effective mass of a steel diaphragm is, however, relatively large and necessitates a comparatively high stiffness to secure the desired resonant frequency. From the standpoint of securing maximum sensitivity of the microphone, i.e. displacement of the diaphragm per unit force, it is of course important to make the stiffness as low as possible and employ as small a value of mechanical resistance as is consistent with the degree of damping required. An improvement in both respects can be effected by decreasing the mass of the diaphragm for with a reduced mass a given resonant frequency can be obtained with lower values of stiffness and the desired damping constant secured with less mechanical resistance.

The aluminum alloys have therefore replaced steel in the diaphragms of most of the condenser microphones in use today. A typical example of such a microphone is the Western Electric Company's instrument (394-type) shown in the photograph, Fig. 1, and the cross-sectional view, Fig. 2. The diaphragm of this instrument is made from aluminum alloy sheet .0011 inch in thickness. The edges are clamped securely between threaded rings, gaskets of softer aluminum being provided to prevent damage at the clamping surfaces. The requisite stiffness is obtained by advancing the stretching ring until a resonant frequency of 5,000 cycles is obtained. The method of determining the resonant frequency of the diaphragm is as follows. The diaphragm assembly to be tested is coupled to a condenser microphone which is provided with a suitable circuit for measuring its output. A special telephone receiver is placed in contact with the diaphragm on the side opposite to the coupler. Current from a vacuum tube oscillator is then passed through the winding of the receiver, setting up eddy currents in the diaphragm under test. The forces which are developed as a result of the reaction of the magnetic field produced by the eddy currents and that of the permanent magnet of the receiver set the test diaphragm in motion. The resonant frequency is determined by noting the frequency at which the output from the condenser microphone is a maximum.

In the early Wente microphone the damping plate was a continuous surface. Subsequent work by I. B. Crandall[5] showed that the required amount of damping at the resonant frequency could be obtained without adding unduly to the impedance at other frequencies by cutting grooves in the plate. This reduced the stiffness introduced by the air film and decreased the irregularity in response at low frequencies previously mentioned. The grooves in the damping plate of the

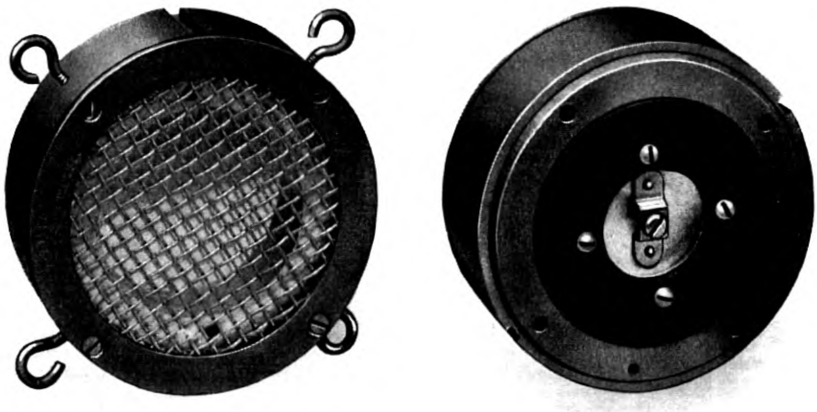


Fig. 1—Western Electric Company's 394-type condenser microphone.

Western Electric Company's 394-type microphone are cut at right angles. Holes, tapered at the outer end to reduce resonant effects, are bored through the plate at the intersection of the grooves to form connecting passages between the air film at the front and the cavity at the back. In order to prevent the resonance which would result if the grooves extended into the portion of the chamber surrounding the damping plate, the outer ends are closed by an annular ring which is pressed over a shoulder on the plate. The surface of the damping plate is plane within 8×10^{-5} inch. The departure from a plane in any individual case is determined commercially by the interference pattern developed when an optically flat plate is placed over the damping plate under test.

[5] "The Air Damped Vibratory System: Theoretical Calibration of the Condenser Transmitter," I. B. Crandall, *Physical Review*, June 1918, pp. 449–460.

A duralumin spacing ring .001 inch in thickness separates the damping plate from the diaphragm. It is essential that all dust and dirt be excluded from this space. To prevent foreign material from entering through the holes in the plate a piece of silk is fastened over the outer surface. The assembly of the diaphragm and the damping plate is made in a dust-proof glass cabinet.

If the back wall of the condenser microphone were rigid, changes in the separation between the damping plate and the diaphragm of sufficient magnitude to affect not only the sensitivity of the instrument but also its frequency response characteristic would result from variations in barometric pressure. Complete compensation for these

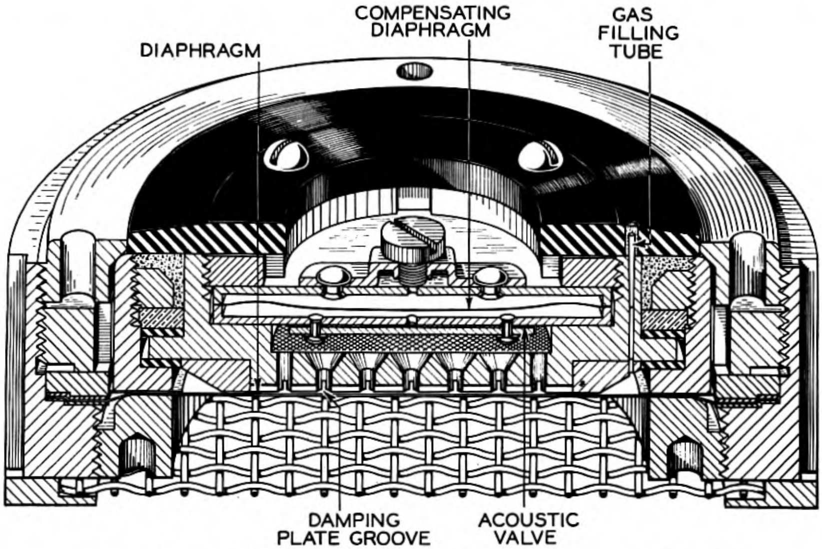


Fig. 2—Cross-sectional view of the 394-type condenser microphone.

changes in pressure can only be obtained by permitting free interchange of air between both sides of the microphone diaphragm. This is, however, objectionable owing to the fact that sufficient moisture is likely to be introduced to start corrosion and affect the insulation between the damping plate and the diaphragm. A compensating diaphragm of organic material has therefore been introduced which prevents this undesirable effect of humidity but is sufficiently low in stiffness to equalize the changes in pressure encountered in the normal use of the microphone.

In order to prevent transmission losses at voice frequencies due to the presence of the compensating diaphragm, an acoustic valve is inserted between the damping plate and this diaphragm. This valve consists of a disc of silk clamped between two aluminum plates of unequal diameters. Gas in passing from the damping plate to the compensating diaphragm moves laterally from the edge of the smaller plate through the silk to a hole in the center of the larger plate. The impedance of this path is high at voice frequencies but low enough for steadily applied pressure differences to permit compensation for changes in barometric pressure.

After the damping plate and diaphragm are assembled the space between the clamping rings is filled with beeswax to make the joints gas-tight and exclude moisture. A hole is, however, provided for filling the microphone with nitrogen. The purpose of the nitrogen is to prevent corrosion of the damping plate and diaphragm surfaces and eliminate any reduction in pressure due to oxidation of the sealing compound.

It has been customary for some time to determine the response characteristics of a condenser microphone by the thermophone method.[6] In making this measurement the diaphragm of the microphone is coupled acoustically to the thermophone in the manner shown in Fig. 3. The thermophone consists of two strips of gold foil

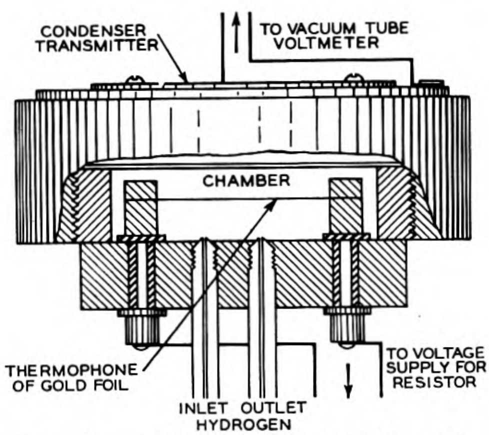


Fig. 3—Cross-sectional view of the thermophone and the condenser microphone.

which are mounted on a plate and fit into the recess in the front of the microphone. Capillary tubes are provided for filling the space enclosed between the plate and the microphone diaphragm with

[6] "The Thermophone as a Precision Source of Sound," H. D. Arnold and I. B. Crandall, *Physical Review*, July 1917, pp. 22–38. "The Thermophone," E. C. Wente, *Physical Review*, April 1922, pp. 333–345. "Speech and Hearing," H. Fletcher, 1929, Appendix A.

hydrogen. This is done in order to make the wave-length of the sound developed in the recess as large as possible compared with dimensions of the chamber. If this were not the case the sound pressure at different positions in the chamber would not be in phase and the conditions on which the computations of the magnitude of the sound pressure are based would not be met. A direct current of known value is passed through the foil. Superimposed upon the direct current is an alternating current of the desired frequency which causes fluctuations in the temperature of the foil and in the gas immediately surrounding it. These fluctuations in temperature in turn cause changes in the pressure on the microphone diaphragm. The magnitude of the pressure developed on the diaphragm can be computed from the constants of the thermophone and the coupling cavity, and the voltage developed by the microphone for a given pressure determined with suitable measuring circuits.[7] Obviously, such a calibration affords a measure of the response of the microphone in terms of the actual pressure developed on the diaphragm and is independent of the external dimensions of the instrument. Hence, it does not take into account any effect which the microphone may have on the sound field when used as a pick-up instrument for recording or broadcasting purposes. The thermophone calibration is often referred to as a "pressure" calibration and the response obtained by placing the instrument in a sound field of constant pressure, a "field" calibration. A thermophone calibration of a representative Western Electric 394-type condenser microphone is shown on Fig. 4.

For many of the uses to which the condenser microphone is put, for example the calibration of head type telephone receivers, the conditions under which it operates agree with those under which the thermophone calibration is made. There are, however, cases where this agreement does not exist, for when a microphone is inserted in a sound field of uniform intensity the pressure on the diaphragm may depart rather widely from a constant value in certain frequency ranges. Several articles [8] have recently appeared calling attention to this discrepancy between the pressure and field calibrations and pointing out that a pressure calibration of a microphone may not be entirely representative of its performance under the conditions which exist in a studio.

[7] "Master Reference System for Telephone Transmission," W. H. Martin and C. H. G. Gray, *Bell System Technical Journal*, July 1929, pp. 556–559.

[8] "The Use of a Wente Condenser Transmitter to Measure Sound Pressures in Absolute Terms," A. J. Aldridge, *P. O. E. E. Journal*, Oct. 1928, pp. 223–225. "Effect of the Diffraction Around the Microphone in Sound Measurements," S. Ballantine, *Physical Review*, Dec. 1928, pp. 988–992. "Measurements of Sound Pressure on an Obstacle," W. West, *Inst. Elec. Eng. Journal*, 1929, pp. 1137–1142.

The difference between the pressure and field calibrations is due to several factors. In the first place the sound is diffracted around the microphone differently at different frequencies. At frequencies where the wave-length is large as compared with its external dimensions the pressure is the same as that of the undisturbed wave. At the higher frequencies where the microphone is large in comparison with the wave-length of the sound, the pressure is twice that developed at the lower frequencies. In the 394-type microphone the effect of diffraction

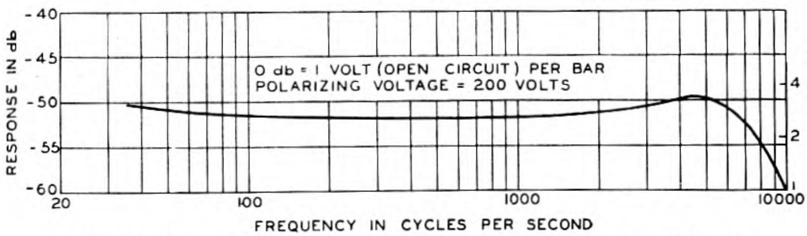


Fig. 4—Pressure calibration of the 394 type condenser microphone.

first becomes noticeable in the region of 1200 cycles and reaches a maximum of 6 db at approximately 2200 cycles. The second factor which causes a difference between the pressure and field calibrations is acoustic resonance in the shallow cavity in front of the microphone. This causes the pressure actuating the diaphragm to be higher than that of the incident sound wave in the frequency region of 1500 to 5500 cycles. The maximum increase in pressure occurs at approximately 3500 cycles. If the sound source is so located relative to the

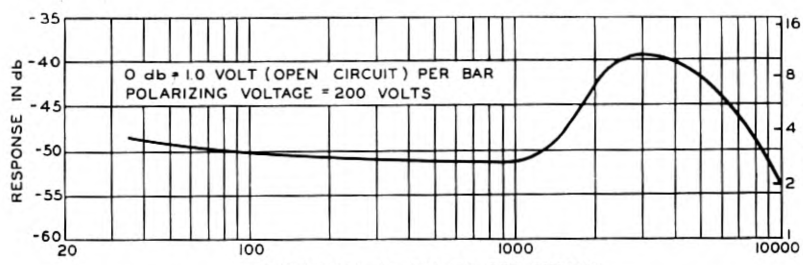


Fig. 5—Field calibration of the 394-type condenser microphone for a direction of approach of sound normal to the diaphragm.

microphone that the waves approach from a direction normal to the diaphragm and reflection from surrounding walls and objects is negligible, the combined effect of diffraction and resonance is to produce a maximum departure from flatness of approximately 12 db as is shown by the field calibration Fig. 5.[9] If the sound wave travels

[9] These curves are taken from unpublished work of P. B. Flanders of the Bell Telephone Laboratories, Inc.

along the diaphragm the effective pressure is reduced at the higher frequencies due to difference in phase. Hence, if the direction of approach of the sound wave is parallel to the plane of the diaphragm, the departure from flatness is materially reduced. This is brought out quite clearly by the field calibration for sound approaching from a direction parallel to the diaphragm, Fig. 6.[9]

The discrepancy between the pressure and field calibrations of the condenser microphone involves two important assumptions, namely, a plane sound wave and no reflection from walls or surrounding objects. When the microphone is used in a studio much of the sound reaches the diaphragm by way of reflection from the walls of the room. The requirement of no reflection is therefore not met and the influence of the acoustic properties of the reflecting surfaces is added to the characteristics of the microphone. The effect of the diffusion of the

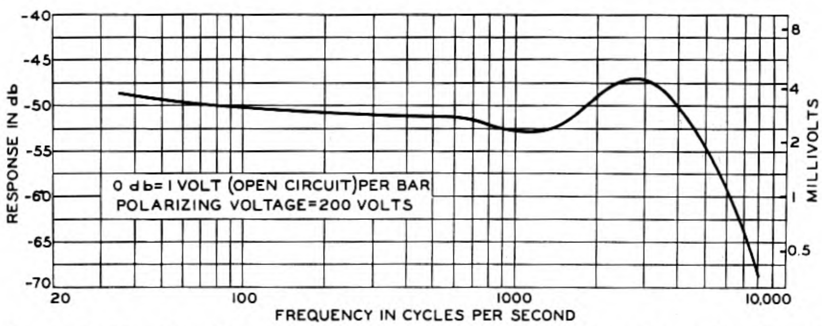


Fig. 6—Field calibration of the 394-type condenser microphone for a direction of approach of sound parallel to the diaphragm.

sound field and the tendency for most materials to be more absorbent for sounds of high frequency appears to cause the response under studio conditions to be more nearly like that obtained when the sound approaches in a direction parallel to the diaphragm and make the departures from the pressure calibration less marked than the field calibration for a direction normal to the diaphragm would indicate. This perhaps accounts in part at least for the instances in which a corrective network designed to compensate for the field calibration normal to the diaphragm failed to effect a material improvement in quality.

The acoustic conditions under which a microphone is used cover a wide range. It would therefore be difficult if not impossible to adopt a set of conditions for use in connection with a field calibration of the condenser microphone, which would be known to be representative of those encountered in practice. The pressure method of calibration on the other hand is definite, simple, and capable of being accurately duplicated in different laboratories. In view of this situation it would seem advisable to retain, at least for the present, the thermophone or pressure method of calibration for general use. In cases where precise quantitative measurements are required a field calibration of the microphone should of course be secured under the conditions of actual use. Various methods of making such a calibration have been proposed. The Rayleigh disc has been used extensively in this work thus far but there are certain very definite limitations to the extent to which it can be applied. An interesting discussion of the use of the Rayleigh disc may be found in papers by E. J. Barnes and W. West,[10] and L. J. Sivian.[11]

It would seem reasonable to expect that future design work would be directed toward reducing transition, resonance and phase difference effects to a minimum. The results of work along this line have been reported by S. Ballantine [12] and D. A. Oliver.[13] In both instances the mechanical design is such that the resonant cavity in front of the diaphragm is eliminated and the housing is spherical or streamline to reduce the diffraction effect. There has as yet been little opportunity to determine the extent of the practical improvement effected by these changes in design and the whole discussion continues to be somewhat academic in character.

Carbon Microphone

Bell's original microphone was essentially a generator and hence was limited in its output to the maximum speech power available at its diaphragm. The demand for telephonic communication over longer distances led to the early introduction of a carbon microphone. In this instrument the resistance of the carbon element is caused to vary in response to the sound pressure on the diaphragm and produces changes in the current supplied from an external source of electrical potential, which are fairly faithful copies of the pressure changes which constitute the sound wave. The carbon microphone is therefore in general an amplifier in which a local source of power is controlled by the acoustic power of the sound wave.

The carbon element or "button" of the first microphones (Edison, 1877) was made from plumbago compressed into cylindrical form.

[10] "The Calibration and Performance of the Rayleigh Disc," E. J. Barnes and W. West, *Inst. of Elec. Eng. Journal*, 1927, Vol. 65, pp. 871–880.

[11] "Rayleigh Disc Method for Measuring Sound Intensities," L. J. Sivian, *Philosophical Magazine*, March 1928, pp. 615–620.

[12] Contributions from the Radio Frequency Laboratories No. 18, S. Ballantine, April 15, 1930.

[13] "An Improved Microphone for Sound Pressure Measurements," D. A. Oliver, *Journal of Scientific Instruments*, April, pp. 113–119.

This type of button was relatively insensitive and shortly after its introduction the suggestion (Hunnings, 1878) was made that the space between the diaphragm and the fixed electrode be "partially filled with pulverized engine coke," [14] in order to increase the number of contact points and render them more susceptible to the forces developed by the motion of the diaphragm. When at its best the Hunnings transmitter was fairly efficient but at times was erratic in its performance due in part to the nature of the microphonic material. In 1886 Edison [15] proposed the use of granules of hard coal which had been heat treated. This was an important advance, for carbon made from anthracite coal is used not only in the microphones which are being considered in this paper but in commercial telephone transmitters as well.

As in the case of the condenser microphone, the displacement of the diaphragm of the carbon microphone must be substantially constant at all frequencies if uniform response is to be obtained. In the early microphones of the carbon type, diaphragm resonance introduced rather prominent irregularities in response. Air damped stretched diaphragms offered one solution of this problem. During the World War instruments of this type were developed and applied to the problem of locating airplanes. In 1921 double button stretched diaphragm microphones were made available for use with the public address equipment installed for the inaugural address of President Harding and the excercises at Arlington on Armistice Day.[16] The carbon microphones employed in sound picture recording are of the stretched diaphragm double button type. The electrical output from this type of microphone is not only of substantially uniform intensity over a wide frequency range but due to the "push-pull" arrangement of the buttons is comparatively free from harmonics. A typical example of the present day carbon microphone is shown in the photograph, Fig. 7. Fig. 8 is a cross-sectional view of the same type of microphone.

The diaphragm is made from duralumin .0017 inch in thickness and is clamped securely at its outer edge. The clamping surfaces are corrugated and emery cloth gaskets are provided to prevent slipping. The stretching of the diaphragm is done in two steps. The initial stretching ring is first advanced by means of six equally spaced screws until the diaphragm is smooth and free from irregularities. The inner or final stretching ring is then adjusted to a position which gives the

[14] "Beginnings of Telephony," F. L. Rhodes, p. 79, 1929.
[15] U. S. Patent No. 406,567, 1889.
[16] "Public Address Systems," I. W. Green and J. P. Maxfield, *A. I. E. E. Journal*, April 1923, pp. 347–358.

diaphragm a resonant frequency of 5700 cycles per second. The method employed in making the determination of the resonant frequency is substantially the same as that used in connection with the assembly of the condenser microphone, with the exception that the

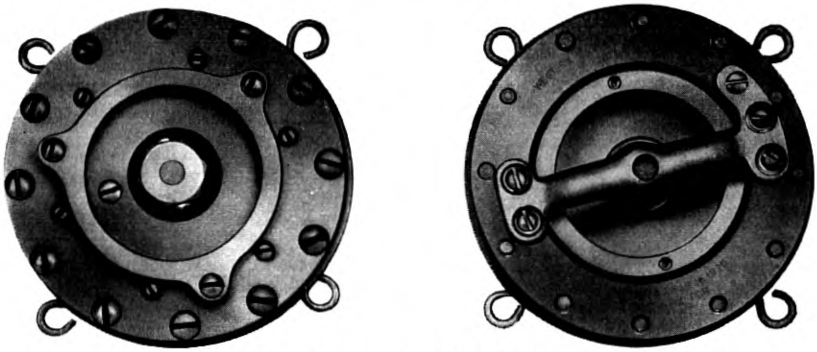


Fig. 7—Western Electric Company's 387-type carbon microphone.

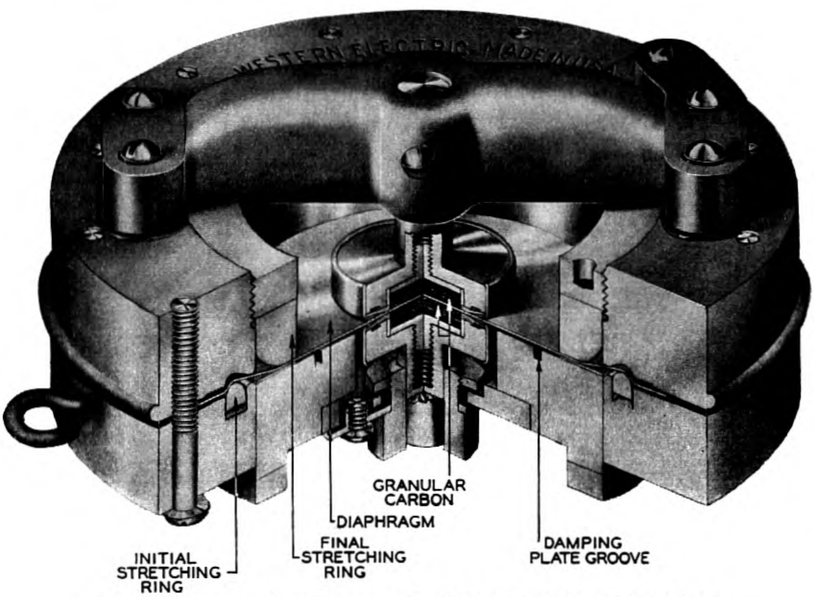


Fig. 8—Cross-sectional view of the 387-type carbon microphone.

frequency at which the maximum output occurs is usually determined by ear rather than by the coupler method previously described. In order to insure a uniformly low contact resistance the portions of the diaphragm which are in contact with the granular carbon are covered with a film of gold deposited by cathode sputtering.

A spacing washer .001 inch in thickness separates the diaphragm from the damping plate. A single concentric groove is provided in the damping plate.

The buttons are of the conventional cylindrical type but are provided with a novel form of closure to prevent carbon leakage at the point where they make contact with the diaphragm. The closure consists of twenty-seven rings of .0004 inch paper clamped firmly together at the outer edge and spreading apart at the inner edge to form a structure which effectively seals the junction between the diaphragm and the buttons without adding materially to the mechanical impedance.

As has already been pointed out the granular carbon is made from selected anthracite coal. The size of the granules is such that they will pass through a screen having 60 meshes per inch but will be retained on a screen having 80 meshes per inch. Before heat treatment the raw material is treated with hydrofluoric and hydrochloric acids to reduce the ash content. Each button contains .060 cc. of carbon, i.e., about 3000 granules.

The bridge which supports the button on the front of the diaphragm partially closes the acoustic cavity on that side. It is essential, therefore, that it be so proportioned as to have a minimum reaction on the response of the microphone and yet provide the required degree of rigidity. It was this consideration that led to the smooth stream line contour now employed.

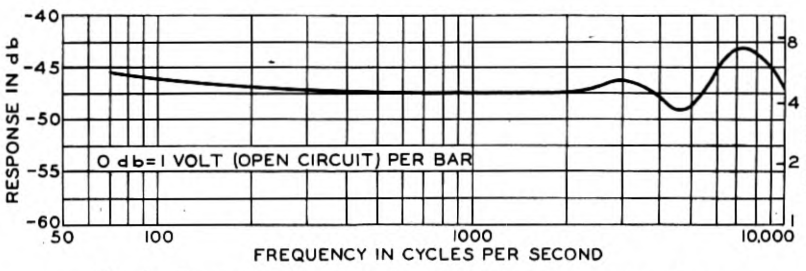


Fig. 9—Pressure calibration of the 387-type carbon microphone.

Referring to Fig. 9 it will be observed that the adoption of an air damped stretched duralumin diaphragm for the carbon microphone has resulted in an instrument having a substantially uniform response over a wide range of frequencies. The arrangement of the apparatus employed in securing the data from which this curve was plotted is shown in the photograph, Fig. 10. The microphone under test was mounted in a highly damped room at a distance of six to eight feet from a source of sound which consisted of two loud speaking receivers.

One of the receivers was the conventional form of moving coil direct radiator and was used to provide sound in the lower frequency range. The other was a special moving coil receiver with a short horn so designed as to serve as an efficient source of sound up to 10,000 cycles.[17] To reduce the effect of standing waves the mounting for the receivers was so constructed that they could be rotated through a circle approximately five feet in diameter and always face the microphone under test. Before starting the test of the carbon microphone the receivers were calibrated by placing a calibrated condenser micro-

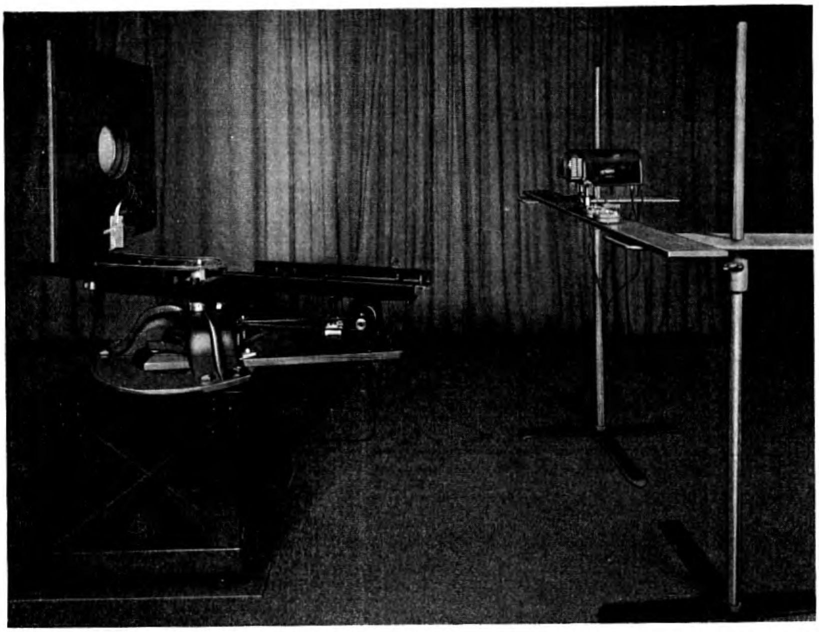


Fig. 10—Apparatus employed in calibrating the 387-type carbon microphone.

phone at the point where the test instrument was to be located and determining the receiver current required to produce a pressure of one bar (one dyne per square centimeter) on the microphone diaphragm. The condenser microphone was then removed and the test microphone substituted. The open circuit voltage developed by the microphone when supplied with a direct current of .025 ampere per button was then measured. The data obtained in this way are essentially a "pressure calibration" of the microphone and in interpreting them in terms of "field" performance the same factors must be taken into account

[17] "An Efficient Loud Speaker at the Higher Audible Frequencies," L. G. Bostwick, *Journal of the Acoustical Society*, Oct. 1930, pp. 242–250.

which have been discussed in considerable detail in connection with the condenser microphone.

The circuit employed in measuring the response of the carbon microphone is shown on Fig. 11. Two steps are involved in the calibration of the sound source. With the output terminals of the microphone circuit and the sound source short circuited and the polarizing voltage for the condenser microphone removed, the attenuator is adjusted until the voltage applied to the measuring circuit is that developed by the condenser microphone when a sound pressure of one bar is impressed on its diaphragm. A record is made of the reading of the

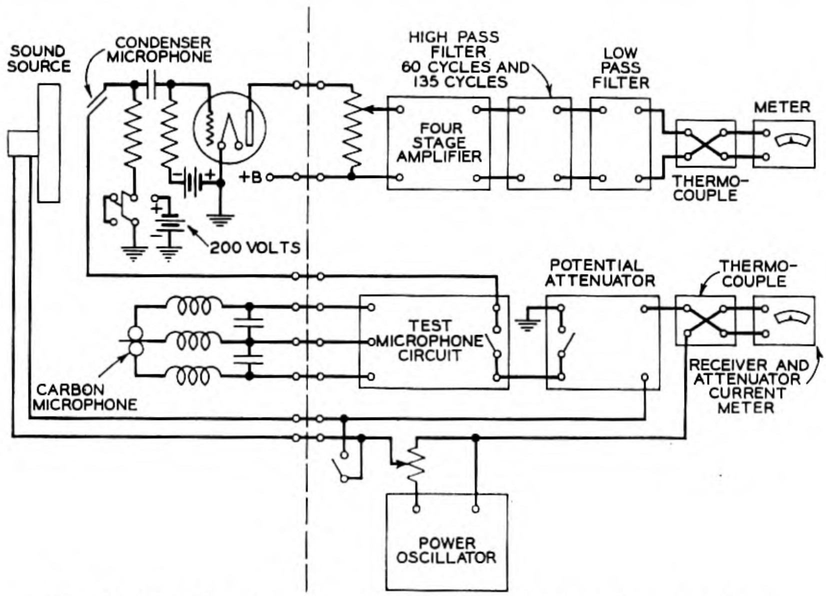


Fig. 11—Circuit employed in calibrating the 387-type carbon microphone.

output meter in the measuring circuit. The polarizing voltage is then applied to the condenser microphone. After the output terminals of the attenuator have been short circuited an alternating current of a known frequency is supplied to the sound source and the magnitude of this current adjusted until the meter reading is the same as that previously obtained with the attenuator. This completes the calibration of the sound source for that frequency. After the carbon microphone has been placed in the position previously occupied by the condenser microphone, the polarizing voltage is once more removed from the condenser microphone and the output from the carbon microphone circuit impressed on the measuring circuit. The reading of the output meter is recorded. The sound source and carbon microphone circuit are then short circuited and the output from the attenuator again applied to the measuring circuit. The attenuator is adjusted until the reading of the output meter is the same as was previously obtained with the carbon microphone in circuit. In this way the voltage applied to the measuring circuit when the carbon microphone is in operation is determined. The open circuit voltage developed by the carbon microphone may then be computed from the voltage and the constants of the microphone circuit. At the locations where these measurements were made a certain amount of interference from 60-cycle circuits and low frequency acoustic disturbances was encountered. The high-pass filter in the measuring circuit was introduced to facilitate the measurements under these conditions. The adjustable low-pass filter was used to confine the measurements to the fundamental frequency. Only that portion of the apparatus to the left of the dotted line was mounted in the damped room.

The two buttons of the carbon microphone are identical in their dimensions and if the granular carbon is in the same mechanical state have substantially the same electrical characteristics. They are also practically free from the cyclic variations in resistance known as "breathing" which result from the temperature changes caused by the power dissipated in the granular carbon. It is, however, a matter of every day experience that a given mass of granular material will occupy different volumes, depending upon the configuration of the particles. In the case of microphone carbon this change in configuration of the granules results in changes in the contact forces of sufficient magnitude to affect the resistance and sensitivity. If these changes occur in unequal amounts in the buttons electrical unbalance results. When complete balance exists the electrical output is free from all harmonics introduced by the circuit. Hence, in using the microphone care should be taken to see that a fair degree of balance between the buttons is maintained.

The performance of a carbon microphone may be affected adversely by cohering of the granules. Severe cohering is accompanied by a serious reduction in resistance and sensitivity which persists for an extended period unless the instrument is tapped or agitated mechanically. One of the common causes of cohering is breaking the circuit when current is flowing through the microphone. Experiment has shown that the insertion of a simple filter consisting of two .02 mf. condensers and three coupled retardation coils each having a self-inductance .0014 henry, will effectively protect the microphone button from cohering influences without introducing an appreciable transmission loss. This filter may be located in the base of the mounting or in a container fastened to the back of the microphone.

Aging of granular carbon may result from changes in the contact surface caused either by mechanical abrasion or overheating due to excessive contact potentials. Aging is usually accompanied by an increase in resistance and loss in sensitivity. Care should therefore be exercised in the use of the carbon microphone that it is not subjected to unnecessary vibration which would cause the granules to move relative to one another and abrade the surfaces. The use of abnormally high voltages should also be avoided.

The quality of transmission obtained with the double button carbon microphone compares favorably with that secured with a condenser microphone. The carbon microphone also requires less amplification. There is, however, one characteristic which limits its use, namely carbon noise. The level of the noise is much higher than that due to thermal agitation within the carbon granules [18] and appears to be caused by heating at the contacts between the granules. A certain amount of gas is contained in the pores in the contact surfaces. When current passes through the button, a sufficient increase in contact temperature takes place to cause a portion of this gas to be driven off and produce the non-periodic changes in resistance which give rise to carbon noise.

In conclusion it may be stated that the condenser and carbon types of microphones have been developed to a point where there is little to choose between them from the standpoint of quality of transmission. The design from a mechanical standpoint has also been carried to a point where little difficulty should be experienced in their use if reasonable precautions are exercised. Although requiring less amplification than the condenser microphone the extent to which the carbon microphone is used at present is limited by the higher noise level obtained. The condenser type of microphone has therefore been adopted for most of the recording work in the sound picture field.

[18] "Thermal Agitation of Electricity in Conductors," J. B. Johnson, *Physical Review*, July 1928, pp. 97–109.

Certain Factors Affecting the Gain of Directive Antennas *

By G. C. SOUTHWORTH

This paper analyzes the performance of antenna arrays as influenced by certain variables within the control of the designing engineer. It starts with an extremely simple analysis of the interfering effects produced by two sources of waves of the same amplitude. This is followed by a short discussion of a paper by Ronald Foster, which considers two antennas and also 16 antennas when arranged in linear array. Two antennas separated in space by ¼ wave-length and in phase by ¼ period give sensibly more radiation in one direction than in the opposite. This, for convenience, has been called a unidirectional couplet. A number of these couplets may be arranged in linear array, thereby giving an extremely useful directive system. Diagrams are shown for such arrays as affected by the number and spacings of the individual couplets. The gains from such arrays are calculated and data are given showing fair agreement between calculation and observation.

Directional diagrams for arrays of coaxial antennas indicate that somewhat less gain may be expected from this form than when the elements are spaced laterally. Combinations of these two types of arrays give marked directional properties in both their horizontal and vertical planes of reference. This principle has been used rather generally in short-wave communication. This paper also discusses effects resulting from combining two or more arrays. In one case the space between two arrays tends to emphasize spurious lobes. The directional diagram of such a combination may be rotated within limits by changing the phasing between adjacent arrays or sections of an array. In all of the above cases the influence of the earth is ignored.

A mathematical appendix gives general equations for calculating directional diagrams of linear arrays. Special cases of these equations apply to the figures included in the main part of the text. General equations are also given for calculating the gains of arrays. Similar equations permit the areas of diagrams to be calculated. An extended bibliography on antenna arrays is appended.

Introduction

THROUGHOUT the development of radio communication the engineer has aspired to a directive system whereby radiation might be projected from one point to another with a maximum of efficiency and a minimum of interference with adjacent stations. Also, he has aimed at similar directivity at the receiver to improve the signal-to-noise ratio and otherwise discriminate against undesirable signals. It was recognized at a very early date that directive radio based on wave interference was feasible provided sufficiently short waves could be utilized, and as a result many interesting suggestions to this end were made. However, as is well known, the early development of the radio spectrum proceeded in the direction of long

* Presented at Convention of I. R. E., Toronto, Ont., Canada, Aug. 19, 1930. *Proc., I. R. E.*, Sept. 1930.

waves rather than short waves, thereby deferring many of the applications of these suggestions.

The principle of wave interference on which most short-wave systems of directive radio are based has probably been known for several centuries. However, the first thorough treatment of this subject was by Sir Thomas Young,[1] who, together with Fresnel, securely established the wave theory of light in the early part of the last century. Even Hooke and Huygens, who had offered the wave theory over a century earlier, failed to recognize the full significance of interference.

When Hertz started his celebrated experiments to verify Maxwell's theory he was, of course, in full knowledge of these phenomena and their explanation, and invoked their use in proving the existence of electric waves. It is interesting that in some of his experiments he made use of parabolic mirrors for both transmitting and receiving, having directional characteristics very similar to those sometimes used in present day radio practice. It is also of interest that he found that parallel wires stretched over a frame were quite as effective as a reflector as a continuous sheet of metal of similar dimensions, provided the wires were kept parallel to the lines of electric force of the arriving wave. He apparently did not investigate the effect of varying the spacing nor the length of the parallel wires, nor did his subsequent experiments otherwise tend toward the present day antenna array technique.

This paper treats in an elementary way certain aspects of the antenna array problem, principally as regards the manner in which calculated directivity is affected by the number and spacing of the individual antennas which go to make up the array. The theory is applicable only to those forms of directive antennas which may be resolved into a series of individual sources. It does not apply to the so-called wave antenna. However, principles are included which have for some time been in general use in combining two or more such antennas.

Extensive study has been given to directive antenna systems for use in transoceanic radiotelephony. Papers dealing with this general subject have appeared from time to time.[2] Further work is in progress. Papers by E. J. Sterba and also by E. E. Bruce and H. T. Friis of the Bell Telephone Laboratories are in preparation which will include

[1] *Phil. Trans. of Royal Soc.*, **92**, 12; 1802.

[2] R. M. Foster, "Directive diagrams of antenna arrays," *Bell Sys. Tech. Jour.*, 292, May, 1926. Austin Bailey, S. W. Dean, and W. T. Wintringham, "Receiving system for long-wave transatlantic radiotelephony, *Proc. I. R. E.*, **16**, 1694, December, 1928. J. C. Schelleng, "Some problems in short-wave radiotelephone transmission," *Proc. I. R. E.*, **18**, 913; June, 1930.

certain calculated data similar to those contained in the present paper, and also experimental results obtained from tests on actual antennas of various sizes and proportions.

In the early part of the following discussion each antenna is considered as a spherical source of waves which radiates equal power in all directions. Furthermore, it assumes that the current in each

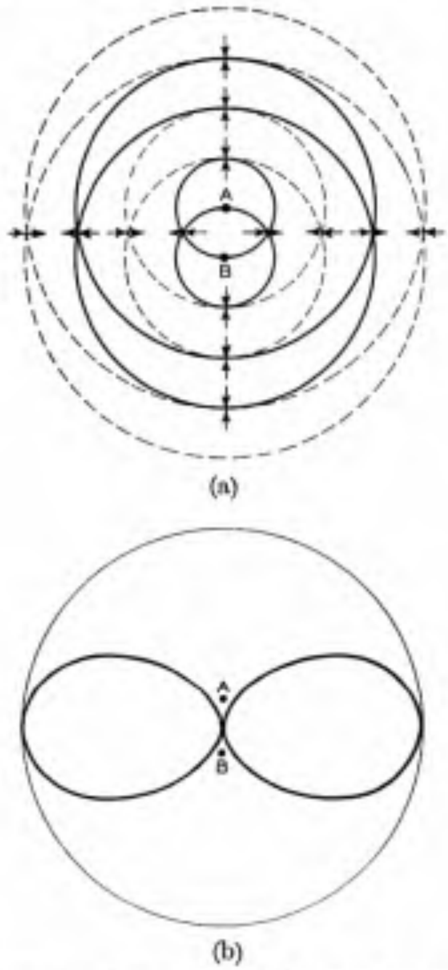


Fig. 1—Interference pattern. Two equiphased sources spaced one-half wave-length.

individual source, in a given array, is the same and is not materially affected in either magnitude or phase by its proximity to other sources. The fair approximation to which these calculated results are realized in practice bespeaks the justification of these assumptions.

The various steps by which present day directional radio has been developed are extremely interesting, but they are so involved in the development of radio itself that their enumeration is considered outside the scope of this paper. However, bibliographies are cited below covering some of their important phases.

Elementary Principles

The interference patterns resulting from a number of individual sources of waves, such as antennas, are dependent on both their spacial arrangement and the magnitudes and relative phases of their forces. This makes possible an almost unlimited number of combinations of which only a portion have thus far found use in com-

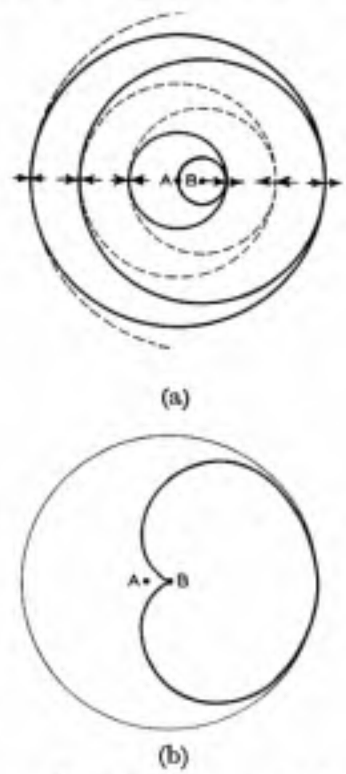


Fig. 2—Interference pattern. Two sources separated in space by one-fourth wavelength and in time by one-fourth period.

munication. This paper will restrict itself mainly to some cases which are already finding general application. As a suitable introduction to this subject, a very simple case of wave interference is discussed in the following paragraph.

Figs. 1a and 2a depict in a rough way the interference resulting from two independent sources of spherical waves of the same amplitude. In the first case they are spaced ½ wave-length but are assumed to be oscillating in phase. In the second case the two sources are separated in space by ¼ wave-length and in phase by ¼ period. Crests and troughs are represented respectively by solid and dotted lines. At points where either two crests or two troughs arrive simultaneously the resultant wave is greatly enhanced, whereas at certain other points crests and troughs arrive together, thereby neutralizing each other's effects. At certain intermediate points these interfering effects are only partially complete. Accompanying each figure is a directive diagram (1b and 2b), plotted in polar coordinates, which shows the effectiveness of the wave in each direction. The circle drawn outside each diagram indicates the effect if the radiation had proceeded from a single non-directional source similar to each of the above. The ratio between the areas of the circle and the inscribed diagram gives roughly the power improvement of such a device as manifested in the intensity of the radiated wave. A more exact calculation of this improvement requires an integration of the force components over a unit sphere.

Linear Antenna Arrays

Most directive antenna systems now in general use for short waves may be regarded as special applications of the linear array. This type consists of two or more antennas all having currents of equal amplitude, equispaced along the same straight line. The properties of such arrays have been treated very generally by Foster,[2] whose paper included several hundred directive diagrams, taken in a bisecting plane perpendicular to the axis of each antenna of the array, and typical of the results which may be expected from two antennas and from arrays consisting of 16 antennas. A portion of these diagrams have been reproduced in Figs. 3 and 4 below. The same principles are applicable to both transmission and reception.

In Fig. 3 are shown diagrams resulting from two antennas as the separation is increased from 0 to 1 wave-length in steps of $\frac{1}{8}$ wave-length and the phase increased from 0 to $\frac{1}{2}$ period in steps of $\frac{1}{8}$ period. The line or axis of the array is assumed to be horizontal and the specified phase difference is such that the current in the right-hand antenna is lagging for a transmitting system and leading for a receiver. It will be noted that for phase differences of both 0 and $\frac{1}{2}T$ the diagrams are symmetrical about both the horizontal and vertical axes of the figure, whereas for other phases the figures are asymmetrical about the vertical axis except for certain limiting cases. Of these asymmetrical diagrams, that corresponding to phase and spacial separations both of $\frac{1}{4}$ (Fig. 3b) is of particular importance and forms the basis of the so-called reflector effect. This particular combination of two sources is referred to later as a unidirectional couplet.[3] In

[2] Loc. cit.

[3] In this, and in other cases in this paper, radiation is referred to as unidirectional when sensibly more power is propagated in one direction than in others.

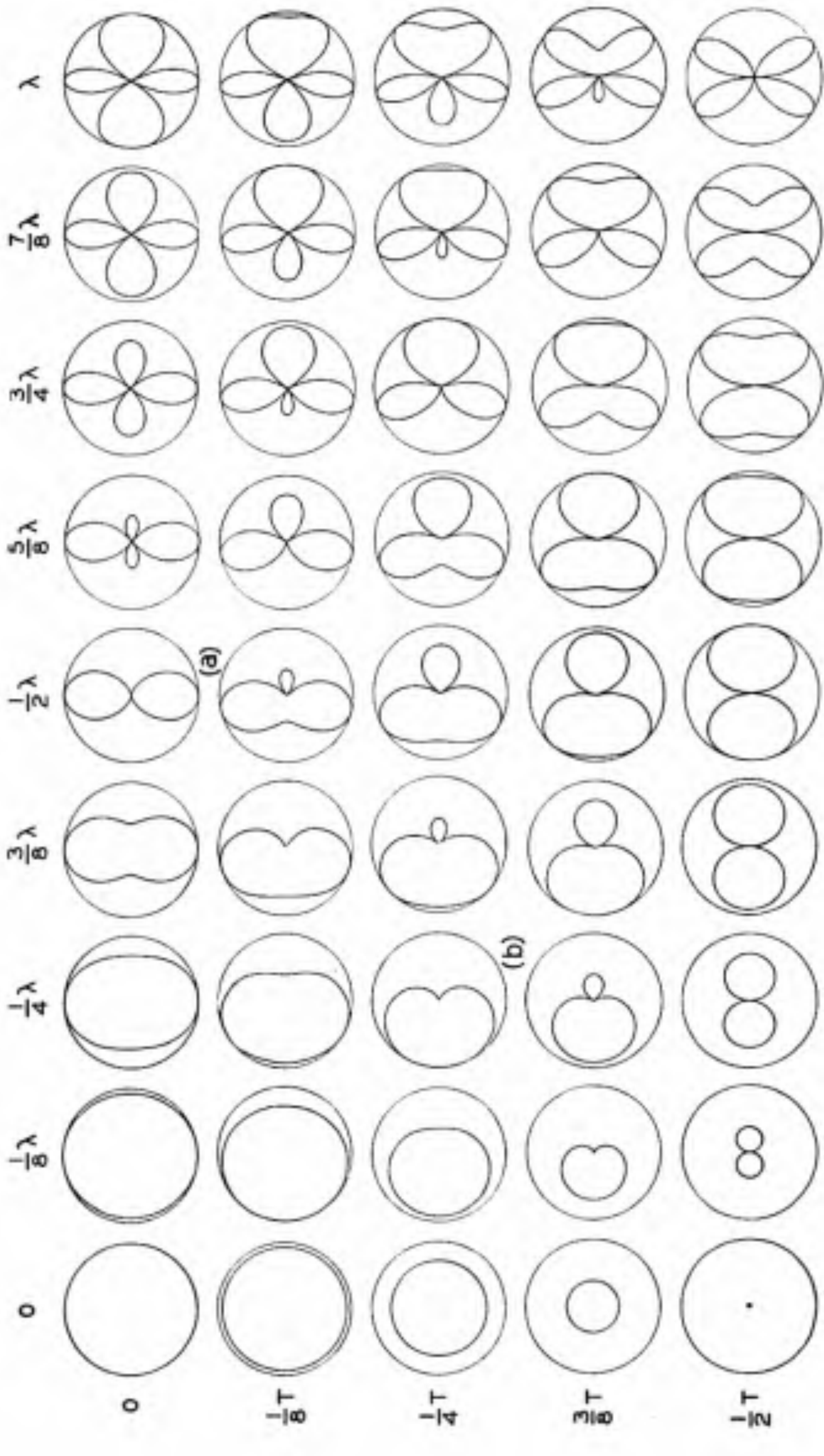


Fig. 3—Directive amplitude diagrams for an array of two antennas. Separation in wave-lengths (λ) along the top. Phase difference in periods (T) at the left.

passing it is also of interest to note that the diagram of the coil or frame aerial as generally used is intermediate between Figs. 3c and 3d. Its diagram would not differ essentially from its neighbors, Figs. 3d, 3e, or 3f, except for scale. This scale may conveniently be regarded as a measure of the impedance of the device, or possibly its radiation efficiency, but not necessarily a measure of its usefulness.

Fig. 4 shows similar diagrams resulting from 16 antennas for various phase and space relations. As in Fig. 3, diagrams in the top and bottom rows corresponding respectively to phases of 0 and ½ T are symmetrical about both the horizontal and vertical axes. The diagrams in the top row are in general bidirectional, while the bottom row has one bidirectional diagram corresponding to phase and space differences both equal to ½. It is of interest that for the most part cases where the phase and space separations are numerically equal correspond to unidirectional diagrams. However, these diagrams are only moderately sharp and thus far such arrays have not been used extensively in practice.

Referring again to the diagrams in the top row corresponding to 16 antennas all driven in phase, we note that directivity becomes progressively sharper as the spacing is increased until in the vicinity of $15/16\lambda$ appendages develop which soon surpass in magnitude the desired lobes. This effect is present in the commercial array, and limits, as we shall later see, the gain that may be derived from a given number of elements. The diagrams shown in Fig. 4 for 16 antennas are typical of others where the number of antennas in linear array is fairly large.

The Linear Array and Reflector

One type of array now in commercial use consists of two parallel linear arrays of equiphased elements where the two parallel arrays are spaced ¼ wave-length and differ in relative phase by ¼ period. It is convenient to regard such a device either as two independent linear arrays, each having a directional characteristic as shown in the top row of Fig. 4, or as an array of couplets, each couplet of which has by itself a heart-shaped characteristic. Both antennas of the couplet may be independently driven at their prescribed phase separation of ¼ period, or one may derive its power from that radiated by the other, in which case the proper phase relation is automatically approximated [4] and the same practical result is obtained. In the latter case one is

[4] The problem of the reflecting antenna has been considered by Wilmotte and McPetrie, *Jour. I. E. E.*, **66**, 949, Englund and Crawford, *Proc. I. R. E.*, **17**, 1277; August, 1928, and Palmer and Honeyball, *Jour. I. E. E.*, **67**, 1045. Their conclusions indicate that the optimum separation between a single antenna and its reflector to give maximum forward radiation is roughly $\lambda/3$. However, it appears that when several antennas and reflectors are involved a separation more nearly $\lambda/4$ is optimum.

frequently known as the driven antenna and the other the reflector. This viewpoint is perhaps only a convenience and may not be altogether correct. An array of the above type transmits and receives best in a direction at right angles to its principal dimension. This type is, therefore, frequently known as a broadside array.

Directive Diagrams from Arrays and Reflectors

In Fig. 5 is plotted a series of diagrams in a bisecting plane normal to the axis of each antenna of the array for different broadside arrangements such as are used commercially. They are systematically arranged horizontally in the order of the number of couplets in the array, and vertically with the increased spacing between adjacent couplets.

Several different forms of such directive diagrams are possible, which may be plotted in either polar or rectangular coordinates. In one form all diagrams are roughly of constant area and relative gains from various antenna systems are expressed in terms of the principal radius vector. In the second form the length of the principal radius vector remains constant and the relative gain is roughly inversely proportional to the area of the diagram. The second of these forms has been adopted in this paper largely because of the relative simplicity of the equation of the diagram and the facility with which properties of antennas may be determined.

In the lower left-hand corner of Fig. 5 will be found a plan showing the arrangement of the elements relative to the important direction of transmission. At its right is the general equation of these diagrams. This formula is also given as equation (14) of the appendix where the analytical theory of arrays is developed. Below each diagram is the ratio of the area of the circumscribed unit circle to the area of the horizontal diagram. Here also will be found the ratio of the area of the subordinate loops to the area of the main loop. The total area may be measured approximately with a planimeter or calculated more accurately by equation (32) in the mathematical appendix. In making up Fig. 5 each diagram was accurately plotted on standard polar coordinate paper from perhaps a hundred calculated points. This was then reduced photographically and the several diagrams were assembled.[5]

Inspection of the diagrams shows that increasing the number of couplets increases in all cases the sharpness of the main loop and hence the gain of the array. However, increasing the separation be-

[5] The diagrams used in this paper were calculated by a group of the Department of Development and Research of the American Telephone and Telegraph Company, under the direction of Miss E. M. Baldwin. Most of the material was checked by Mrs. Isabel Bemis, who assembled it in its present form and prepared the attached bibliography.

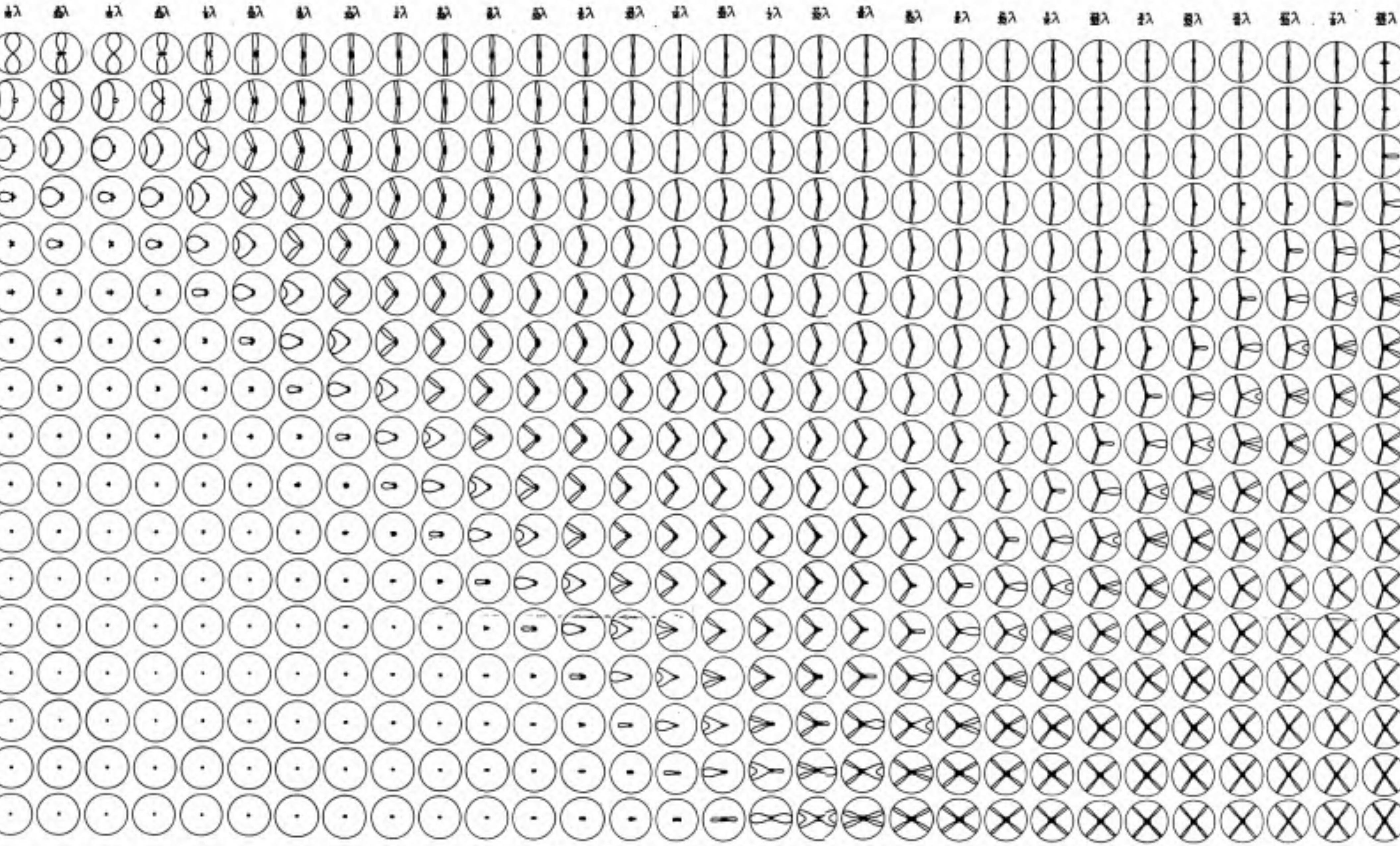


Fig. 6—Directive amplitude diagrams for an array of sixteen antennas. Separation in wave-lengths (λ) along the top. Phase difference in periods (T) at the left.

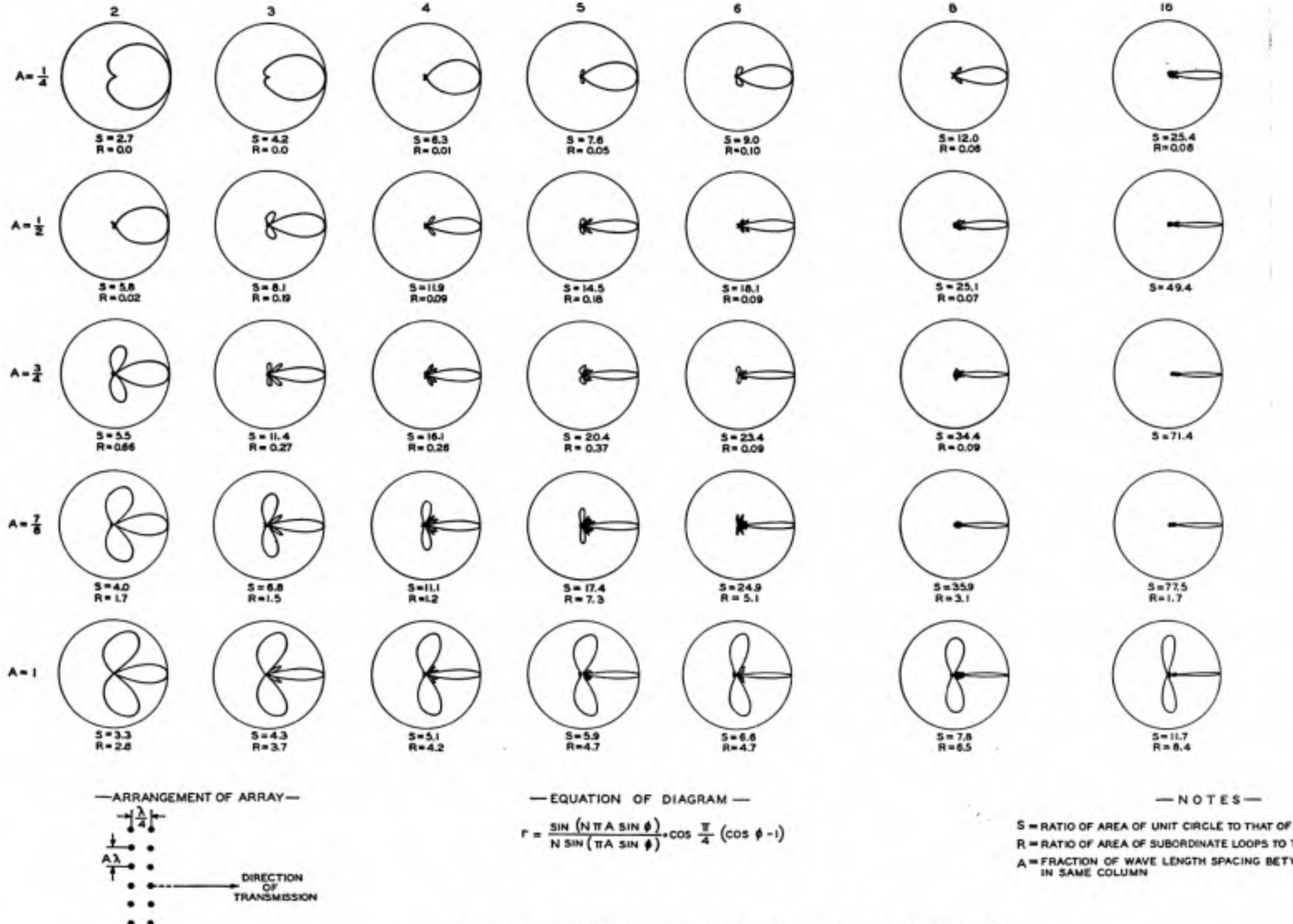


Fig. 5—Horizontal plane diagrams—number of couplets versus separation in wave-lengths.

tween couplets increases the gain only up to a certain point, after which the formation of parasitic lobes decreases the effectiveness of the array. The trend of these gains may be illustrated more effectively in graphical form.

In Fig. 6 calculated gain ratio is plotted against number of couplets giving one graph for each separation considered. These ratios are not based on the data given in Fig. 5, but were obtained from the integration of the equation of the directional diagram over an arbitrary sphere by use of equation (27) below. It may be noted that for many conditions the difference between these methods of calculating gain is only moderate. These power ratios are for the most part linear,

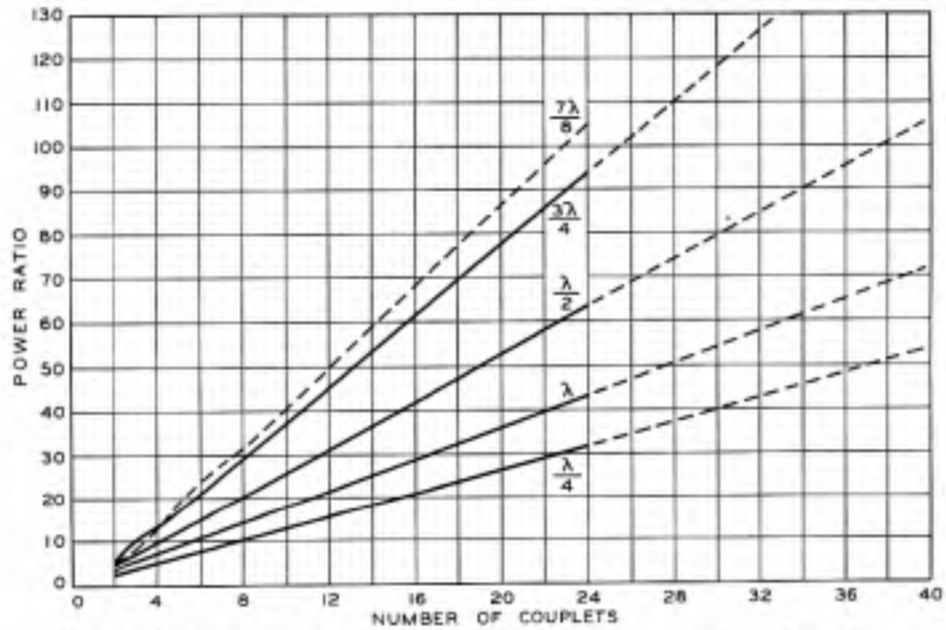


Fig. 6—Antenna arrays. Calculated power ratios vs. number of couplets.

indicating that such gains are proportional to the length of the array. This is in keeping with the view that a receiving antenna can intercept wave power more or less in proportion to its dimensions. It is also interesting to note that the slope of the curve of $\lambda/2$ is approximately twice that for $\lambda/4$, so that 16 couplets spaced $\frac{1}{4}$ wave-length give approximately the same gain as eight couplets spaced $\frac{1}{2}$ wave-length. This again shows that the length of the array is the most important criterion in determining its gain. In Fig. 7 the same data have been plotted in decibels.

In Fig. 8 gains expressed in decibels are plotted against the separation between elements. This shows more definitely the trend of the antenna gain to a maximum, after which spurious lobes become of importance. Fig. 8 suggests that the spacing, giving optimum gain, would be the desideratum in antenna design. However, this is not

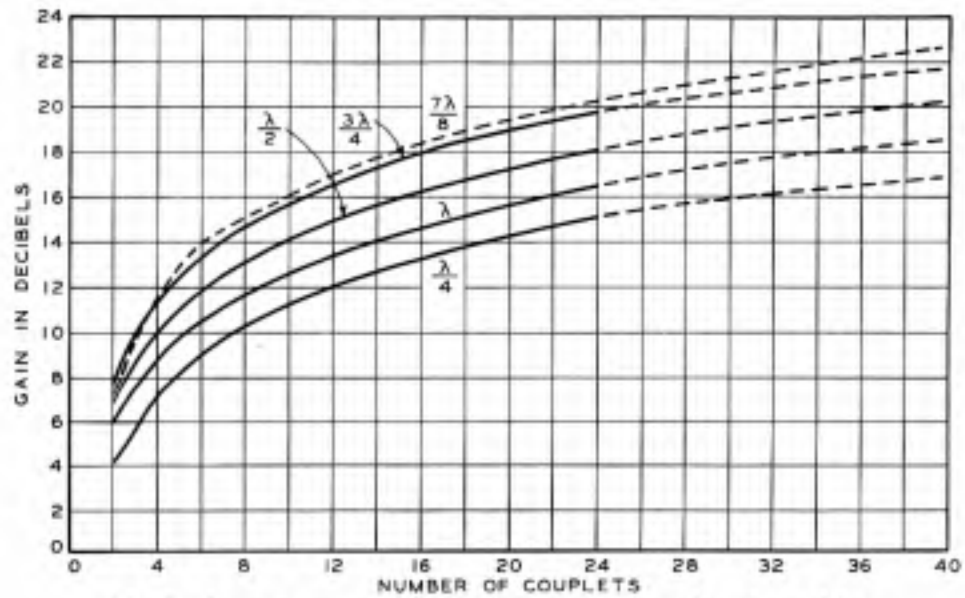


Fig. 7—Antenna arrays. Calculated gains vs. number of couplets.

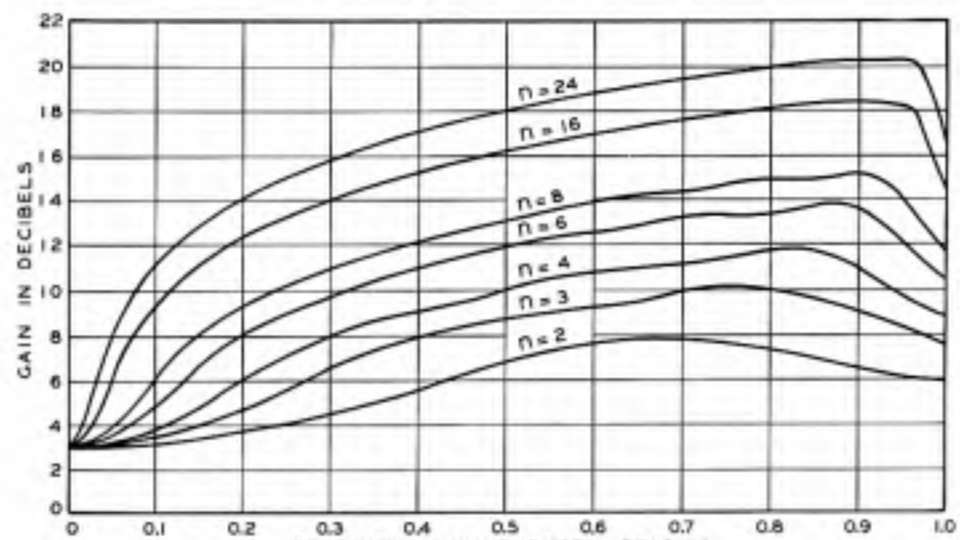


Fig. 8—Antenna arrays. Calculated gains vs. lateral spacing between couplets.

necessarily the case, as we shall presently see. It has already been pointed out that the over-all length of array, rather than the spacing or the number of conductors per unit length, constitutes the most important factor in determining the gain. Furthermore, minimum area diagrams are frequently attended by fairly large spurious lobes which are undesirable particularly on receiving antennas. Also the

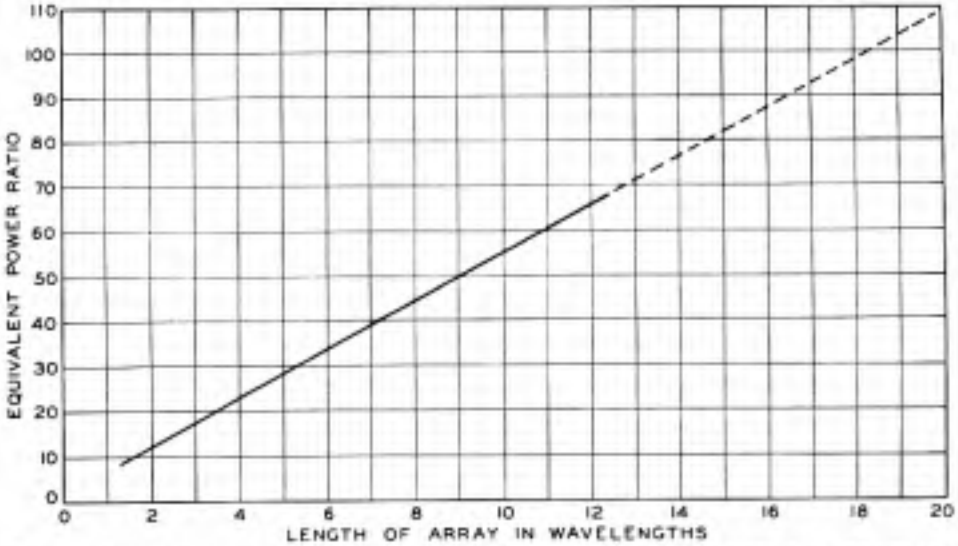


Fig. 9—Approximate gains to be expected from arrays of couplets for spacings of approximately $\lambda/4$ and $\lambda/2$.

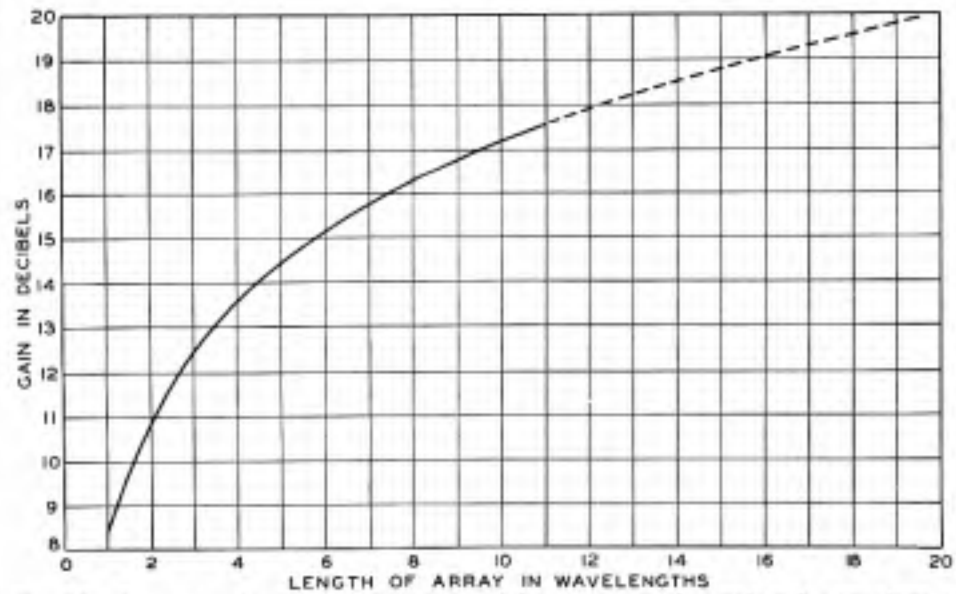


Fig. 10—Approximate gains to be expected from arrays of couplets for spacings of approximately $\lambda/4$ and $\lambda/2$.

cost of an antenna system of a given height is more or less proportional to its length, and in many cases is not materially affected by the number of conductors present. These considerations, together with the fact that proper phases may often be most readily accomplished with intervals of either ¼ wave-length or ½ wave-length, have led to a rather general adoption of these closer spacings.

In Fig. 9, approximate gain ratios from arrays of various lengths have been plotted. These are most applicable for separations in the vicinity of ¼ and ½ wave-length. Fig. 10 shows the same data plotted in decibels. Within these limits, it appears that the gain ratio may be expressed by the simple formula $G = KL$, where L is the array length in wave-lengths and K is approximately 5.6. The result expressed in decibels is $G' = 10 \log_{10}(KL)$.

Measured Antenna Gains

The degree to which the gains calculated above are approximated in practice is indicated by the data given in the diagrams of Figs. 11 and 12 and in Table I.

TABLE I

Array Designation	Nominal Operating Frequency Megacycles	Number Couplets	Spacing	Measured Gain Over Similar Single Element db	Calculated Gain db	Difference db
1-A.....	18	24	$\lambda/4$	15.3	15.0	+ 0.3
2-A.....	18	24	$\lambda/4$	15.2	15.0	+ 0.2
3-A.....	18	24	$\lambda/4$	15.0	15.0	0.0
1-B.....	12	24	$\lambda/4$	15.6	15.0	+ 0.6
2-B.....	12	24	$\lambda/4$	14.5	15.0	− 0.5
3-B.....	15	24	$\lambda/4$	13.6	15.0	− 1.4
4-B.....	15	24	$\lambda/4$	16.6	15.0	+ 1.6
2-C.....	10	24	$\lambda/4$	16.3	15.0	+ 1.3
3-C.....	10	24	$\lambda/4$	15.5	15.0	+ 0.5
1-C.....	9	18	$\lambda/4$	13.6	13.8	− 0.2
D *.....	14	9	$\lambda/2$	13.0	13.7	− 0.7

* This antenna actually consisted of two arrays of four couplets each spaced laterally by one wave-length. The resultant diagram of such an array is for all practical purposes the same as that produced by a continuous array of nine couplets.

Fig. 11 shows a calculated diagram corresponding to certain receiving arrays used in the transatlantic telephone service between America and England. Several points are plotted on this diagram which correspond to the relative strengths of signals received at various angles. These points were obtained by observing the relative received signal voltage, measured on a standard field-strength measuring set connected to the array as an electric oscillator of constant amplitude was carried around the array at a distance of perhaps 20 wave-lengths. The plotted data correspond to the case where the reflector was "floating." Although this arrangement most nearly corresponds to the conditions assumed in the calculated curve, it is not necessarily the most desirable adjustment to minimize noise arriving from the rear. This diagram corresponds to the antennas designated as 1-A, 2-A, and 3-A in Table I. These antennas consist effectively of 24 vertical couplets spaced horizontally at intervals of ¼ wave-length.

In this table are given further data on the strength of signals received on arrays, as compared with those received simultaneously on a single element of similar structure and height above earth. The different antennas represented involve varying conditions of wave-

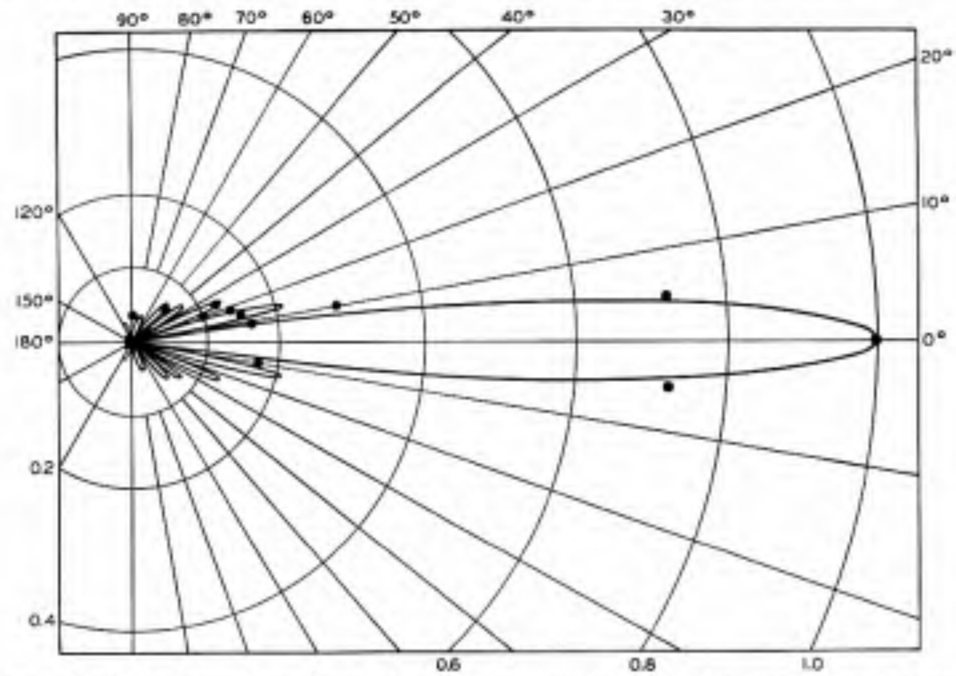


Fig. 11—Calculated directional diagram. Twenty-four couplets spaced one-fourth wave-length. Circles indicate experimental points.

length, height above earth, adjacent terrain, and types of support. These details are not believed to be of sufficient importance for discussion here. Two different array lengths are represented. The relative gains were substantially the same when observed on a local source of waves and when the signal came from a distant station. The last array represented in Table I was one used for transmitting. To effect the test, equal power was transmitted alternately from the array and from a single element while comparative measurements of electric field strength were made at a distance of approximately 3500 miles. The datum given is the mean of perhaps 100 observations extending over a total of eight hours on three different days. Two errors are involved in the data of Table I. One is due to the doubtful magnitude of a correction necessary to account for the various heights at which the arrays were located above the earth and the second is the error of measurement of gain as compared with the reference antenna. These errors are approximately equal and together amount to ± 1 db.

In order to test further the agreement between measured gains and those calculated from the simple assumptions above, a receiving array was assembled step by step and corresponding measurements made. Certain precautions, such as to maintain impedance matches at points of coupling, were observed. The resulting data were plotted as points in Fig. 12. A smooth curve represents the corresponding calculated

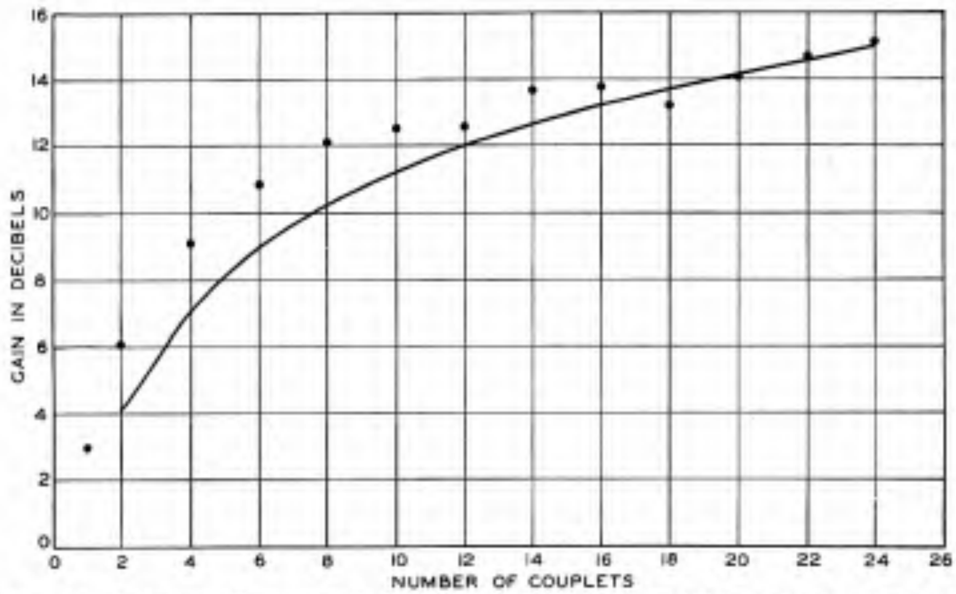


Fig. 12—Relation of measured to calculated gain of receiving antenna array at 14,350 kc.

data. It will be observed that the measured values are consistently higher than those calculated at the lower end of the curve, and in this region the agreement can hardly be regarded as satisfactory. However, limited time prevented a thorough study of the errors of measurement. Consequently these limited data may not be regarded as any adequate test of the theory.

Combinations of Arrays

It may be shown that two or more similar directive systems may be combined to give a total directive effect, represented by the product of the individual effect, multiplied by the group effect. This principle is partially covered by equation (35) of the mathematical appendix.

Two cases are of special interest. First, it is sometimes desirable to divide an array into two or more bays, in order to make room for a supporting structure. This, of course, gives rise to a definite discontinuity in the over-all array.

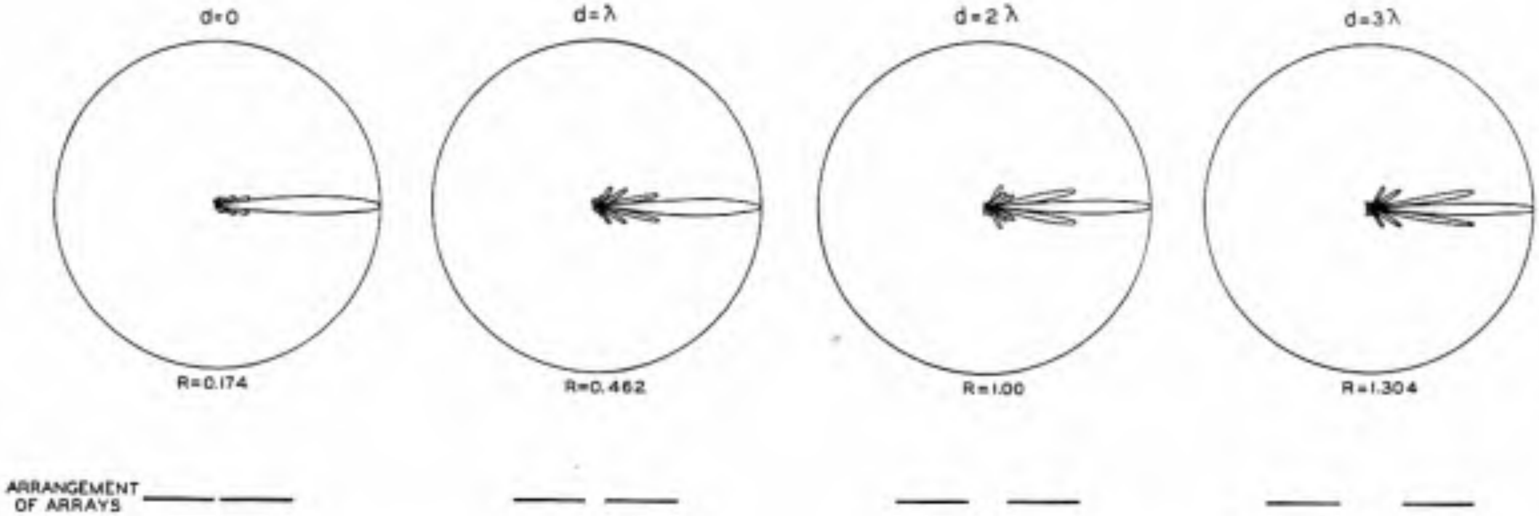
Fig. 13 shows a series of diagrams resulting from a typical case of two such arrays, each having a length of 2½ wave-lengths but separated variously from 0 to 2 wave-lengths in steps as noted. These diagrams, of course, do not take into consideration the reaction resulting from proximity to an antenna mast, located in such an opening. The most important result is to emphasize the spurious lobes, as the spacing between arrays is increased.

A second effect of grouping which is of considerable interest is that of varying the direction of transmission by altering the respective phases between two or more arrays or between sections of the same array. In Fig. 14 a series of diagrams is shown for a typical case of two 3½ wave-length arrays, spaced one wave-length. All elements in the same array are driven in phase, but the two arrays differ in phase by various amounts, as noted. It will be observed that the possible rotational effect is very limited. The general equation for this diagram is given by formula (36) of the mathematical appendix.

This effect was investigated further by assuming a continuous array 7½ wave-lengths long, made up of 16 couplets spaced at intervals of ½ wave-length. The results are depicted in Fig. 15. The top row assumes that the array is divided into two sections of eight couplets each. This gives similar but not exactly the same results as those of Fig. 14. The array, however, might have been divided into other sections for purposes of phasing. The various possible combinations are tabulated below:

Number of Sections	Number of Couplets per Section
2	8
4	4
8	2
16	1

Diagrams in rows two, three, and four show that, as the array continues to be divided into smaller sections, the direction of transmission is capable of greater variation without sensible loss of sharpness. If the array be divided into two sections this range is limited to perhaps 3 deg. as in the case depicted in Fig. 14. Although this is very moderate, it is extremely useful in correcting for any errors in the orientation of the supporting structure or possibly correcting for deviation of the projected radiation caused by peculiarities of the adjacent terrain.



— EQUATION OF SPACED ARRAY —

$$r = \frac{\sin(3\pi \sin\phi)}{6 \sin\left(\frac{\pi}{2}\sin\phi\right)} \cos\frac{\pi}{4}(\cos\phi - 1)\cos\left[(3+d)\pi \sin\phi\right]$$

— NOTES —

R = RATIO OF AREA OF SUBORDINATE LOOPS TO THAT OF MAIN LOOP
A = FRACTION OF WAVE LENGTH SPACING BETWEEN ELEMENTS IN SAME COLUMN
d = SEPARATION BETWEEN TWO SIMILAR ARRAYS

Fig. 13—Directional diagrams in horizontal plane of two arrays spaced laterally as noted. Case N = 6 and A = ½.

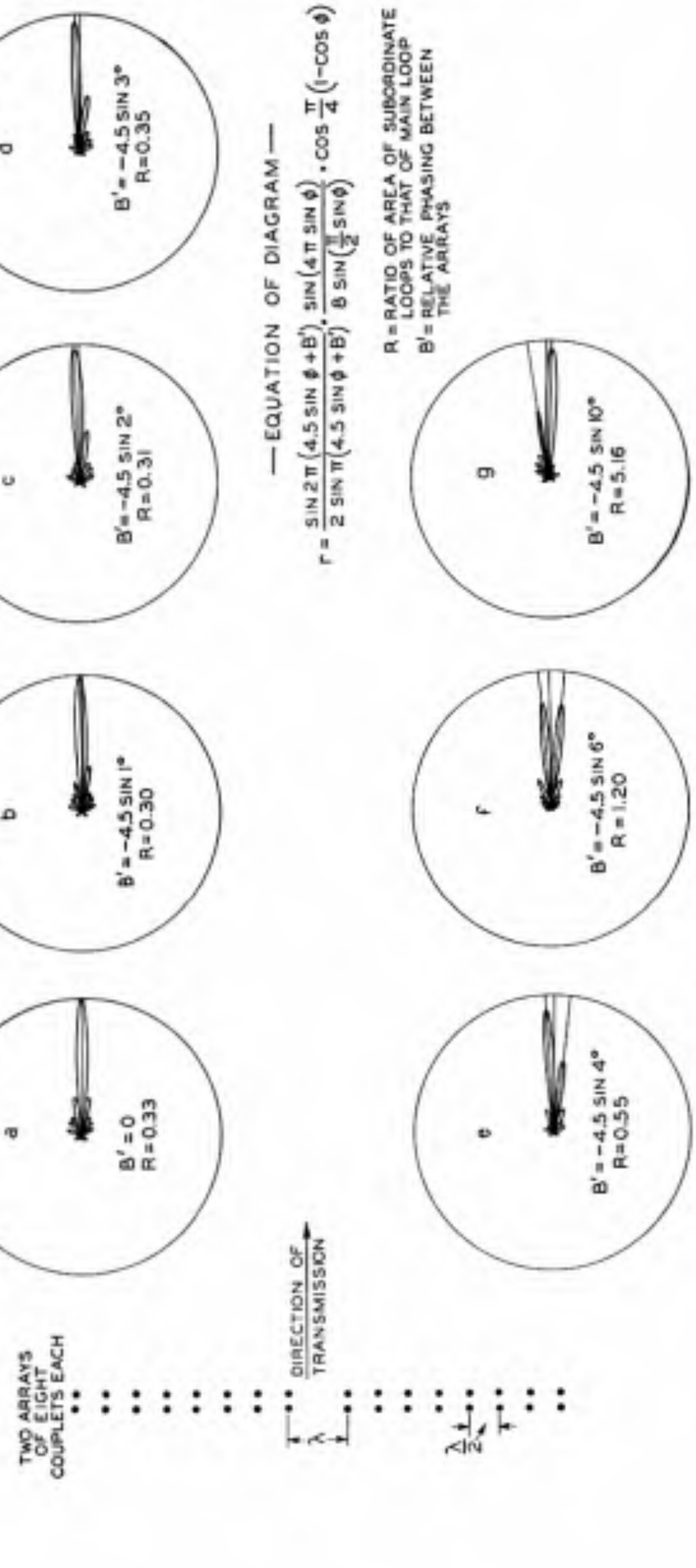


Fig. 14—Effect of phasing between two arrays.

If the array is divided into four sections the rotation may extend over a range of perhaps 9 deg., while for eight sections it may be 15 deg. The final case of 16 sections of one couplet each permits of considerable flexibility such as would be useful in operating with several distant stations in the same general direction. It should be pointed out, however, that the problem of making 16 phase adjustments each time a station wishes to change its direction of transmission is of considerable magnitude. For the particular case illustrated above it appears that the maximum rotation of the projected radiation is more or less proportional to the number of sections into which the array is divided. It may readily be seen from the two top rows of diagrams in Fig. 15 that continued addition of phasing amounts effectively to negative rotation. This may also be seen from an analysis of the equation of the diagram.

Fields of Linear Arrays

The successful use of an array of couplets to give unidirectivity suggests that the use of more than two parallel linear arrays might further be employed to advantage.[6] Obviously many such combinations are possible, but one of some interest has been investigated below. As a concrete example of this variation of gain with arrangement of arrays, a series of diagrams for 36 elements has been plotted in Fig. 16. The condition of spacing and phase intervals between columns of each of $\frac{1}{4}\lambda$ has been chosen. The horizontal characteristic is given for separations between rows of both $\frac{1}{2}$ and $\frac{1}{4}$ wavelength. The vertical characteristic common to these two separations is also shown. The equation of the diagram is given in formula (17) of the mathematical appendix below.

It will be observed from Fig. 16 that the horizontal directivity is for the most part only moderate, but approaches a maximum for the condition where a long broadside array prevails, whereas the vertical directivity is increased by increasing the number of columns in the field. A substantial loop will be found near the rear of diagrams corresponding to an odd number of columns. It is of further interest that, as far as horizontal directivity alone is concerned, the optimum may be derived either from a single array of 36 elements or from 18 couplets. Considerations of both minimum interference and total gain, however, make the latter preferable. These conclusions may also be reached by more direct analysis.[7]

[6] U. S. Patent 1,643,323, John Stone Stone, September 27, 1927.

[7] Wilmotte, "General considerations of the directivity of beam systems," *Jour. I. E. E.*, **66**, 955.

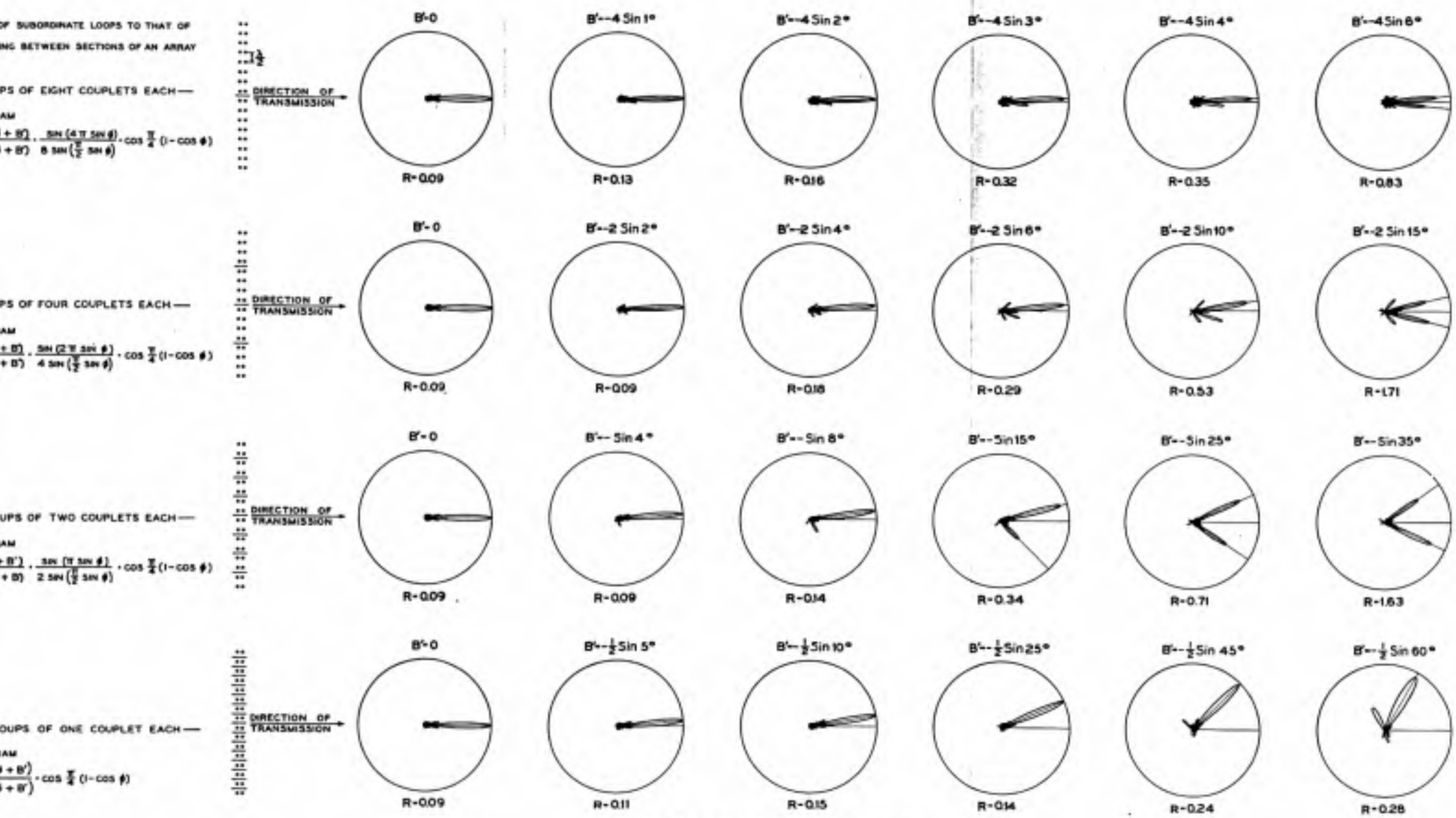


Fig. 15—Effect of phasing between sections of an array.

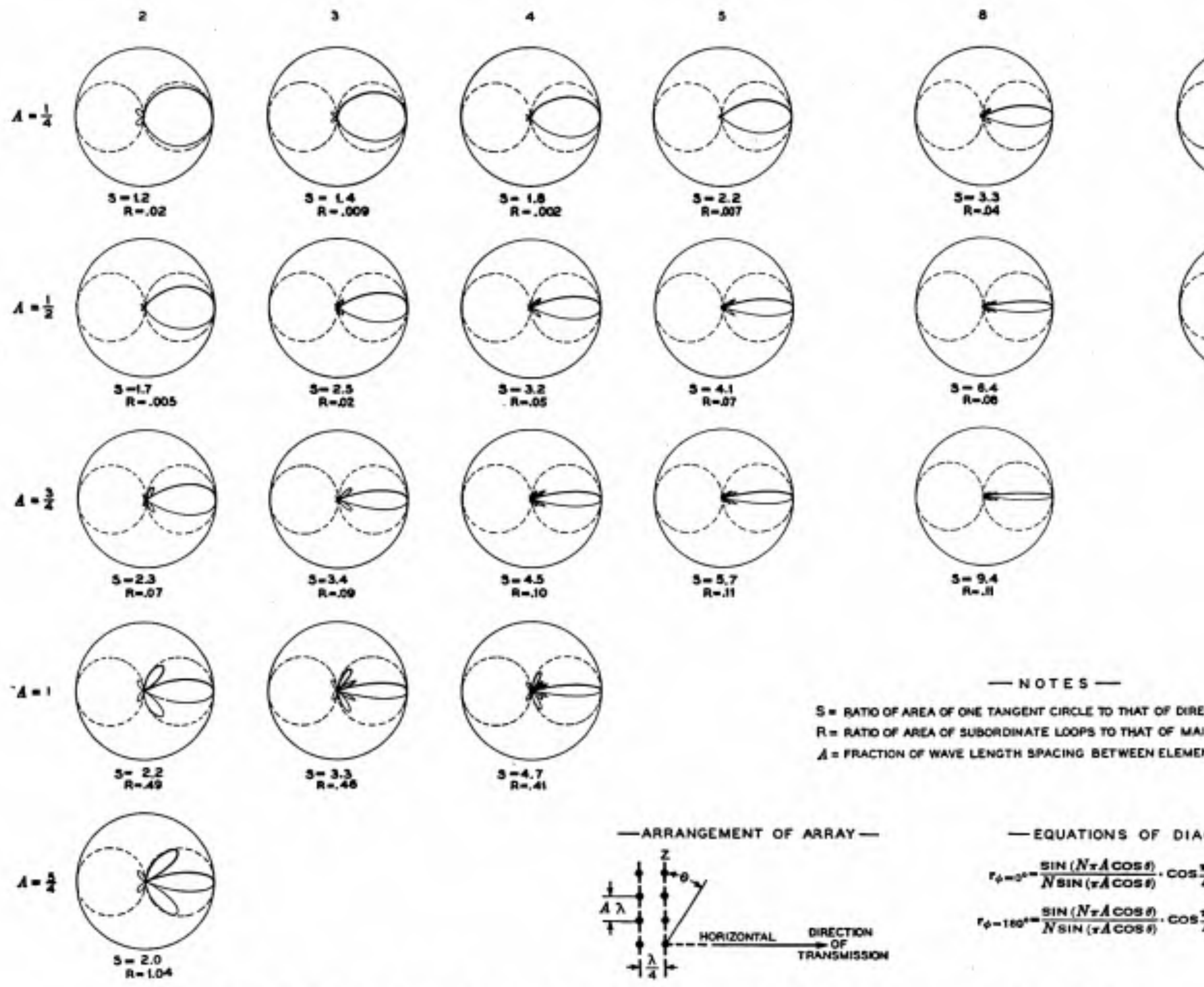
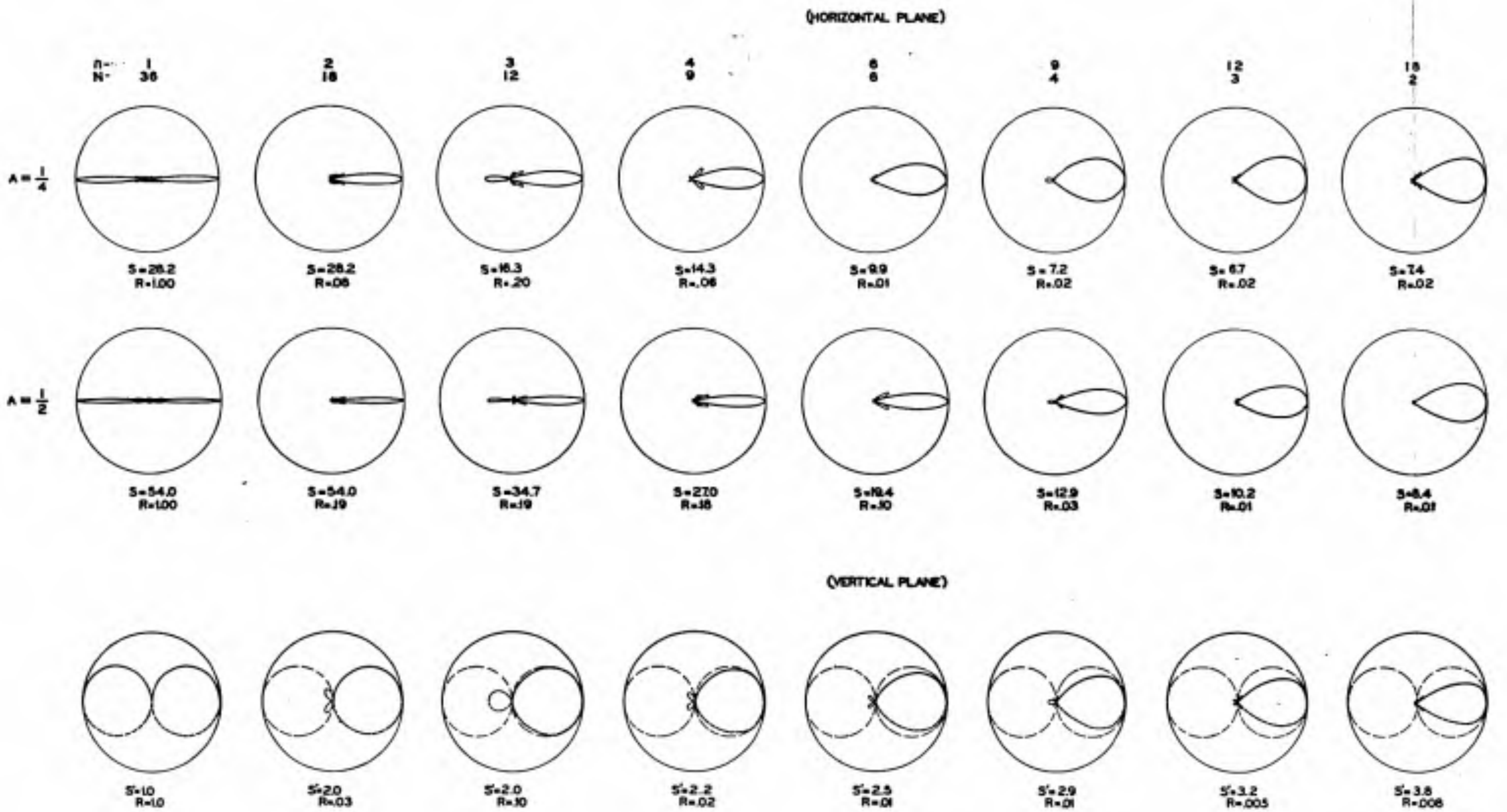
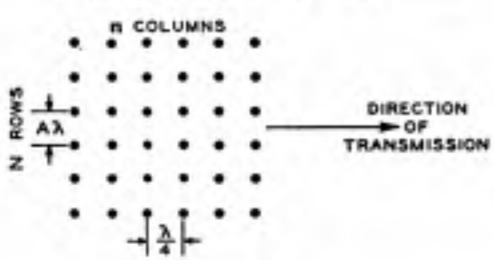


Fig. 17—Vertical plane diagrams due to couplets of coaxial antennas—number of couplets versus separation in wave-lengths.



— ARRANGEMENT OF ARRAY —



— EQUATIONS OF DIAGRAMS —

$$r_{\theta=0} = \frac{\sin(N\pi A \sin\phi)}{N \sin(\pi A \sin\phi)} \cdot \frac{\sin\left(n\frac{\pi}{4}[\cos\phi - 1]\right)}{n\sin\left(\frac{\pi}{4}[\cos\phi - 1]\right)}$$

$$r_{\phi=0} = \frac{\sin(N\pi A \sin\theta)}{N \sin(\pi A \sin\theta)} \cdot \frac{\sin\left(n\frac{\pi}{4}[\sin\theta - 1]\right)}{n\sin\left(\frac{\pi}{4}[\sin\theta - 1]\right)} \cdot \sin\theta$$

$$r_{\phi=180^\circ} = \frac{\sin(N\pi A \sin\theta)}{N \sin(\pi A \sin\theta)} \cdot \frac{\sin\left(n\frac{\pi}{4}[\sin\theta + 1]\right)}{n\sin\left(\frac{\pi}{4}[\sin\theta + 1]\right)} \cdot \sin\theta$$

— NOTES —

S = RATIO OF AREA OF UNIT CIRCLE TO THAT OF DIRECTION
S' = RATIO OF AREA OF TANGENT CIRCLES TO THAT OF DIRECT
R = RATIO OF AREA OF SUBORDINATE LOOPS TO THAT OF MA
A = FRACTION OF WAVE LENGTH SPACING BETWEEN ELEME

Fig. 16—Directional diagrams due to a field of thirty-six antennas.

Stacked Antennas

Thus far the discussion has centered mainly around directivity produced by placing vertical antennas in horizontal array. Added gain may be had also by incorporating directivity in a vertical plane.[8] This is frequently accomplished by arranging individual antennas one above another with their axes collinear, and is sometimes known as stacking. The fundamental principles of analysis are the same as those already utilized. However, an approximate correction must be allowed to account for the fact that the radiation from a linear oscillator increases from zero along the axis to a maximum in a plane perpendicular to the axis. The directional characteristic in planes passed through and parallel to such a radiator is approximated by two tangent circles.

Fig. 17 shows a series of directional diagrams indicating the results of stacking unidirectional couplets. The diagrams shown refer to the plane passed through the axes of the two linear oscillators comprising the couplet. On each diagram is a unit circle corresponding to a single point source. Inscribed are the two tangent circles, representing the vertical directional characteristic of a single linear source. Inside one of the tangent circles is the final directional diagram of the stacked array. The ratio of the area of the tangent circles to that of the characteristic diagram is given under each figure. This may be regarded as a rough measure of the relative gain. These diagrams are arranged horizontally in order of increasing number of couplets and vertically in order of separation. It frequently happens in practice that each radiator is approximately ½ wave-length long so it is convenient to utilize a vertical spacing interval also of ½ wave-length. Consequently the second row of diagrams is probably of greatest practical interest. In calculating these diagrams earth effects have been ignored.

In Figs. 18 and 19, the gain in decibels to be expected from stacking couplets has been plotted against number of couplets and fractional wave-length spacing. These values, like those for Figs. 7 and 8 above, were calculated by integrating the equation of diagram over a sphere of arbitrary radius. This was accomplished by use of equation (30) below. On account of the limited data at hand, Figs. 18 and 19 should be regarded only as a convenient method of illustrating the trend of the variables. These indicate that somewhat lower corresponding improvements result from stacking than from increasing the length of an array.

[8] U. S. Patent 1,683,739, John Stone Stone, September 11, 1928.

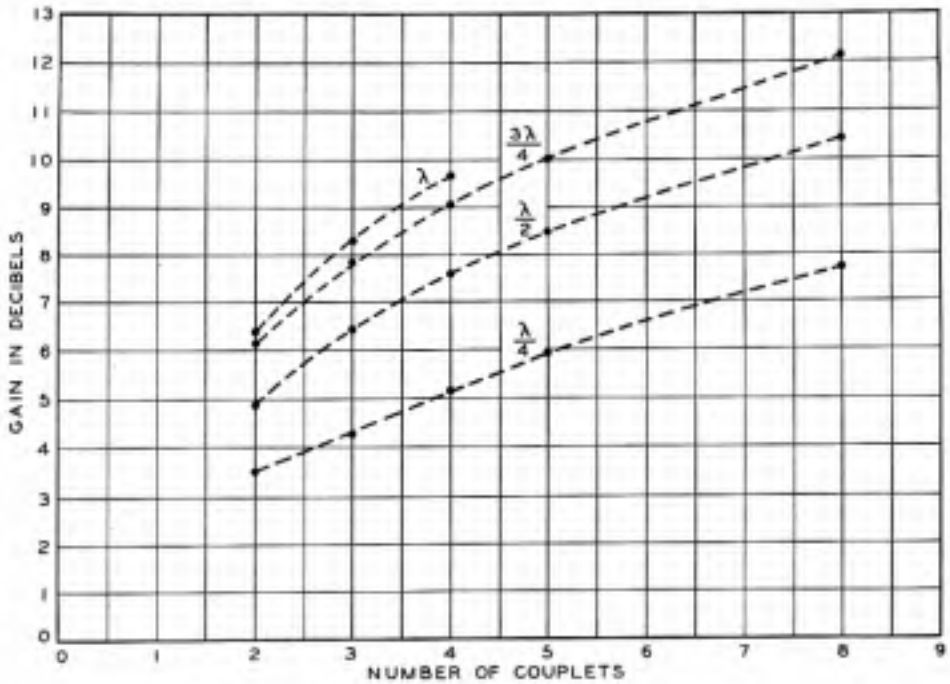


Fig. 18—Calculated gains from stacked antennas.

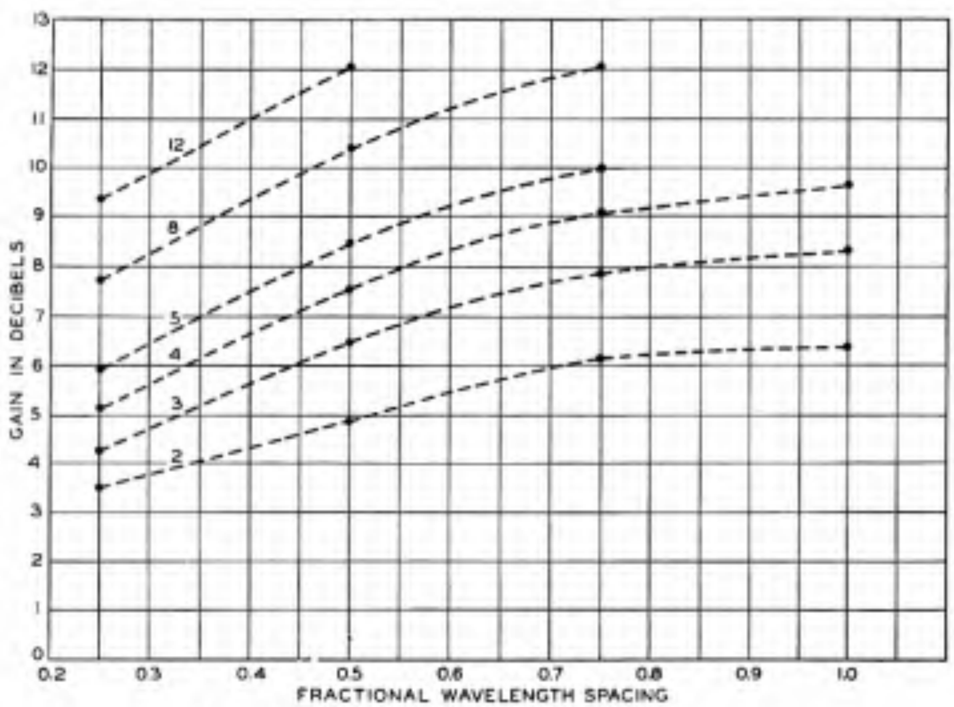


Fig. 19—Calculated gains from stacked antennas.

Arrays Incorporating Both Horizontal and Vertical Directivity

The gains of arrays combining both horizontal and vertical directivity may not be simply calculated by adding the gains (expressed in decibels) corresponding to elements arranged respectively along the two principal coordinate axes. However, they may be calculated except for earth effects by means of equation (26) below. Some calculations of this kind have been made and the data are tabulated below. They assume a total of 36 couplets which are arranged variously as noted. In the first case all 36 couplets are arranged as a simple horizontal array. The second case assumes that they are arranged in a

TABLE II

Number of Couplets Along Horizontal Axis	Number of Couplets Along Vertical Axis	Gain over Single Half-Wave Element Decibels
N	N	G
36	1	19.7
18	2	19.0
12	3	18.9
9	4	18.8
6	6	18.7
4	9	18.6
1	36	17.5

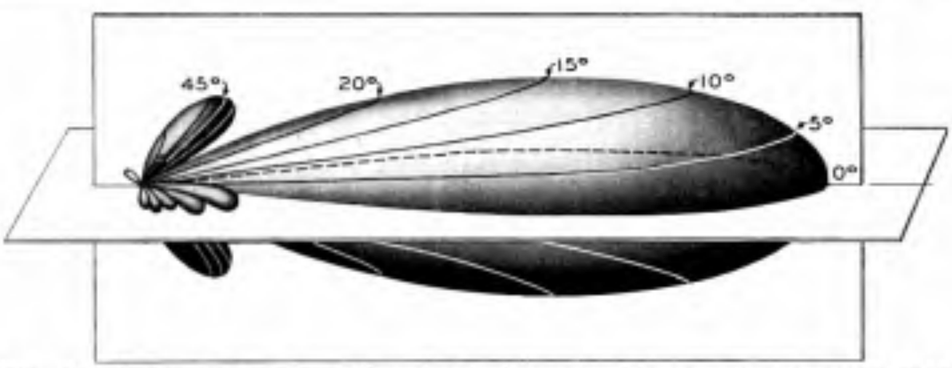


Fig. 20—Approximate three-dimensional diagram. Linear antenna array with reflector. Aperture two wave-lengths by eight wave-lengths.

broadside rectangle two elements high and 18 elements wide. This combination may be regarded as two arrays of 18 couplets arranged one above the other. The third case similarly assumes three arrays of 12 couplets each. A separation between couplets of ½ wave-length has been assumed throughout. The most economical arrangement of such an array depends not only on the relative costs of real estate and towers but also on feed-line losses and effects due to the proximity of the earth. The latter have specifically been omitted in this discussion.

Fig. 20 shows roughly the calculated directional characteristics of a typical stacked array incorporating both horizontal and vertical directivity. The planes passed through the diagram serve only as convenient references to assist in visualizing the horizontal and vertical diagrams. Earth effects of course, have been ignored.

Appendix

A general case of linear arrays which includes those used extensively in short-wave radio work, consists of a number of sources equispaced and equiphased along each of the three principal coordinate axes such that the space between sources is made up of rectangular parallelopipeds with the individual sources located at each corner. This may be regarded as $\boldsymbol{N}$ parallel planes each made up of N parallel columns where each column is made up of n individual radiating elements. The arrangement is made more evident by Fig. 21. The

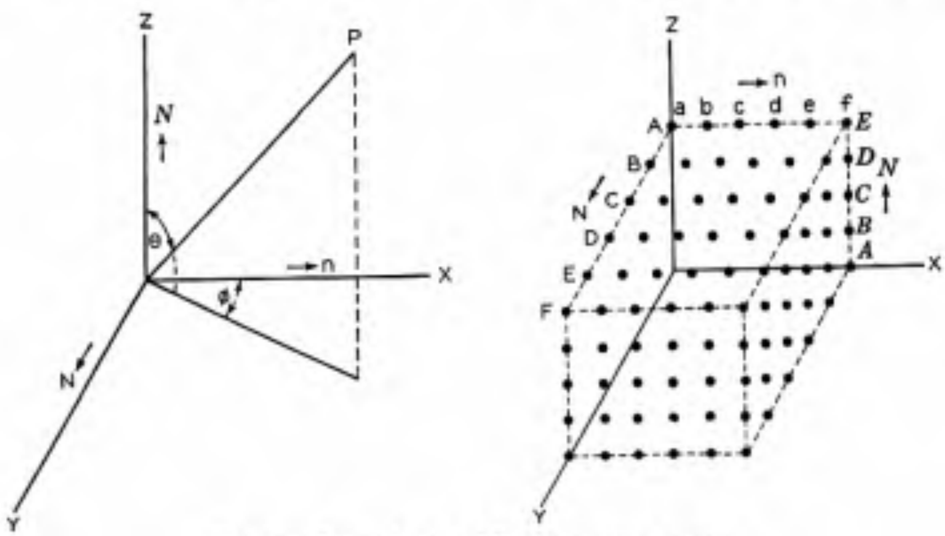


Fig. 21—General case of linear antenna arrays.

usual conventions for representing three-dimensional space have been adopted. We may designate the spacing between elements along the x, y, and z axes, respectively, by $a\lambda$, $A\lambda$, and $\boldsymbol{A}\lambda$ and their corresponding phase displacements between adjacent elements along the three principal axes by bT, BT and $\boldsymbol{B}T$.

The distance from any point in space to a particular radiator is

$$R_{nN\boldsymbol{N}} \doteq R - (\boldsymbol{N} - 1)\boldsymbol{A}\lambda \cos \theta - (N - 1)A\lambda \cos \phi \sin \theta - (n - 1)a\lambda \sin \theta \sin \phi. \tag{1}$$

Similarly the time phase of any particular element relative to the origin is

$$\delta_{nNN} = [(n-1)b + (N-1)B + (\boldsymbol{N}-1)\boldsymbol{B}]T. \tag{2}$$

The instantaneous value of the electric field at any remote point P due to one of these sources is given by

$$E_{n'} = A \cos \frac{2\pi}{\lambda}(ct - R_{n'}) + \delta_{n'} = A \cos \psi_{n'}, \tag{3}$$

where $n' = nN$.

The resultant interfering effect at a point P due to n' such sources all of equal amplitude is given by

$$\begin{aligned} E^2 = n'E_0^2 + 2E_0^2[&\cos(\psi_1 - \psi_2) + \cos(\psi_1 - \psi_3) + \cos(\psi_1 - \psi_4) + \cdots \text{ etc.} \\ &+ \cos(\psi_2 - \psi_3) + \cos(\psi_2 - \psi_4) + \cos(\psi_2 - \psi_5) + \cdots \text{ etc.} \\ &+ \cos(\psi_3 - \psi_4) + \cos(\psi_3 - \psi_5) + \cdots \text{ etc.} \\ &+ \cos(\psi_{n'-1} - \psi_{n'})]. \end{aligned} \tag{4}$$

The summation above gives rise to three series as follows:

$$\begin{aligned} S_x = {} & (n-1)\cos 2\pi(a \sin\theta \cdot \sin\phi + b) \\ & + (n-2)\cos 2\cdot 2\pi(a \sin\theta \cdot \sin\phi + b) \\ & + (n-3)\cos 3\cdot 2\pi(a \sin\theta \cdot \sin\phi + b) + \cdots \\ & + \cos(n-1)\cdot 2\pi(a \sin\theta \cdot \sin\phi + b), \end{aligned} \tag{5}$$

$$\begin{aligned} S_y = {} & (N-1)\cos 2\pi(A \sin\theta \cdot \cos\phi + B) \\ & + (N-2)\cos 2\cdot 2\pi(A \sin\theta \cdot \cos\phi + B) \\ & + (N-3)\cos 3\cdot 2\pi(A \sin\theta \cdot \cos\phi + B) + \cdots \\ & + \cos(N-1)\cdot 2\pi(A \sin\theta \cdot \cos\phi + B), \end{aligned} \tag{6}$$

$$\begin{aligned} S_z = {} & (\boldsymbol{N}-1)\cos 2\pi(\boldsymbol{A}\cos\theta + \boldsymbol{B}) \\ & + (\boldsymbol{N}-2)\cos 2\cdot 2\pi(\boldsymbol{A}\cos\theta + \boldsymbol{B}) \\ & + (\boldsymbol{N}-3)\cos 3\cdot 2\pi(\boldsymbol{A}\cos\theta + \boldsymbol{B}) + \cdots \\ & + \cos(\boldsymbol{N}-1)\cdot 2\pi(\boldsymbol{A}\cos\theta + \boldsymbol{B}), \end{aligned} \tag{7}$$

such that

$$E^2 = E_0^2(n + 2S_x)(N + 2S_y)(\boldsymbol{N} + 2S_z). \tag{8}$$

Each series is of the type

$$\begin{aligned} S = (n-1)\cos x + (n-2)\cos 2x \\ + (n-3)\cos 3x + \cdots + \cos(n-1)x' \end{aligned} \tag{9}$$

which is readily summed giving

$$n + 2S = \frac{(\cos nx - 1)}{(\cos x - 1)} = \frac{\sin^2 \frac{nx}{2}}{\sin^2 \frac{x}{2}}, \tag{10}$$

so

$$E = E_0 \frac{\sin n\pi(a \cos\phi \cdot \sin\theta + b)}{\sin \pi(a \cos\phi \cdot \sin\theta + b)} \cdot \frac{\sin N\pi(A \sin\phi \cdot \sin\theta + B)}{\sin \pi(A \sin\phi \cdot \sin\theta + B)} \cdot \frac{\sin \mathbf{N}\pi(\mathbf{A} \cos\theta + \mathbf{B})}{\sin \pi(\mathbf{A} \cos\theta + \mathbf{B})} \cdot \tag{11}$$

Reducing to common voltage level and including a term $\sin\theta$ to cover the case of radiation from linear oscillators we have for the equation of the directional diagram

$$r = \frac{\sin n\pi(a \cos\phi \cdot \sin\theta + b)}{n \sin \pi(a \cos\phi \cdot \sin\theta + b)} \cdot \frac{\sin N\pi(A \sin\phi \cdot \sin\theta + B)}{N \sin \pi(A \sin\phi \cdot \sin\theta + B)} \cdot \frac{\sin \mathbf{N}\pi(\mathbf{A} \cos\theta + \mathbf{B})}{\mathbf{N} \sin \pi(\mathbf{A} \cos\theta + \mathbf{B})} \cdot \sin\theta. \tag{12}$$

It will be recognized that this equation is made up of four factors. The first three account for the effects of the disposition of elements along the x, y, and z axes, respectively, while the fourth, of course, accounts for the direction of radiation from a linear oscillator. This is an equation giving magnitudes only. In plotting polar diagrams from this equation negative signs have no physical significance, and are plotted in a positive sense.

An examination of this equation shows that there are many possibilities which allow radiation in preferred directions, and at the same time limit it in others. Some of these are discussed below.

Special Cases

If we assume $n = 2$, $a = \frac{1}{4}$, $b = -\frac{1}{4}$ and $B = \mathbf{B} = 0$

$$r = \frac{\sin(N\pi A \sin\phi \cdot \sin\theta)}{N \sin(\pi A \sin\phi \cdot \sin\theta)} \cdot \frac{\sin(\mathbf{N}\pi\mathbf{A} \cos\theta)}{\mathbf{N} \sin(\pi\mathbf{A} \cos\theta)} \cdot \cos\frac{\pi}{4}(\cos\phi \cdot \sin\theta - 1) \cdot \sin\theta. \tag{13}$$

This corresponds to the practical case of transmission along the x axis from an antenna curtain and reflector made up of N vertical columns of $\mathbf{N}$ elements each.

The equation for the diagram in the (XY) plane may be had by placing $\theta = \pi/2$ giving

$$r = \frac{\sin (N\pi A \sin \phi)}{N \sin (\pi A \sin \phi)} \cos \frac{\pi}{4} (\cos \phi - 1), \tag{14}$$

which is the equation of the diagrams in Fig. 5 above. The corresponding equation for the principal vertical section may be had by placing $\phi = 0$ and $\phi = \pi$ giving

$$\left.\begin{aligned} r &= \frac{\sin (N\pi A \cos \theta)}{N \sin (\pi A \cos \theta)} \cos \frac{\pi}{4} (\sin \theta - 1) \sin \theta \\ \text{and} \quad r &= \frac{\sin (N\pi A \cos \theta)}{N \sin (\pi A \cos \theta)} \cos \frac{\pi}{4} (\sin \theta + 1) \sin \theta \end{aligned}\right\}, \tag{15}$$

which is the equation for the diagrams of Fig. 17.

The diagram of a single linear array of point sources is specified by the first term of equation (12) where $\theta = \pi/2$ or

$$r = \frac{\sin n\pi(a \cos \phi + b)}{n \sin \pi(a \cos \phi + b)} \cdot \tag{16}$$

The diagrams of Figs. 3 and 4 above may be calculated from equation (16) by placing $n = 2$ and $n = 16$, respectively. This also agrees with Foster's equation (1), page 367.[2]

The diagram of a field of coplanar linear arrays such as depicted in Fig. 16 above follows from equation (12) by placing $N = 1$, $a = \frac{1}{4}$, $b = -\frac{1}{4}$ and $B = 0$.

If the diagram is to be restricted to the (XY) plane, $\theta = \pi/2$ and

$$r = \frac{\sin (N\pi A \sin \phi)}{N \sin (\pi A \sin \phi)} \cdot \frac{\sin \left(n \frac{\pi}{4} (\cos \phi - 1) \right)}{n \sin \left(\frac{\pi}{4} (\cos \phi - 1) \right)} \cdot \tag{17}$$

Calculated Gains from Arrays

The flow of power through each unit area due to an advancing electric wave is given by the Poynting vector as

$$s = \frac{c}{4\pi} E \times H, \tag{18}$$

where E and H are vectors representing respectively, the electric and magnetic components of the advancing wave.

[2] Loc. cit.

For free space $|E| = |H|$ so

$$s = \frac{c}{4\pi} E^2. \tag{19}$$

Now the total power radiated through a sphere enclosing an array of sources is

$$P_1 = \int s d\sigma = \frac{c}{4\pi} \int_0^{\pi} \int_0^{2\pi} E_1^2 \sin \theta d\phi d\theta. \tag{20}$$

A second system would give

$$P_2 = \frac{c}{4\pi} \int_0^{\pi} \int_0^{2\pi} E_2^2 \sin \theta d\phi d\theta. \tag{21}$$

The radiated powers of these two systems might be so adjusted at the source as to give equal fields at any point along a preferred direction. A ratio of these powers, therefore, would be a convenient measure of the relative directional properties of the two arrays. This "test ratio" may conveniently be set up in terms of the equations of the diagrams derived above. In which case

$$T = \frac{\int_0^{\pi} \int_0^{2\pi} r_1^2 \sin \theta d\phi d\theta}{\int_0^{\pi} \int_0^{2\pi} r_2^2 \sin \theta d\phi d\theta} \cdot \tag{22}$$

If we assume all comparisons are to be made with respect to a single linear oscillator the denominator reduces to $8\pi/3$, so

$$T = \frac{3}{8\pi} \int_0^{\pi} \int_0^{2\pi} r_1^2 \sin \theta d\phi d\theta. \tag{23}$$

This ratio may conveniently be expressed in decibels. In which case $G = 10 \log_{10} 1/T$ is sometimes called the gain of an array.

If we are interested in the solid array shown in Fig. 21, where $n \cdot N \cdot \boldsymbol{N}$ linear oscillators, each having respective space and phase separations of $a\lambda$, bT; $A\lambda$, BT; and $\boldsymbol{A}\lambda$, $\boldsymbol{B}T$, are arranged progressively along the three principal coordinate axes, this becomes

$$\begin{aligned} T = \frac{3}{8\pi} \int_0^{\pi} \int_0^{2\pi} & \frac{\sin^2 [n\pi(a \cos \phi \sin \theta + b)]}{n^2 \sin^2 [\pi(a \cos \phi \sin \theta + b)]} \\ & \cdot \frac{\sin^2 [N\pi(A \sin \phi \sin \theta + B)]}{N^2 \sin^2 [\pi(A \sin \phi \sin \theta + B)]} \\ & \cdot \frac{\sin^2 [\boldsymbol{N}\pi(\boldsymbol{A} \cos \theta + \boldsymbol{B})]}{\boldsymbol{N}^2 \sin^2 [\pi(\boldsymbol{A} \cos \theta + \boldsymbol{B})]} \cdot \sin^3 \theta d\phi d\theta. \end{aligned} \tag{24}$$

This integration has been carried out by R. M. Foster who has very kindly placed the results at the writer's disposal. Only the final result is given herewith:

$$
\begin{aligned}
T = \frac{1}{nN\mathbf{N}} &+ \frac{3}{n^2N\mathbf{N}}\sum_{k=1}^{n-1}(n-k)\cdot\cos(2\pi kb)\cdot Q(2\pi ka,\,0) \\
&+ \frac{3}{nN^2\mathbf{N}}\sum_{K=1}^{N-1}(N-K)\cdot\cos(2\pi KB)\cdot Q(2\pi KA,\,0) \\
&+ \frac{3}{nN\mathbf{N}^2}\sum_{\mathbf{K}=1}^{\mathbf{N}-1}(\mathbf{N}-\mathbf{K})\cdot\cos(2\pi\mathbf{KB})\cdot Q(0,\,2\pi\mathbf{KA}) \\
&+ \frac{6}{n^2N^2\mathbf{N}}\sum_{k=1}^{n-1}\sum_{K=1}^{N-1}(n-k)(N-K)\cdot\cos(2\pi KB) \\
&\qquad\qquad \cdot\cos(2\pi kb)\cdot Q(2\pi\sqrt{k^2a^2+K^2A^2},\,0) \\
&+ \frac{6}{nN^2\mathbf{N}^2}\sum_{K=1}^{N-1}\sum_{\mathbf{K}=1}^{\mathbf{N}-1}(N-K)(\mathbf{N}-\mathbf{K})\cdot\cos(2\pi KB) \\
&\qquad\qquad \cdot\cos(2\pi\mathbf{KB})\cdot Q(2\pi KA,\,2\pi\mathbf{KA}) \\
&+ \frac{6}{n^2N\mathbf{N}^2}\sum_{k=1}^{n-1}\sum_{\mathbf{K}=1}^{\mathbf{N}-1}(n-k)(\mathbf{N}-\mathbf{K})\cdot\cos(2\pi kb) \\
&\qquad\qquad \cdot\cos(2\pi\mathbf{KB})\cdot Q(2\pi ka,\,2\pi\mathbf{KA}) \\
&+ \frac{12}{n^2N^2\mathbf{N}^2}\sum_{k=1}^{n-1}\sum_{K=1}^{N-1}\sum_{\mathbf{K}=1}^{\mathbf{N}-1}(n-k)(N-K)(\mathbf{N}-\mathbf{K}) \\
&\qquad\qquad \cdot\cos(2\pi kb)\cdot\cos(2\pi KB)\cdot\cos(2\pi\mathbf{KB}) \\
&\qquad\qquad \cdot Q(2\pi\sqrt{k^2a^2+K^2A^2},\,2\pi\mathbf{KA}). \qquad (25)
\end{aligned}
$$

Where the function

$$
\begin{aligned}
Q(x,y) = \frac{x^2}{(x^2+y^2)^{3/2}}\sin(\sqrt{x^2+y^2}) &+ \frac{x^2-2y^2}{(x^2+y^2)^2}\cos(\sqrt{x^2+y^2}) \\
&- \frac{x^2-2y^2}{(x^2+y^2)^{5/2}}\sin(\sqrt{x^2+y^2}). \qquad (25a)
\end{aligned}
$$

In particular

$$Q(x,0) = \frac{\sin x}{x} + \frac{\cos x}{x^2} - \frac{\sin x}{x^3} \qquad (25b)$$

and

$$Q(0,x) = -\frac{2\cos x}{x^2} + \frac{2\sin x}{x^3}\,. \qquad (25c)$$

Special Cases

(1) If we assume $n = 2$, $a = \frac{1}{4}$, $b = -\frac{1}{4}$ and $B = \mathbf{B} = 0$, the test ratio is given by

$$T_1 = \frac{1}{2N\mathbf{N}} + \frac{3}{2N^2\mathbf{N}} \sum_{K=}^{N-1} (N - K)\cdot Q(2\pi KA,\, 0)$$
$$+ \frac{3}{2N\mathbf{N}^2} \sum_{\mathbf{K}=1}^{\mathbf{N}-1} (\mathbf{N} - \mathbf{K})\cdot Q(0,\, 2\pi \mathbf{K}\mathbf{A})$$
$$+ \frac{3}{N^2\mathbf{N}^2} \sum_{K=1}^{N-1} \sum_{\mathbf{K}=1}^{\mathbf{N}-1} N - K)(\mathbf{N} - \mathbf{K})\cdot Q(2\pi KA,\, 2\pi \mathbf{K}\mathbf{A}). \quad (26)$$

This, like equation (13), corresponds to the practical case of transmission from an antenna curtain and reflector each made up of N vertical columns of $\mathbf{N}$ elements, all driven in the same phase.

(2) If we assume that no stacking is involved, then $\mathbf{N} = 1$ and we have for the test ratio for N couplets

$$T_2 = \frac{1}{2N} + \frac{3}{2N^2} \sum_{K=1}^{N-1} (N - K)\cdot Q(2\pi KA,\, 0)$$
$$= \frac{1}{2N} + \frac{3}{2N^2} \sum_{K=1}^{N-1} (N - K)\cdot \left[\frac{\sin 2\pi KA}{2\pi KA} + \frac{\cos 2\pi KA}{(2\pi KA)^2} - \frac{\sin 2\pi KA}{(2\pi KA)^3} \right]. \quad (27)$$

This equation was used in the calculation of the data given in Figs. 6, 7, and 8.

(3) If we wish to apply equation (25) to the case of a single array of N linear oscillators driven in phase we have $n = \mathbf{N} = 1$ and $B = 0$, so

$$T_3 = \frac{1}{N} + \frac{3}{N^2} \sum_{K=1}^{N-1} (N - K)\cdot Q(2\pi KA,\, 0), \quad (28)$$

which differs from equation (27) by a factor of two. This indicates that an array of N equiphased linear couplets gives twice the field in the preferred direction as received from N equiphased linear elements radiating the same power.

(4) Applying equation (25) to the extremely simple case of one couplet, $n = 2$, $a = \frac{1}{4}$, $b = -\frac{1}{4}$ and $N = \mathbf{N} = 1$ and

$$T_4 = \tfrac{1}{2}. \quad (29)$$

(5) We may calculate the test ratio for a single stack of linear couplets (earth effects not considered) by placing $N = 1$, $n = 2$, $a = \frac{1}{4}$, $b = -\frac{1}{4}$, and $\mathbf{B} = 0$ and get

$$T_5 = \frac{1}{2\mathbf{N}} + \frac{3}{2\mathbf{N}^2} \sum_{\mathbf{K}=1}^{\mathbf{N}-1} (\mathbf{N} - \mathbf{K})\cdot Q(0,\, 2\pi \mathbf{K}\mathbf{A})$$
$$= \frac{1}{2\mathbf{N}} - \frac{3}{\mathbf{N}^2} \sum_{\mathbf{K}=1}^{\mathbf{N}-1} (\mathbf{N} - \mathbf{K}) \left[\frac{\cos (2\pi \mathbf{K}\mathbf{A})}{(2\pi \mathbf{K}\mathbf{A})^2} - \frac{\sin (2\pi \mathbf{K}\mathbf{A})}{(2\pi \mathbf{K}\mathbf{A})^3} \right]. \quad (30)$$

This equation was used in calculating the data given in Figs. 18 and 19.

(6) The test ratio for the case of the rectangular array of nN elements discussed in connection with Fig. 16 may be calculated by placing $N = 1$, $a = \frac{1}{4}$, $b = -\frac{1}{4}$ and $B = 0$. In which case

$$T_6 = \frac{1}{nN} + \frac{3}{nN^2}\sum_{K=1}^{N-1}(N-K)\cdot Q(2\pi KA, 0) + \frac{3}{Nn^2}\sum_{k=1}^{n-1}(n-k)\cdot\cos\frac{(k\pi)}{2}\cdot Q\left(\frac{k\pi}{2}, 0\right) + \frac{6}{n^2N^2}\sum_{K=1}^{N-1}\sum_{k=1}^{n-1}(n-k)(N-K)\cdot\cos\left(\frac{k\pi}{2}\right) \cdot Q\left(2\pi\sqrt{\frac{k^2}{16}+K^2A^2}, 0\right). \quad (31)$$

Areas of Directional Diagrams

In general, the areas of directional diagrams may be calculated from their equations by the usual integration methods. The special case of N couplets in horizontal array, such as used rather generally in practice and shown in Fig. 5 above, is of sufficient importance to be given here. The area of the diagram in the (XY) plane is

$$S = \frac{1}{N^2}\left[\frac{N}{2} + \sum_{K=1}^{N-1}(N-K)\cdot J_0(2\pi KA)\cdot\cos 2\pi KB\right]. \quad (32)$$

This equation was used in calculating the data given in Fig. 5.

The area of diagrams in the horizontal plane due to a single array of N oscillators is given by the equation:

$$S = \frac{2}{N^2}\left[\frac{N}{2} + \sum_{K=1}^{N-1}(N-K)\cdot J_0(2\pi KA)\cdot\cos 2\pi KB\right].^* \quad (33)$$

This differs from equation (32) by a factor of two and indicates that regardless of whether the gain is reckoned by an integration over a unit sphere or in terms of the area of the horizontal diagram the effect of the reflector is to double the radiated field in the preferred direction.

Placing $N = 1$ in equation (32)

$$S = \tfrac{1}{2}. \quad (34)$$

This is analogous to equation (29) above.

* R. M. Foster, "Directive diagrams of antenna arrays," *Bell Sys. Tech. Jour.*, **5**, 307; 1926.

Arrays of Arrays

Each element of a generalized linear array, such as shown in Fig. 21, may be replaced by a generalized array, thereby producing an array of arrays.[9] It may be shown that the resultant is given by an array factor, representing the characteristics of individual arrays, times other factors representing the relative position of the individual arrays in the array of arrays. A derivation analogous to that beginning on page 22 results in the equation

$$R = r \cdot \frac{\sin n'\pi(a' \sin \phi + b')}{n' \sin \pi(a' \sin \phi + b')} \cdot \frac{\sin N'\pi(A' \sin \phi + B')}{N' \sin \pi(A' \sin \phi + B')} \cdot \frac{\sin \boldsymbol{N}'\pi(\boldsymbol{A}' \sin \phi + \boldsymbol{B}')}{\boldsymbol{N}' \sin \pi(\boldsymbol{A}' \sin \phi + \boldsymbol{B}')}, \quad (35)$$

where $a'\lambda$, $A'\lambda$ and $\boldsymbol{A}'\lambda$ are the coordinate spacings between arrays and $b'T$, $B'T$, and $\boldsymbol{B}'T$ are the corresponding phase intervals, and r represents the characteristics of one of the individual arrays. If each array is of the type shown in Fig. 5, r is given by equation (14) above. Placing $n' = \boldsymbol{N}' = 1$ and $N' = 2$ also $n = 2$ and $B = 0$, the above equation reduces to

$$R = \frac{\sin N'\pi(A' \sin \phi + B')}{N' \sin \pi(A' \sin \phi + B')} \cdot \frac{\sin N\pi(A \sin \phi)}{N \sin \pi(A \sin \phi)} \cos \frac{\pi}{4}(1 - \cos \phi), \quad (36)$$

which is that made use of in calculating the diagrams in Figs. 14 and 15.

Bibliography

Sources with extensive bibliographies:

Walter, L. H., "Directive wireless telegraphy," 119–121, 1921.

Beverage, H. H., Rice, C. W., and Kellogg, E. W., "The wave antenna," *Trans. A. I. E. E.*, **42**, 215–266; February, 1923.

Zenneck, J. and Rukop, H., "Drahtlose Telegraphie," 486–508, 1925.

Smith-Rose, R. L., "A study of radio direction finding," Radio Research Board Special Rept. No. 5, 1927.

Keen, R., "Wireless direction finding and directional reception," 451–467, 1927 (2d Edition).

Smith-Rose, R. L., "Radio direction finding by transmission and reception," *Proc. I. R. E.* **17**, 425–478; March, 1929.

1926

Bellini, E., "La possibilité de la télégraphie sans fil dirigée à grande concentration," *L'Onde Élect.*, **5**, 475–483; September, 1926.

Bontsch-Bruewitsch von M. A., "Die Strahlung der Komplizierten Rechtwinkeligen Antennen mit Gleichbeschaffenen Vibratoren," *Ann. d. Physik*, **81**, 425–453; October 18, 1926.

Catterson-Smith, J., "The characteristics of beam transmitting aerials," *Jour. Indian Inst. Sci.*, 9B Part 2, 9–19, 1926.

Chireix, H., "Transmission en ondes courtes," *L'Onde Élect.*, **5**, 237–262; June, 1926.

[9] Bailey, Dean, and Wintringham, *Proc. I. R. E.*, **16**, 1694; December, 1928.

Esau, A., "Richtcharakteristiken von Antennenkombinationen," *Zeits. f. Hochf.*, **27**, 142–150; May, 1926; **28**, 1–12; July, 1926; **28**, 147–156; December, 1926.
Foster, R. M., "Directive diagrams of antenna arrays," *Bell Sys. Tech. Jour.*, **5**, 292–307; April, 1926.
Meissner, A., "Über Raumstrahlung," *Zeits. f. Hochf.*, **28**, 78–82; September, 1926.
Murphy, W. H., "Space characteristics of antennae," *Jour. Franklin Inst.*, 201, 411–429; April, 1926.
Tatarinoff, W., "Zur Konstruktion der Radiospiegel," *Zeits. f. Hochf.*, **28**, 117–120; October, 1926.
Uda, S., "On the wireless beam of short electric waves," *Jour. I. E. E.* (Japan), No. 452, Part I, 273–282; March, 1926; No. 453, Part II, 335–351; April, 1926; No. 456, Part III, 712–724; July, 1926.
Yagi, H. and Uda, S., "Projector of sharpest beam of electric waves," *Proc. Imp. Acad.* (Tokio), **2**, 49–52; February, 1926.
—— "Imperial wireless communication," *Electrician*, **96**, 62–63; January 15, 1926.
—— "Imperial wireless 'beam' communication," *El. Rev.*, **99**, 709–712; October 29, 1926; **99**, 749–751; November 5, 1926.

1927

Blondel, A., "Électricité—Sur les procédés de répérage d'alignement par les ondes hertziennes et sur les radiophares d'alignement," *Comptes Rendus*, **184**, 561–565; March 7, 1927.
Blondel, A., "Électricité—Remarque au sujet des émissions hertziennes dirigées," *Comptes Rendus*, **184**, 923–925; April 11, 1927.
Bouthillon, L., "Inclinaison des ondes et systèmes dirigés," *Comptes Rendus*, **184**, 190–192; January 24, 1927.
Chireix, H., "Nouvelle antenne directive simple pour l'onde courte," *Q. S. T. Français*, **8**, 43–46; April, 1927.
Eckersley, T. L., English patent No. 305,733, "Improvements in or relating to aerial systems for wireless signaling." Application date, November 18, 1927.
Esau, A., "Vergrösserung des Empfangsbereiches bei Doppelrahmen und Doppelcardioidenanordnungen durch Goniometer," *Zeits. f. Hochf.*, **30**, 141–151; November, 1927.
Fleming, J. A., "Approximate theory of the flat projector (Franklin) aerial used in the Marconi beam system of wireless telegraphy," *Exp. Wireless*, **4**, 387–392; July, 1927.
Green, E., "Calculation of the polar curves of extended aerial systems," *Exp. Wireless*, **4**, 587–594; October, 1927.
Hémardinquer, P., "Transmissions radioélectriques par ondes dirigées," *Nature* (Paris), **55**, no. 2760, 407–413; May 1, 1927.
Lee, A. G., "Atmospherics and transatlantic telephony—A new directional polar curve," *Exp. Wireless*, **4**, 757–759; December, 1927.
Meissner, A., "Richtstrahlung mit horizontalen Antennen," *Zeits. f. Hochf.*, **30**, 77–79; September, 1927. "Directional radiation with horizontal antennas," *Proc. I. R. E.*, **15**, 928–934; November, 1927.
Meissner, A., "Raumstrahlung von Horizontal—Antennen," *E. N. T.*, **4**, 482–486; November, 1927.
Mesny, R., "Electromagnetic radiation," *Tijds. Nederland. Radiogenootschap.*, **3**, 49–66; February, 1927.
Mesny, R., "Émissions dirigées par rideaux d'antennes, antennes en grecque," *L'Onde Élect.*, **6**, 181–199; May, 1927.
Murphy, W. H., "Space characteristics of antennae," *Jour. Franklin Inst.*, **203**, 289–311; February, 1927.
Plendl, H., "Berechnung von Richstrahl—Antennen," *Zeits. f. Hochf.*, **30**, 80–82; September, 1927.
Standard Telephones and Cables Ltd., English patent No. 307,446, "Improvements in aerial systems." Application date, December 7, 1927.
Stone, J. S., U. S. patent No. 1,643,323, "Directive antenna array," September 27, 1927.
Uda, S., "Wireless beam of short electric waves," *Jour. I. E. E.* (Japan), No. 462, Part IV, 26–51; January, 1927; No. 462, Part V, 52–62; January, 1927; No. 465, Part VI, 396–403; April, 1927; No. 467, Part VII, 623–634; June,

1927; No. 470, Part VIII, 1092–1100; September, 1927; No. 472, Part IX, 1209–1219; November, 1927. Written in Japanese with English abstract.

Uda, S., "High-angle radiation of short electric waves," Tohoku Univ. Technol. Reports, **7**, 25–32; 1927; *Proc. I. R. E.*, **15**, 377–385; May, 1927.

—— "Short-wave beam transmission—Equipment of the Marconi stations at Grimsby and Skegness," *Electrician*, **98**, 319–320; March 25, 1927; **98**, 378–379; April 8, 1927.

1928

d'Ailly, G. H., "Théorie du rayonnement de la beam antenne," *Q. S. T. Français*, **9**, 14–19; June, 1928; **9**, 36–39; July, 1928.

Bailey, Austin, Dean, S. W., and Wintringham, W. T., "The receiving system for long-wave transatlantic radio telephony," *Proc. I. R. E.*, **16**, 1645–1705; December, 1928.

Böhm, O., "Die Bündelung der Energie kurzer Wellen," *E. N. T.*, **5**, 413–421; November, 1928.

Bouthillon, L., "La direction des ondes radioélectriques; Idées et réalisations récentes," *Bull. de la Soc. Franç. des Élect.*, **8**, 657–679; July, 1928.

Bouthillon, L., "La direction des ondes radioélectriques," *Le Génie Civil*, **92**, 623; June 23, 1928.

Burnett, D., "Directional properties of wireless receiving aerials," *Proc. Cambridge Phil. Soc.*, **24**, 521–530; October, 1928.

Chireix, H., "Un système français d'émission à ondes courtes projetées," *L'Onde Élect.*, **7**, 169–195; May, 1928.

Chireix, H., "Liaisons radiotéléphoniques à grande distance par ondes courtes projetées," *Bull. de la Soc. Franç. des Élect.*, **8**, 680–691; July, 1928.

Clapp, J. K. and Chinn, H. A., "Directional properties of transmitting and receiving aerials," *Q. S. T.*, **12**, 17–30; March, 1928.

Dieckmann, Max, "Strahlungsdichte und Empfangsfläche," *Zeits. f. Hochf.*, **31**, 8–15; January, 1928.

Franklin, C. S., English patent No. 311,449, "Improvements in or relating to aerial systems." Application date, February 11, 1928.

Franklin, C. S., English patent No. 310,451, "Improvements in or relating to wireless telegraphy and telephony and aerial systems therefor." Application date, January 26, 1928.

Galetti, R. C., German patent No. 460,270, "Reflektor für elektromagnetische Wellen," May 29, 1928.

Gothe, A., "Über Drahtreflektoren," *E. N. T.* **5**, 427–430; November, 1928.

Gresky, G., "Die Wirkungsweise von Reflektoren bei kurzen elektrischen Wellen," *Zeits. f. Hochf.*, **32**, 149–162; November, 1928.

Kato, Y., "Directivity of the saw-tooth antenna," *Jour. I. E. E.* (Japan), No. 480, 706–711; July, 1928.

Koomans, N., English patent No. 298,131, "Improvements in or relating to directive aerials." Application date, September 29, 1928.

Marconi, G., "Radio communication," *Proc. I. R. E.*, **16**, 40–69; January, 1928.

Noël, Robert, "La radiotéléphonie par ondes courtes projetées—Les premières communications entre Paris et Alger," *Le Génie Civil*, **92**, 373–379; April 21, 1928.

Pistolkors, A., "On the calculation of the radiation of directional antennae and on the radiation of a simple antenna in the presence of a reflecting wire," *Teleg. i. Telef. b. Prov.*, **10**, 540; October, 1928.

Radio Corporation of America, French patent No. 648,548, "Perfectionnements aux systèmes pour la réception d'énergie radiante," August 14, 1928.

Radio Corporation of America, French patent No. 655,778, "Perfectionnements aux systèmes pour la transmission d'énergie radiante," December 2, 1928.

Standard Telephones and Cables Ltd., English patent No. 319,055, "Improvements in aerial systems." Application date, June 15, 1928.

Stone, J. S., U. S. patent No. 1,683,739, "Directive antenna array," September 11, 1928.

Turlyghin, S. I., "Transmitting aerials for beam stations," *Vestnik Elektrotech* (Moscow), p. 69, February, 1928.

Uda, S., "On the wireless beam of short electric waves: High-angle radiation of horizontally polarized waves," *Jour. I. E. E.* (Japan), No. 477, 395–405; April, 1928.

Uda, S., "On the wireless beam of short electric waves," *Jour. I. E. E.* (Japan) (Reprint No. 20), July, 1928.
Walmsley, T., "Polar diagrams due to plane aerial reflector systems," *Exp. Wireless*, **5**, 575–577; October, 1928.
Wells, N., "Beam wireless telegraphy," *El. Rev.*, 102, Part I, 898–902; May 25, 1928; 102, Part II, 940–943; June 1, 1928.
Wilmotte, R. M., "General considerations of the directivity of beam systems," *Jour. I. E. E.*, **66**, 955–961; September, 1928.
Wilmotte, R. M., "The nature of the field in the neighborhood of an antenna," *Jour. I. E. E.*, **66**, 961–967; September, 1928.
Wilmotte, R. M. and McPetrie, J. S., "A theoretical investigation of the phase relations in beam systems," *Jour. I. E. E.*, **66**, 949–954; September, 1928.
Yagi, H., "Beam transmission of ultra-short waves," *Proc. I. R. E.*, **16**, 715–741; June, 1928.
—— "Émissions radiotélégraphiques dirigées," *L'Industrie Élect.*, **37**, 341–346; August 10, 1928; **37**, 372–376; August 25, 1928.

1929

Beauvais, G., "Les ondes électriques très courtes (15 à 20 centimètres)," *Rev. Gén. de l'Élect.*, **25**, 393–394; March 16, 1929.
Bechmann, von R., "Berechnung der Strahlungsdiagramme von Antennenkombinationen," *Telefunken Zeit.*, No. 53, 54–60; December, 1929.
Campbell, G. A., U. S. patent No. 1,738,522, "Electromagnetic wave signaling system," December 10, 1929.
Chireix, H., "French beam system for short waves," *Bull. de la Soc. Franç. Radiotélégraphique*, **3**, 79; May, 1929.
Chireix, H., "French system of directional aerials for transmission on short waves," *Exp. Wireless*, **6**, 235–244; May, 1929.
Gresky, G., "Richtcharakteristiken von Antennenkombinationen deren einzelne Elemente in Oberschwingungen erregt werden," *Zeits. f. Hochf.*, **34**, 132–140; October, 1929; **34**, 178–183; November, 1929.
Hahnemann, W., German patent No. 474,123, "Einrichtung zum gerichteten Senden und Empfangen mittels elektrischer Wellen," March 27, 1929.
Koomans, N., French patent No. 660,639, "Antenne directive," February 19, 1929.
Mathieu, G. A., "The Marconi-Mathieu method of multiplex-signaling," *Marconi Rev.*, **7**, 1; April, 1929.
Mesny, R., "Les ondes dirigées et leurs applications," *Revue Scientifique*, No. 19, 577–585; October 12, 1929.
Moser, W., "Versuche über Richtantennen bei kurzen Wellen," *Zeits. f. Hochf.*, **34**, 19–26; July, 1929.
Ostroumov, G. A., "A directional untuned short-wave receiving antenna," *Teleg. i. Telef. b. Prov.*, **10**, 111, 1929.
Palmer, L. S. and Honeyball, L. L. K., "The action of a reflecting antenna," *Jour. I. E. E.*, **67**, 1045–1051; August, 1929.
Pistolkors, A., "Calculation of radiation resistance of antennae composed of perpendicular oscillators," *Teleg. i. Telef. b. Prov.*, **10**, 33, 1929.
Pistolkors, A., "Radiation resistance of beam antennas," *Proc. I. R. E.*, **17**, 562–579; March, 1929.
Sammer, von F., "Die Wirkungsweise von Drahtreflektoren," *Telefunken Zeit.*, No. 53, 61–71; December, 1929.
Stenzel, H., Über die Richtcharakteristik von in einer Ebene angeordneten Strahlern," *E. N. T.*, **6**, 165–181; May, 1929.
Strutt, M. J. O., "Strahlung von Antennen unter dem Einfluss der Erbodeneigenschaften," *Ann. d. Physik*, Series 5, **1**, 721–750 and 751–772; April 6, 1929; Series 5, **4**, 1–16; January 18, 1930.
Villem, M. R., "La liaison radiotéléphonique Paris—Buenos Aires par ondes courtes projetées," *Bull. de la Soc. Franç. des Élect.*, **9**, No. 98, 1107–1145; October, 1929.
Yagi, H., German patent No. 475,293, "Einrichtung zum Richtsenden oder Richtempfangen," April 25, 1929.

Absolute Calibration of Condenser Transmitters

By L. J. SIVIAN

Several methods have been used or proposed for the calibration of the Wente condenser transmitter. The methods falling under the two classifications conveniently designated "constant pressure" or "pressure" calibration and "constant field" or "field" calibration are most useful and amenable to measurement. Which of these two calibrations is more significant depends on the particular use made of the transmitter. In the following pages the methods now used or proposed are reviewed and the advantages or disadvantages of each from the standpoint of transmitter application are discussed.

IN the original design of the Wente[1] transmitter the effective diaphragm resonance was well above 10,000 c.p.s. The new design (Western Electric No. 394-Type), developed by Wente, has an effective resonance at approximately 5,000 c.p.s. It is about ten times more sensitive (on a voltage-pressure basis), and more immune from effects of humidity and of barometric changes. The important external dimensions of the instrument are shown in Fig. 1*A*.

The response of the transmitter is defined as the ratio of the electromotive force generated to the acoustic pressure acting on the trans-

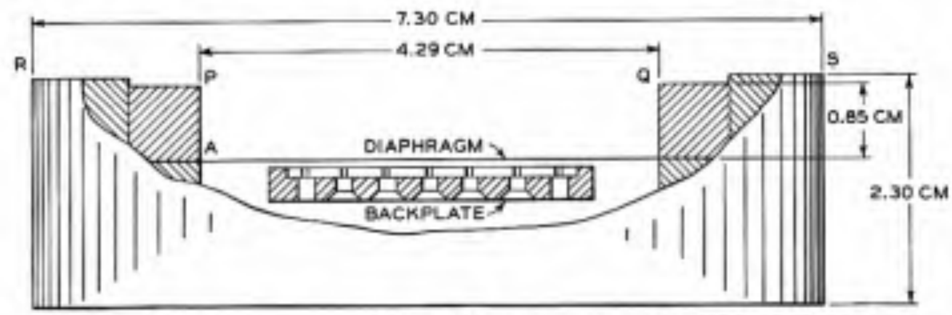


Fig. 1*A*—Contour dimensions of No. 394-type condenser transmitter.

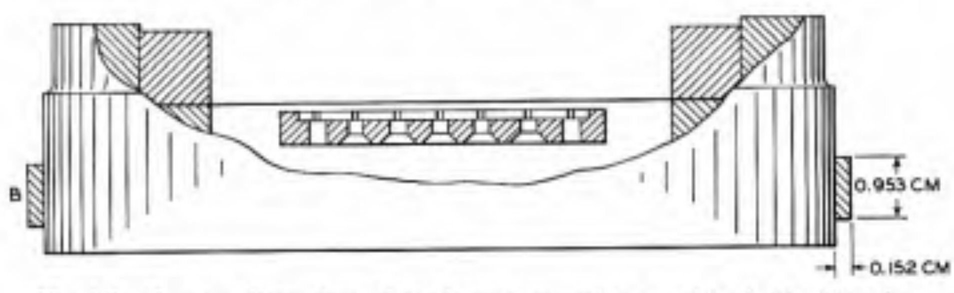


Fig. 1*B*—Contour dimensions of condenser transmitter used for field calibration.

[1] See bibliography.

mitter. That ratio $[R(f) = e/p]$, as a function of frequency, gives the calibration. Where and how is the acoustic pressure to be measured? This can be done in any one of a number of ways, all of which in general lead to different calibrations. The two calibrations most useful and amenable to measurement are when the pressure is uniform over the diaphragm and measured at the diaphragm and when the pressure is the pressure in a progressive plane wave, undistorted by the transmitter or any other obstacles; when the electromotive force is measured the distortion of the sound field must be due to the transmitter alone.

It is convenient to designate the former as "constant pressure" or "pressure" calibration, the latter as the "constant field" or "field" calibration. In general the field calibration will depend on the angle of wave incidence. Incidence normal to the diaphragm gives the "normal field" calibration. Where no confusion can arise, "field" calibration will be used to imply normal incidence. The pressure and field calibrations tend to coincide when the transmitter dimensions are small compared to the sound wave-length and when there are no appreciable impedances between the diaphragm and the sound field in front of it. Neither condition obtains for the No. 394-Type Transmitter, except at very low frequencies.

Which of the two calibrations—"pressure" or "field"—is more significant depends on the particular use made of the transmitter. Thus in the receiver testing machine, where the sound is substantially uniform throughout a small chamber closed by the transmitter diaphragm and by the receiver under test, the pressure calibration is important. When the transmitter is used to pick up sound in the open air at a distance from the source, the field calibration applies. For other cases, neither calibration is directly applicable, this being discussed at the end of the paper.

Constant Pressure Calibrations

For the several methods available for constant pressure calibration, the pressure may be applied either acoustically or electrically. In the acoustical group are the following methods:

1. Thermophone.[2]
2. Pistonphone.[1, 2]
3. Resonating tube.[3]
4. Compensation methods.
    a. Electrodynamic compensation for acoustic pressure.[4]
    b. Electrostatic compensation for acoustic pressure.[5]
5. Membranephone.

In the electrical group for pressure calibration are the following methods:

6. The back electrode (backplate) serving as the driving electrode.[5, 6]
7. An auxiliary third electrode driving the diaphragm.

All but two of the above methods have been described in detail in the articles to which references have been given so that only brief descriptions of the methods are given in the following paragraphs.

1. *Thermophone.*—The alternating pressure generated in the chamber of which diaphragm D (see Fig. 2) is one wall, is computed from the physical constants of the thermophone T, and of the gas (hydrogen) filling the chamber. A computation similar to that in reference[2] is discussed in Appendix I and II. The difference is in the manner in which the heat conductivity of the walls is taken into account. Also a slight correction for the yielding of the diaphragm is introduced, which was superfluous with the earlier, less sensitive model. An important advantage of the thermophone method is that it is not necessary to have the heating element parallel to the diaphragm. This makes it applicable to transmitters with curved or corrugated diaphragms. In such cases it is difficult to provide the accurately parallel and narrow spacing between the diaphragm and driving or compensating electrode, required in electrostatic methods.

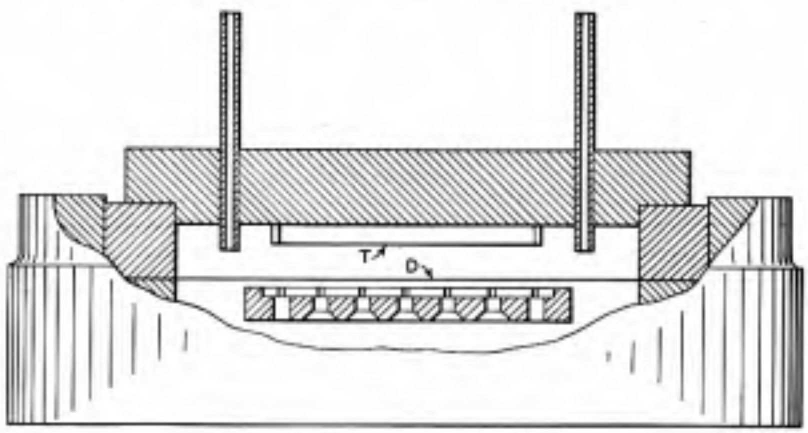


Fig. 2—Thermophone method.

2. *Pistonphone.*—The pressure is generated by means of a reciprocating motor-driven rigid piston as shown in Fig. 3. The piston amplitude is computed from the dimensions and the angular velocity of the cam driving it. The motor drive makes the method suitable for relatively low frequencies, up to about 200 c.p.s.

3. *Resonating Tube.*—The pressure at the diaphragm end of the tube (see Fig. 4) is computed from a measurement of the air particle velocity at a pressure node. That velocity is obtained by observing the deflection of a Rayleigh disk, R. D., placed in the tube. The sound source R is shown as a moving çoil receiver.

4. *Compensation Methods.*—The pressure in the chamber is determined by measuring the force required to prevent motion of a

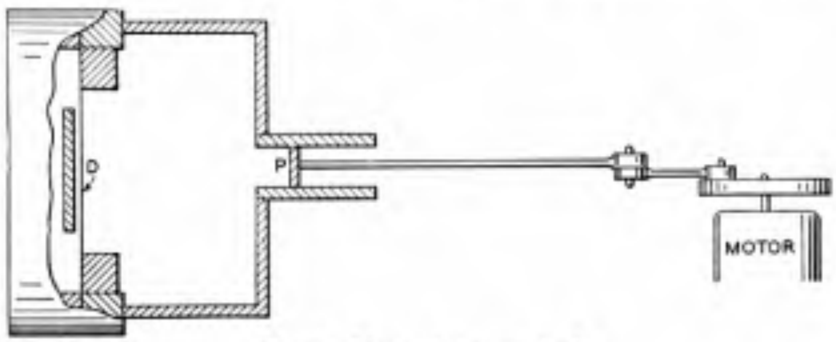


Fig. 3—Pistonphone method.

small auxiliary diaphragm D_2, Fig. 5. With the sound pressure so determined the corresponding electromotive force of the transmitter is measured. The rest condition of D_2 is indicated by absence of sound in an exploring tube communicating with the space back of D_2 or by absence of frequency variation in a high frequency circuit in which D_2 is made one plate of a condenser controlling the oscillation frequency.

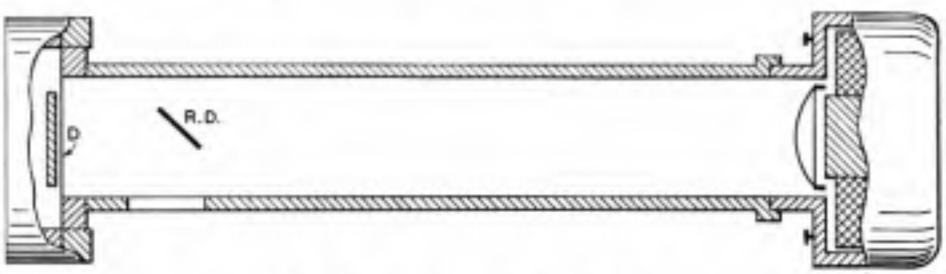


Fig. 4—Resonating tube method.

4*a*. *Electrodynamic Compensation for Acoustic Pressure.*—The compensating pressure is provided by sending a current of adjustable frequency, amplitude and phase through D_2 placed in a steady magnetic field.

4*b*. *Electrostatic Compensation for Acoustic Pressure.*—The same end is attained with a potential difference of adjustable frequency, amplitude and phase applied between D_2 and a fixed electrode parallel to it.

In particular the transmitter diaphragm and backplate may serve as D_2 and the fixed electrode. This, however, requires caution. The air gap is so small (approximately 2.5×10^{-3} cm.) that unavoidable variations in its value will in general cause appreciable variations in the value of the electric driving force over different parts of the diaphragm. The non-uniformity of the air gap is due to mechanical imperfections and to the electrostatic pull of the polarizing voltage. Furthermore, in transmitters of the type here considered, the backplate diameter is substantially smaller than that of the diaphragm, and hence the compensating electric force is not effective in a peripheral portion of the diaphragm.

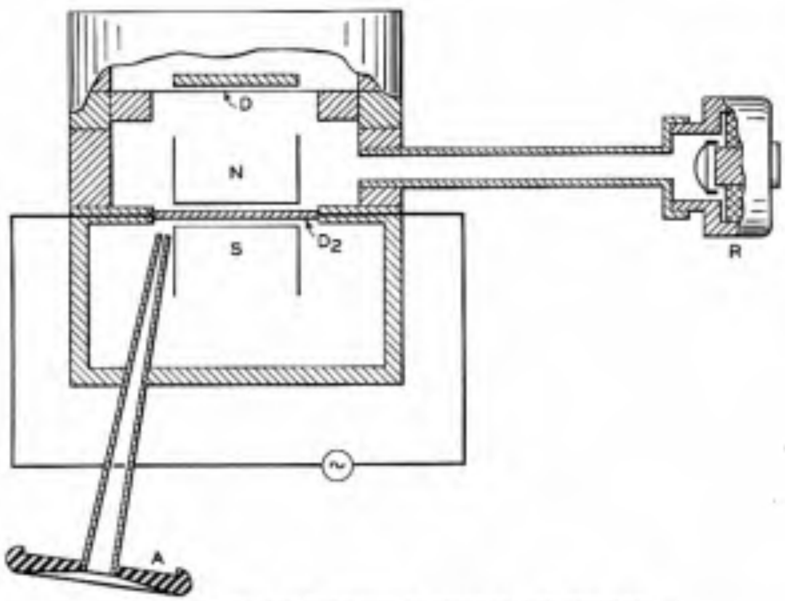


Fig. 5—Electrodynamic compensation method.

The electric force in this case is provided by inserting between the diaphragm and the backplate a steady potential difference, V_0, and a much smaller alternating potential difference, $V_1 \sin \omega t$, in series. One of the resultant force components is $\alpha V_0 V_1 \sin \omega t$ which has the same frequency as the sound source (e.g. a thermophone). The amplitude and phase of the electric force are adjusted until it balances the acoustic pressure on the diaphragm. This gives the value of the acoustic pressure, provided α is known. The compensating electric force is then removed, and the output of the transmitter due to the acoustic pressure is measured. Thus the pressure calibration is obtained. The value of α is given by a measurement of the value of V_0 required to balance a known static gas pressure established at the face of the diaphragm. This must be done for each instrument to be calibrated.

5. *Membranephone.*—In principle this method is similar to the pistonphone. An acoustically driven membrane M (see Fig. 6) replaces the motor-driven piston. From the volume displacement, ΔV, of M the pressure on the transmitter diaphragm D is computed. The value of ΔV is given by a measurement of the alternating variation in capacitance between M and an auxiliary perforated electrode G. The range of the method is from the lowest frequencies up to those at which the linear dimensions of the chamber become comparable with the sound wave-length (λ). As with the thermophone, that upper limit can be extended through the use of hydrogen instead of air.

The computation of ΔV is given in Appendix III. It will be noted that the computation is independent of the mode in which the membrane vibrates. However, for frequencies above the first resonance of

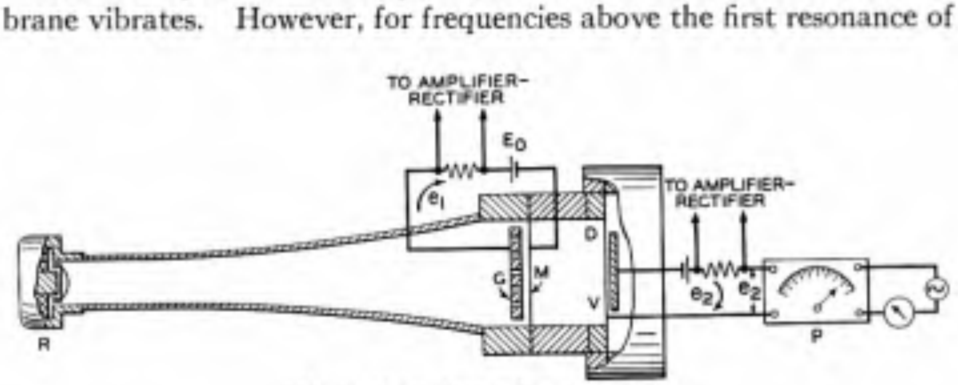


Fig. 6—Membranephone method.

the membrane the requirement as to smallness of chamber dimensions relative to λ, becomes much more stringent than in the thermophone case.

Methods Employing Electrical Drive.—Since the driving forces in this group are electric the pressure on the diaphragm is affected by the acoustic load on the front face of the diaphragm. To obtain the true pressure calibration that acoustic load must be known. Practically this is taken care of by making that load sufficiently small, rather than accurately determining its value.

6. *The Back Electrode Serving as the Driving Electrode.*—The alternating potential difference, $V_1 \sin \omega t$, is impressed in series with the steady potential V_0, see Fig. 7. This gives a driving force component $\alpha V_0 V_1 \sin \omega t$. The corresponding alternating variation in the transmitter capacitance is determined by having that capacitance control the frequency of a high frequency oscillator circuit. Absolute values are obtained by means of a static pressure calibration as in Method 4.

In this case, however, that does not give the force acting on the diaphragm unless the air impedance between the diaphragm and backplate is negligible in comparison with that of the diaphragm itself. Hence the method does not apply to the No. 394-Type Transmitter. The same consideration as to non-uniformity of the driving force over the area of the diaphragm which was mentioned in connection with Method 4*b*, applies to this case.

7. *Auxiliary Third Electrode Driving the Diaphragm.*—Here an auxiliary electrode *M* and a circular metal screen furnishes the electrostatic drive (see Fig. 8). It has nearly the same diameter as *D* and is parallel to it. The gap between *M* and *D* is about thirty times greater

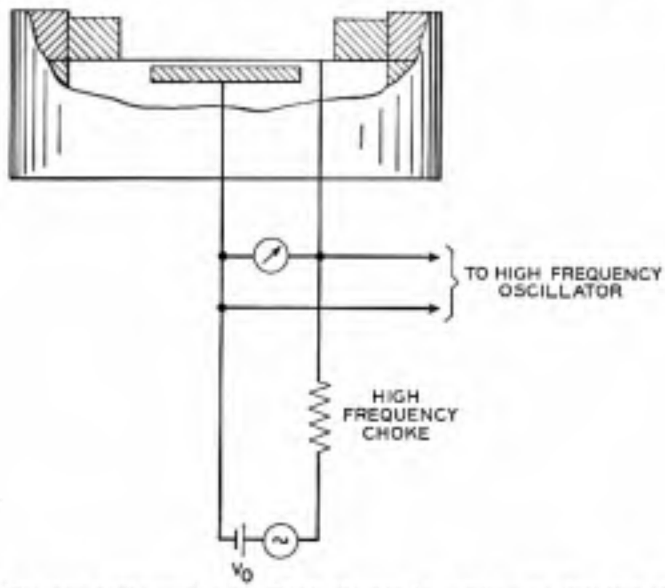


Fig. 7—Electrostatic method—Back electrode serving as driving electrode.

than between *D* and the backplate. Hence the electric force on *D* is uniform over the surface of *D*, and its absolute value can be computed with some accuracy. The calculation is given in Appendix IV. Care must be taken to avoid acoustic loading of *D* in a manner that would materially change its impedance. With this possibility guarded against, this method admits of an absolute transmitter calibration from 20 to 20,000 c.p.s. A comparison of a calibration so obtained with that given by a thermophone for the same transmitter,* is shown in Fig. 9. The two are quite independent. The discrepancy between the two up to about 6,000 c.p.s. is regarded as being within limits of experimental error. The acoustic load imposed on the diaphragm by

* This particular instrument happened to be about 4 db less efficient than the average No. 394-Type Transmitter.

the calibrating apparatus, while relatively small in either case, is not the same for both methods. At higher frequencies other factors contribute. At the highest frequencies, say above 10,000 c.p.s., the pressure on the diaphragm probably is more uniform in the present method than in Method 1.

Constant Field Calibration

For constant field calibration methods it is difficult to provide a plane progressive wave over a sufficiently large wavefront. Instead a small source in a chamber lined with highly absorbing material is used. The resultant progressive spherical wave, at sufficient distance from

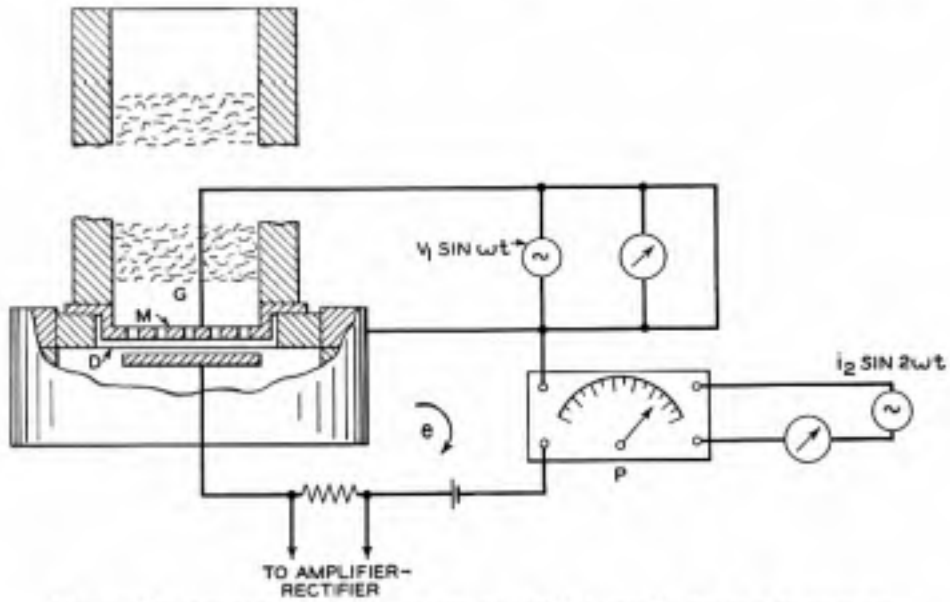


Fig. 8—Electrostatic method—auxiliary third electrode driving diaphragm.

the source, gives approximately the desired sound field. The measuring device must give the absolute value of the undistorted field intensity. We shall not consider the thermal, optical and sound radiation pressure methods possible, on account of the experimental difficulty which they present. One other absolute method is more readily available:

The Rayleigh Disc, which on certain assumptions gives the absolute value of the particle velocity in the sound wave. In the sound field presupposed for the field calibrations, the corresponding sound pressure is easily computed.[8]

Another procedure is to measure the sound pressure with the aid of a "search transmitter." This is a transmitter whose dimensions are so small relative to the sound wave-length that its pressure calibration, as obtained say by Method 1, may be taken to coincide with its field calibration.

The normal field calibration of a No. 394-Type Transmitter is shown in Fig. 10. The contour of the particular instrument used is shown in Fig. 1*B*. It was suspended from a thin rod clamped to the metal band *B*. The measurements were made with a Rayleigh disc (0.5 cm. diameter, 2.46 second period), using the modulated sound method.[9] The transmitter was placed 32 cm. from the sound source, a 1-cm. diameter tube attached to a loud-speaking receiver. The data obtained for frequencies below 500 c.p.s., are believed to be not so reliable as the rest because of appreciable reflections from the chamber walls.

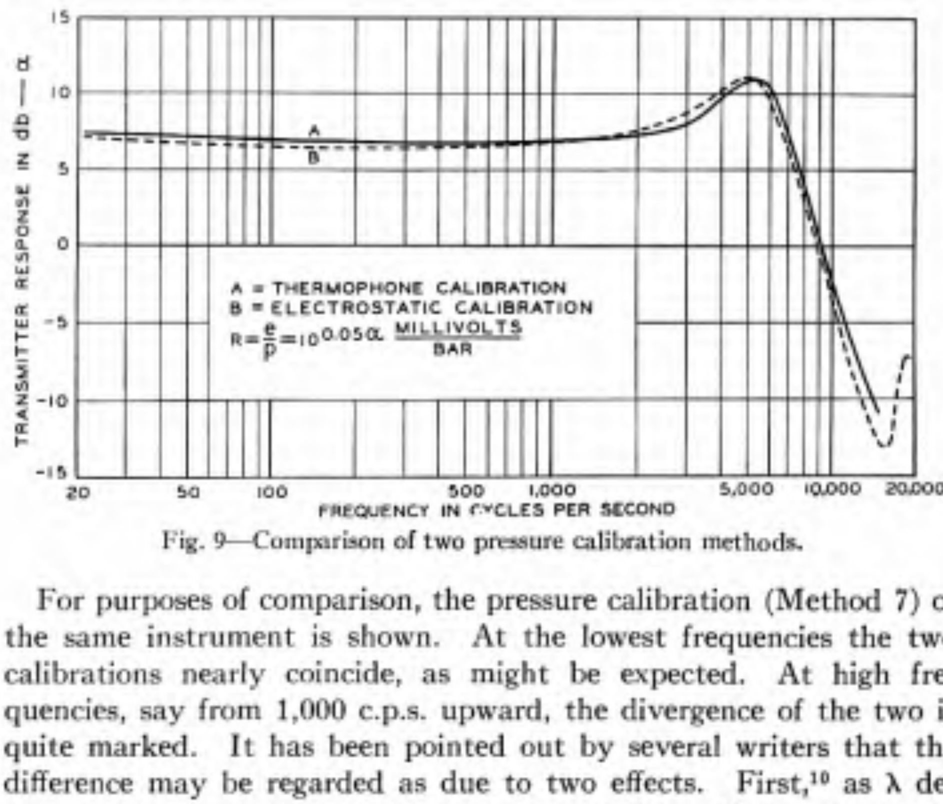


Fig. 9—Comparison of two pressure calibration methods.

For purposes of comparison, the pressure calibration (Method 7) of the same instrument is shown. At the lowest frequencies the two calibrations nearly coincide, as might be expected. At high frequencies, say from 1,000 c.p.s. upward, the divergence of the two is quite marked. It has been pointed out by several writers that the difference may be regarded as due to two effects. First,[10] as λ decreases, the transmitter tends to cause a doubling of the pressure in front of it as would a rigid wall. Second,[11] the recess in front of the diaphragm (Fig. 1) introduces a broad resonance which has its maximum approximately at 3,500 c.p.s. An estimate of this effect is given in Appendix V.

The observed differences between the field and pressure calibrations, from 500 to 8,000 c.p.s. are in fair agreement with those computed for the two effects given above. The computations are based on assumptions as to the transmitter contour which are quite removed from the actual case. Thus for the first effect it has been suggested that the transmitter may be replaced by an "equivalent" rigid sphere of equal

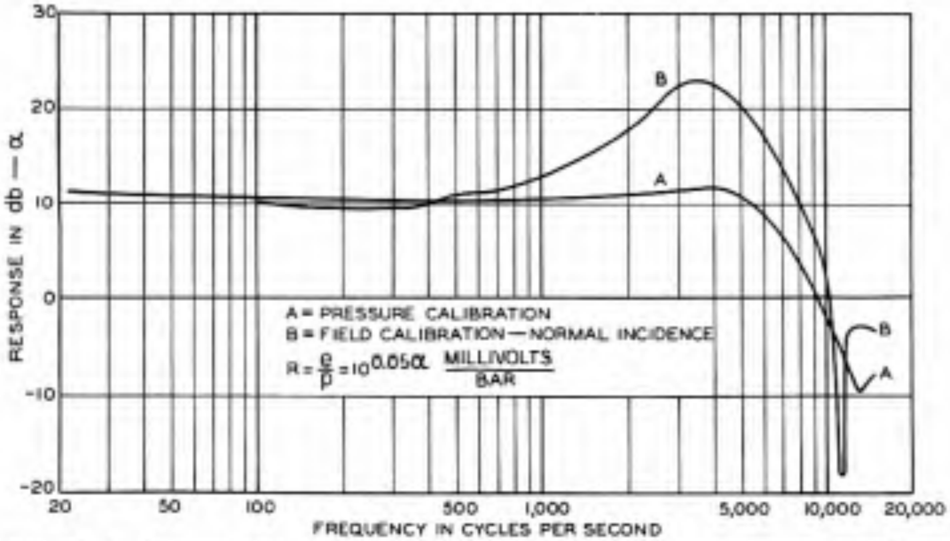


Fig. 10*A*—Pressure and field calibrations of No. 394-type condenser transmitter.

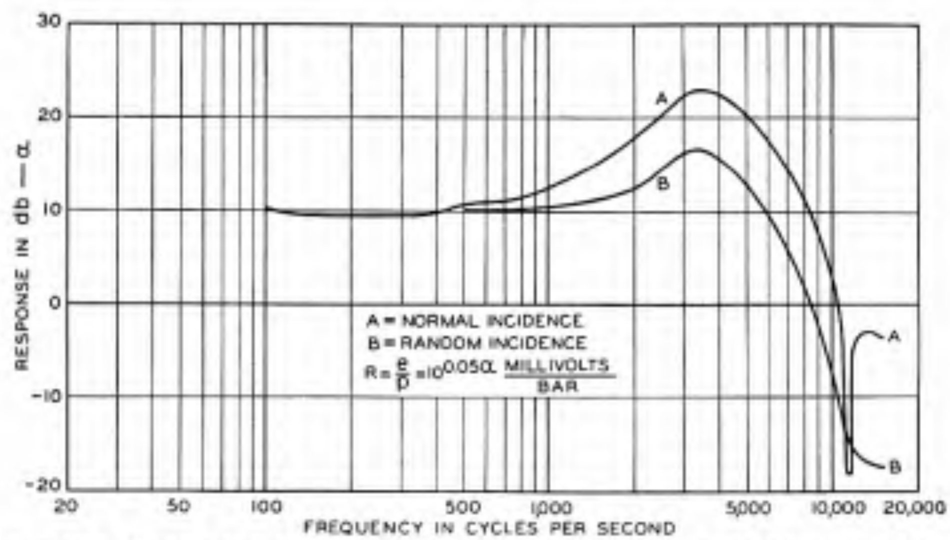


Fig. 10*B*—Field calibrations of No. 394-type condenser transmitter for normal and random incidence.

volume [12] or of equal diameter.[13] The data in Fig. 10*A* are best fitted by assuming a sphere of 9 cm. diameter, i.e., a diameter even larger than that of the transmitter. For the second effect the assumption is made that the face of the transmitter acts as an infinite wall, and that the air particles in the recess aperture all move in phase and normally to the diaphragm.

At still higher frequencies the doubled pressure effect largely persists and superposed on it are a number of rather complicated diffraction effects. These involve radial wave propagation across the diaphragm recess while the above two effects are due to normal plane waves. The marked dip at 11,200 c.p.s. corresponds to a sound wave-length such that

$$\sqrt{\left(\frac{1}{2}PQ\right)^2 + (PA)^2} - PA = \frac{1}{2}\lambda$$

(see Fig. 1*A*).

So far normal incidence of the sound wave has been assumed. For other directions of arrival, substantially different field calibrations are obtained. Since the transmitter is symmetrical about any diaphragm diameter, the effect of direction may be given in terms of the azimuth angle of incidence. A set of azimuth curves for various frequencies are given in Fig. 11, all expressed relative to the normal field calibration. In general, the higher the frequency the greater the effect of azimuth. For a large range of angles that effect is as great as or greater than the difference between the pressure and the normal field calibrations. It is interesting to note that the anomalous azimuth curve at 11,200 c.p.s. corresponds to a pronounced dip at that frequency in the normal field curve.

Relation of Field Calibration to Actual Transmitter Performance

We now consider the bearing of field calibrations upon the response of the No. 394-Type Transmitter under one or two conditions of actual use.

First, consider the case of a person speaking directly toward the diaphragm. The normal field calibration approximately applies, provided the distance is not great enough for reflected waves to be comparable with the direct wave and the distance is not so small that the transmitter reacts back on the source (the voice), or that pronounced standing waves are set up between the transmitter and the head. Outdoors and in a well damped room distances ranging say from 6 inches to 3 feet are likely to be within the above limits for the important voice frequencies.

On the other hand, for much of indoor work the distances from the microphone to the source and to the several reflecting surfaces are such that waves reaching the microphone by reflections are comparable with and often predominate over the direct sound. Besides, the microphone often is so placed that the direct sound strikes it more nearly at a 45-degree or 60-degree angle rather than normally. In a 29-foot × 29-foot × 13-foot room having a reverberation time of 1

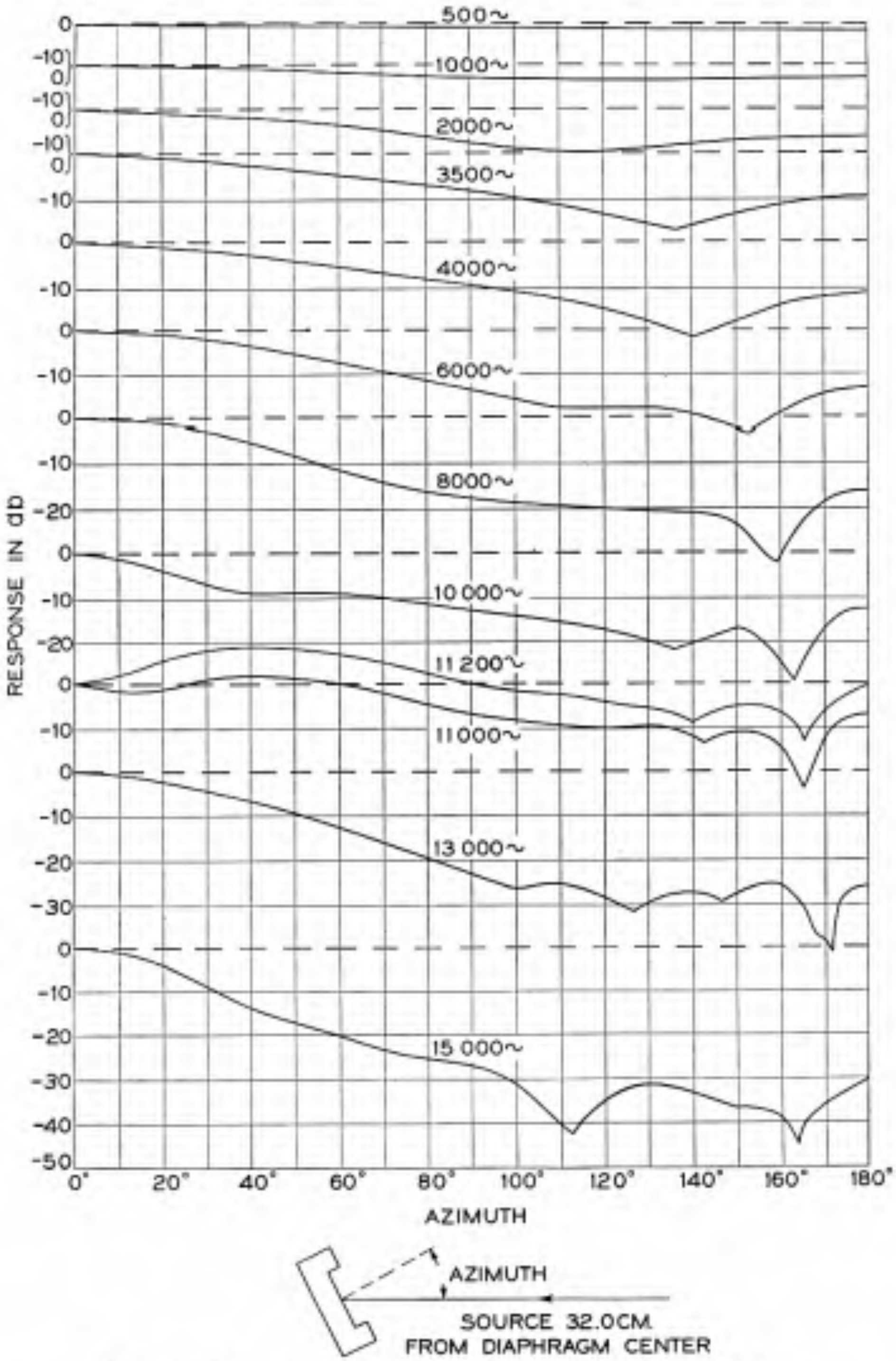


Fig. 11—Azimuth response of No. 394-type condenser transmitter.

second, the reflected waves reaching the microphone at 12 feet from a small source contribute much more to the microphone output than does the direct sound. To illustrate the effect of these reflections, the curve (*b*) in Fig. 10*B* has been plotted. It is based on the data of Fig. 10 and Fig. 11, and assumes that the transmitter is acted upon by progressive plane waves arriving with equal intensity from all directions in space. Their phases are taken to have random distribution. At any one frequency the response of the transmitter is then proportional to

$$\sqrt{\int_0^{\Pi} [A(\theta)]^2 \cdot \sin\theta \cdot d\theta},$$

where $A(^\theta)$ is the azimuth factor taken from Fig. 11. The result is seen to be intermediate between the pressure and the normal field calibrations, for frequencies up to about 8,000 c.p.s. Under these circumstances it is immaterial which way the diaphragm faces, but this holds only for sustained sounds. For sounds of short duration, the peak amplitudes in the microphone output often are of particular interest. They will be more nearly given by that single field curve corresponding to the azimuth with respect to the sound source in which the transmitter happens to be.

The above discussion of directional effects is simplified by the fact that the No. 394-Type Transmitter is symmetrical about any diaphragm diameter. Hence a single parameter—azimuth angle—is sufficient. The amplifier mounting cases usually employed destroy that symmetry. The directional effect becomes much more complicated since it involves two parameters, e.g. two direction cosines of the diaphragm axis. It has been suggested [13] that this complication can be done away with by placing the transmitter and its amplifier case in a rigid hollow sphere, only the transmitter front being exposed. If the front contour of the instrument be designed closely to conform to the rest of the sphere, and if the diaphragm subtend a sufficiently small angle at the center of the sphere, the directional effect can be computed.[14]

The simplest directional properties, i.e. uniform response for all directions of incidence, require a transmitter whose linear dimensions are small (say $< \frac{1}{4}\lambda$) relative to the shortest sound wave-length to be picked up. For a frequency range extending to 10,000 c.p.s., this means a transmitter less than 0.85 cm. in diameter. In general such restriction on the permissible size adds to the difficulties of construction and operation of the instrument. It is not intended to imply that nondirectivity of the transmitter is always desirable for pickup systems of highest quality.

A complete description of the performance of the microphone as an electro-acoustic converter is extremely complex. It involves the microphone, the sound source, their relative positions, and the surrounding acoustic configuration. Furthermore, it is limited to sound sustained long enough to allow the reflection pattern to attain a steady state. Therefore, in order to obtain a reasonably simple and useful statement of the transmitter response, the field calibration is made under the ideal acoustic conditions stated in part *A*. Even then the field calibration (including, of course, the azimuth measurements) is far more difficult and laborious than the corresponding pressure calibration. For some important purposes the pressure calibration is sufficient, even though the transmitter be intended for use in an "open" sound field. An instance is the specification and comparison of instruments having similar contours. The difference between the field and pressure calibrations, once determined for an individual instrument, applies to all others. That is, provided the acoustic impedances of their diaphragms are not too widely different, which usually is the case. Therefore the response of any instrument, as a function of frequency, age, barometer pressure, temperature, etc., is given by the pressure calibration. The thermophone method (Method 1) is particularly suitable for rapid and reproducible determinations of the pressure calibration. That is the method employed for the specification of No. 394-Type Transmitters, and of others having similar contours, in the Master Reference Systems [15] for Telephone Transmission in Europe and in this country.

I am indebted to Messrs. R. T. Jenkins, H. T. O'Neil and E. M. Little of Bell Telephone Laboratories for much of the material used in this paper.

Appendix I

The pressure generated by the thermophone is slightly reduced by the heat conductivity of the chamber walls. That conductivity is so great as compared with that of the gas, that zero temperature variation at the walls may be taken as one of the boundary conditions of the problem. This results in a solution nearly identical with that of eq. (7), p. 336, in the original derivation.[2] The correction factor given there on p. 340, which takes care of the wall conductivity, is now found to be more nearly unity. The difference between the two solutions is shown in Fig. 12 for a special case typical of condenser transmitter calibrations. As might be expected, it is greater the lower the frequency.

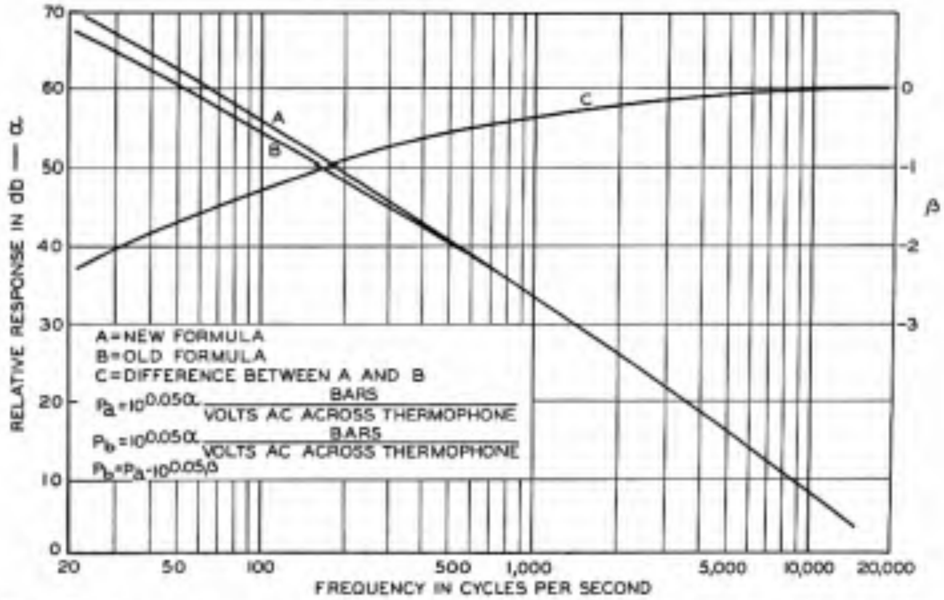


Fig. 12—Pressure generated by a thermophone in a transmitter calibration chamber.

Appendix II

In the thermophone theory the walls of the chamber were treated as being rigid. Actually the transmitter diaphragm presents a small but finite admittance in shunt with the elastic admittance of the gas in the chamber. The correction factor M due to this, is approximately

$$M = \frac{1}{\sqrt{1 + \frac{(\gamma p_0 v)^2}{V_0} + 2\,\frac{\gamma p_0 v}{V_0} \cdot \cos\theta}},$$

where

$$\gamma = 1.4 = \frac{C_p}{C_v},$$

assuming adiabatic conditions

$p_0 = 10^6$ bars atmospheric pressure,
$V_0 =$ volume of thermophone chamber,
$v =$ volume displacement of diaphragm per bar,
$\theta =$ phase angle of above displacement with respect to the pressure on the diaphragm.

At low frequencies $\cos\theta$ may be taken as nearly unity, and v can be approximately computed as below

$$v = \tfrac{1}{2}\Pi a_1^2 y,$$

$$y = \frac{\Delta C}{C}(h - y_1)\left[\frac{1 - \frac{1}{2}\frac{a_2^2}{a_1^2}\cdot\frac{y_1}{h - y_1}}{1 - \frac{1}{2}\frac{a_2^2}{a_1^2}\left(1 + \frac{2y_1}{h}\right) + \frac{1}{3}\frac{a_2^4}{a_1^4}\cdot\frac{2y_1}{h - y_1}}\right],$$

$$\frac{\Delta C}{C} = \frac{C_3}{C_2}\cdot\frac{e_1}{E_0}, \qquad \text{and} \qquad y_1 = \frac{h\left(\frac{C_2}{C_1} - 1\right)}{1 - \frac{1}{2}\frac{a_2^2}{a_1^2}},$$

where h = separation between diaphragm and back plate without polarizing voltage.

C_1 = capacity between diaphragm and back plate without polarizing voltage.

C_2 = above capacity in presence of polarizing voltage.

C_3 = total transmitter capacity, with polarizing voltage.

E_0 = polarizing voltage.

e_1 = transmitter e.m.f. per bar, uncorrected for yielding of diaphragm.

a_1 = diaphragm radius; a_2 = back plate radius.

For the 394-Type Transmitter, up to about 2,500 c.p.s., M is nearly 0.92. Above that the correction decreases owing to decreasing $\cos\theta$, and becomes negligible at 5,000 c.p.s. For still higher frequencies the correction becomes negative but remains small due to the increasing diaphragm impedance.

Appendix III

Schematically the membrane phone is shown in Fig. 6. D is the diaphragm of the transmitter to be calibrated; M, a stretched membrane acoustically driven from the receiver R; G, a perforated plate. Let V = volume between D and M; y_0 = normal separation between G and M; C_0 = normal capacitance between G and M.

Then, if $y_0[1 + K(S)\cdot \sin\omega t]$ represents the GM separation when M is driven by R, the resultant capacitance variation is:

$$\Delta C = \sin\omega t\cdot\frac{1}{4\Pi y_0}\cdot\int\frac{K(S)}{1 + K(S)\sin\omega t}\cdot dS$$

and

$$\Delta V = \sin\omega t\cdot y_0\int K(S)\cdot dS,$$

the integration extending over the entire area of M.

Taking $K(S) << 1$, but without restrictions on the variation of $K(S)$ over the surface of M,

$$\Delta V = 4\Pi y_0^2 \cdot \Delta C_0.$$

Hence the transmitter sensitivity is given by

$$R = \frac{e_2}{p} = \frac{e_2 V E_0}{\gamma p_0 \cdot 4\Pi y_0^2 C_0 \cdot e_1} \text{ volts/bar.}$$

The above presupposes: (1) $V/S << \lambda$, $\sqrt{S} << \lambda$; (2) acoustic admittance of D is very small compared with that of V; (3) adiabatic compression. If necessary, corrections for deviations from (2) can be made in accordance with Appendix II. The correction for (3) is found to reduce the pressure in the ratio

$$R' = R \cdot \frac{1}{1 + (\gamma - 1) \cdot \dfrac{\tanh \beta a}{\beta a}},$$

where

$$\beta = (1 + i)\sqrt{\frac{\rho \omega C_p}{2K}},$$

when C = specific heat at instant pressure,
K = thermal conductivity of the gas,
ρ = density,
$\gamma = \dfrac{C_p}{C_v}$.

The upper frequency limit imposed by condition (1) can be raised by filling V with hydrogen. For the No. 394-Type Transmitter, and with R a No. 555-W Western Electric Receiver, an air-gap $y_0 = 0.075$ cm. corresponds to easily measurable values of e_1 and e_2. M was a 0.001 inch duralumin diaphragm, stretched to 5,000 c.p.s. resonance frequency. It was found that the upper frequency limit of the method is determined by M breaking up when vibrating in one of its higher natural modes. This tends to produce a non-uniform pressure on D, and the above condition must be met much more perfectly than in the thermophone case.

Appendix IV

The particular electrostatic calibration described below, employs a separate driving electrode and a sinusoidal driving voltage which produces a sinusoidal driving force of double frequency. The latter has the advantage of adding frequency selectivity to shielding as the means for keeping the relatively large driving voltage out of the transmitter output circuit.

In terms of Fig. 8 the sensitivity of the transmitter is given by

$$R = \frac{e}{p} \text{ volts/bar}, \quad e = i_a r_a \cdot 10^{-0.05\alpha}, \quad p = \frac{V^2}{9 \times 10^4 \times 8\sqrt{2} \cdot \Pi h^2},$$

where $V\sqrt{2} = V_1$, measured in volts, $i_a\sqrt{2} = i_1$, h = separation between M and D. The e.m.f. e is measured by means of the potential attenuator P, carrying a known current i_a, and having an input resistance r_a. At any one frequency two quantities must be measured: V, say with an electrostatic voltmeter, and α, the setting of the attenuator in decibels. The current i_a must be known but can readily be kept constant at all frequencies if a heterodyne oscillator be used as the source.

Two corrections must be applied. First, the auxiliary electrode is perforated. Hence not all of its area is electrostatically effective. Second, p in the above is the electrostatic force per unit area, rather than the acoustic pressure on the diaphragm. The two are different, in general, because of the acoustic load (Z_d) on the front face of the diaphragm. The value of Z_d is affected by form of C, the auxiliary electrode, and by the acoustic impedance beyond it in the chamber G.

It is best to have Z_d as small and as free from reactance as possible. This is accomplished by using stretched fine metal gauze. Copper gauze, 300-inch mesh, is quite good. It terminates in the tube TT, ¼-inch iron wall, which is filled with several layers of loose cotton batting and hairfelt. The effectiveness of the arrangement was judged by the fact that altering the size of G did not appreciably affect the calibration.

While the screen electrode provides a practically uniform electrostatic pressure over the surface of D, it is rather complicated to compute the effective absolute values of h and of the electrostatic area. This is more easily done by comparing it at low frequencies (say at 100 c.p.s.) with a steel plate electrode in which the perforations take up about 12 per cent of the total area. The surface facing D is carefully machined so that h is uniform and known within less than $\pm$ 2 per cent. This is for absolute values of h in the range 0.075–0.080 cm. The acoustic load which this electrode imposes on D, with G removed, is negligible at low frequencies. A lower limit on the electrostatic correction for the perforations is made by adapting the calculation given by Maxwell ("El. Mag.," 3d ed.) for rectangular grooves in one plate of a parallel plate condenser. The above value of R is corrected to

$$R' = R \cdot \frac{1}{1 - \frac{S_1}{S} \cdot \frac{g}{h+g} - \frac{S_1}{S} \cdot \frac{gh}{(h+g)2}},$$

when

$$g = \frac{\sqrt{S_1} \cdot \log_e 2}{\Pi}, \qquad S_1 = \text{area of perforation}, \qquad S = \text{total area}.$$

An upper limit on the above correction is given by:

$$R' = R \cdot \frac{1}{1 - \frac{S_1}{S}} \cdot$$

The R' actually used was the mean of the above two values. The value of R obtained with the screen electrode is shifted up or down to make it coincide with R' given by the perforated electrode, at 100 c.p.s.

Appendix V

For frequencies below about 5,000 c.p.s. the difference between the pressure and normal field calibrations is mainly due to two effects: (1) reflection from the transmitter face and from the diaphragm; (2) air resonance caused by the recess in front of the diaphragm.

Consider Fig. 1*A*. Assume that in the circular aperture PQ, the air particles are all moving in phase and parallel to AP. Then we may treat PQ as a rigid massless piston in the wall RS. If RS/λ is large enough, the pressure on PQ held motionless will be double that of the field pressure. The motional impedance of PQ imposed by the air above PQ is given by Rayleigh (Sound, vol. II, § 302). Per unit area it is

$$Z_1 = \rho C(a + ib),$$

where

$$a = 1 - \frac{J_1(2kR)}{kR}; \quad b = \frac{\omega\rho\Pi}{2k^3} K_1(2kR); \quad k = \frac{2\Pi}{\lambda}; \quad 2R = PQ.$$

Let R_F/R_P represent the ratio of " field " to " pressure " calibration. Using the expression for plane wave propagation in a tube (e.g. Crandall, Theory of Vibrating Systems and Sound, p. 99) we have at once:

$$\frac{R_F}{R_P} = 2 \cdot \frac{1}{[\cos kl + i(a+ib) \cdot \sin kl] + \frac{\rho C}{Z_A}[(a+ib)\cos kl + i \sin kl]},$$

where Z_A is the equivalent impedance per unit area of the transmitter diaphragm, and $l = AP$. On substituting numerical values, R_F/R_P is found to have a maximum value of nearly 3.3 at $\omega/2\Pi = 3500$ c.p.s. This means that the air resonance adds a factor of 1.65 to the ordinary doubling of pressure caused by a plane wall. A substantially similar calculation has been given by W. West.[13]

The observed R_F/R_P is a maximum at 3,500 c.p.s. but its value is somewhat larger—nearly 3.65.

Bibliography

1. H. D. Arnold and I. B. Crandall, *Phys. Rev.*, X, 22 (1917). E. C. Wente, *Physical Review*, July 1917 and June 1922.
2. E. C. Wente, *Physical Review*, April 1922.
3. W. West, *Jl. I. E. E.*, Sept. 1929.
4. E. Gerlach, *Wiss. Veroff. Siemens-Konz.*, III, 1923.
5. E. Meyer, *Zs. F. Tech. Phys.*, p. 609, 1926; *E. N. T.*, *4* 1927; C. A. Hartmann, *E. N. T.*, 7, H. 3, 1930.
6. M. Grutzmacher u. E. Meyer, *E. N. T.*, *4*, 1927.
7. W. Zernow, *Ann. d. Physik*, 26, 1908.
8. Mallett and Dutton, *Jl. I. E. E.*, May 1925.
9. L. J. Sivian, *Phil. Mag.*, March 1928.
10. I. B. Crandall and D. Mackenzie, *Phys. Rev.*, March 1922.
11. A. J. Aldridge, *P. O. E. E. Jl.*, October 1928.
12. S. Ballantine, *Phys. Rev.*, 32, 988, 1929; *Proc. I. R. E.*, July 1930.
13. W. West, *Jl. I. E. E.*, April 1930.
14. Rayleigh, Theory of Sound, Vol. II.
15. L. J. Sivian, A Telephone Transmission Reference System, *Elect. Comm.*, Oct. 1924. W. H. Master and C. H. G. Gray, Master Reference System for Telephone Transmission, *Bell Sys. Tech. Journ.*, July 1929.

Rating the Transmission Performance of Telephone Circuits

By W. H. MARTIN

This paper discusses the rating of the transmission performance of telephone circuits on the basis of the rate of repetitions in telephone conversations and presents the rating method set up on this basis, which is being adopted in the Bell System for determining and expressing the data for the transmission design of the telephone plant.

A METHOD of rating the transmission performance of telephone circuits is of course an essential in specifying the grades of transmission service to be furnished, in designing, constructing and maintaining telephone systems to provide the desired grades of service economically and in the development of the various elements of the telephone system which affect its transmission. As the art of telephone transmission has developed and greater refinements have become possible and desirable, changes have naturally been made in the methods of specifying and rating transmission performance. Since many such changes have been made in recent years, it seems opportune to discuss the rating of transmission performance and to set forth the rating method which is now being adopted in the Bell System for determining and expressing the data for the transmission design of the plant. In this connection, various methods which have been employed for measuring the transmission performance of telephone circuits will be discussed to indicate their application and also their relation to the new rating method. It is the purpose here to discuss this rating matter primarily from the qualitative standpoint rather than to present in quantitative detail the various relations involved in rating telephone transmission. Obviously, the determination of many of these relations presents sufficient material for separate treatment.

In carrying on a telephone conversation three major functionings are involved, namely, that of the talker in formulating his ideas and uttering words to convey these ideas, that of the telephone circuit in taking the sounds of these words and reproducing them at another point, and, lastly, that of the listener in hearing and recognizing these reproduced sounds and in comprehending the ideas which they are intended to convey. It is evident that all three of these functionings affect the success of the telephone conversation. Since, however, the functionings of the talker and listener are common to both direct and telephone conversations it might seem that the consideration of the transmission performance of a telephone circuit could be limited to the functioning of the circuit itself. In line with this, there has been some tendency to confine such considerations to relations between the sounds reproduced by the telephone circuit and the sounds impressed upon it. The performances of the talker and listener, however, are materially affected in certain important respects by the telephone circuit, and determinations of the relative merits of the transmission performances of different telephone circuits, must therefore go farther than the performances of the circuits themselves and take account of the combined action of the talker, circuit and listener.

Characteristics of Conversation

In view of this reaction of the telephone circuit on the talker and listener, attention is directed to the pertinent characteristics of their performances in both face-to-face and telephone conversations.

Direct Conversation

In direct or face-to-face conversations both the talker and listener more or less subconsciously adjust their actions in many respects to each other and to their circumstances. The loudness of talking is placed initially at a level which experience has shown to be suitable for the conditions and for the particular listener. If the listener indicates verbally or by his expression that he is understanding easily or with difficulty, some further adjustment may be made in the loudness of talking. Since the talker judges his own talking level largely by the loudness with which he hears his own voice, this level will be a function of the amount of reverberation in the place in which he is talking. Apparently for a wide variation of loudness in the customary talking range, the speaker is not in general conscious of the amount of energy which he is expending. Noise at the place of conversation also plays a part in determining the talking levels since it makes louder talking necessary in order to permit a given degree of understanding on the part of the listener.

Along with an adjustment in talking level, the talker may improve his enunciation if difficulty of understanding is expected or indicated by the listener. There may also be a change in the manner of expressing the ideas in avoiding words which experience has shown are difficult to understand and an idea may be stated in more than one way in order to insure its comprehension. Other adjustments on the part of the talker may be determined by his opinion of the mental acuity of the listener, by the familiarity of the listener with the matter under discussion and by the interest in it. These factors affect the way of expressing an idea, the kind and number of words used and hence the time taken.

The listener also adjusts himself to the conditions by an amount which is determined somewhat by his interest in the matter under discussion. He may strive to comprehend the transmitted ideas and require few repetitions by the speaker or he may refrain from exerting himself and so tend to evoke greater effort on the part of the talker. At times he may pretend not to understand in order to get confirmation of a statement or to gain time in replying to a question.

In view of these factors and of the normal variations of different talkers and listeners in all these respects, the portion of the questions and statements of conversation which is correctly understood and the time required to interchange certain ideas may vary widely for different conversations even when they are carried on under a fixed set of local conditions. If it were desired to determine a measure of the conversational satisfactoriness of these conditions, in addition to some quantitative method for rating each conversation, there would be required, therefore, observations on a large number of conversations between different people in order to take account of the variables due to the material of conversation, the people, and their abilities and desires to accommodate themselves to the conditions.

Telephone Conversations

In telephone conversations, there are adjustments between talker and listener as is the case in direct conversations, but there are certain definite differences in this regard because of the interposition of the telephone circuit between the participants. Here also, the speaker tends to adjust his talking level to the loudness with which he hears his own voice. In this case, however, he hears his own voice not only through the air path, but also through the "sidetone" of the telephone set, that is, through the electrical path from his own transmitter to his own receiver. When this electrical path is more efficient than the acoustic path, the sidetone will tend to control the talking level. It has been found that varying the sidetone of the set has on the average a definite effect on the talking volume of the speaker, the talking volume being lowered as the efficiency of the sidetone path becomes greater.

In a telephone conversation, there is also a tendency for a person in talking to adjust his volume on the basis of the loudness with which he hears the person at the other end of the circuit. If the voice of the other person comes through weakly, he judges that the connection requires loud talking and acts accordingly. If the listener indicates that he is not understanding, the talker may talk more loudly or closer to his transmitter, and also make such adjustments in enunciation and in setting forth his ideas as in the case of direct conversation. Also, the loudness of talking may be affected by the room noise at the location of the speaker, which noise incidentally may not reach the listener and so play no part in his reception. Aside from cases where the room noise at the far end is severe enough to be heard over the telephone circuit, the speaker does not have definite knowledge of the room noise at the listener's end and therefore is not in a position to adjust his manner of talking to this condition except in so far as the listener may indicate difficulty in understanding.

In listening, the result is of course dependent upon the position of the receiver with respect to the ear. The local room noise reaches the ear to which the receiver is held both by the path between the ear cap of the receiver and the ear, and also through the sidetone path of his telephone set. Some telephone users have learned that this effect may be reduced by holding the receiver tightly to the ear and by covering the mouthpiece of the transmitter when they are listening.

It is evident that the success of telephone conversations depends not only upon the performance of the telephone but also upon the performances of its users, the material of their conversations, the way in which they talk into the transmitters and hold the receivers to their ears and the room noise conditions. In addition, it is seen that the performance of the telephone affects the performances of the users in such important respects as the loudness of talking, the manner of presenting the ideas and the amount of effort exerted to understand. Also, the effect of the room noise is a function of the circuit characteristics. Furthermore, the reactions of the circuit performance on those of the users are not constant but may vary from person to person and from conversation to conversation. In view of the random nature of these factors, which are beyond the control of those who design and operate the telephone system, the service performance rating of a telephone circuit should be on a basis which takes adequate account of their ranges and combinations in practice. This points to a rating based on a statistical analysis of results obtained under service conditions.

To determine and specify these factors so that it may be known how to duplicate the range of service conditions in laboratory investigations would be a prodigious task. Moreover, the duplication of these conditions under control is bound to introduce a large element of artificiality which would vitiate the results or at least raise serious questions as to their dependability.

The practical solution is to get sufficient data regarding the results obtained over telephone circuits of different performance characteristics by their normal users in carrying on regular conversations. This requires a suitable quantitative method of rating conversations and observations on a sufficient number of conversations over each circuit condition to be investigated to constitute a reliable sample. This does not mean necessarily that all the practicable circuit conditions have to be observed in this manner but rather that sufficient data be so obtained for the establishment of correlations with performance measurements which are susceptible to laboratory determination. The fundamental point is that service performance ratings need to be based on service results in order to take proper account of all the factors involved.

Transmission Performance of Circuits

The distinction has been made between two kinds of transmission performance of a telephone circuit, namely, that indicated by relations between output and input sounds and that indicated by the results obtained by the users of the telephone in carrying on their conversations under service conditions. Performance indications of the first kind will be referred to as "transmission characteristics" of the circuit. The second kind of performance may be termed "transmission service performance." The distinction between these two kinds of performance is an important one and should be kept clearly in mind.

The output sounds dealt with in transmission characteristics are not only the reproduced sounds which correspond to the input sounds but also the accompanying extraneous sounds which are delivered by the circuit. Also, the output sounds to be investigated cover not only those delivered by the receiver at the far end of the circuit but also those reproduced by the receiver in the station set containing the transmitter energized by the input sounds. The sounds from the near receiver include both those transmitted through the sidetone path of the set and those returned to the sending end by reflection at impedance irregularities in the circuit. Due to the time required for propagation over the circuit these latter sounds may be delayed with respect to the sidetone and hence appear as echoes. Likewise, echo sounds may be delivered at the far receiver.

Transmission characteristics do not in themselves show the service performance as realized by the users of the telephone but are essentially indications of the functioning of the circuit in reproducing sounds. They provide, therefore, a means for investigating and specifying the performance of a telephone circuit without involving many, and in some kinds of transmission characteristics any, of the actions of the talker and the listener in conversation. With the establishment of proper correlations between transmission service performance and transmission characteristics, these latter can of course be used to indicate service performance.

In addition to specifying any kind or grade of circuit performance on the basis of performance results there is the method, which has had important practical application, of indicating performance in terms of types of instruments and circuits and of the conditions of their use. For example, a statement of the types of transmitters, receivers, station sets and cord circuits and of the length and types of loops and trunk, together with specific conditions of use, provides an indirect specification of a performance. This method, which is extensively used in many fields, may be termed the "instrumentality designation method." An outstanding application of this method in telephone transmission work is the Standard Cable Reference System [1] which was so widely employed to provide a scale of performance. This method has many present applications where physically determined characteristics are unavailable or are difficult of definite determination and specification. Also, the designation of instrumentalities is convenient in many cases because it provides a ready means of specifying a practical combination of various kinds of transmission characteristics. While this method is often expedient practically, taken by itself, it is inherently cumbersome for the development of improved instrumentalities because of the lack of physical indication of the features to be investigated.

Transmission Characteristics

The usefulness to the listener of the speech sounds reproduced over a telephone circuit is a function of their loudness, of their distortion or degree of departure from facsimile reproduction, and the magnitude and character of the extraneous sounds or noise which accompany them. Transmission characteristics are therefore directed primarily to indications of the effects of the circuit and its parts on the reproduction of sounds in these three respects. As already indicated, transmission characteristics are determined not only for the path from transmitter at one end to receiver at the other, but also for the sidetone and echo paths.

Speech sound transmission characteristics, that is, expressions of the relations between impressed and reproduced speech sounds, while they have been extensively used, present some difficulty in quantitative determination and specification because of their complex nature. Also, the human element is involved in the persons used as generators

[1] "Master Reference System for Telephone Transmission," Martin and Gray, *Bell System Technical Journal*, July 1929.

of the speech sounds to be investigated and as observers to give indications of loudness and distortion and of their effects on the recognition of the reproduced sounds. Two outstanding kinds of relations of this type are those given by volume tests and articulation tests, which will be discussed later. It has therefore been of great convenience to take a further step and to study and specify the performance of telephone circuits and their parts in terms of their functioning for single-frequency sounds and currents. In this procedure, this functioning is investigated for a number of different single-frequency sounds and currents, so taken as to cover the range of frequencies transmitted by the circuit. In the single-frequency transmission characteristics, the personal element is eliminated and the measurements are made entirely on a physical basis.

A great deal of attention has been given to the correlation of speech sound and single-frequency transmission characteristics so as to enable the former to be derived from the latter and so extend the application of the type which is more readily susceptible to quantitative determination. Also, use has been made of easily specifiable multi-frequency sounds and currents to permit the physical measurement to approach more nearly speech sound conditions, of phonographic reproduction to reduce the personal factor in the generation of speech sounds for measurement purposes and of meter arrangements to simulate the ear ratings of sounds, particularly from the standpoint of relative loudness.

As a result of the correlation of speech and single frequency characteristics, extensive use has been made of determinations at selected typical single frequencies to check the design, installation and maintenance of lines and other associated circuit elements.

The widely used volume test is essentially a means of specifying the action of a telephone circuit or its parts, on the relation between the reproduced and impressed sounds from the standpoint of their relative loudness. In this test use has been made for many years of the Standard Cable Reference System and recently of the Master Reference System for Telephone Transmission [2] as references for comparison. These reference circuits with their adjustable trunks provide a means of obtaining different loudness ratios between input and output sounds. By talking alternately over the reference circuit and the one being investigated and adjusting the trunk of the reference system until the output sounds of the two circuits are judged to be equally loud, a specification of the loudness reproduction ratio is obtained of the circuit under investigation in terms of the length of the trunk in the reference system. The effect of a change in the

[2] See Reference (1).

telephone circuit, such as the replacement of one receiver by another, is measured in terms of the change required in the reference trunk to give a loudness balance for the second condition. In this way, measurements are also obtained of the effect on the loudness reproduction ratio of the various parts of telephone circuits. When the circuits used commercially consisted of apparatus and lines similar to those in the Standard Cable Reference System and the major controllable factor was the loudness reproduction ratio, such measurements constituted reasonably adequate means for indicating the comparative functioning of circuits and apparatus.

The noise on a telephone circuit may be measured in various ways. The method which has been most generally used is that of comparing it with the controllable output of a fixed source of a complex wave shape and adjusting this output until it and the line noise are judged to have equal interfering effects.

With the availability of circuits and apparatus having widely different distortion effects, the volume ratings became insufficient for indicating the relative performances of commercial circuits. The earliest method used in rating distortion effects was one in which observers listening to transmission over the circuits, gave judgments as to their relative merits. By so comparing various kinds and amounts of distortion, two at a time, relative ratings can be established for placing them in order of merit. This procedure was particularly useful in the early days in working out the designs of transmitters and receivers, especially from the standpoint of the location in the frequency range of their points of maximum response. While such a judgment method has the shortcoming of not providing quantitative ratings it has been found that experienced observers can in general obtain results which are relatively consistent with the results of more definite measuring methods. Such judgment comparisons of distortion effects are frequently used, particularly in exploratory work, and are still more or less necessarily relied upon in setting limitations on circuit properties which primarily affect the naturalness of reproduction.

To provide for the need of a method for measuring the relation between the reproduced and impressed sounds from the standpoint of effects of different kinds of distortion, use has been made of the articulation testing method.[3] In this method, which has been widely used in recent years, lists of syllables, usually meaningless monosyllables, are called over the circuits to be rated and the percentage of syllables correctly understood is taken as a measure of the circuit performance.

[3] "Articulation Testing Methods," Fletcher and Steinberg, *Bell Sys. Tech. Jour.*, Oct., 1929.

This testing method thus offers a means of indicating the distortion effects of circuits in terms of the recognizability of the reproduced sounds of speech. Probably one of its first applications[4] was in determining the cutoff frequency to be used in the design of coil loaded circuits.

The articulation testing method provides, of course, quantitative measures in terms of the recognizability of the reproduced sounds of speech not only of distortion effects, but also of the effects of the loudness of these sounds and of the noise which may accompany them. This method has provided a very powerful tool for investigating the effects of changes in the reproduction characteristics of telephone circuits on the recognition of the reproduced sounds and has been particularly useful in indicating the lines to be followed in reducing causes of distortion in circuits and apparatus and in evaluating the impairment caused by noise on telephone circuits. It has been recognized, however, that while such measurements indicate the capabilities of the circuits in reproducing recognizable speech sounds, they do not in themselves give direct measures of the degree of success which the users of the telephone obtain in conversations where their actions are free from the control which is necessary in articulation testing and where the contextural relation of the words plays such a large part in their recognition. To make the results of this type of testing approach more nearly the conversational results, words and sentences have been used in place of the meaningless syllables but it is evident that even with sentences, the control on the actions of the testers and on the ideas to be communicated presents a condition which is quite different from those of regular conversations.

All these ways of investigating and measuring the performance of telephone circuits in reproducing sounds have useful applications in present day transmission work. Frequently it is convenient to use different methods for the various parts of a circuit in specifying the complete functioning of the circuit in reproducing sounds.

Transmission Service Performance

From the standpoint of the users of the telephone circuit, the transmission performance is measured by the success which they have in carrying on conversations over the circuit. Different degrees of success in this process may be taken as being indicated by the number of failures to understand the ideas transmitted over the telephone and by the amount of effort required on the part of the users to impart and receive these ideas. Service performance is of course affected also by

[4] "Telephonic Intelligibility," Campbell, *Phil. Mag.*, Jan., 1910.

accidental irregularities in circuit conditions such as interruptions and cutoffs, but from the standpoint of transmission design, attention can be concentrated on the results obtained when the circuit is in normal operating condition. Since failures to understand and exertion of effort are experienced also in direct conversations, their occurrence in telephone conversations obviously cannot be entirely ascribed to the functioning of the circuit. Variations in these factors for different types of circuits can, however, be used as a measure of the effect of the differences in the transmission characteristics of these circuits.

The repetitions required in a conversation can be noted but a determination of the effort factor presents difficulties. There is undoubtedly the tendency in carrying on conversations, as in other activities, to exert no more effort than is necessary to obtain what the participants consider to be satisfactory results. This effort, however, will in general be increased as the difficulty of conversing becomes greater and so bears a relation to the increase in repetitions. Also, it is probable that two dissimilar circuits which cause the same rate of repetitions when used for the same service, will, on the average, call for the same amount of effort by their users.

In line with this, the rate of occurrence of repetitions requested by the users of a particular telephone circuit in carrying on their regular telephone conversations can be used as a direct measure of the service performance of the circuit. By determining the repetition rate for a large enough number of different people at the two ends to take account of the variability of their personal characteristics in talking and listening to the telephone and of the conversational material and conditions, a rating can be placed on the service afforded. By making such observations on connections having different transmission characteristics, relative ratings can be established for these various transmission characteristics.

It should be recognized, however, that while the rate of repetitions required can be used for relative ratings of the transmission service performance of different circuits, such ratings in themselves do not give a complete picture of the service from the users' standpoint because they do not show directly the amount of effort required. Some idea of the effort exerted can be formed by the observers who are noting the repetitions but this cannot be quantitative. In addition to the repetition rate and effort there is undoubtedly another factor which affects the users' opinion of the service. In conversing over a circuit having a poor transmission performance, annoyance or irritation may be felt by the users because the amount of effort required may be considered by them to be unreasonable. These factors, by their smallness or largeness, may lead the users in the course of their conversations to make favorable or adverse comments regarding the circuit performance. These comments can be noted by the repetition observers and used, together with any notations on effort and annoyance, to supplement the repetition rating in arriving at a better picture of the service.

Effective Transmission Ratings for Plant Design

To provide for the transmission design of the telephone plant along the lines of the previous discussion, ratings, termed "effective transmission" ratings, are being determined which are based on the repetition rate in normal conversations. Circuits of different transmission characteristics are considered to have the same effective transmission if their repetition rates are equal when they are used for the same kind of service. Furthermore, two changes in the transmission characteristics of a circuit are taken as equivalent on the same basis. The effects of such changes, however, are a function of the initial transmission characteristics and it is therefore desirable to take as a basis for rating such changes, a circuit which has characteristics typical in the various respects of the ranges encountered in practice.

As a standard reference circuit for determining an expressing effective transmission ratings, it is proposed to use a modification of the Master Reference System, inserting in this certain amounts of distortion, sidetone and noise to give it transmission characteristics comparable to those of present commercial circuits. Pending the development of this standard reference circuit, however, use will be made of a circuit consisting of station sets and instruments of kinds in general use, loops of typical length and construction connected by typical cord circuits to a trunk of specified transmission characteristics. For this latter it is convenient to assume a trunk having a cutoff typical of the loading systems in use and having a frequency characteristic which is flat below the cutoff point. It is also convenient to assume that the attenuation of this trunk can be varied uniformly for all frequencies below the cutoff point. This circuit may also be assumed to deliver at the two ends a typical amount of line noise and to have typical room noise at the terminals. Such a circuit then specifies a complete combination of transmission characteristics which are typical of the telephone plant in commercial use and may be considered as a working reference circuit. The transmission service performance of such a circuit in commercial use can be changed by varying the attenuation of the trunk and this attenuation, expressed in decibels with respect to some reference value, can thus be taken as constituting a scale for expressing different grades of service performance.

Starting with such a circuit, changes can be made in its transmission characteristics such as varying the attenuation of the trunk and its cutoff, varying the length and type of the subscribers' loops, using different types of transmitters and receivers in order to get different efficiencies and kinds of distortion and changing the type of station circuit to get different amounts of sidetone. By using circuits of these various characteristics in commercial service and determining the repetition rates obtained, a relation can be established between grade of service and transmission characteristics both for different overall circuit combinations and also for the various changes which can be made in such a circuit. An outstanding advantage of selecting the type of circuit which has been indicated, as a working reference circuit, is that it readily permits direct comparisons of the service performance of the working reference circuit, or of circuits having closely similar characteristics, with the service performances of various types of commercial circuits.

It is desirable to go one step further and to express the effects of changes in various transmission characteristics all in terms of changes in some one characteristic of the circuit. For this latter has been chosen the attenuation of the trunk. In accordance with this, then, the effect of such things as differences or changes in cutoff of the trunk, line noise, room noise, transmitter and receiver volume efficiencies and distortions, sidetone, and, in fact, of any transmission characteristics of any part of the circuit can each be expressed in terms of an equivalent change in the attenuation of the trunk on the basis of equality of effect on service performance. Thus the ratings of all such effects can be placed on a basis which makes them readily comparable. For the practical range of variations in these factors it has been found that in general the effects so expressed can be added together with a good degree of approximation. Where this is not the case, interrelated sets of effective transmission ratings can be supplied to cover the various typical combinations which are likely to be found in practice. This places the application of the ratings given by this method on a comparable basis with the application of the old volume ratings, that is, the assignment of a number to each part of the circuit, which numbers can be combined by algebraic summation in arriving at an overall rating for any particular circuit.

In line with this, the effective transmission of a trunk, for example, is rated in terms of an attenuation loss of so many db plus a rating in db which expresses the effect of the range of frequencies transmitted with respect to some range selected as standard, plus another rating expressed also in db to take account of the noise on this trunk. Similarly, loop loss curves can be drawn up for the combination of instruments, set, loop and cord circuit such as has been used in the past on a volume basis. On the new basis, these curves will include not only the ratings of volume losses but also the ratings for the distortions in the loop and instruments and the effect of the sidetone on transmitting and receiving. In this manner, the transmission design of the plant can be carried out in about the same manner as it has been on the volume rating basis but the effects of distortion, noise and sidetone can all be included in these effective transmission ratings which are based directly on service performance.

This in outline is the method of determining effective transmission ratings which is now being worked to, its method of formulation and its application. The complete discussion and description of these matters involves innumerable details which, as already stated, it is not the purpose to set forth here. From this outline it is seen that this method provides the following outstanding things:

1. A scale for indicating different grades of effective transmission, which scale is expressed in decibels and is directly correlated with service performance by means of a typical circuit selected as a reference. This permits the specification of grades of service.
2. The use of this same scale as a means of assigning to each element of practical telephone circuits an index, expressed in decibels, which measures its contribution to the effective transmission of the circuit, these indices being of such a nature that those corresponding to the elements in a circuit can be combined in a simple way to give an overall performance index for that circuit. Such a system of indices is necessary for plant design.
3. A means of correlating effective transmission service and circuit transmission characteristics. This correlation is advantageous in setting up the indices of (2) and in development and design work in determining the desirability of possible changes in the performance of the various elements.

The selection for the present of the typical practical circuit described above, as a working reference circuit, has two important advantages, which will be restated. First, by using a reference circuit having typical transmission characteristics, the indices established for changes in the various characteristics within the range of practical interest, are directly applicable to the present plant and can be combined in a simple manner to provide an overall circuit index. Second, and by no means of minor importance in the earlier stages of the application of the rating method here described, by using as a reference a practical circuit, it is possible and practicable to make direct comparisons of the service performance of the reference circuit, or circuits having closely similar characteristics, with the performances of various commercial circuits.

The maintenance of the first advantage will require, however, changes in the working reference circuit as material improvements are made in the transmission characteristics of the commercial circuits. To obtain the second advantage means the use at present of carbon transmitters in the working reference circuit. These are open to the same objection here as they were in the Standard Cable Reference System, namely, the difficulty of exactly specifying their performance raises questions as to the reproducibility of their performance from time to time. This was one of the major reasons for the replacement of the Standard Cable Reference System as the basis for volume ratings by the Master Reference System for Telephone Transmission with its specifiable performance. To preserve the first advantage mentioned and at the same time to obtain a reference system whose reproducibility can be assured, it is the purpose, as more complete correlations are obtained between transmission characteristics and service performance, to associate with the Master Reference System, the means to make its transmission characteristics meet the requirements necessary to retain the first advantage. Meanwhile the Master Reference System will continue its function as a reference for volume ratings.

Determination of Ratings

To provide the basis for such a system of effective transmission ratings as has been outlined, several series of tests have been made, the most comprehensive of which has been under way for more than a year between several hundred stations in the American Telephone and Telegraph Company headquarters building and a similar number of stations of the Bell Telephone Laboratories, between which there is a large amount of intercommunication. The connections between these stations are handled over special trunks in which the attenuation, cutoff frequency and line noise can be varied. At the stations, different types of instruments and station circuits have been employed. Observers are connected to each of these trunks who monitor the conversations over them and note the number of repetitions requested in each conversation and also the duration of the conversation. In this way is determined the repetition rate for a number of conversations between a number of different people for the various combinations of circuit characteristics so provided. Thus ratings are established directly of such effects as those of trunk cutoff, noise on the trunk, different types of transmitters and receivers and of variation of sidetone in the station set. In addition to the observations of repetitions, measurements are made of the talking levels on the trunks by means of volume indicators to determine the reaction of the circuit performance on talking levels.

An illustration of the results of such observation is given in Fig. 1. The curve shows the variation of the repetition rate with change in trunk attenuation for connections having the same kinds of terminal sets and loops at both ends. This then provides a means of expressing different grades of service performance in terms of trunk attenuation in this circuit.

On this figure is shown also the repetition rate obtained for trunks of two different effective upper cutoff frequencies. The change in trunk

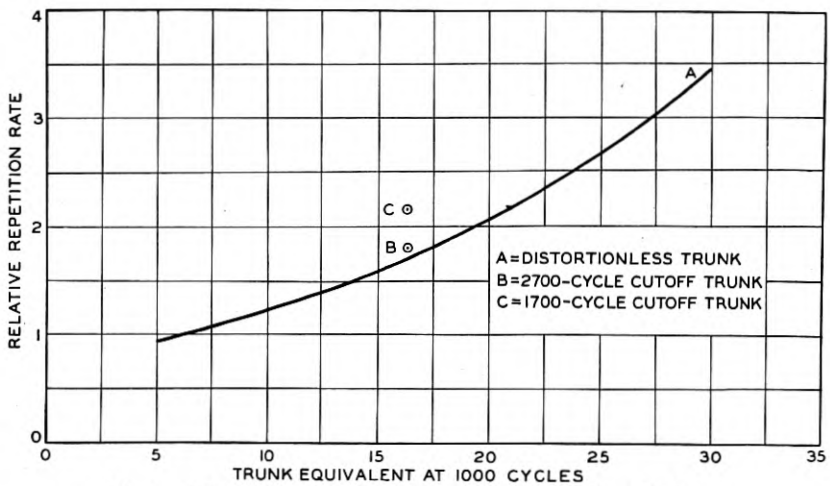


Fig. 1—Relation between repetition rate and trunk equivalent.

attenuation required to produce the same increase in repetition rate as that obtained in going to point C, for example, from the corresponding 1,000-cycle attenuation point on Curve A, is taken as the rating in db of the lower cutoff frequency represented by point C. This rating is about 5 db. The rating of point B with respect to A is obtained in a like manner to be about 1 db and correspondingly the rating of the cut-off frequency of C with respect to B is about 4 db. This illustrates the manner of obtaining effective transmission ratings for any change from the characteristics of the circuit of Curve A.

It is obviously laborious to cover the ranges of all the transmission characteristics of circuits of this kind. The idea is to establish points which will cover the practical range and to use the results of articulation tests and other similar measurements for interpolating between the points established by the repetition method. In this way it is planned to put the rating of transmission on a basis which indicates the effect on service of changes in the various parts of the circuit.

Conclusion

In concluding, it may be restated that the primary purpose here has been to discuss the rating of the transmission performance of telephone circuits and the method which is being adopted in the Bell System for determining and expressing effective transmission ratings for the design of the plant. The salient features of this method which should be emphasized are the following:

1. In establishing the rating of the transmission performance of a telephone circuit, its performance is taken as that obtained when the circuit forms part of the combination of talker, circuit and listener, where the talker and listener represent the users of the telephone system in commercial service.
2. The ratings of the effective transmission of circuits are based on the rate of repetitions required.
3. The ratings of effective transmission will eventually be referred to a modification of the Master Reference System arranged with typical distortion, sidetone and noise. For a working reference circuit, use is made of a circuit which has transmission characteristics typical of those encountered in service. The trunk of this circuit is taken as adjustable in attenuation for the purpose of providing a scale for specifying different grades of overall transmission performance and also for expressing ratings of the effect on transmission performance of the various transmission characteristics of circuits and their parts.

Paragutta, A New Insulating Material for Submarine Cables *

By A. R. KEMP

Gutta percha and balata have proven eminently suitable for the insulation of long deep sea telegraph cables, but their dielectric losses are too high to meet the requirements of submarine telephone cables designed to operate over long distances or of shorter cables employing carrier currents.

This paper describes a new material called paragutta which has been developed to meet the present needs. It consists essentially of the purified hydrocarbons of balata (or gutta percha) and of rubber together with minor quantities of waxes to modify the mechanical characteristics. The purification of rubber particularly with respect to nitrogenous constituents is necessary to effect electrical stability in water. A commercially usable method of purifying rubber is described.

Evidence is furnished that paragutta has all of the desirable thermoplastic and mechanical properties of gutta percha while possessing such superior insulation characteristics as to make it suitable for use on long cables designed for transoceanic telephony. Its use is also advantageous on shorter deep sea cables designed for carrier telephony as well as for ocean telegraphs.

FORMERLY deep sea cables were used exclusively for telegraph purposes but in recent years there has been an increasing use of this type of cable for telephone service. Telephonic communication requires cables of very much superior transmission quality to that needed for telegraph. At the higher frequencies of voice transmission the energy losses in the insulating material become a serious factor and a radical improvement in submarine insulation is called for.

The longest existing deep sea cables operating at voice frequency only slightly exceed 100 miles and the construction of a transoceanic telephone cable with standard materials has been regarded as beyond the practical limits of feasibility.

The installation and rapid expansion of transatlantic radio telephony during the past few years have created a need for a deep sea telephone cable to supplement this service, particularly during periods of atmospheric disturbances. In addition the development of carrier telephony offers possibilities for increasing the traffic over shorter submarine cables. For the shorter cable, the still higher frequencies of carrier telephony make demands upon the insulating material similar to those of long cables operating at voice frequency.

In view of these circumstances an extended study was undertaken of the causes of losses and other electrical weaknesses of submarine insulation and a search has been made for better materials. As a

* *Jour. Franklin Institute*, Jan., 1931.

result of this investigation an insulation called paragutta has been developed which, as the name suggests, is derived essentially from rubber and gutta percha. It is the purpose of this paper to describe this material and give an account of the tests to which it has been subjected to determine its suitability for the purpose.

By virtue of its superior electrical properties, the use of paragutta in place of gutta percha for the insulation of telephone and telegraph cables offers advantages either from the standpoint of improved transmission or the economies in materials of construction which can be made as a result of modified design.

Gutta percha and balata have been the standard materials for the insulation of deep sea cables since the inception of the submarine cable industry some seventy-five years ago. Although these substances are inadequate for modern telephone needs as regards their electrical characteristics, their mechanical properties are peculiarly adapted to submarine insulation. This is so much the case that gutta percha can fairly be taken as a model which must be closely imitated in respect to mechanical characteristics by any successful substitute. This is fortunate since the use of any substitute which differs radically from gutta percha would mean discarding large existing investments in special technique, equipment and trained personnel, and would involve serious risks as to the integrity of cables made with the new material. It may be remarked in passing that no manufacturing process requires a higher degree of insurance against occasional defects than does the submarine insulation art, a fact that has engendered a strong conservatism in the industry.

Because of its almost ideal mechanical properties, the requirements for submarine cable insulation may conveniently be described by reference to gutta percha. Gutta percha insulation, which often includes more or less balata in its composition, is made of raw materials carefully selected for quality, which are thoroughly washed and extremely uniformly blended. The thermoplasticity of the material is of great service in these operations and further permits it to be readily extruded onto a conductor in multiple layers in a continuous sheath with great exactness and freedom from mechanical defects. After being forced around the conductor the material quickly sets to a hard, tough covering when drawn through cold water. Its firmness and toughness are essential to resist subsequent handling operations in the factory, as well as those involved in laying, picking up and repairing. The warm, soft material adheres readily to the conductor and is well adapted to the making of joints in the insulation between core lengths both in the factory and on the cable ship.

In addition to those excellent and unique mechanical properties, gutta percha possesses electrical characteristics peculiarly adapted to submarine cable construction. Its outstanding electrical merit consists in the fact that its electrical characteristics are stable under sea bottom conditions over a great many years.

Gutta percha is obtained from the latex of a large number of species of trees growing wild in the forests of the Malay Peninsula and the East Indian Islands. The products of the various species of trees are by no means of equal value, varying as they do in the content of hydrocarbon, resins, moisture and other substances. Since the material is gathered and worked up upon the spot by primitive people a great deal of carelessness as well as deliberate adulteration is practiced and the material comes upon the market in a dirty condition and in a bewildering variety of forms which almost prohibit effective inspection, standardization and grading.

The essential constituent of gutta percha is an unsaturated hydrocarbon of colloidal nature which is similar in its chemistry to rubber. It is this constituent which makes gutta percha plastic when warm and tough when cold, and which contributes most conspicuously to its electrical excellence as an insulator. The usual gutta percha insulation is the result of blending and washing various grades of crude gutta percha to remove dirt and water soluble components. The hydrocarbon, resin, dirt and moisture contents as determined by analysis of the crude material together with the electrical and mechanical properties after washing are the principal characteristics used to determine whether or not a particular grade of crude gutta percha is suitable for use as submarine cable insulation. The hydrocarbon content of gutta percha insulation when applied to the conductor is usually about 60 per cent, the remainder being mostly the natural resins together with small amounts of very finely divided dirt (humus) and residual moisture. The proteins or albumens in crude gutta percha and balata are almost completely removed by simple washing.

Balata comes from two species of trees of the same general botanical family as gutta percha, but is native to the forest regions of upper South America and is unknown in the gutta percha producing area of the Far East. The latex of the balata tree is more fluid than that of gutta percha, which permits the trees to be tapped and the fluid to be collected at a central point in the forest, where the product from various trees is mixed for recovery of the gum. Because of the small number of species involved and the transportability of the fluid latex, balata is produced in a much more limited number of grades and is cleaner and more dependable as to uniformity of quality. Its essential constituent is the same hydrocarbon which gutta percha contains. In addition to the hydrocarbon, there is present in balata some 40 per cent of resins and amounts of dirt, moisture and other impurities which usually total about 15 per cent. The resins of balata are softer than those of gutta percha and make the product in its raw state a little less desirable than the better grades of gutta percha from the mechanical standpoint. Balata, however, contains a smaller amount of finely divided dirt or humus than gutta percha, which is reflected in its superior electrical characteristics and lower water absorption.

The resins of both gums have been usually included with the hydrocarbon in making submarine insulation. Sometimes, however, a portion of the resins are removed, partly to increase the toughness and partly to improve the electrical characteristics.

There are several methods which may be used for preparing gutta hydrocarbon nearly free from resinous substances. One of these methods involves dissolving the balata or gutta percha in warm petroleum naphtha, filtering the solution from dirt and precipitating the gutta hydrocarbon from solution by refrigeration, leaving most of the resins in solution. A simpler and less expensive method, however, is that of leaching out the resins by simply soaking the sheeted or finely cut material in a suitable grade of petroleum naphtha at ordinary temperature, followed by draining off the solution of resins and finally evaporating the residual solvent from the extracted material.

The completely deresinated hydrocarbon from either source is not suitable for use alone as submarine cable insulation because insufficiently plastic at safe working temperatures, as well as prohibitively expensive. Otherwise the complete deresination of these products would be highly advantageous as, for example, is indicated by the superior electrical characteristics of deresinated balata shown in Table I. A substantial amount of experimentation upon the methods of refining balata has been necessary to secure the excellent electrical characteristics therein indicated but no revolutionary innovation has been necessary.

TABLE I

EFFECT OF RESIN CONTENT ON THE ELECTRICAL CHARACTERISTICS OF BALATA

Material	Electrical Characteristics 0° C., 1 Atm., 2000 Cycles Dielectric Constant	Specific Conductance Unit = 10^{-12} mho. cm.
Balata	3.1	66
Deresinated Balata	2.6	3
Balata Resins	3.3	52

In attempting to develop a new insulating material for deep sea cables it seemed best to begin with gutta hydrocarbon as a basis, since its mechanical properties are so unique, rather than to attempt to synthesize a new chemical compound which would imitate it. In order to overcome the excessive stiffness of the pure gutta hydrocarbon, as well as its prohibitive cost, it was determined to attempt to blend large quantities of rubber with it, since rubber is the nearest kindred material and is commercially available at low cost. There resulted thermoplastic products of fairly good mechanical characteristics which, however, proved to be insufficiently stable electrically.

Meanwhile, a thorough study was being made of the electrical and physical characteristics of rubber and particularly of the causes of its electrical instability upon prolonged immersion in water. Our hope that such a study would not only reveal the nature of the defects of rubber but also suggest means for remedying them has been realized to a gratifying degree.

Rubber, as is well known, is also derived from the latex of certain trees, chiefly *Hevea Brasiliensis*. This tree has been cultivated in large areas on the plantations in the Far East and the product is obtainable commercially in excellently standardized grades. Its principal constituent is a hydrocarbon scarcely distinguishable from that of gutta percha by chemical means, but radically different from it in physical properties, notably in that it has but a slight degree of thermoplasticity and is far more distensible in the cold state. Aside from the hydrocarbon, rubber also contains small amounts of resins, proteins and other impurities, but the aggregate non-hydro-carbon constituents in the better grades are usually less than 10 per cent in contrast to 50 per cent or thereabouts for gutta percha and balata.

Rubber is used almost exclusively in industry in a vulcanized form, that is, in combination with a small percentage of sulphur. In this form rubber has also been used to a limited extent for submarine cable insulation, but has long been recognized as lacking sufficient electrical stability for deep sea cables designed to carry a heavy traffic. It is still used to a considerable extent with a fair degree of success for insulation on short cables where the electrical requirements are not severe. In tropical waters it has the advantage over gutta percha of greater resistance to teredo attack and to damage by high temperature.

Some years ago an extended study [1] was made of the causes of the electrical instability of vulcanized rubber, which led to the conclusion that the water soluble impurities are largely responsible. These impurities can be removed comparatively readily and satisfactorily in the process of manufacture, and a submarine insulation of a fair degree of stability is thereby attained.

[1] Williams and Kemp, *Jour. Franklin Inst.*, **230**, 35 (1927).

Even so, vulcanized rubber is very inferior to gutta percha for submarine insulation as the necessary manufacturing operations are more difficult and likely to lead to defects. The removal of mechanical impurities is by no means simple because the raw stock is not plastic enough for thorough straining. The lack of plasticity also interferes with multiple covering of conductors, and the process of heating to bring about vulcanization is liable to result in deformation of the insulating layers. The joining and repairing of core lengths insulated with rubber is also more of a problem than with gutta percha, which can be so readily remolded in case imperfections appear in the course of the process.

The methods of electrical stabilization of vulcanizable rubber compositions are only partially effective in the absence of vulcanization and it was therefore necessary to extend the study in an effort to secure the desired electrical properties in rubber in the raw state. It might be supposed that mere admixture of raw rubber with gutta hydrocarbon would produce the necessary stability. This is true only to a limited extent. When the proportions of rubber are high enough to meet the mechanical and economic requirements, the electrical stability is impaired.

Effect of Proteins on Electrical Stability of Crude Rubber Immersed in Water

It has been previously shown that crude rubber contains considerable water soluble impurities and that their removal results in a large reduction in water absorption.[1,2] Rubber so prepared absorbs no more water than good cable gutta percha but in a raw state when immersed in water, it fails sooner or later as an insulator, often suddenly and completely.

To determine the reason for this electrical instability of crude rubber in water, samples of very pure rubber hydrocarbon completely freed from proteins, resins and other impurities were prepared and tested. It was found that this material not only absorbed very little water but showed practically no change in electrical characteristics as a result of prolonged immersion in water. The impurities natural to rubber therefore seem to be responsible for its instability.

It has been known for many years that crude rubber contains proteins, ordinary plantation rubber containing about 3 per cent. Previous investigators have postulated and shown considerable indirect evidence to the effect that the rubber globules in rubber latex have an adsorbed film of protein around them and that this condition

[2] Lowry and Kohman, *Jour. Phys. Chem.*, **31**, 23 (1927).

also exists in crude rubber. It is also known that latex serum contains a substantial quantity of protein in solution. The preparation of crude rubber from latex by addition of acid or by processes of evaporation of the water by heat undoubtedly results in the precipitation of considerable quantities of this protein which becomes entrapped between the globules as they coalesce. It is easy then to visualize that in crude rubber there exists a continuous phase of protein or a protein network which, acting like most protein matter, absorbs large quantities of water, resulting in paths through which electrical conduction occurs.

Removal of Nitrogen Containing Bodies from Rubber

The problem of developing a suitable commercial method for preparing rubber free from nitrogenous matter offered many apparent difficulties. The proteins are colloidal in nature and in the presence of water form gelatinous masses rather than true solutions. On this account they often cannot be removed by simple washing as can be done in the case of gutta percha and balata. It has been known for some time that proteins can be broken down to water soluble products by boiling with dilute hydrochloric or sulphuric acids. This treatment did not produce satisfactory results in the case of rubber. As a result of many experiments involving a variety of methods, it was found that heating rubber in an autoclave at an elevated temperature in the presence of water alone fairly rapidly brought about the desired hydrolysis of the rubber proteins, converting them to water soluble materials. As a result of subsequent washing, it was found that the nitrogenous bodies had been almost completely eliminated without deleterious effect on the rubber hydrocarbon.

Rubber either in the form of sheets immersed in water or as an aqueous rubber dispersion such as latex can be employed in the process. The treatment of latex, however, results in a more rapid hydrolysis of the proteins. Considerable latitude exists in the choice of conditions, but the following example will suffice to describe one method of carrying out the process: ammonia preserved latex is diluted 1 to 5 with pure water. The latex is then heated in an autoclave for approximately ten hours at 150° C. After cooling it is coagulated with acetic acid and thoroughly washed. As a result of this treatment the nitrogen content of the rubber is found to be less than 0.10 per cent, which is about one fourth that of ordinary plantation crude rubber. Figures 1, 2, 3 and 4 illustrate the relative water absorption and electrical stability of deproteinized rubber as compared with the ordinary crude product. Vulcanized deproteinized rubber was also found to be somewhat superior to ordinary vulcanized crude rubber as regards its electrical stability in water.

In addition to the superior electrical stability of deproteinized rubber, it was found to be more readily plasticized and mixed with gutta

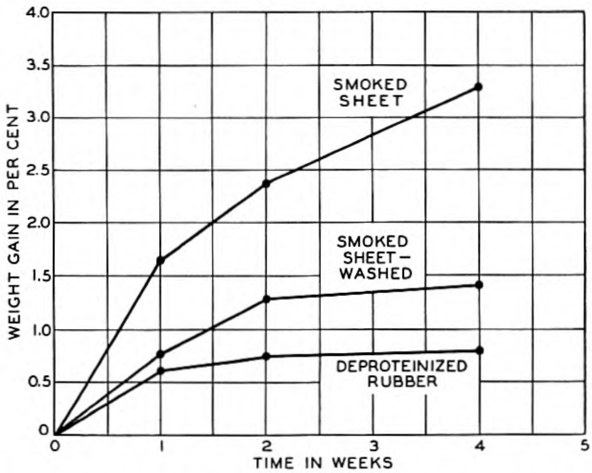


Fig. 1—Effect of washing and removal of protein on the water absorption of crude rubber when immersed in 3.5 per cent NaCl solution at room temperature.

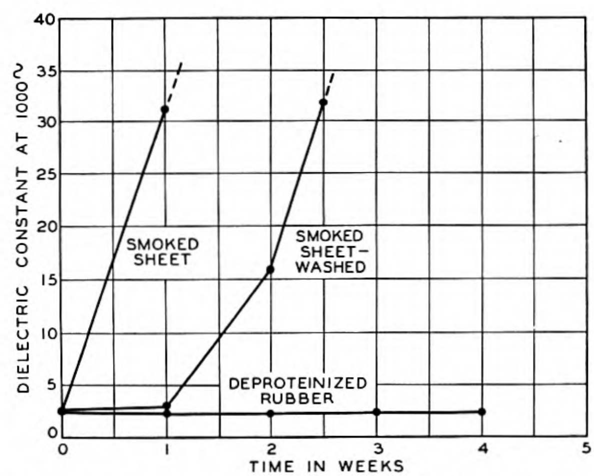


Fig. 2—Effect of washing and removal of protein on the dielectric constant of crude rubber when immersed in 3.5 per cent NaCl solution at room temperature.

than is the case with crude rubber, thereby yielding a product with better thermoplastic properties.

Preparation of Paragutta

As previously stated, the principal constituents of paragutta are deproteinized rubber and purified gutta hydrocarbon. Specially treated hydrocarbon or montan waxes may also be added as a third constituent to modify mechanical properties and reduce cost. The proportions of these constituents may be varied over a wide range to achieve the desired characteristics, but in general rubber and gutta

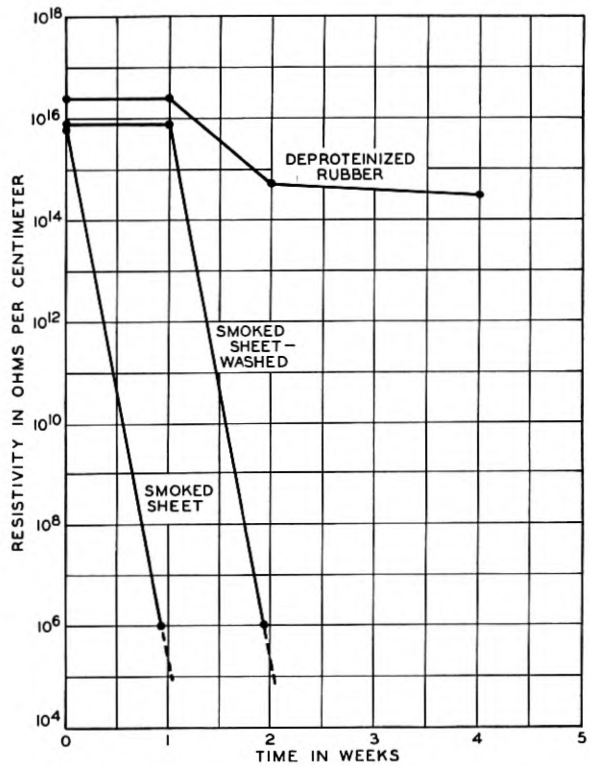


Fig. 3—Effect of washing and removal of proteins on resistivity of crude rubber, when immersed in 3.5 per cent NaCl solution at room temperature.

are used in about equal proportions and purified montan wax may be added up to about 40 per cent. Superior electrical properties, however, result from the use of hydrocarbon waxes, which may be added in amounts up to about 20 per cent. By the proper blending of these materials, a thermoplastic insulation is obtained which closely approximates gutta percha in mechanical properties and is fully its equal as to electrical stability in water. Its specific electrical characteristics represent a substantial improvement over those of the classical insulation compounds and its cost is lower.

The final steps in processing paragutta are very similar to those used for gutta percha and involve blending and washing the deproteinized rubber and deresinated balata or gutta together, masticating to remove excessive water and at the same time incorporating such waxes as are found necessary. The material is then strained through fine sieves under hydraulic pressure to remove adventitious impurities, kneaded to remove air and finally placed on the covering machine rolls to be forced around the conductor. The machinery in use for pro-

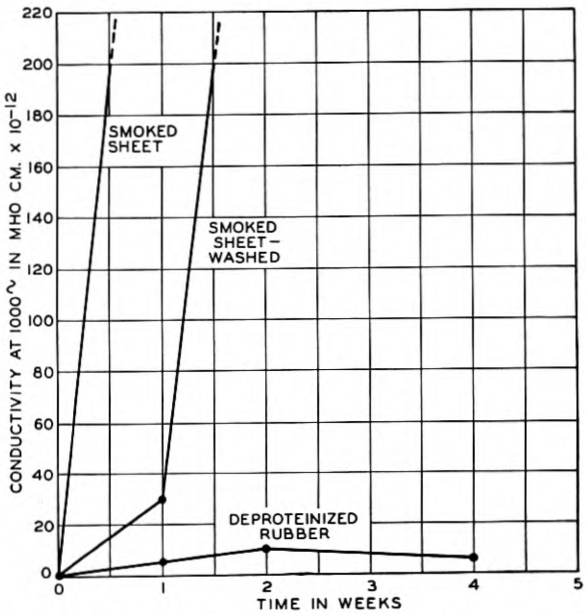


Fig. 4—Effect of washing and removal of proteins on conductivity of crude rubber when immersed in 3.5 per cent NaCl solution at room temperature.

cessing gutta percha is suitable for handling paragutta in these operations.

Comparative Properties of Paragutta and Gutta Percha

Tensile Properties: Although submarine insulation is not subjected to tensile deformation in practice, tensile properties indicate to some degree the relative mechanical suitability of a given material for the purpose. Figure 5 shows the stress-strain characteristics of paragutta and gutta percha submarine cable insulation. These results show that paragutta has tensile properties equal to cable gutta percha although its gutta content is substantially lower.

Compression Properties: The insulated submarine cable conductor commonly known as the core is frequently subjected to uneven compression stresses during manufacture, laying and repairing. The insulation must therefore be capable of withstanding these stresses without appreciable deformation. To determine the relative merits of paragutta and gutta percha in this respect their comparative stress-strain characteristics under compression have been measured, using a special compression machine,[3] and are shown in Fig. 6. In this test

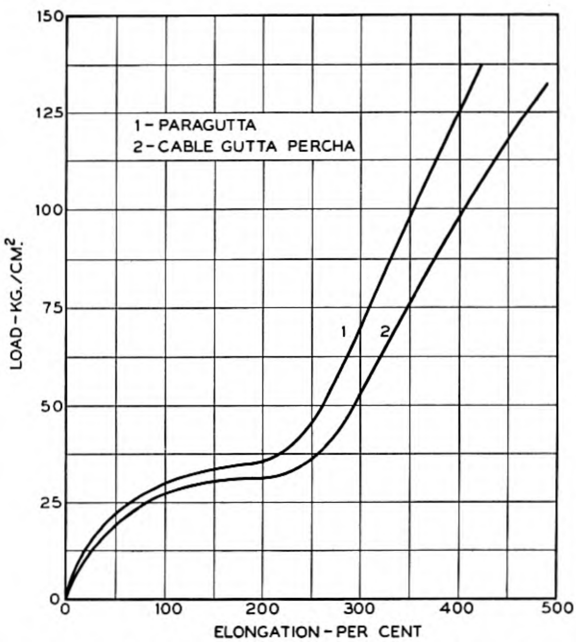


Fig. 5—Comparative tensile properties of paragutta and gutta percha at 25° C.

a steel rod 1.6 cm. in diameter was forced endwise into a sheet of the material .375 cm. in thickness at a rate of about 4 cm. per minute while simultaneously recording the deformation and load. These results show that very little difference exists between these materials in this test, and factory handling of cores confirms the general conclusion.

Flexibility: The flexibility of submarine cable insulation is important because the core is subjected to considerable flexing during manu-

[3] Hippensteel, *Bell Laboratories Record,* **5**, 153 (1928).

facture, laying and repairing and possibly at times during use, especially where tidal currents may cause movement in the cable. Paragutta and gutta percha cores have been subjected to slow and continuous flexing at 0° and 25° C. for long periods and it was found that both materials will withstand millions of repeated flexures at small amplitudes without failure. When the amplitude of flexure was increased to strain the conductor slightly beyond its elastic limit, the conductor always failed in advance of the insulation.

Plasticity Tests: Laboratory tests were made to determine the relative plasticity of paragutta and gutta percha, using both the Williams [4]

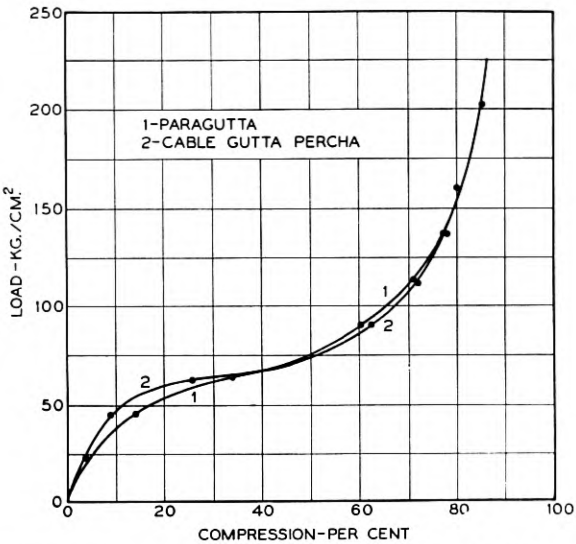


Fig. 6—Comparative compression properties of paragutta and gutta percha at 25° C.

and the Marzetti [5] type of plastometers. These tests are valuable guides but the final judgment of a material as regards thermoplasticity was made by determining its workability on commercial gutta percha insulating machines. Paragutta is somewhat more resistant to flow than gutta percha at temperatures ranging from about 40° to 70° C. When applied to the conductor, however, its greater resistance to flow at elevated temperatures can be taken as an advantage as it lessens the danger of faults occurring if the core should be accidentally exposed to elevated temperatures or to conditions which might exist in connection with cable used in the tropics.

[4] Williams, *Jour. Ind. & Engg. Chem.*, **16**, 262 (1924).
[5] Marzetti, *Giorn. Chim. Ind. Applicata*, **5**, 342 (1923).

Figure 7 shows the relative plasticities of cable gutta percha and paragutta at several temperatures as determined by the Williams [4] method, which can be taken to indicate the relative plasticities of these materials at working temperatures.

Brittle Temperature: It is extremely important that the temperature at which submarine cable insulation becomes brittle should be far below the range of sea bottom temperatures to be encountered in use. This is one of the properties in which rubber and gutta percha greatly excel any other available insulating material. Kohman and Peek [6]

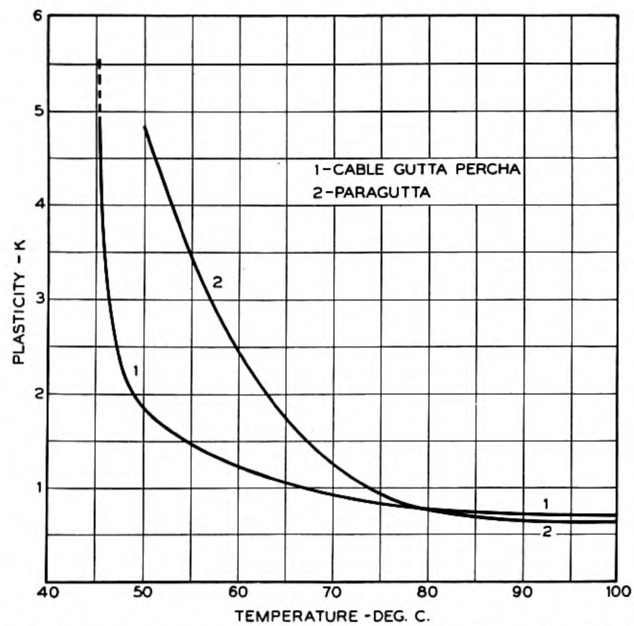


Fig. 7—Effect of temperature on the plasticity of cable gutta percha and paragutta.

have described an apparatus for accurately determining this temperature. The brittle temperature of paragutta is somewhat lower than cable gutta percha, as can be seen from the results in Table II, which give the range of brittle temperature values found for different samples of several materials.

Water Absorption—Electrical Stability

The amount of water absorbed by rubber and gutta percha when immersed in water is the result of a complicated mechanism. The quantity and nature of water soluble or water absorbing impurities

[6] Kohman and Peek, *Jour. Ind. & Engg. Chem.*, **20**, 8 (1928).

TABLE II

BRITTLE TEMPERATURE OF PARAGUTTA AND OTHER INSULATING MATERIALS

Material	Brittle Temperature ° C.
Gutta Percha (Cable Insulation)	−23 to −36
Paragutta	−45 to −61
Balata (Washed)	−44 to −52
Balata (Washed and Deresinated)	−62 to −67
Crude Rubber	−57 to −58
Vulcanized Rubber (Soft)	−53 to −58

in the rubber or gutta percha and the salt concentration of the water in which the samples are immersed are controlling factors. The enormous increase in the quantity of water absorbed by ordinary rubber when immersed in distilled water as compared with its absorption in salt solutions has been explained on the basis of osmotic theory.[1] In accordance with this theory rubber acts as a semi-permeable membrane. Water soluble crystalloids or hydrophillic colloids (proteins) attract the water which enters the rubber by diffusion. When immersed in distilled water these impurities tend to reach infinite dilution with water, being opposed in this by the resistance of the rubber itself to swelling. In salt solutions the amount of water absorbed is finite and depends on the equalization of osmotic pressures of the internal and external solutions. The change in water absorption of pure rubber hydrocarbon with the salt concentration of the external solution is small over the whole range, which indicates that the water enters by a process of solution. This has also been found to be the case for gutta hydrocarbon and is more or less true for paragutta and gutta percha. The water absorption in distilled water can therefore be taken as a measure of the freedom from water soluble or water absorbing impurities. Figure 8 shows the effect of NaCl concentration in the immersion solution on the quantity of water absorbed by samples of rubber, paragutta and gutta percha at room temperature. Samples of rubber containing water soluble matter or proteins do not readily reach an equilibrium water content in distilled water. Crude rubber has been found to absorb more than 100 per cent water in distilled water at ordinary temperature without reaching equilibrium.[1] Gutta percha, paragutta and pure rubber hydrocarbon on the other hand reach a definite and lower equilibrium water content in distilled water, which shows their greater freedom from water soluble or water absorbing matter.

As the electrical stability of paragutta in sea water is of paramount importance an exhaustive study has been made on a large number of specimens as regards their changes in electrical values over long periods of immersion in 3.5 per cent salt solution. Gutta percha insulation contains about one per cent water when at equilibrium with sea water whereas paragutta contains somewhat less than this amount. These values have been determined by testing samples made up with various water contents below and above equilibrium values and determining the water content after prolonged immersion in 3.5 per cent NaCl solution, as seen in Figure 9. The equilibrium value is practically the same when equilibrium is approached from either direction.

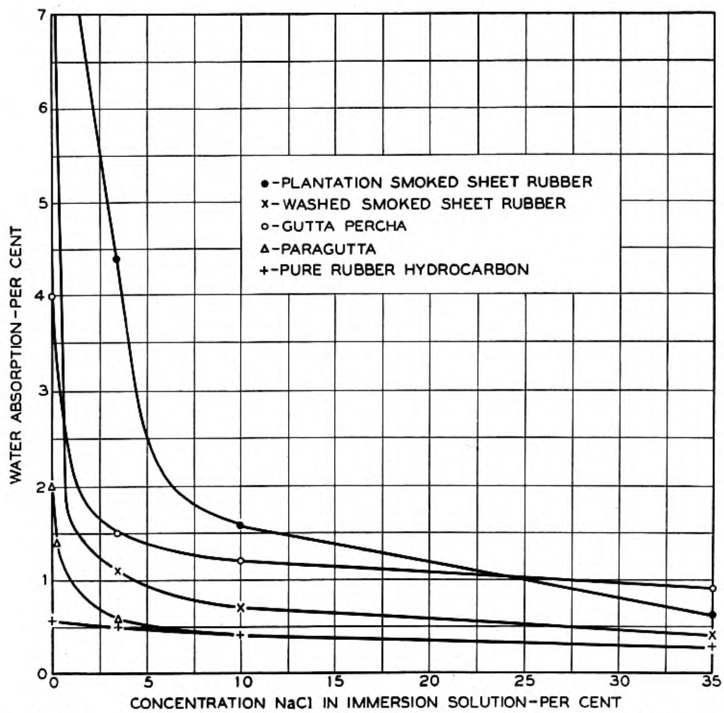


Fig. 8—Relation of water absorption to salt concentration in immersion solution.

The overall quantity of water absorbed, however, cannot be used as a final criterion by which to judge insulation for it has been previously shown (Figs. 1 to 4) that washed crude rubber completely fails as an insulator after absorbing less than one per cent water. The mode of distribution of water absorbing impurities in an insulating material has been found to be of utmost importance as regards the magnitude of the effect of moisture in various insulating materials. Examples where large effects on insulating properties are caused as a result of moisture absorption by localized impurities are found in the above case of proteins in crude rubber, water soluble salts associated with fillers in vulcanized rubber [1] and hygroscopic salts on the surfaces of textile fibers.[7]

On the other hand, the electrical properties of paragutta or gutta percha are not impaired when several times their equilibrium water content is incorporated with them. Gutta percha, however, does show an increase in capacitance of about 10 per cent as a result of water absorbed by a completely dried specimen, but as it is always the practice to apply it to the conductor in a wet condition this

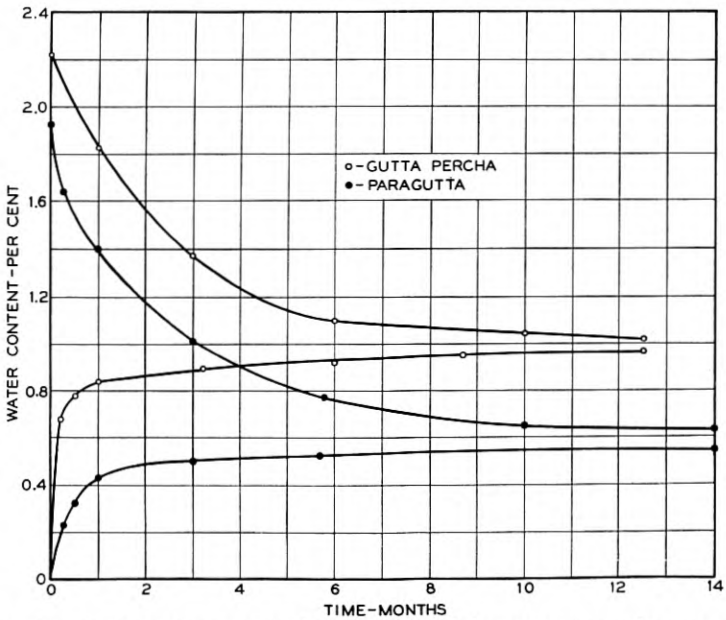


Fig. 9—Changes in water content of 50 mil wet and dry paragutta and gutta percha sheets when immersed in 3.5 per cent NaCl at room temperature.

change is not of practical significance. The electrical properties of paragutta on the other hand show practically no changes as a result of moisture absorption by a dry sample. These facts are taken to be the best evidence of the electrical stability of paragutta in contact with water.

Hundreds of specimens of paragutta and gutta percha have been studied as regards changes taking place in electrical characteristics after long periods of continuous immersion in 3.5 per cent salt solution. These tests, some of which have been for periods of three to five years, show that paragutta is fully equal to gutta percha as regards its

[7] Williams and Murphy, *Bell Sys. Tech. Jour.*, **8**, 225 (1929).

stability. When properly prepared both of these materials show practically negligible changes in electrical properties as a result of prolonged submergence in water. Sea bottom conditions are even less likely to affect these materials than those existing in the laboratory. This is because of the absence of light, limited oxygen supply and low temperature, all of which reduce the tendency of materials such as paragutta or gutta percha to oxidize or otherwise deteriorate. It has also been shown [2] that the low temperature and high pressure existing at sea bottom reduce the rate of water absorption but do not materially affect the amount absorbed.

Electrical Characteristics: The electrical properties of paragutta depend upon the particular composition chosen, the quality of the raw materials and the care exercised in processing them. For long telephone cable insulation, it is necessary to exercise the utmost care to obtain a material having dielectric constant and specific conductance values sufficiently low to reduce to the minimum its effect on the attenuation. On the other hand, for ordinary telegraph cables these values are less critical and it may be advantageous to modify the practice for purposes of economy. Representative values for the electrical properties of a superior grade of paragutta and typical cable gutta percha under sea bottom conditions are given in Table III. It will be seen in this table that paragutta has a 20 per cent lower dielectric constant and a specific conductance one-thirtieth that of ordinary cable gutta percha under sea bottom conditions. The insulation resistance and dielectric strength of the two materials are practically the same.

TABLE III

Comparative Electrical Properties of Paragutta and Cable Gutta Percha at Sea Bottom Conditions

	Specific Inductive Capacity 2° C., 400 Atm., 2000 Cycles	Effective A-C Conductivity 2° C., 400 Atm., 2000 Cycles Unit = 10^{-12} mho. cm.
Cable Gutta Percha	3.3	90
Paragutta	2.6	3

Acknowledgment

The author wishes to acknowledge his indebtedness to Mr. R. R. Williams for counsel and assistance during the prosecution of the work and writing of the paper.

Abstracts of Technical Articles From Bell System Sources.

An Efficient Loud Speaker at the Higher Audible Frequencies.[1] L. G. Bostwick. This paper describes a loud speaker designed for use as an adjunct to existing types for the purpose of extending the range of efficient performance to 11,000 or 12,000 cycles. A moving coil piston diaphragm structure is used in conjunction with a 2000–cycle cutoff exponential horn having a mouth diameter of about 2 inches. Motional impedance measurements on this loud speaker indicate an average absolute efficiency of about 20 per cent within the frequency range from 3000 to 11,000 cycles. The variation in response within this band does not exceed 5 db. By using a high-frequency loud speaker of this type the efficiency and power capacity of the associated low-frequency loud speaker can be improved and a uniform response-frequency curve from 50 to 12,000 cycles can be obtained.

Results of Noise Surveys. Part I. Noise Out-of-Doors.[2] Rogers H. Galt. The purpose of a noise survey of a locality is to study the space and time distribution of noise intensity, the frequency composition of the noise, the contributions of various noise sources, the relation between the annoyance effect of the noise and its physical and auditory characteristics, and the effectiveness of methods of noise reduction. The extent to which each of these phases of the noise problem has been investigated heretofore has depended upon the point of view of the investigator and upon the apparatus employed. From one standpoint or another, any audible sound may fall within the category of noise; hence the variety of possible noise surveys is almost unlimited. Not many such surveys have been carried out, however, partly because the appropriate apparatus is of recent development; nor has any extensive comparison been published between the results obtained in different places and with different instruments. It has therefore seemed worth while to assemble such previously published results as are available, and certain new observations, in the present series of papers, of which this paper deals with noise out-of-doors.

Microphonic Action in Telephone Transmitters.[3] F. S. Goucher. This semi-technical article gives a brief resumé of the theories of microphonic action and describes the results of some experiments on the

[1] *Jour. Acous. Soc. Amer.*, July, 1930.
[2] *Jour. Acous. Soc. Amer.*, July, 1930.
[3] *Science*, Nov. 7, 1930.

contact behavior of granular carbon of the type used in commercial microphones.

A technique is described whereby contacts—either singly or in groups—may be studied under contact forces of the order of 1 dyne.

Through a study of the temperature coefficients of resistance of such contacts it is possible to conclude that the conducting portions of the contact junctions are of the nature of carbon and that new contact points are established or broken when the resistance is varied in a reversible resistance force cycle.

The experiments show that for such reversible cycles the relation between the resistance and force is of the approximate form $R = K(F)^{-n}$. The exponent n varies considerably from cycle to cycle but its average value depends on the force limits. The largest values of n are obtained with the aggregates of granules under such conditions of force limits that the elastic strains must be relatively large. A maximum mean value substantially independent of the force limits over a wide range closely approximates the value 7/9.

This value 7/9 is the maximum given by a theory of contact resistance worked out by F. Gray, assuming that the contact is made between two spheres of conducting material having surface roughness equivalent to an assembly of minute spherical hills. On account of the elasticity of the material both the microscopic area of contact between the spheres and the microscopic areas of contact between the hills increase with contact force. A strained aggregate of granules may therefore be made to behave like an ideal single contact between spheres having a rough surface.

For single contacts and for aggregates at small strains the value of n falls below the minimum value 1/3 which is accounted for by theory. This is associated with internal contact forces, or cohesion, which render the contacts relatively insensitive to changes in the applied force. The existence of cohesion is readily demonstrated by the fact that contacts always require a finite force to break them even when no current has passed through the contact.

The Architecture of Living Cells—Recent Advances in Methods of Biological Research—Optical Sectioning with the Ultra-Violet Microscope.[4] F. F. LUCAS. In previous papers of the past few years the development and application of the ultra-violet microscope to the science of metallography have been described.

Metallography, at first thought, appears wholly unrelated to histology or other branches of biology but the two branches of science do

[4] *Proc. Nat. Acad. of Sciences*, Sept., 1930.

have many points in common. Both deal in the last analysis with the structure of matter and, in each, the microscope is an indispensable tool. Improvements in microscopic vision which enlarge the world of vision in one branch of science inevitably have a reflection in the other.

It is not the purpose of this paper to enter into a discussion of structures of living cells as revealed by the ultra-violet microscope. More particularly, the object is to present a tool for biological research; a tool which enables us to photograph the structure at different planes or levels within a single cell or group of cells; one which enables us to see the living cell with a degree of precision and clarity not heretofore possible by any other known means and with a potential resolving ability at least twice that of the best apochromatic system using visible light.

Production of Plastic Molded Telephone Parts.[5] A. M. LYNN. The Western Electric Company now manufactures for Bell System apparatus a large number of different phenol-plastic, shellac, and hard-rubber molded parts, the output of which varies from a few thousand to several million per year. The majority of these molded parts are produced in comparatively small quantities, but certain of them, such as the phenol-plastic molded parts used in the hand-set type of telephone, a new molded subscriber's set housing, and the receiver shell, cap, and mouthpiece used on the older type of desk-stand telephone, are heavy-running parts. The tools and press equipment used in the production of these parts are described in this paper.

Variation of the Inductance of Coils Due to the Magnetic Shielding Effect of Eddy Currents in the Cores.[6] K. L. SCOTT. An analysis is made of the shielding effect of eddy currents on the flux in the interior of cores of cylindrical or flat sheet material. It is shown that the counter voltage of self inductance of an iron-cored coil is due only to the component of flux in the core which is in phase with the flux at the surface of the core. Expressions are obtained and curves plotted showing the variations of inductance of a coil with frequency, or with the conductivity and permeability of the core material. Sample calculations and some experimental results are given. The results show that the inductances at high frequencies are actually less than the predicted values, which leads to the suspicion that some factor other than eddy currents causes the flux in the interior of the cores to decrease with increasing frequency.

[5] *Mech. Engg.*, Oct., 1930.
[6] *Proc. I. R. E.*, Oct., 1930.

Results of Noise Surveys. Part II. Noise in Buildings.[7] R. S. Tucker. Noise experienced indoors is in one sense more important than that experienced outdoors, for, with the growth of our industrial civilization, increasing numbers of people are spending most of their waking hours indoors. They are thus exposed to indoor noise for a large part of the time, including the hours of work when noise has its opportunity to impair their working efficiency.

Certain typical values for noise in various locations in buildings have been published, and are summarized in this paper. Our knowledge of indoor noise levels is far from complete, however. Further information has been obtained in a survey of room noise in New York City and the surrounding area which was made in 1929 by the National Electric Light Association and the American Telephone and Telegraph Company in the course of the work of their Joint Subcommittee on Development and Research. Some results of the New York City measurements are given. About 70 test locations are included. It will be realized that this is only a small sample of the total number of places where indoor noise is experienced in New York City alone. The conclusions given must therefore be regarded only as suggestive rather than as holding true in any general sense.

[7] *Jour. Acous. Soc. America*, July, 1930.

Contributors to this Issue

Charles B. Aiken, B.S., Tulane University, 1923; M.S. in Electrical Communication Engineering, Harvard University, 1924; M.A. in Physics, 1925. Bell Telephone Laboratories, 1928–. Mr. Aiken has been engaged on work in connection with aircraft communication and more recently with the design of broadcast radio receiver equipment.

F. E. Haworth, A.B., University of Oregon, 1924; M.A., Columbia University, 1929; Bell Telephone Laboratories, 1925–. Mr. Haworth's work has been in crystal analysis by means of X-rays, magnetic materials, and more recently in studies of dielectrics.

Herbert E. Ives, B.S., University of Pennsylvania, 1905; Ph.D., Johns Hopkins, 1908; assistant and assistant physicist, Bureau of Standards, 1908–09; physicist, Nela Research Laboratory, Cleveland, 1909–12; physicist, United Gas Improvement Company, Philadelphia, 1912–18; U. S. Army Air Service, 1918–19; research engineer, Western Electric Company and Bell Telephone Laboratories, 1919 to date. Dr. Ives' work has had to do principally with the production, measurement and utilization of light.

W. C. Jones, B.S. in E.E., Colorado College, 1913; Western Electric Company, 1913–25; Bell Telephone Laboratories, 1925–. As Transmission Instruments Development Engineer, Mr. Jones has specialized in the development and application of instruments for the transmission of speech and music.

A. R. Kemp, B.S., California Institute of Technology, 1917, M.S., 1918; Engineering Department, Western Electric Company, 1918–25; Bell Telephone Laboratories, 1925–. Mr. Kemp has been engaged in chemical research on rubber and allied materials used for submarine and other types of insulation.

W. H. Martin, A.B., Johns Hopkins University, 1909; B.S., Massachusetts Institute of Technology, 1911; American Telephone and Telegraph Company, Engineering Department, 1911–19; Department of Development and Research, 1919–. As Local Transmission Engineer, Mr. Martin has been engaged in development work on the transmission of telephone sets and local exchange circuits, transmission quality and loading.

L. J. Sivian, A.B., Cornell University, 1916; Engineering Department, Western Electric Company, 1917–19 and 1920–25; Bell Telephone Laboratories, 1925–. Mr. Sivian's work is in acoustics, chiefly in connection with methods of electroacoustic measurements.

George C. Southworth, B.S., Grove City College, 1914, M.S., 1916; Ph.D., Yale, 1923; assistant physicist, Bureau of Standards, 1917–18; instructor, Yale University, 1918–23; Information Department, American Telephone and Telegraph Company, 1923–24; Department of Development and Research, 1924–. Mr. Southworth's work in the Bell System has been concerned chiefly with the development of short-wave radiotelephony. He is the author of several papers on radio-frequency phenomena.