The Receiving System For Long-Wave Transatlantic Radio Telephony[1]

By AUSTIN BAILEY, S. W. DEAN, and W. T. WINTRINGHAM

Transmission considerations and practical limitations indicate that in the lower frequency range, frequencies near 60 kc are best suited for transatlantic radio-telephone transmission. A radio receiving location in Maine gives a signal-to-noise ratio improvement over a New York location equivalent to increasing the power of the British transmitter about 50 times.

Various types of receiving antennas are briefly discussed. The wave-antenna is selected as being most suitable for long-wave radio telephony. The various factors affecting wave-antenna performance and methods for measuring the physical constants of wave-antennas are discussed in detail. High-frequency ground conductivities determined from wave-antenna measurements are given. Combination of several antennas to form arrays is found to be a desirable means of decreasing interference. The use of a wave-antenna array in Maine decreases the received noise power by an additional 400 times. If the receiving were to be accomplished near New York using a loop antenna, we would have to increase the power of the British transmitting station 20,000 times to obtain the same signal-to-noise ratio. Comparisons of calculated and observed directional diagrams of wave-antennas and wave-antenna arrays are presented and discussed.

The transmission considerations governing the design of a radio receiver for commercial telephone reception are outlined.

Mathematical discussions of the wave-antenna, antenna arrays, quasi-tilt angle, and probability of simultaneous occurrence of telegraph interference are given in the appendices.

EARLY in October, 1915, engineers of the Bell System stationed in Paris heard the words "good night Shreeve," which had been transmitted from Arlington. That date then marks the inception of transatlantic radio-telephone receiving. The progress which has been made in the radio-telephone receiving art since these first experiments is demonstrated by contrasting the homodyne receiver and the non-directional antenna then used with the present commercial receiving system employing double-demodulation of single side band signals and an extensive array of wave-antennas forming a highly directional system. In the pages which follow we shall endeavor to give some of the engineering considerations upon which the design of the present receiving system was based.

Choice of Frequency

In the early development of long-distance radio telegraphy, the strength of the received signal was the principal factor upon which the selection of the operating frequency was based. After the development of the vacuum-tube amplifier, however, the following considerations each became important, especially so for a telephone circuit:

[1] Published in *Proc. of I. R. E.*, Dec., 1928, pp. 1645–1705.

1. The signal-to-noise ratio at the receiving location, which in turn is dependent upon four factors—
   (*a*) The efficiency of the transmitting set,
   (*b*) The efficiency of the transmitting antenna,
   (*c*) Attenuation in the radio path,
   (*d*) Variation of radio noise with frequency;
2. Band width of the transmitting antenna;
3. Receiving antenna efficiency;
4. Available space in the frequency spectrum.

1. *Signal-to-Noise Ratio at the Receiving Location.* At the time that the transatlantic radio-telephone development was undertaken, engineers of the Western Electric Company Engineering Department (now Bell Telephone Laboratories) had developed a form of water-cooled vacuum tube capable of generating efficiently large amounts of power at any frequency up to perhaps several hundred kilocycles.[2] Therefore transmitter efficiency, although a major problem in itself, imposed no restriction on the frequency for the telephone circuit.

For transmission over a given path, utilizing a particular transmitting antenna with constant power supplied to it, there will be, in general, a frequency at which the greatest signal-to-noise ratio is obtained. To illustrate this point, we have chosen the problem of transmission from an antenna of the type used at the Rocky Point station of the Radio Corporation of America in U. S. A. to a receiving station in England, a distance of approximately 5,000 kilometers. The approximate variation with frequency of loss resistance, radiation resistance, and efficiency of this antenna is shown in Fig. 1. The loss resistance at 60 kilocycles was determined by engineers of Bell Telephone Laboratories, while the data in the lower frequency range were published by Alexanderson, Reoch, and Taylor.[3] The radiation resistance was calculated from the measured effective height of the antenna. It is seen in Fig. 1 that the antenna efficiency increases with frequency throughout the range we are considering, first rapidly and then more slowly.

For a constant power radiated, radio attenuation tends to cause a decrease in the average received signal strength as the frequency is increased. This effect is in the opposite direction to the effect of antenna efficiency, so that for a given power supplied to the antenna the field strength at a given distance will be a maximum at a certain

[2] W. Wilson, "A New Type of High Power Vacuum Tube," *Bell System Tech. Jour.*, *1*, 4; July, 1922. *Elec. Comm.*, *1*, 15; August, 1922.
[3] E. F. W. Alexanderson, A. E. Reoch, and C. H. Taylor, "The Electrical Plant of Transocean Radio Telegraphy," *Trans. A. I. E. E.*, *42*, 707; July, 1923.

frequency. In Fig. 2 we have shown the calculated field strength at 5,000 kilometers for a power of 85.9 kilowatts supplied to the Rocky Point antenna, using efficiency data of Fig. 1 and the radio transmission formula given by Espenschied, Anderson, and Bailey.[4] Since this curve reaches a maximum near 18.5 kilocycles, the reason for the operation of early transatlantic radio-telegraph circuits in the range 10 to 30 kilocycles becomes apparent in light of the limitation then placed on the receiving systems.

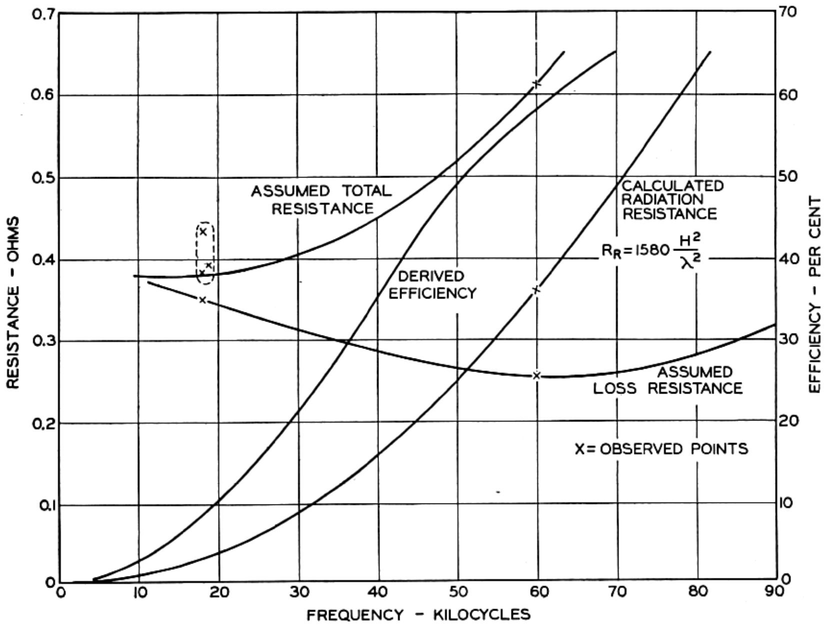


Fig. 1—Assumed resistance and efficiency of Rocky Point antenna. (Effective height 75 meters.)

Systematic measurements of radio noise by the warbler method,[5] begun early in 1923, have yielded important information on the variation of noise with frequency.[4] From measurements begun by engineers of Bell Telephone Laboratories and continued by engineers of the International Western Electric Company at New Southgate, England, during 1923 and 1924, the average daylight noise curve, in Fig. 2, was obtained. It is seen that the noise decreases with increasing

[4] Lloyd Espenschied, C. N. Anderson, and Austin Bailey, "Transatlantic Radio Telephone Transmission," *Bell System Tech. Jour.*, *4*, 459; July, 1925. *Proc. I. R. E.*, *14*, 7; Feb., 1926.

[5] Ralph Bown, C. R. Englund, and H. T. Friis, "Radio Transmission Measurements," *Proc. I. R. E.*, *11*, 115; April, 1923.

frequency, at first rapidly and then more slowly, being almost constant after passing the frequency of 40 kilocycles.

From the values of signal and noise so obtained, the signal-to-noise ratio has been computed, and is also plotted in Fig. 2. The curve of signal-to-noise ratio reaches a maximum near 44 kilocycles which would seem to be the optimum frequency for daylight transmission from the Rocky Point station to England. This is not strictly the case, however, since there is some evidence that a phenomenon exists which makes frequencies in the vicinity of 40 kilocycles particularly poor for the transatlantic path. Data published by Anderson [6] tend

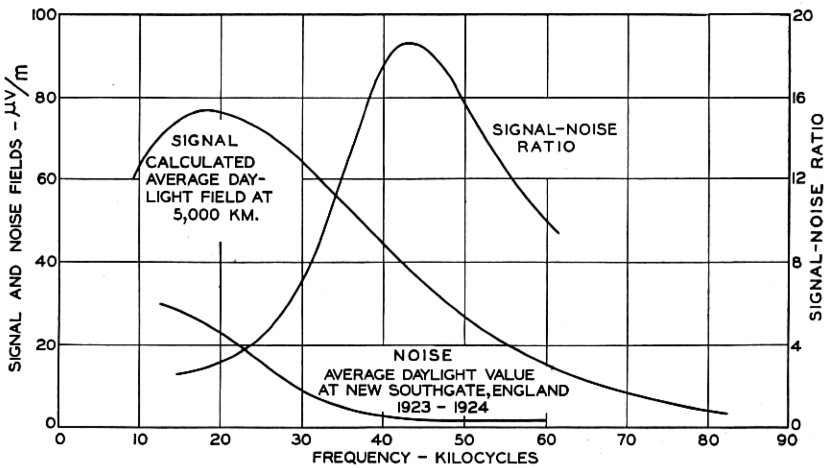


Fig. 2—Variation of signal, noise, and signal-noise ratio with frequency. Transmission from U. S. A. to England. 85.9 kw. supplied to antenna of Rocky Point characteristics.

to show that the field strength is distinctly subnormal in the vicinity of 44 kilocycles and remains approximately constant from that frequency up to about 60 kilocycles, where the observed values agree fairly well with the calculations. (See later in this paper.)

2. *Band Width of the Transmitting Antenna.* Since the output of the transmitting set is at a high power level, the circuits coupling it to the antenna must be of the simplest type to reduce the loss to a minimum. In view of this requirement, the antenna constants largely determine the band width of the antenna system. At frequencies much lower than 60 kilocycles it was not possible to secure a sufficient width of band even for commercial telephony from the Rocky Point

[6] C. N. Anderson, "Correlation of Long Wave Transatlantic Radio Transmission with other Factors Affected by Solar Activity," *Proc. I. R. E.*, *16*, 297; March, 1928. In connection with reference above see Fig. 19, p. 315.

antenna, but at this frequency reasonably satisfactory results are obtained.

3. Receiving Antenna Efficiency. The use of directional receiving antennas is essential to satisfactory and economic results over such distances as the transatlantic radio path (see later in this paper). The directivity of an antenna system of a given kind, size, and cost in general increases with frequency, since the directivity is a direct function of the ratio of the dimensions of the antenna system to the wave-length employed.

4. Available Space in the Frequency Spectrum. Each of the above factors operates to make the frequency of 60 kilocycles about the best which could be used in the present state of the art for this transmission path. Fortunately this frequency was so located in the radio spectrum that a band of the desired width free from interference could be obtained.

It has been noted that the radio noise as shown in Fig. 2 varies very little with frequency above 40 kilocycles. There is some doubt as to whether or not this accurately represents the actual state of affairs, since the measurement sets used for measuring the noise would not satisfactorily measure much below one microvolt per meter on account of tube noise. At frequencies of 40 kilocycles and above, especially in the winter, there are many days during which the radio noise is practically absent. On these days the measurements tended to approach the minimum determined by the set noise. The fact that many such readings were incorporated in the average probably tends to mask the true variations of radio noise with frequency in this range. On the other hand, however, they indicate a very real limitation which tends to operate against the use of frequencies higher than about 60 kilocycles unless fields were increased by increase in transmitting power. This would be particularly true during the sunset and sunrise dips and during periods of abnormally poor transmission when the fields fall much below the average. If the set noise limitation could be removed it is quite possible that frequencies above 60 kilocycles would become more useful. Higher frequencies for radio telephone use would be particularly advantageous because of the greater band width which could be obtained from the transmitting antenna and because of the greater directivity which could be obtained in the receiving system at the same cost.

Selection of a Satisfactory Receiving Location

The selection of a suitable receiving location is based upon three major considerations; namely, maximum received signal-to-noise ratio, reasonably suitable terrain for receiving antenna construction, and adequate wire connection facilities between the location and the more densely populated areas.

Since about 10 per cent of the populations of the United States and the British Isles are located within a radius of 40 miles of New York and London, respectively,[7] it was natural to decide upon making those cities the terminal points. It would hence be desirable to locate the receiving stations near and with good wire circuits to those cities.

Very early in the history of radio communication [8] it was, however, realized that in the United States a decrease of radio noise was obtained by a northerly location of the receiving station and, for receiving from European stations, the northern location is further advantageous, since higher field strengths result from the reduced transmission distance. The Radio Corporation of America had already taken advantage of this improvement by locating a receiving station at Belfast, Maine.

To obtain quantitative information on this matter, the American Telephone and Telegraph Company made comparative measurements of noise as received on loop antennas at Riverhead, New York; Green Harbor, Massachusetts; and Belfast, Maine; the loops were so oriented as to give maximum receptivity in the direction of England. Although these tests were only continued for a few months at each location, they left no doubt that the absolute level of the noise was less at the northerly locations.

In Fig. 3, there is shown the diurnal variation of improvement in noise conditions (in TU) for average days of each month at Belfast over Riverhead. The average hourly improvement was determined by averaging the ratios of practically simultaneous observations of noise at the two locations for each hour during any one month and taking a three-hour moving average of the result to reduce the effect of purely local phenomena at either of the two stations. The data for the two half years were taken on slightly different frequencies as is indicated on the figure. Unfortunately, during the month of July only two weeks data were taken on each of the frequencies, namely, 52 and 65 kilocycles, and these data were taken a year apart, namely in 1924 and in 1925. In order to give some idea of the location noise improvement for the month of July we have averaged in the same way the four weeks data thus obtained, and plotted the result as a broken line. Fortunately, the improvement of the more northerly location is, in general, large during the overlapping business day of England and the United States.

[7] "New York's New 10,000,000 Zone," *Literary Digest, 95*, 12, p. 14; Dec. 17, 1927.

[8] G. W. Pickard, "Static Elimination by Directional Reception," *Proc. I. R. E., 8*, 358; October, 1920.

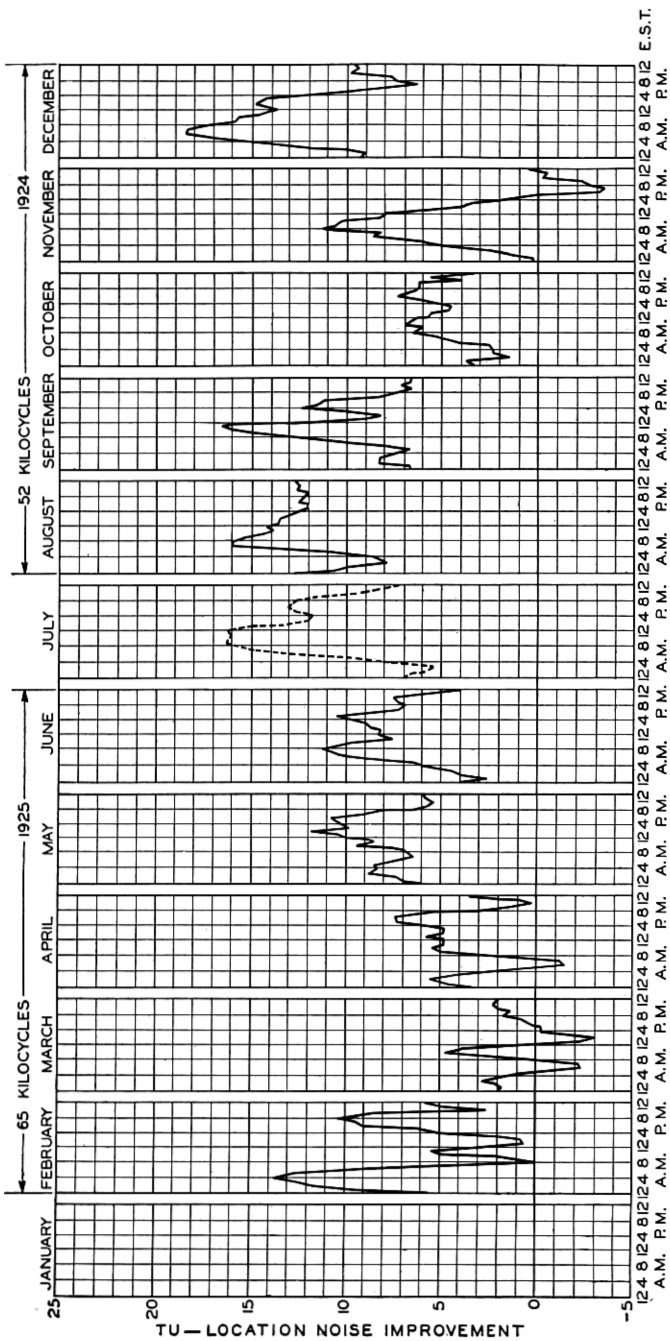


Fig. 3—Transatlantic radio noise measurements. Diurnal variations of location noise improvement (in transmission units) of Belfast, Maine, over Riverhead, New York. Three-hour moving averages of simultaneous observations.

It is apparent that the improvement is a maximum in the middle of the summer when the noise is high, and in the middle of the winter when the field strengths are usually abnormally low. This is important, since the greatest improvement is needed at each of these times.

The monthly averages of variations of noise and of signal have previously been published,[4, 6, 9] and the generalizations given above can be confirmed by reference to these articles.

For calculating daylight radio transmission, several formulas have been proposed.[10, 11, 12] In Fig. 4 the heavy curve was calculated

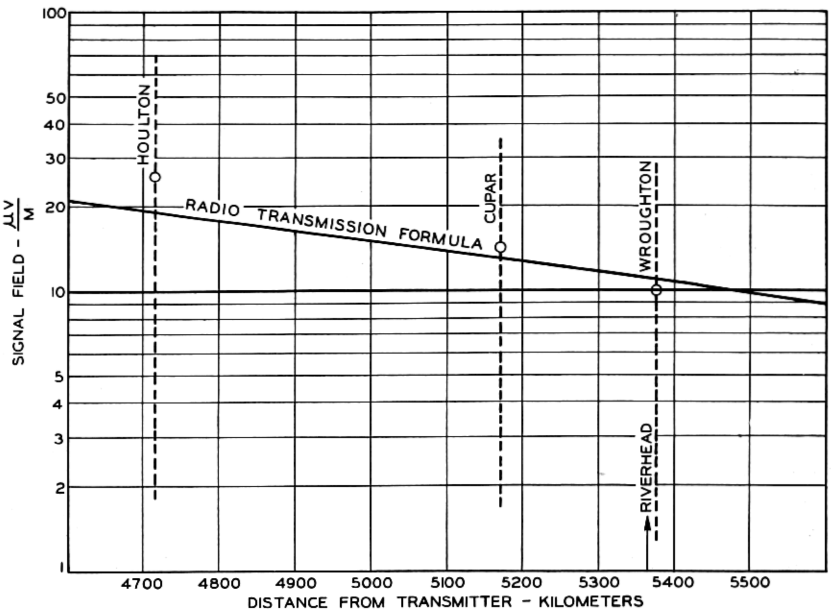


Fig. 4—Transatlantic radio daylight field strength. Average of hourly observations during 1927. Corrected to 50 kw. radiated power—frequency 60 kilocycles

from the empirical formula given by Espenschied, Anderson and Bailey,[4] and assumes a radiated power of 50 kilowatts. The great circle distance from the transmitting stations used in the transatlantic radio-telephone circuit to various receiving stations is indicated by the name of the receiving station. The average of daily averages

[9] Ralph Bown, "Some Recent Measurements on Transatlantic Radio Transmission," *Proc. Natl. Acad. of Sci.*, *9*, 221; July, 1923.

[10] A. Sommerfeld, "Ueber die Ausbreitung der Wellen in der drahtlosen Telegraphie," *Ann. d. Phys.*, *28*, 665; 1909.

[11] L. F. Fuller, "Continuous Waves in Long-Distance Radio Telegraphy," *Trans. A. I. E. E.*, *34*, pt. 1, 809; 1915.

[12] L. W. Austin, "Quantitative Experiments in Radiotelegraphic Transmission," *Bull. Bureau of Std.*, *11*, 69; Nov. 15, 1914.

of hourly measurements of the field strength made at Houlton and Wroughton during the time that the transatlantic path was entirely in daylight during 1927 is indicated by points on this figure. The data for Cupar are less complete since this station was not in regular daily operation until May, 1927. The range of variation between the maximum daily average and the minimum daily average for each receiving location is given by the limits of the dotted vertical line. (It is interesting to note that at a frequency of 60 kilocycles and for distances in the order of 5,000 kilometers any of the radio-transmission formulas referred to above will give a computed value lying within the range of variation of average daylight readings.)

The improvement in signal-to-noise ratio obtained by locating the receiving station in Maine instead of in New York is easily seen by reference to Figs. 3 and 4. The improvement due to decrease of noise, during that time of year when improvements are most needed on account of high noise values, is about 10 TU. The improvement due to increase of the average received daylight signal by decrease of the distance is calculated to be 5 TU. During 1927, this improvement was actually observed to be 8 TU. We may, therefore, state in round numbers that the total improvement realized by locating the receiving station in Maine instead of New York was equivalent to a fifty-fold increase of the power radiated by the British transmitting station.

The British General Post Office, during 1926, carried out a set of measurements of field and noise at various locations in the United Kingdom. Those tests led them to the same conclusions as regards the advantage to be obtained by locating their receiving station at some more northerly point.[13] They decided upon a location near Cupar, Scotland, and comparisons made daily from 1230 to 2300 GMT indicate that this location is better for receiving than Wroughton, England. The geometric mean of the improvement in signal-to-noise ratio for the more northerly location during the months May to September, 1927, inclusive, and for the daily period given above is 6.4 TU. This is equivalent to an increase of between four and five times in power from the American transmitting station.

Since such relatively large improvements were to be obtained by northerly locations of the receiving station it seemed best to take advantage of this fact and locate the receiving station in America at some place in the state of Maine. This decision led to further consideration of two factors mentioned above, namely, reliable wire

[13] A. G. Lee, "Wireless Section: Chairman's Address, *Jour. I. E. E.*, *66*, 12; Dec., 1927.

connections to New York and a suitable terrain for antenna construction. The first of these factors required a location along one of the main telephone trunk routes in Maine and the second, since we had decided upon the use of a wave-antenna [14] for reasons which will be given in the following section, demanded a rather large and reasonably flat land area available for pole-line construction. A location, although not altogether ideal, was decided upon near Houlton, Maine, about six miles from the Canadian border.

Choice of Receiving Antenna Systems

The number of fundamental types of receiving antennas that may be employed for long-wave reception is quite definitely limited. In fact all of the known practical receiving antennas may be considered as falling into one of three principal classes of structure; i.e., the vertical antenna, the loop or coil antenna, and the wave-antenna. The selection of the proper receiving antenna system quite evidently becomes a problem—first, of choosing the best type of antenna from one of these three classes and, second, of choosing a particular antenna structure in the class which is found to be best.

The factors governing the choice of a receiving antenna are as follows:

1. Directional Discrimination Against Static. Inasmuch as the signal to be received has a definite average value, the receiving system can only better the circuit in the amount that it improves the signal-to-noise ratio. A directional antenna system affords a means of reducing the received noise in relation to the desired signal.[8, 14] The directional characteristics of the principal antenna types are shown in Fig. 5.

A measure of the directional discrimination of the various antenna types is the Noise Reception Factor (abbreviated *NRF*) which is defined as the ratio of the total noise current received from the antenna in question to that received from a vertical antenna under the conditions of continuous, constant distribution of noise sources about the antenna and of equal output currents for signals from the direction of maximum receptivity. The back end *NRF* is the noise reception factor for the arc between 90 degrees and 270 degrees from the direction of maximum receptivity.

On this basis, the choice rests quite unmistakably with the wave-antenna.

2. Transmission-Frequency Characteristic. Since the receiving antenna is to be used on a system for communication by speech, necessi-

[14] H. H. Beverage, C. W. Rice and E. W. Kellog, "The Wave Antenna," *Trans. A. I. E. E.*, *42*, 215; 1923.

tating the transmission of a relatively wide band of frequencies, it must pass such a band without undue discrimination against any frequency contained therein. To utilize the vertical and the loop antennas

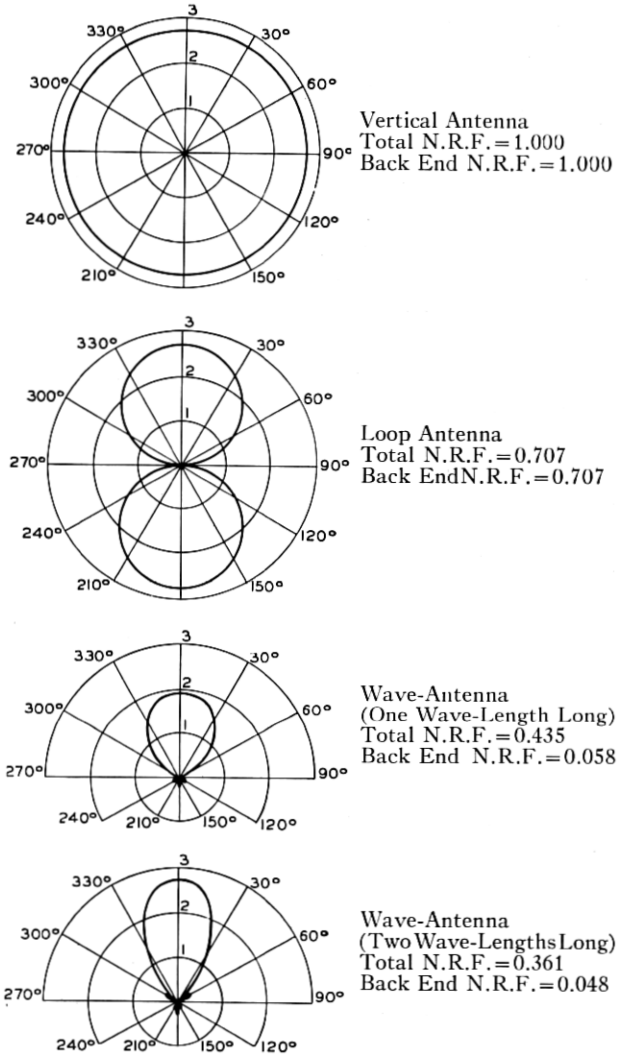


Fig. 5—Comparison of polar diagrams of simple antennas. (The unit for the radii is output current into the same resistance for a constant impressed field.)

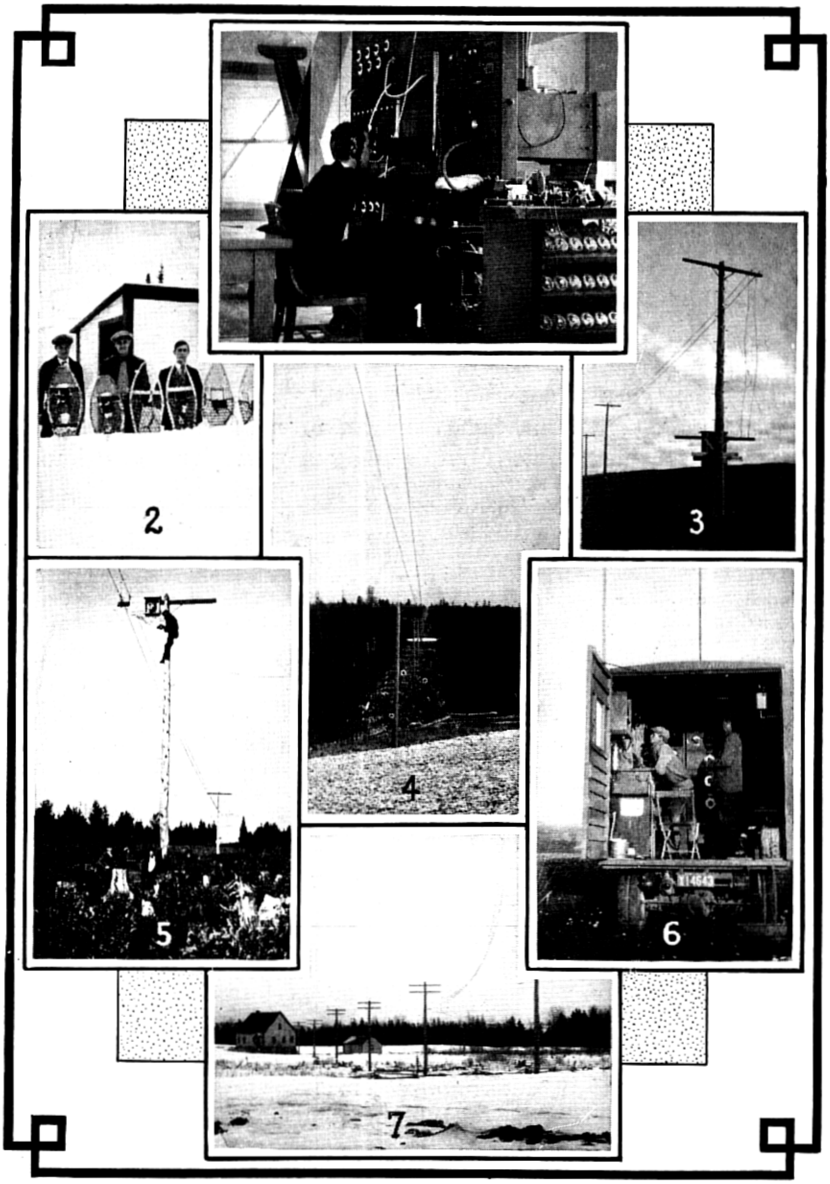
efficiently, it is desirable that they be tuned, introducing the frequency discrimination of a tuned circuit. If these types of antennas be used it is necessary, therefore, that the resonance characteristic be studied and means provided to eliminate excessive frequency discrimination within the desired band. On the other hand, the wave-antenna is an aperiodic structure and, in consequence, its transmission-frequency characteristic is so flat that it need not be considered.

3. Sensitivity. There are two factors which require that the output from the receiving antenna for a given field strength be as large as possible. First, if the receiving station be located at any position other than that at the terminal of the antenna, which is necessarily the case if more than one antenna be used in an array, the signal on the transmission line from the antenna to the station must be much greater than the noise currents induced into the transmission lines. If the antenna output be excessively small, it is impossible to balance the transmission lines so completely that this requirement is met. Second, the amount of gain that can possibly be used at the radio receiver is ultimately limited by the noise produced in an amplifier. (This is discussed more fully under "Power Output Required from the Radio Receiver" later in this paper.) To the first approximation, the sensitivity of each of the antenna classes under consideration is a direct function of its physical dimensions. There is, however, a limit to the sensitivity of each antenna class, for mechanical limits govern the maximum size of a vertical antenna, distributed capacity and mechanical considerations limit the loop, and in the wave-antenna a restriction occurs because of the peculiarity that the sensitivity reaches maximum values at definite lengths.

Since cost is likewise a factor governing the ultimate selection of an antenna system, the sensitivities may well be compared for antennas of equal cost. On this basis, a loop or a vertical antenna of effective height of fifty meters is directly comparable with a wave-antenna one wave-length long. By reference to Fig. 5, where the scale is the same for all the directional diagrams, it becomes evident that the sensitivities of all three classes of antennas are of the same order of magnitude, being slightly greater for the vertical antenna and the loop than for the one-wave-length wave-antenna.

4. Stability. The sensitivity and frequency-transmission characteristics of the antenna must be substantially constant during changes of weather and seasonal conditions. The antenna classes which require tuning are slightly poorer than the wave-antenna in this respect.

5. Reproducibility. Further improvement in directional discrimination against noise is obtained by using several similar antennas in an array. The loop and the vertical antennas probably are best for



1—Measuring field strength.
2—Outside an antenna terminal hut.
3—Pole box for reflection transformer.
4—The wave-antenna *A* at Houlton.
5—Measuring ground connection impedance at a temporary location.
6—The sixty kilocycle portable transmitting station.
7—Transmission line O-B with receiving station in background.

combining in arrays because several of either type of antenna can be made identical with one another. Wave-antennas combined in an array, however, give satisfactory results.

Although each of these factors governing the choice of the receiving antenna system is important, their relative importance is indicated by the order in which they have been presented. In view of the low noise reception factor of the wave-antenna, its lack of frequency discrimination, and its inherent stability, the wave-antenna was selected for the fundamental type of antenna to be used at the receiving station at Houlton.

The Wave-Antenna

Among the types of antennas which may be considered for use in long-wave radio communication, the wave-antenna [14] possesses several characteristics which single it out as being unique. The most important of these are:

1. The length of a wave-antenna is directly comparable to and of the same order of magnitude as the wave-length of the signals for which it is designed.
2. Considering the straight horizontal wire comprising the wave-antenna as a grounded transmission line, a termination, equal to the characteristic impedance, is applied to each end of that line. The wave-antenna then becomes an essentially aperiodic antenna.
3. The major response of a properly designed wave-antenna is to the horizontal component of the impressed electric field. The propagated electric wave must therefore have an electric component parallel to the surface over which the wave-antenna is constructed.
4. On the basis of the preceding consideration, the design of a wave-antenna definitely excludes elevation of the antenna above ground to any extent greater (a) than is physically necessary to provide safe clearance and (b) than that height where the loss in the antenna considered as a transmission line reaches a nominal value. Practically, the wave-antenna is constructed as a high-grade telephone line, on 30-foot poles.

It is evident that the major electrical characteristics which distinguish the wave-antenna are intimately connected with the character of the surface over which the antenna is built, and with the details of construction of the wave-antenna. The performance of a wave-antenna at any specified location then can only be determined by constructing such an antenna and measuring its constants. The measurements made in determining the characteristics of any particular wave-antenna are outlined in the following paragraphs.

1. Ground-Connection Impedance. It is shown in Appendix 1 that the wave-antenna is considered to be a smooth line with uniformly distributed constants. This assumption is met to a sufficient degree in practice, but, unfortunately, it is impossible to connect to the four terminals of the practical line, since the connections to the ground side of the line must be made by burying wires in the earth rather than connecting to a discrete terminal which is the real ground. As is shown in Fig. 6a, the actual wave-antenna may still be considered as a smooth line, but between the terminals of the wave-antenna and the terminals that are available at the physical ends of the wave-antenna ground-connection impedances exist. To determine the constants of the wave-antenna, these impedances must be evaluated and taken into account as follows: In Fig. 6a, an impedance Z is applied to the avail-

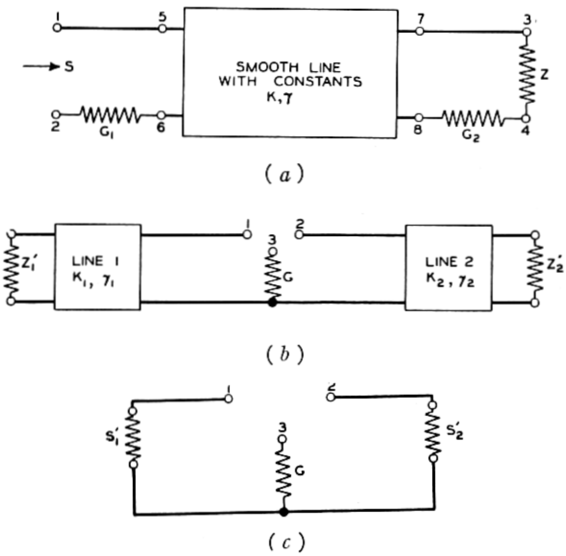


Fig. 6.

able terminals of the wave antenna 3–4 and the impedance S measured at the available terminals 1–2; under this condition, the actual terminal and input impedances of the wave-antenna are respectively:

$$Z' = Z + G_2 \tag{1}$$

$$S' = S - G_1 \tag{2}$$

where G_1 and G_2 are the ground-connection impedances at the two ends of the antenna.

Figs. 6b and 6c illustrate the method that was used to determine the ground-connection impedance. In Fig. 6b, lines 1 and 2 represent two smooth ground-return transmission lines extending in opposite directions from the ground connection for ½ kilometer or more, the lines being terminated at the distant end in impedances Z_1' and Z_2', respectively. In practice one of these lines was the wave-antenna and the other a temporary line of insulated wire laid along the surface of the ground.

For the purpose of analysis, each of the lines may be replaced by its input impedance. This simplification is shown in Fig. 6c, where

$$S' = \frac{K \tanh \gamma s + Z'}{1 + \frac{Z'}{K} \tanh \gamma s}. \tag{3}$$

The impedance between terminals 1 and 3 is:

$$S_1 = S_1' + G. \tag{4}$$

The impedance between terminals 2 and 3 is:

$$S_2 = S_2' + G. \tag{5}$$

The impedance measured between terminals 1 and 2 in parallel and terminal 3 is:

$$S_0 = G + \frac{S_1' S_2'}{S_1' + S_2'}. \tag{6}$$

Eliminating S_1' and S_2' fiom equations (4), (5), and (6) and solving for G:

$$G = S_0 - \sqrt{\frac{1}{2}[(S_1 - S_0)^2 + (S_2 - S_0)^2 - (S_1 - S_2)^2]}. \tag{7}$$

By building out either line 1 or line 2 with added series impedances until

$$S_1 = S_2 = S_{12} \tag{8}$$

the expression for the ground-connection impedance simplifies greatly, and incidentally the precision of the determination becomes greater because the number of measurements involved is less. Under this condition

$$G = 2S_0 - S_{12}. \tag{9}$$

This latter case is the one that was actually used in measuring the ground impedances.

Since the distribution of ground currents about the buried ground may be different under each of the three conditions that are measured, there is undoubtedly some error in measuring the ground-connection impedance by this method. This error is a second-order effect, however, so that the values determined are reliable within the precision that the method allows, involving as it does, differences between measurements of high-frequency impedance.

All of the impedance measurements were made using a high-frequency bridge designed and constructed by Mr. C. R. Englund of Bell Telephone Laboratories. This bridge is similar to that described by Shackelton [15] except that the standards used consist of a calibrated condenser and a decade resistance. Impedances having capacitive reactance are measured by direct comparison with the standards, while impedances having inductive reactance are tuned with the standard condenser to parallel resonance and the resonant combination compared with the decade resistance. Impedances involving extremely small reactances, either positive or negative, are built out with a condenser in parallel to a value that may be measured conveniently.

2. *Characteristic Impedance and Propagation Constant.* Since the early days of transmission line study, the characteristic impedance and the propagation constant have been determined by two impedance measurements at the near end of the line with the far end of the line open- and short-circuited, respectively.[16] For two reasons, this method has not been used in our determination of the fundamental antenna constants: first, it is impossible to apply a short to the real terminals of the wave-antenna due to the presence of the ground-connection impedance; and, second, with lines multiple quarter wave-lengths long the input impedance, as a result of resonance in the line when it is open-circuited or grounded, attains either extremely large or extremely small values which could not be measured accurately with the available testing equipment.

To obviate these difficulties, Mr. C. R. Englund, of Bell Telephone Laboratories, developed a method of determining the characteristic impedance and the propagation constant of the wave-antenna by measuring the input impedance with two known finite terminations at the far end. Under this condition it may be shown that the characteristic impedance is given by the expression:

$$K = \sqrt{\frac{(S_1 - G_1)(S_2 - G_1)(Z_1 - Z_2) + (Z_1 + G_2)(Z_2 + G_2)(S_2 - S_1)}{(S_2 - S_1) + (Z_1 - Z_2)}} \quad (10)$$

and that the propagation constant is given by:

[15] W. J. Shackelton, "A Shielded Bridge for Inductive Impedance Measurements at Speech and Carrier Frequencies," *Bell System Tech. Jour.*, *6*, 142; Jan., 1927.

[16] Bela Gati, "On the Measurement of the Constants of Telephone Lines," *The Electrician*, *58*, 81, Nov. 2, 1906.

$$\gamma = \frac{1}{s}\tanh^{-1}\left[K\frac{(S_2 - S_1) + (Z_1 - Z_2)}{(Z_1 + G_2)(S_1 - G_1) - (Z_2 + G_2)(S_2 - G_1)}\right], \quad (11)$$

where the symbols in equations (10) and (11) have the following meanings:

Z_1 = the first termination applied to the available terminals at the far end of the line (ohms)

Z_2 = the second termination applied to the available terminals at the far end of the line (ohms)

S_1 = the impedance measured at the available near-end terminals corresponding to the termination Z_1 (ohms)

S_2 = the impedance measured at the available near-end terminals corresponding to the termination Z_2 (ohms)

G_1 = the ground-connection impedance at the near end of the line (ohms)

G_2 = the ground-connection impedance at the far end of the line (ohms)

s = length of the line (kilometers)

K = characteristic impedance (ohms)

γ = propagation constant (hyps per kilometer)

3. *Effective Height.* The effective height of a wave-antenna is defined as the ratio of the voltage produced at any specified point in the antenna to the potential gradient of the electromagnetic field producing that voltage. If the constants of the antenna system are known, the effective height at any point in the antenna system may be calculated from the value at any other point in the system.

A convenient way to measure an effective height of a wave-antenna and obtain a value which may be easily correlated with wave-antenna theory is to introduce in series with the initial-end terminating impedance a voltage which produces the same output current from the antenna as is produced by an electromagnetic wave. The ratio of this induced voltage to the potential gradient of the electromagnetic field has been called "the effective height referred to the characteristic impedance." For small values of the quasi-tilt angle, the total potential gradient of the electric field is very closely equal to the vertical component of the electric field, so that within the precision of measurement we may write:

$$H_\theta = \left|\frac{E_K}{E'}\right|, \quad (12)$$

where

H_θ = Effective height of the wave-antenna referred to the characteristic impedance (kilometers)

E' = The potential gradient of the vertical component of the impressed field (volts per kilometer)

E_K = The electromotive force introduced in series with the characteristic impedance at the initial end of the wave-antenna producing the same current at the distant end as the impressed field (volts)

4. Quasi-tilt Angle and Ground Resistivity. The measured effective height of a wave-antenna is a function of four constants:

1. The length of the antenna;
2. The height of the antenna;
3. The propagation constant of the antenna;
4. The ratio of the component of the electric wave parallel to the surface over which the antenna is constructed to the vertical component of the electric wave.

In general, the first three of these constants are different in value for antennas constructed at different locations, but they may be varied over a limited range by changing the construction and dimensions of the wave-antenna. The comparison of effective heights, therefore, does not readily yield information regarding the relative suitability of various locations for wave-antenna systems. The ratio of the horizontal component to the vertical component of the impressed field is, however, a constant whose value is dependent solely upon the ground conditions at the location (assuming a fixed frequency for the comparison).

In case the time phase between the horizontal component and the vertical component of the impressed field were zero, the ratio of these two components would represent the tangent of the angle of forward inclination of the propagated wave front. In general, the phase angle between the two components is not zero, so that such a simple relation does not hold. It is convenient, however, to call the ratio of the two components of the impressed field the tangent of the "quasi-tilt angle," where the "quasi-tilt angle" becomes the real tilt angle in the limiting case.

In terms of the effective height, the antenna constants, and the vertical component of the impressed field, the current produced at the far end of the wave-antenna is (using the nomenclature of Appendix 1 and to the same degree of approximation as equation (12)):

$$|I_\theta| = \left|\frac{H_\theta E' \epsilon^{-\gamma S\lambda'}}{2K}\right| \tag{13}$$

or

$$|I_\theta| = H_\theta \epsilon^{-\alpha S\lambda'} \left|\frac{E'}{2K}\right|. \tag{14}$$

In terms of antenna constants alone, it is shown in Appendix 1 that the current produced at the far end of the wave-antenna is:

$$|I_\theta| = |I_{E'\theta} + I_{F'\theta}|, \tag{125}$$

where

$$I_{E'\theta} = -\frac{1}{\epsilon^{j\delta} \tan T} \frac{S\lambda' F}{2K}(a + jb), \tag{15}$$

$$I_{F'\theta} = \frac{S\lambda' F'}{2K}(c + jd), \tag{16}$$

also

$$-\frac{F'}{E'} = \epsilon^{j\delta} \tan T. \tag{301}$$

In equations (15) and (16), $(a + jb)$ and $(c + jd)$ are abbreviations defined as follows:

$$(a + jb) = \frac{h}{S\lambda'}\left(1 - \epsilon^{-[\alpha S\lambda' + j2\pi S(m - \cos\theta)]}\right)\epsilon^{-j2\pi S \cos\theta} \tag{17}$$

and

$$(c + jd) = \cos\theta \frac{1 - \epsilon^{-[\alpha S\lambda' + j2\pi S(m - \cos\theta)]}}{\alpha S\lambda' + j2\pi S(m - \cos\theta)} \epsilon^{-j2\pi S \cos\theta}. \tag{18}$$

If we equate expressions (14) and (125), and solve for $\tan T$:

$$\tan T = \frac{ac + bd + \sqrt{(ac + bd)^2 - (c^2 + d^2)(a^2 + b^2 - g^2)}}{c^2 + d^2}, \tag{19}$$

where

$$g = \frac{H_\theta}{S\lambda'} \epsilon^{-\alpha S\lambda'}. \tag{20}$$

It is pointed out in Appendix 3 that the phase angle δ may be expressed as a function of the quasi-tilt angle T and the dielectric constant k. For that reason, the determination of T must be made in two steps. The procedure is as follows: first, it is assumed that the component of the total received current due to the vertical component of the impressed field is zero, i.e.,

$$(a + jb) \equiv 0.$$

Under this condition:

$$T = \tan^{-1}\frac{g}{\sqrt{c^2 + d^2}}. \tag{21}$$

Using Fig. 20 of Appendix 3, the value of δ corresponding to this value of T is determined (generally $\delta = \pi/4$). Second, T is revaluated from (19) using the value of δ so obtained.

The ground resistivity is evaluated from the value of the quasi-tilt angle by using Fig. 20 of Appendix 3.

5. *Directional Characteristics.* The measurement of the directional characteristics of a wave-antenna or a wave-antenna system consists entirely of measuring the effective height of the antenna for several directions of wave propagation, and determining the relative directional receptivity of the antenna in these directions by dividing the effective height for each direction by the value obtained for the direction of the axis of the antenna. For this purpose, the effective height at the output of the antenna system is most convenient to measure and use. This constant is defined as the ratio of the voltage at the input of the radio receiver to the field strength producing this voltage. It is exactly related to the effective height referred to the characteristic impedance (defined in the preceding subsection of this paper) by the real part of the total transfer constant between the termination at the initial end of the antenna and the input terminals of the radio receiver, and an additional factor of one-half because the voltage at the radio receiver is measured across the proper termination.

In certain receiving station locations, it is possible to utilize for determining the relative directional receptivity the regular transmission from existing radio transmitters operating at or very close to the frequency for which the directional characteristic is desired. At sites less favorably located with regard to existing transmitters, the directional characteristic may be measured by transmitting test signals from a portable transmitter, located successively in the several directions for which data are desired, and at least 15 wave-lengths from the antenna system.

A distinctly different method of measuring the directional characteristics of an antenna is based on a statistical study of the reduction of noise obtained by its use. While it is difficult to evaluate the directional characteristic exactly by this method, data showing the comparative decrease in noise with the wave-antenna as against a loop or a vertical antenna are of great value in predicting the improvement in a radio circuit to be obtained by its use. As a converse to these results, the statistical combination of the improvement given by the wave-antenna, and a measured directional diagram, yields information on the direction of arrival of static.

Data on wave-antenna characteristics have been taken at several widely separated locations. Two antenna systems have been constructed by the British General Post Office—one at Wroughton, Wiltshire, in southern England, and one at Cupar, Fifeshire, in southeastern Scotland. We likewise have data on our antenna system at Houlton, Maine. The character of the earth under each of these antenna systems is different, resulting in widely different quasi-tilt angles and antenna directional characteristics.

The probable geological formations under individual antennas at each of the three antenna sites mentioned in the preceding paragraph are shown in Fig. 7. The data for the British locations were compiled from the published reports of geological surveys conducted by the British Government, and the data for the Houlton location were taken from the "Soil Survey of the Aroostook Area, Maine," published by the U. S. Department of Agriculture. In Table I, the ground constants are given for these three locations, determined by the method given in Section 4, "Quasi-tilt Angle and Ground Resistivity":

TABLE I

Location	Characteristic Sub Soil	Quasi-Tilt Angle at 60 kc Radians	Ground Resistivity Ohms per cm^3
Wroughton	Chalk	0.011	3630.
Cupar	Sandstone	0.017	8670.
Houlton	Limestone	0.047	66300.

Fig. 8 shows the directional characteristics, calculated by the method given in Appendix 1, of wave-antennas erected over the geological formations shown in Fig. 7. In these directional characteristics, it is important to notice that a decrease in quasi-tilt angle increases the relative importance of the component of the received current due to the vertical component of the impressed field (abbreviated to the "vertical effect"). It is evident from Fig. 8 that the relative directional receptivity for the arc between $\Theta = 90$ degrees and $\Theta = 270$ degrees is smaller and that the effective height is much greater for the antenna at Houlton, for which the ground resistivity is higher than for the other two antennas.

Measured relative directional receptivities are also shown in Fig. 8. The values for the Cupar antenna were determined by using trans-

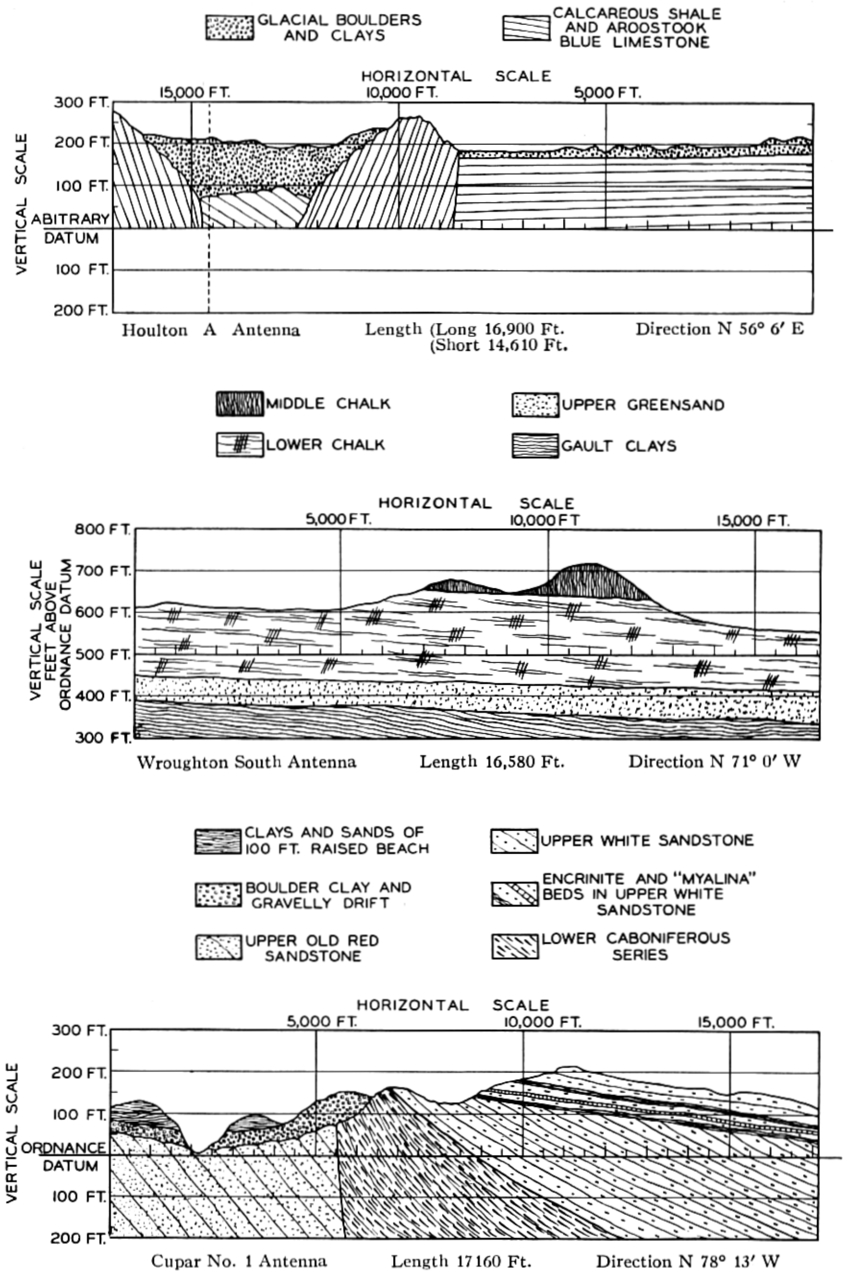


Fig. 7—Cross section of probable geological formation under several wave-antennas.

mission from the several European transmitting stations which are designated on this figure. The measurements on the Houlton antenna system were made using a portable two-kilowatt transmitter located

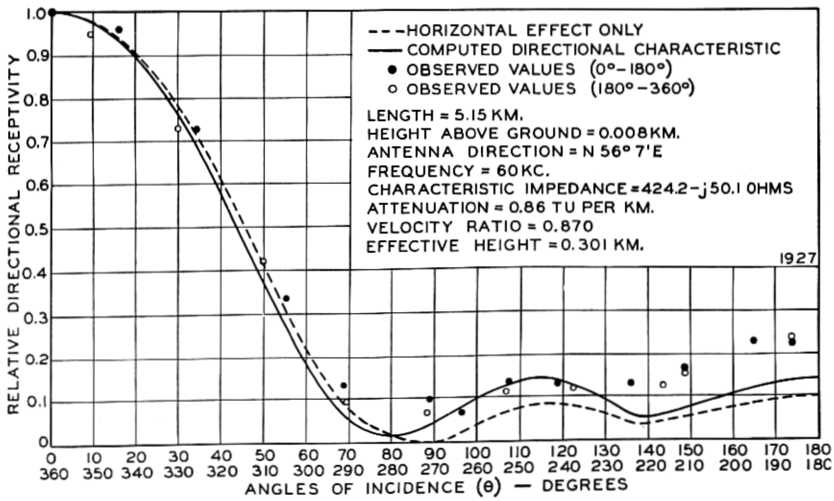


Fig. 8a—Relative Directional Receptivity of Houlton Antenna "A" Uncompensated (Long)

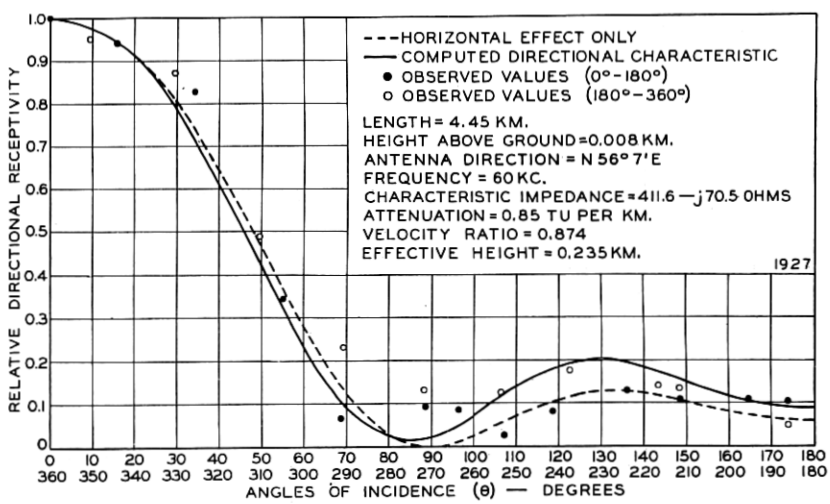


Fig. 8b—Relative Directional Receptivity of Houlton Antenna "A" Uncompensated (Short)

successively at each of the 22 positions shown on the map, Fig. 9. The authors wish to thank Mr. G. D. Gillett for his efficient operation of this transmitter during the summer of 1927.

It is seen that the agreement between the measured and the computed directional characteristic is much better for the shortened Houlton *A* antenna than it is for the same antenna 0.70 kilometer longer.

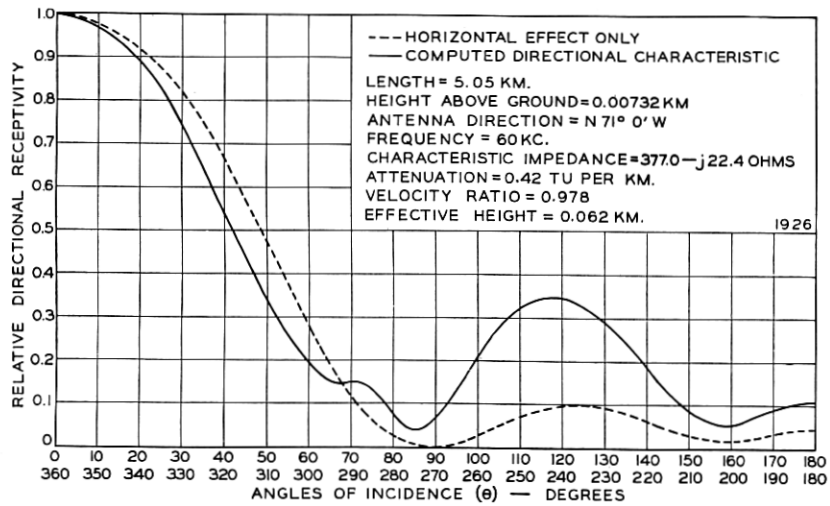


Fig. 8*c*—Relative Directional Receptivity of Wroughton, England—South Antenna

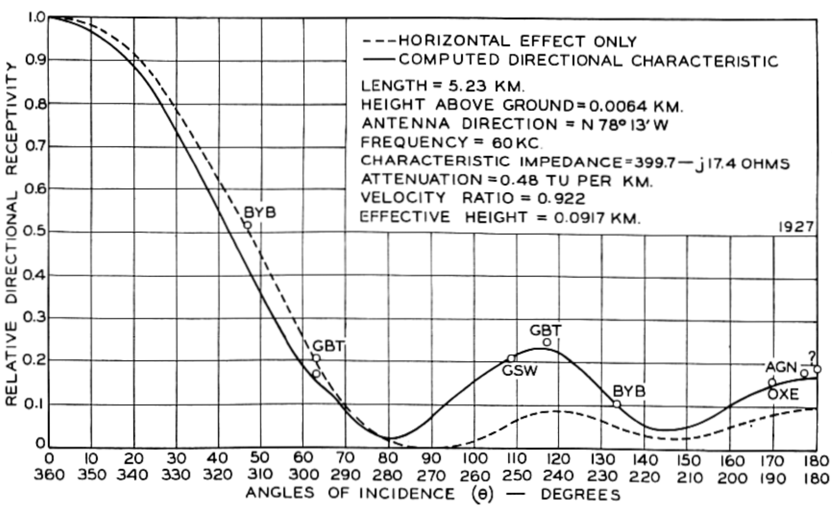


Fig. 8*d*—Relative Directional Receptivity of Cupar, Scotland—Antenna No. 1

The reason for this can be appreciated by reference to Fig. 7, where it is shown that the far end of the long antenna is at the top of a rocky hill; while after shortening, the far end is in a swamp, at the same average elevation as the remainder of the antenna. The elimination of this sharp rise over rocky ground serves principally to remove an irregularity in the constants of the wave-antenna near the end, so that the entire antenna may be considered more nearly a smooth line. This makes the antenna function more satisfactorily as a unit of an array in connection with other antennas constructed nearby.

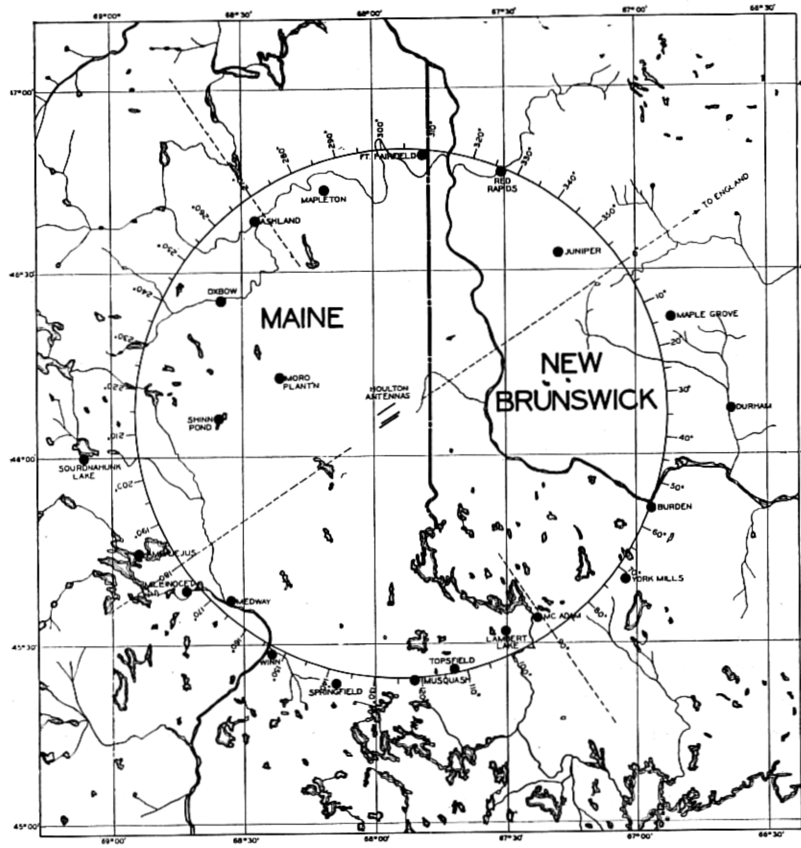


Fig. 9.

6. *Wave-Antenna Arrays.* Since 1899, when S. G. Brown [17] proposed the use of two vertical antennas, separated in space by an appreciable portion of a wave-length and excited at a half-period phase difference, as a means of directional transmission, the use of arrays of antennas

[17] R. M. Foster, "Directive Diagrams of Antenna Arrays," *Bell System Tech. Jour.*, *5*, 292; April, 1926. Also see references listed in Foster's paper.

for directional transmission and reception has become increasingly important. Antenna arrays may be divided into two general classes: (1) arrays of antennas having dissimilar directional characteristics, and (2) arrays of antennas the directional characteristics of which are identical. The array formed by the use of a loop and a vertical antenna to form the familiar "cardioid" is representative of the first class of antenna arrays. Foster [17] has pointed out that the ideal wave-antenna may be considered as an array of an infinite number of loop antennas, extending for the length of the wave-antenna, and hence an antenna array of the second class. (An ideal wave-antenna has no attenuation and a velocity of propagation equal to the velocity of radio propagation in free space.)

An important difference between arrays of dissimilar antennas and arrays of identical antennas lies in the following peculiarity of these two types. In general, the directivity of dissimilar antennas may be increased with no loss in desired signal receptivity by combining them in arrays with little or no separation between the individual antennas. To obtain an increase in directivity by using several identical antennas in an array, however, without too great a sacrifice in desired signal receptivity, the array must cover a space comparable to and of the same order of magnitude as the wave-length of the signals for which it is designed.

It has been stated earlier in this paper that the fundamental form of wave-antenna consists of a single straight horizontal wire, terminated to ground at each end in its characteristic impedance. If the input circuit of a radio receiver be connected across the termination at the end of the antenna most distant from the desired transmitter (the far end of the antenna) this simple form of wave-antenna can be used as a directional receiving system. If arrangements are made to bring the output from the initial end of the wave-antenna to the radio receiver as well as the output from the far end, the simple wave-antenna immediately becomes available for use as two identical antennas in an array. The ends of these two antennas from which the outputs are taken are separated by the length of the antenna and their axes are parallel but in the opposite sense. If before combining these two output currents, that from the initial end of the antenna is changed in phase and magnitude by the proper amount, it is possible to produce a null point of reception in any desired direction. The name "compensation" has been applied to the use of a single wave-antenna to form this array.[14] Since this null point is produced by balancing the back-end current from one antenna of the array (relative to its directional diagram) against the front-end current from the other antenna, the null point does not remain in the directional characteristic over a band of frequencies.

A directional diagram of a single antenna compensated to produce a null point at $\Theta = 161.4$ degrees (the bearing of the Rocky Point transmitter relative to the axis of the antenna) is shown in Fig. 10. This diagram was calculated, by the method outlined in Appendix 2, from the average of the measured constants of Houlton antennas A, B, and D. In this same figure, measured points are indicated, these points being the average of observations on these three antennas.

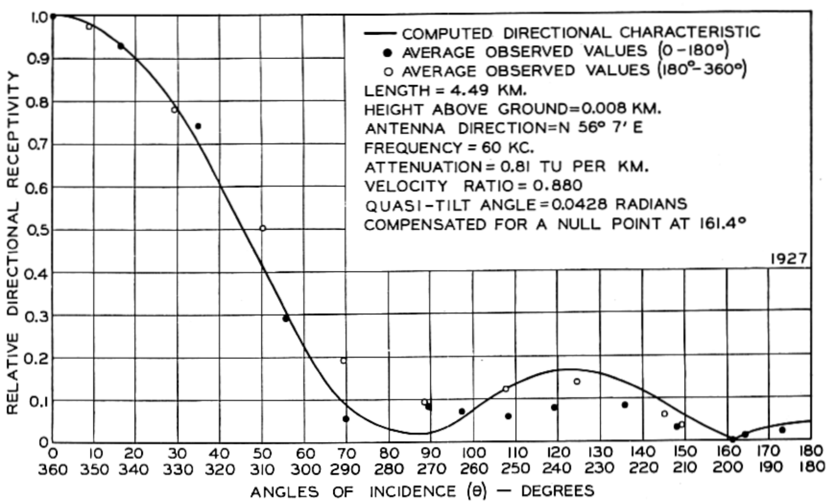


Fig. 10—Wave-antenna directional characteristic. Relative directional receptivity of compensated average Houlton antenna. (Short.)

Beverage, Rice, and Kellog [14] have shown that there are important practical advantages to be gained by constructing the wave-antenna as a two-wire line, and using the metallic circuit acquired thereby as a transmission line to bring the output from one end of the antenna to the radio receiver. The circuits used to bring the output currents from the two ends of the wave-antenna to the radio receiver are shown in Fig. 11. In this case, the radio receiver is located at the initial end of the antenna, so that the predominant desired signal currents are transmitted over the metallic circuit of the wave-antenna to the radio receiver, while the compensation currents are taken directly from the initial end termination when this form of array is used.

To obtain a greater reduction in the Noise Reception Factor (defined under "Directional Discrimination Against Static" earlier in this paper) than is given by compensation, two or more parallel wave-

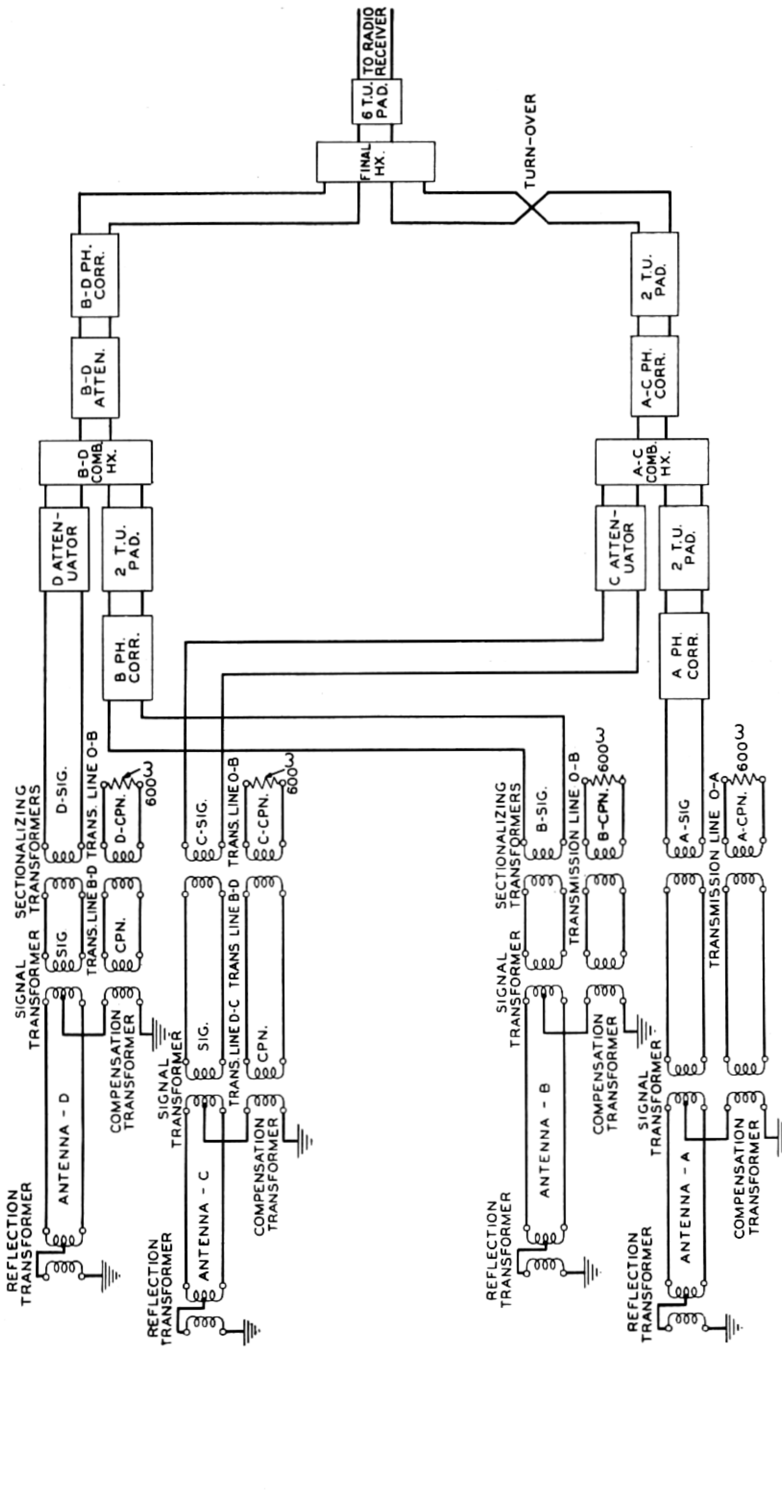


Fig. 11—Houlton antenna system.

antennas are used in either a lateral array, a longitudinal array, or a combination of the two.

In the lateral array, the initial ends of the wave-antennas are spaced in the direction perpendicular to the axes of the antennas. Since it extends over space in the lateral direction, unless there be undue sacrifice in desired signal, the lateral array can only reduce the width of the directional diagram.

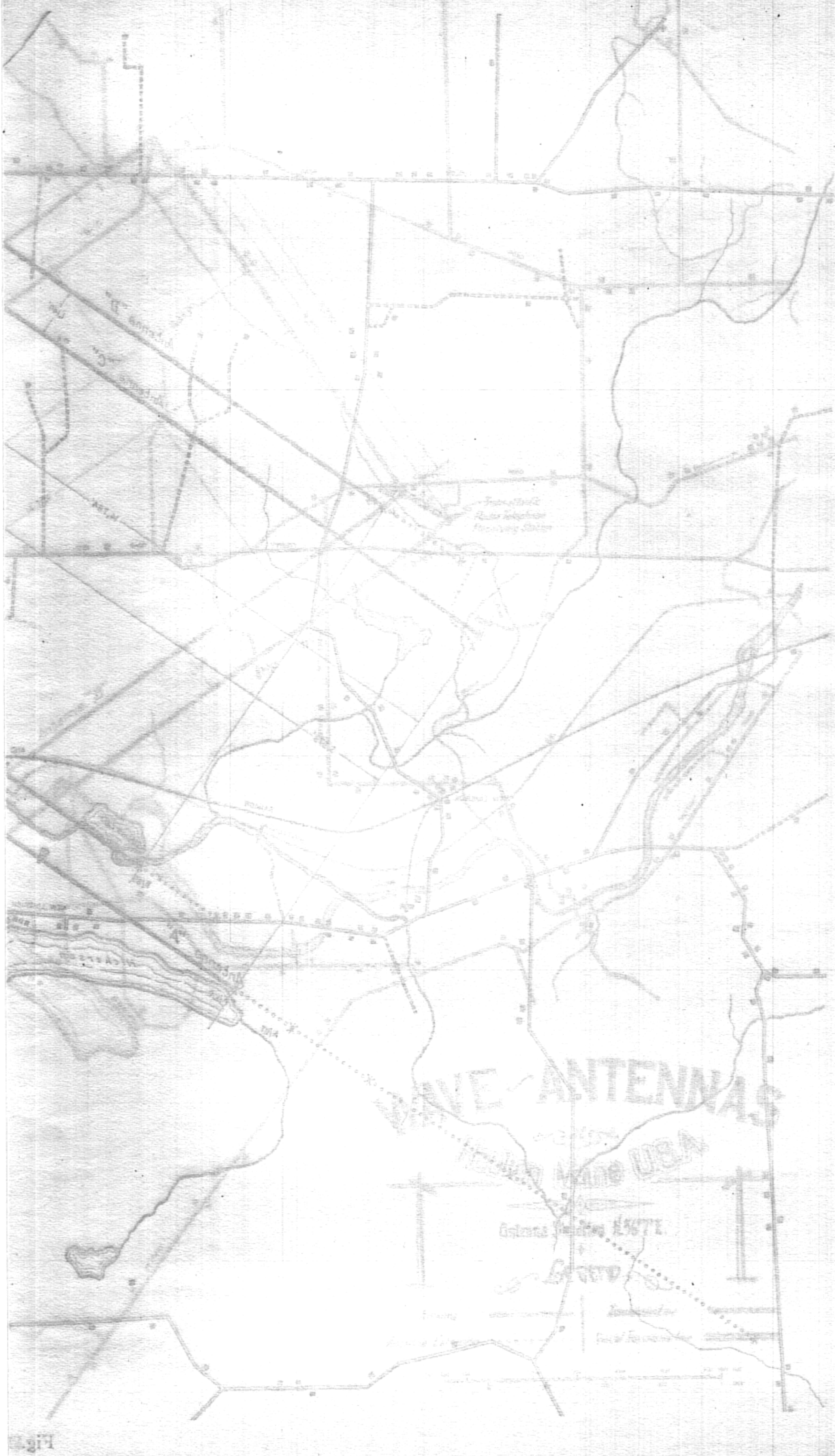
In a true longitudinal array, the antennas are coaxial, but their initial ends are separated by an appreciable fraction of a wave-length. If the wave-antennas forming this type of array overlap one another, then the mutual impedance between them would greatly modify their individual characteristics. In practice, a small amount of lateral spacing between the units of a longitudinal array is necessary to make the mutual impedance negligible. When this type of array is properly designed, the reduction in directional receptivity due to the array is principally in the back-end direction.

The physical layout of the Houlton antenna system is shown in Fig. 12, and the circuits serving to connect the antennas to the radio receiver are shown in Fig. 11. The same letters are used for corresponding line sections in both of these figures. At the time that the directional characteristics of the Houlton antennas were measured, the antenna system comprised only three antennas, A, B, and D. Antenna A at that time extended from pole A-33 to pole A-117. Two arrays could then be used: antennas B and D forming a lateral array, and antennas A and B forming a modified longitudinal array. In normal operation using either of these two arrays, the transducers in the antenna output circuits were adjusted to combine equal amplitudes of the desired signals from the two antennas in phase with one another.

Using as the unit antenna for the arrays a directional diagram derived from the average constants of the antennas A, B, and D, the directional diagrams of these two arrays have been computed. Fig. 13 shows the computed directional characteristic of the lateral array and Fig. 14 the computed directional characteristic of the modified longitudinal array. On each of these figures, the measured points are shown.

The three antennas A, B, and D represented an uneconomical antenna system inasmuch as but two of the antennas could be used simultaneously in an array. To utilize fully these three antennas, at the same time increasing the discrimination against noise, the fourth antenna C has been constructed. To use these four antennas, they are arranged in pairs to form two lateral arrays, and the two lateral arrays arranged in a longitudinal array. The resultant total array


WAVE ANTENNAS
at
Houlton Maine U.S.A.
Antenna "B"


ANTENNAS
USA

quite evidently combines the narrowing of the directional diagram due to the lateral array and the reduction of the back end area of the

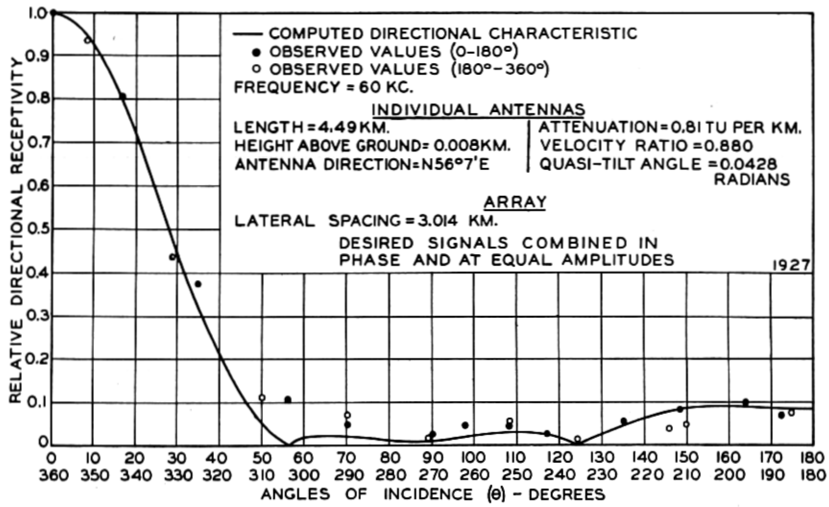


Fig. 13—Wave-antenna array directional characteristic. Relative directional receptivity of lateral array of two Houlton antennas. (Short)

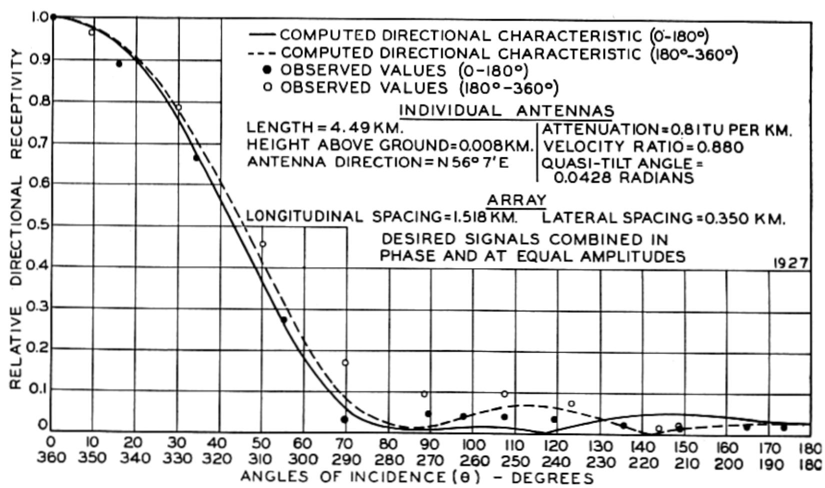


Fig. 14—Wave-antenna array directional characteristic. Relative directional receptivity of modified longitudinal array of two Houlton antennas. (Short)

directional diagram caused by the longitudinal array. A map of this array is shown in Fig. 12.

The circuits for combining four antennas of an array of the type described in the preceding paragraph are shown in Fig. 11. Antennas *A* and *C* form one lateral array; antennas *B* and *D* form the second. Since antennas *C* and *D* are further removed from the station than *A* and *B*, phase correctors are inserted in the circuits from *A* and *B* to compensate for the phase change in the transmission lines from *C* and *D*, so that the desired signals are combined in phase. The combination of the 2 TU fixed pads and the variable attenuators makes it possible to correct for the attenuation in the transmission lines to the more distant antennas. These several output currents are actually combined in hybrid coils, since this method of combination prevents the antennas from reacting one upon another through the combining system.

After the antennas are combined in pairs to form two lateral arrays, the lateral arrays are combined in the longitudinal array.

The change of phase of space waves between one antenna and the next in an array is a linear function of frequency, and that on the metallic transmission lines practically so. By using phase correctors which have a phase change linear with frequency,[18] the outputs of the antennas in the array may then be combined to produce a null point or a reduction in receptivity, as a result of the array, which retains the same position in the directional diagram for every frequency within a finite band. The longitudinal array at Houlton is designed and combined to produce such an invariable null point in the direction 161.4 degrees relative to the axis of the wave-antenna array. At this angle of incidence, it is evident that the space waves arrive at the lateral array of antennas *A* and *C* before arriving at the lateral array of antennas *B* and *D*. To bring these undesired signals in phase, therefore, phase shift must be introduced into the output of the first of these arrays. Part of this phase shift is supplied by the metallic transmission lines and part by the phase correctors in the combining equipment. At this point, the undesired signals remain in phase as the frequency is varied, so that a turn-over (reversal) inserted in the circuit to the lateral array of antennas *A* and *C* before the array is combined produces the null point which is invariable with variation of the frequency. Under these conditions, the phase of combination of the desired signals, incident at zero angle, varies as the frequency of the desired signals varies. To minimize the effect of this change in phase over the desired frequency band, the spacing of the antennas in the longitudinal array must be so chosen that the desired signals combine very nearly in phase at the middle of the frequency band. For that

[18] O. J. Zobel, "Distortion Correction in Electrical Circuits with Constant Resistance Recurrent Networks," *Bell System Tech. Jour.*, 7, 438; July, 1928.

reason, the longitudinal spacing in the modified longitudinal array was decreased at the same time that the fourth antenna was constructed.

At the time that the extension of the antenna system was undertaken, the measured directional characteristics of the antennas A, B,

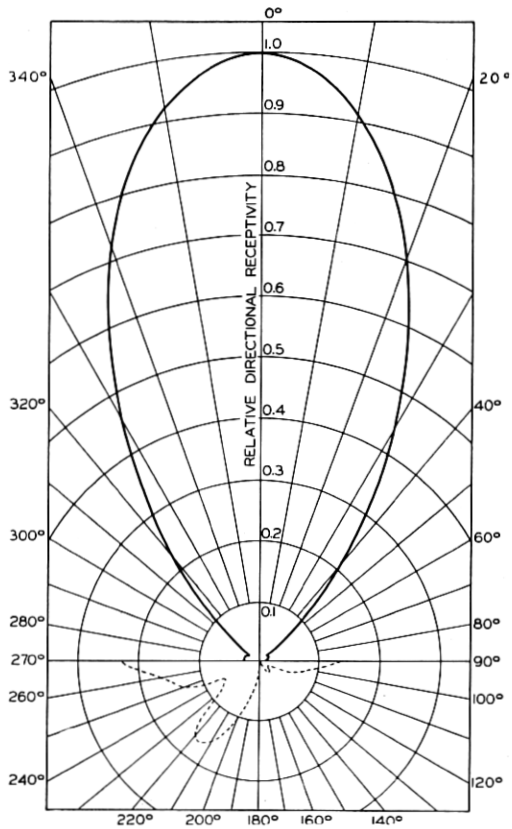


Fig. 15—Wave-antenna array directional characteristic. Calculated relative directional receptivity of array of Houlton antennas A, B, C, and D. (From average measured unit antenna characteristic.) Dotted curve—magnified $\times$ 10.

and D were available, so that the unit directional diagram for use in determining this array characteristic was taken as the average measured characteristic of these three antennas. The calculated directional diagram of the complete Houlton antenna system is shown in Fig. 15. It should be noticed that the scale for the back-end dotted curve is ten times as great as that for the full-line curve for the major lobe.

It is believed that the directional diagram, shown in Fig. 15, represents about the ultimate that can be done economically in a general reduction of back-end area and narrowing of the diagram by means of wave-antennas. Future extensions or redesign of the array at Houlton must be based on the reduction of the relative receptivity in distinct directions determined either by statistical study of the noise received by the antenna system or actual measurements of the direction of arrival of the noise which limits the operation of the transatlantic radio-telephone circuit.

The Radio Receiver

A description of the design and performance of the radio-telephone receiving set will constitute another paper. The radio receiving equipment employed in connection with the antenna systems was developed and constructed by Bell Telephone Laboratories.

The major transmission requirements upon which design must be based are as follows:

1. The limiting values of the signal field to be received;
2. The output power of the receiving antenna for a given signal strength;
3. Power output required from the radio receiver;
4. The type of telephone transmission to be received;
5. The frequency band to be received;
6. The nature and strength of interference from other radio stations and from noise; The selectivity required to reduce undesired modulation
   a. In amplifiers,
   b. In demodulators;
7. Stability of frequency, gain, and transmission-frequency characteristic.

1. Limiting Values of the Signal Field to be Received. The range of daily averages of signal field at 60 kilocycles for all daylight path hours is shown in Fig. 4. The fields, as previously published data indicate,[4, 6, 9] vary diurnally between much wider limits. At night the field frequently approaches, as a maximum, the value calculated on the basis of the inverse distance law. During sunrise and sunset dip periods the field frequently goes to a value less than one microvolt per meter with even 50 kilowatts radiated from a transmitter 5,000 kilometers away. Suppose we take as being approximately correct values, field strengths of 0.4 microvolts per meter as the lower limit and 400 microvolts per meter as the upper limit. We then have determined that the receiving set should have a variation of gain of 60 TU.

2. The Output Power of the Receiving Antenna for a Given Signal Strength. From the observed constants of a Houlton wave-antenna and the assumed value of 0.4 microvolts per meter received at zero degrees to the antenna direction as the lowest field, we calculate, using equations (125), (126), and (127) in Appendix 1, that the power supplied to the reflection transformer terminals is 3.716×10^{-6} microwatts. This power must suffer loss as a result of the transmission back to the receiving station over transmission lines and as a result of the necessity of providing flexibility in the operation of the apparatus used to combine the output of the antenna in question with the output of other antennas before it reaches the input terminals of the radio receiving set. (See Fig. 11.) This loss is such that the power at the input terminals of the radio receiving set from a single antenna and for the minimum signal field is very nearly equal to 3.7×10^{-7} microwatts. With the combining system actually used, the input to the radio receiver from all four antennas will be 12 TU above this value or 5.9×10^{-6} microwatts.

3. Power Output Required from the Radio Receiver. The value of output power required from the radio receiver is really governed by considering the whole radio circuit as a part of a long-distance telephone system. An overall loss of 10 TU has been found satisfactory for long toll circuits. If the telephone lines connecting the circuit terminals to the transmitting and receiving stations have an equivalent of 0 TU then we can place the 10 TU loss in the radio portion of the circuit. If we then supply on a single frequency within the voice-frequency band a power of 1 milliwatt to the input terminals of the radio transmitter, to get a 10 TU equivalent in the radio circuit we must obtain 0.1 milliwatt at the output of the radio receiver.

In the preceding section we determined that the minimum input would be 3.7×10^{-7} microwatts from a single antenna and hence the maximum gain required in the radio receiver to raise this power to the specified 100 microwatts output is 84 TU.

Within amplifiers using three-electrode vacuum tubes, noise is generated in two ways: (*a*) by thermal agitation [19] in the conductor of the input circuit; and (*b*) by "Schottky Effect" [20] in the vacuum

[19] J. B. Johnson, "Thermal Agitation of Electricity in Conductors," *Phys. Rev.*, *32*, 97; July, 1928.
Harry Nyquist, "Thermal Agitation of Electric Charge in Conductors," *Phys. Rev.*, *32*, 110; July, 1928.
J. B. Johnson, "Thermal Agitation of Electricity in Conductors," *Nature*, *119*, 50; Jan. 8, 1927.

[20] Walter Schottky, "Atomare Schwankungsvorgänge an Glühkathodenoberflächen," *Physik. Zeitschr.*, *27*, 701; Nov. 1, 1926.
T. C. Fry, "The Theory of the Schroteffekt," *Jour. Frank. Inst.*, *199*, 203; Feb., 1925.

tubes themselves. Since the transatlantic radio-telephone circuit is so operated that the strength of the voice waves, or "electrical volume," is constant at the output of the radio receiver,[21] the maximum allowable noise at this point in the circuit is likewise constant. Good engineering practice specifies that the continuous "tube noise" should be more than 40 TU below the signal or less than 0.01 microwatt for the specified receiver output of 100 microwatts when using any gain up to the maximum of 84 TU. (With uniformly distributed noise over the voice-frequency band, this is equivalent to about 400 noise units.[22])

4. *The Type of Telephone Transmission to be Received.* The "single-sideband, suppressed-carrier" type of telephone transmission, invented by John R. Carson,[23] has long been used in the Bell System in carrier systems on wire circuits.[24] Since the advantages of single sideband in radio transmission have been described by Hartley,[25] and in the radio transmitter by Heising,[26] we shall only briefly review the benefits arising from its use.

Transmission of two sidebands with the carrier suppressed represents an improvement over the "carrier and two-sideband" method ordinarily used in "broadcasting" since all of the transmitter power may be concentrated in the intelligence-bearing frequencies. By transmitting only one sideband, further advantages are gained since the frequency space occupied is slightly more than halved for the same grade of circuit, the distortion at the output of the receiver is decreased, and practical simplifications may be made at the transmitting and receiving stations.[27] If the radio transmitter radiates equal power in each of the above-mentioned suppressed carrier transmission schemes and if the radio receiver accepts only the intelligence-bearing fre-

J. B. Johnson, "The Schottky Effect in Low Frequency Circuits," *Phys. Rev.*, *26*, 71; July, 1925.

[21] S. B. Wright and H. C. Silent, "The New York-London Telephone Circuit," *Bell System Tech. Jour.*, *6*, 736; October, 1927.

[22] The noise unit is an arbitrary unit used in the Bell System for comparison of any noise with a certain arbitrary source of noise known as a noise standard. The output of the noise standard may be attenuated to produce the same interfering effect on speech as the noise being measured. See W. H. Harden, "Practices in Telephone Transmission Maintenance Work," *Bell System Tech. Jour.*, *4*, 26; Jan. 1925, for details of making such comparisons.

[23] U. S. Patents Nos. 1,343,306 (1920); 1,343,307 (1920); 1,449,382 (1923), to J. R. Carson.

[24] E. H. Colpitts and O. B. Blackwell, "Carrier Current Telephony and Telegraphy," *Trans. A. I. E. E.*, *40*, 205; 1921.

[25] R. V. L. Hartley, "Relation of Carrier and Sideband in Radio Transmission," *Proc. I. R. E.*, *11*, 34; Feb., 1923.

[26] R. A. Heising, "Production of Single Sideband for Transatlantic Radio Telephony," *Proc. I. R. E.*, *13*, 291; June, 1925.

[27] J. R. Carson, "Signal-to-Static Interference Ratio in Radio Telephony," *Proc. I. R. E.*, *11*, 271; June, 1923.

quencies in each case, then the signal-to-noise ratio will be the same,[25] provided the resupplied carrier is in frequency synchronism in both systems and in addition in phase synchronism with the suppressed carrier in the two-sideband system.[27]

When receiving single-sideband transmission the carrier suppressed at the transmitting station is resupplied in the radio receiver. Since this carrier will demodulate both sidebands with equal efficiency, the opposite sideband must be eliminated before demodulation to prevent the noise in this sideband from appearing in the voice-frequency output. If the noise power in either sideband is p, then the noise power without opposite sideband suppression is $2p$ and if we reduce the noise power from the opposite sideband to $0.1p$ the total received noise will be reduced

$$10 \log_{10} \frac{2p}{p + 0.1p} = 2.59 \text{ TU}.$$

The maximum possible reduction in noise is 3.01 TU,[28] so that for engineering purposes a 10 TU suppression of the noise in the opposite sideband may be considered adequate. For other reasons to be brought out later in this paper the opposite sideband loss must be greatly in excess of this value.

Provided the resupplied carrier used to demodulate the single sideband suppressed carrier signals is large relative to the signal magnitude [25] at that point in the circuit where we choose to supply it, the only other requirement is that its frequency be correct. Since a displacement of the resupplied carrier 50 cycles above or 20 cycles below the zero of the equivalent voice-frequency band is sufficient to give an appreciable decrease in speech intelligibility, its frequency should be maintained within the smaller of these two limits or within plus or minus 20 cycles of the correct value. It is interesting to note that an absolute variation of only one-tenth of this amount can be observed on music and that speech naturalness is similarly affected.

5. *Frequency and Frequency Band to be Received.* To utilize the power available at the transmitter most effectively, it is essential to transmit only those frequencies contributing most to received intelligibility. The energy of speech lies largely below 500 cycles while the frequencies most important for intelligibility lie between 400 and 2,600 cycles.[29] By limiting the band transmitted to speech frequencies above 400 cycles some saving is obtained in the transmitter power

[28] J. R. Carson, "Selective Circuits and Static Interference," *Bell System Tech. Jour.*, *4*, 265; April, 1925.

[29] W. H. Martin and Harvey Fletcher, "High Quality Transmission and Reproduction of Speech and Music," *Trans. A. I. E. E.*, *43*, 384; 1924.

required. The range of frequencies transmitted may then extend from 58.9 to 61.1 kilocycles with the suppressed carrier at 58.5 kilocycles, and the radio receiver must be designed to accept this band of frequencies. The transmission-frequency characteristic of the overall radio receiving set should not vary more than $\pm$ 2 TU within the band specified above to give a good telephone communication circuit.

6. Selectivity Requirements. The selectivity required in the receiving set is such that when the desired signal is at the assumed minimum value no deleterious effects will be caused by undesired signals.

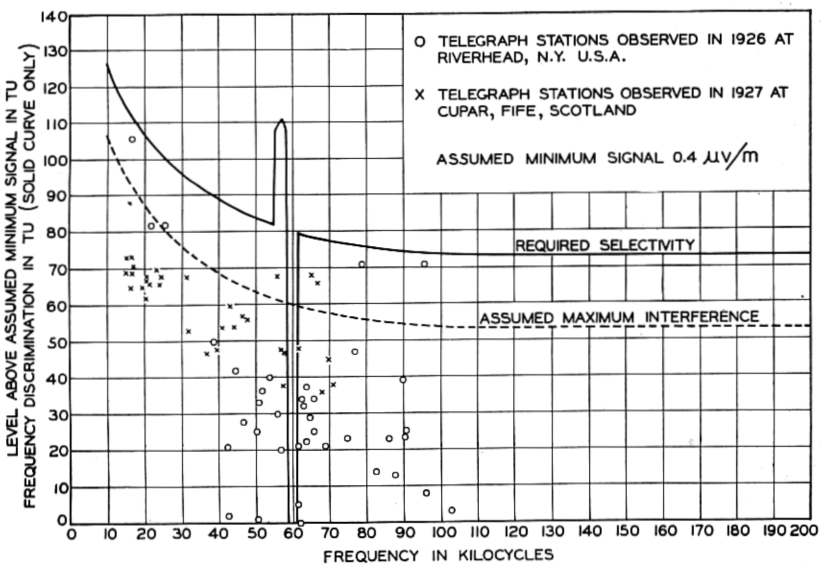


Fig. 16—Selectivity requirements for the long wave transatlantic radio-telephone receiving system.

In Fig. 16 there are shown measured daylight field strengths of various existing radio-telegraph stations as observed at Riverhead, New York, and at Cupar, Scotland. Since these measurements could not be indefinitely extended in frequency nor could they take into account all stations which might exist in this range, they may be considered only as a guide in obtaining a curve of the maximum telegraph interference to be expected. These data, in Fig. 16, have been expressed as ratios (in TU) to the minimum desired signal to be received. It is important to note that the directional selectivity of the receiving antenna system used materially decreases the relative magnitude of many of these interfering signals, particularly at the higher frequencies. The American receiving station is now located in

Houlton, Maine, instead of Riverhead, New York, and this increase in distance from many of the American high-power transmitting stations decreases somewhat their field strengths. In view of these factors the "Assumed Maximum Interference," although it is not greater than any observed station field strength, is, in fact, greater than the interfering signals when the outputs from the actual receiving antennas are used instead of field-strength observations.

6a. Selectivity Requirements Imposed by the Use of Amplifiers. To operate a vacuum tube as an amplifier with negligible distortion the peak voltage applied to its grid must be less than a limiting value so that the tube always operates over the practically linear portion of its characteristic. If no discrimination were provided against unwanted signals, we would be placed in the peculiar situation of having to supply ample tube capacity in the radio receiver to care for the combined load produced by perhaps 100 telegraph stations each of which, on the average, may have a received signal strength 1,000 times the assumed minimum signal. An easy way to decrease the load produced by interference is to insert a filter at the input of the receiving set, which will reduce the required capacity of the first tube. Additional selectivity following the first tube still further reduces the load of undesired signals on the following tubes as more of the capacity of those tubes is used for the desired signals.

Now for design purposes let us assume that the load capacity of each tube is at least 6 TU greater than the capacity required in the tube for the performance of its functions on the desired signal. The undesired signals may then be allowed to produce on the tube grid a voltage equal to that of the desired signal.

Since each of the undesired signals shown in Fig. 16 are about 60 TU stronger than the minimum desired signal, they must be reduced by that amount to make them each no greater than the desired signal.

It is shown in Fig. 21 of Appendix 4 that, as a result of unit random input voltages from 100 operating radio-telegraph stations, a peak voltage will be produced equal to or greater than 10 such units during less than 0.1 per cent of the time. If the undesired telegraph station signals were all of the same magnitude as the desired signal then the voltage which they would produce would be 20 TU above the voltage of the desired signal.

From purely load considerations then, the total required suppression of every interference-bearing frequency outside of the desired signal receiving band will be

$$60 + 20 = 80 \text{ TU}.$$

6b. Selectivity Requirements Imposed by the Use of Demodulators. In all demodulators, the range of desired output frequencies should not be included in the input frequency band because the input frequencies amplified appear in the output of the demodulator as the first order modulation product. The output band of the demodulator should be at a lower frequency than the input band in order to reduce the number of undesired modulation products in the output and in order to obtain the benefit of greater selectivity from circuits operating at lower frequencies. Of course, if all the required selectivity can be conveniently put before the first demodulator, there is no valid reason why multiple demodulation should be used.

In all demodulators except the final demodulator of a radio receiver, the band of frequencies allowed to pass into the demodulator should not be greater in width than the absolute value of the lowest desired frequency in the demodulator output. This requirement is apparent when we consider second order modulation products of interference within the band accepted by the demodulator. Suppose we assume the use of double demodulation and choose 30 kilocycles as the lowest desired frequency in the output of the first demodulator. Then if the band impressed upon the demodulator be more than 30 kilocycles in width, two interfering signals within the band might together give a difference frequency of 30 kilocycles producing load in subsequent stages and possibly tone or noise in the output circuit.

Second order modulation between two signals, one lying within the band accepted by the demodulator and one outside it, may also give rise to interference, due to the difference frequency falling in the output band of the demodulator. Assume that one interfering signal lies within the band accepted by the demodulator and is + 60 TU referred to the minimum desired signal at the grid of the first tube. An equal signal at a frequency outside the band and subject to the selectivity provided for meeting the load requirement will be − 20 TU referred to the minimum desired signal at the same point. Since the second order output from a demodulator is approximately proportional to the product of the grid voltages producing it,[30] we may write (in TU):

$$\text{Relative desired signal} = (0) + (\text{Beating oscillator voltage}),$$
$$\text{Relative interference} = (+\,60) + (-\,20) = +\,40.$$

Tests have shown that an interrupted tone, similar to telegraph interference, which is heard at a frequency of 1,100 cycles in a tele-

[30] J. R. Carson, "A Theoretical Study of the Three-Element Vacuum Tube," *Proc. I. R. E.*, 7, 187; April, 1919.

phone receiver is about the most serious frequency of interference to received speech on a telephone circuit, and that frequencies above and below 1,100 cycles are of somewhat less importance. Signals at the equivalent 1,100-cycle frequency produce a type of interference which good engineering practice requires should be reduced at least 50 TU below the desired signal. (This amount of interference is equal to about 500 noise units at the − 10 TU transmission level.[22])

To satisfy this requirement, the relative desired signal should be 50 TU greater than the relative interference at the output of the demodulator. Assigning a minimum magnitude to the beating oscillator voltage of 90 TU above the minimum desired signal voltage on the grid of the demodulator reduces this type of interference sufficiently. (Using a balanced demodulator arrangement, this value might be reduced some 20 TU).

Since any two signals at frequencies entirely outside the band accepted by the demodulator are suppressed some 80 TU, we need not consider their second-order modulation products.

If we use double demodulation in a radio receiver as is assumed in the first part of this section, then we must consider other products of modulation with the beating oscillator for frequencies distant from the accepted band. Space will not permit us to more than mention these, but since the frequencies to be suppressed are distant from the frequencies to be received, their suppression is relatively simple.

In the final demodulator of a radio receiver we must tolerate a certain amount of distortion due to the intermodulation of input frequencies. By limiting the band width into the final demodulator to the same width as the desired output band, the distortion due to intermodulation with interference lying outside the desired band is eliminated. By supplying a large amount of carrier to the final demodulator and by using a balanced demodulator the amount of noise and distortion due to intermodulation of frequencies lying inside the desired band is reduced.

If the desired signal band extends from 58.9 to 61.1 kilocycles and is an upper sideband corresponding to voice frequencies from 400 to 2,600 cycles then, in effect, we must supply a carrier at 58.5 kilocycles to produce the proper voice frequencies in the output circuit. This carrier frequency will also demodulate the frequencies below it in such a way as to produce audible signals and for this reason ample protection must be supplied against the opposite sideband if stations are likely to exist in that range. Calculations show that this is the case; for if 100 stations are distributed at random over the 190-kilocycle range between 10 and 200 kilocycles, then the probability that at least one station lies between 55.5 and 58.5 kilocycles is 0.796. If the assumed maximum interference at the equivalent 1,100 cycles in the opposite sideband, as indicated by the dashed line in Fig. 16, is 61 TU above the minimum signal and we wish to have it 50 TU below, as previously stated, then we require a selectivity of 61 TU plus 50 TU or 111 TU for this frequency. For other tone-producing frequencies of the opposite sideband similar selectivity requirements have been set up and the resultant for frequencies from 55.5 to 58.5 kilocycles is shown by the solid curve in Fig. 16.

7. *Stability.* As mentioned in Section 4 above, the carrier for a single sideband receiver must be resupplied at the correct frequency. All of the oscillators in the radio link must have sufficient frequency stability to maintain the voice frequencies at the receiver output correct within 20 cycles per second over long periods of time. Suppose we allow 10 cycles per second variation in frequency to exist at the transmitter and an equal amount at the receiver, then the variation in frequency at the receiver must never exceed 0.017 per cent if the resupplied carrier is at 58.5 kilocycles. Certain advantages in stability of the resupplied carrier can be obtained by the use of double demodulation in the receiving set and these will be discussed in another paper.

Variations in the efficiency of the transatlantic radio transmission path for long wave-lengths occur with time of day and season, but during any individual all-daylight transmission period the transmission efficiency of the path is fairly constant. If the gain of the receiver is constant, then, during this important period of the day, the minimum of circuit adjustments will be required. It is hence desirable that the gain of the entire receiving set be made to hold constant within $\pm$ 2 TU for all variations of temperature and of voltage of battery supply, within the operating limits.

It is almost self-evident that the transmission-frequency characteristic through the radio receiver should not vary with temperature and time. Changes of this nature should not exceed 0.5 TU within the transmission band nor 5 TU outside of the transmission band. Design of stable filters and vacuum-tube circuits are essential to produce this result.

The authors have endeavored, in the limited space of the preceding pages, to show what radio transmission considerations must be taken into account in properly designing a receiving system for a commercial radio-telephone circuit. A rather detailed discussion has been necessary to present an accurate picture of the various factors entering into the production of the very essential and highly directional long-wave receiving antenna system employed.

The cooperation of the engineers of the Wireless Section of the British General Post Office, particularly Col. A. G. Lee and Mr. I. J. Cohen, in the measurements made on wave-antennas in England and Scotland, is greatly appreciated and we take this occasion to thank them for having made possible the obtaining of these data. All of our early work in connection with wave-antennas and our initial field trials of lateral and longitudinal arrays of wave-antennas were carried out using wave-antennas located at Belfast, Maine, and Riverhead, New York. These antennas were made available through the courtesy of the Radio Corporation of America, and the authors wish to express to Mr. H. H. Beverage of that organization their appreciation for his interest and assistance during the tests.

APPENDIX 1

The Wave-Antenna

Fundamentally, the wave-antenna consists of a straight horizontal wire, terminated to ground at each end in its characteristic impedance.[14] The determination of the receptivity characteristics of the wave-antenna consists in determining the current flowing in the terminal impedances of the antenna resulting from a field impressed along the antenna.[31]

The wave-antenna is shown in Fig. 17, consisting of a line of length s extending from $x = 0$ to $x = s$. In the nomenclature of the following discussion, letters with no primes refer to the antenna, letters with a single prime (′) to the impressed field, and letters with a double prime (″) to the resultant field. The wave-antenna is in an impressed electromagnetic field which is defined by the quantities ϕ', V', f_w', and f_g' where

ϕ' = impressed magnetic flux between the lower surface of the wire and the surface of the ground (per unit length);
V' = impressed electric force between the wire and the ground;
f_w' = the impressed electric force *along* the lower surface of the wire;
f_g' = the impressed electric force *along* the surface of the ground.

The total field about the antenna is the sum of this impressed field and a secondary field due to the currents and charges produced in the circuit by the impressed field, so that

$$\begin{aligned} \phi'' &= \phi' + \phi, & f_w'' &= f_w' + f_w, \\ V'' &= V' + V, & f_g'' &= f_g' + f_g, \end{aligned} \tag{101}$$

[31] J. R. Carson and R. S. Hoyt, "Propagation of Periodic Currents over a System of Parallel Wires," *Bell. System Tech. Jour.*, *6*, 495; July, 1927.

where ϕ, V, f_w, and f_g are the components of the secondary field set up by the currents and charges in the system and ϕ'', V'', f_w'', and f_g'' represent the resultant field about the system.

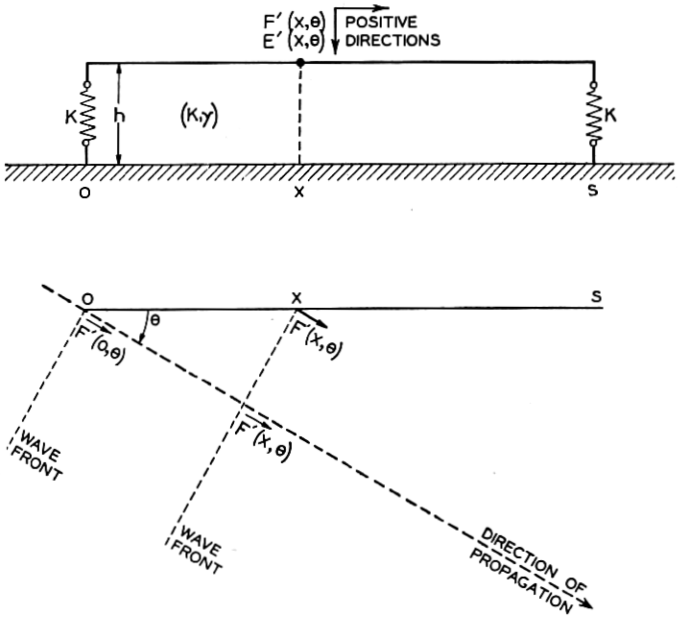


Fig. 17.

As a result of the impressed field, a current I flows in the wire, and a corresponding superposed current distribution is induced in the ground. If the internal impedance of the wire be z_w and that of the ground be z_g, the resultant longitudinal electric force along the wire may be written

$$f_w'' = Iz_w = f_w' + f_w \tag{102}$$

and similarly the resultant longitudinal electric force along the ground is

$$f_g'' = (-Iz_g + f_g') = f_g' + f_g. \tag{103}$$

The second curl law applied to the periphery of the rectangle formed by the vertical at x, the wire, the vertical at $(x + \Delta x)$, and the ground yields

$$zI - f_g' + \frac{dV''}{dx} = -\frac{d\phi''}{dt}, \tag{104}$$

where z is the total series impedance of the wire and the ground circuit and is

$$z = z_g + z_w \tag{105}$$

A summation of the voltages around the above defined rectangle yields

$$f_w' - f_g' + \frac{dV'}{dx} = -\frac{d\phi'}{dt}. \tag{106}$$

Subtracting (106) from (104) we get

$$zI - f_w' + \frac{dV}{dx} = -\frac{d\phi}{dt}. \tag{107}$$

If we write Q as the charge, C as the capacity to ground, and L as the external inductance, each per unit length of the wire, equation (107) becomes

$$zI + L\frac{dI}{dt} + \frac{1}{C}\frac{dQ}{dx} = f_w', \tag{108}$$

but the line current is decreased by the amount of the charging current and the leakage current

$$-\frac{dI}{dx} = \frac{dQ}{dt} + I_Y, \tag{109}$$

where I_Y is the leakage current per unit length of the wire. If the admittance of the leak to ground be designated as Y, the leakage current is

$$I_Y = YV'' = Y(V' + V). \tag{110}$$

Since we are interested only in the steady state, the operator d/dt may be replaced by $j\omega$. Substituting the expression (110) for I_Y into (109) and differentiating with respect to x yields

$$-\frac{d^2I}{dx^2} = \frac{dQ}{dx}j\omega + Y\frac{dV'}{dx} + \frac{Y}{C}\frac{dQ}{dx}. \tag{111}$$

By means of (111) we may eliminate Q from (108)

$$(z + jL\omega)I - \frac{1}{Y + jC\omega}\frac{d^2I}{dx^2} = f_w' + \frac{Y}{Y + jC\omega}\frac{dV'}{dx} \tag{112}$$

and if

$$K = \sqrt{\frac{z + jL\omega}{Y + jC\omega}}, \tag{113}$$

$$\gamma = \sqrt{(z + jL\omega)(Y + jC\omega)}, \tag{114}$$

where K is the characteristic impedance and γ the propagation constant of the antenna circuit, equation (112) may be written

$$\frac{K}{\gamma}\left(\gamma^2 - \frac{d^2}{dx^2}\right) I = f_w' + \frac{YK}{\gamma}\frac{dV'}{dx}. \tag{115}$$

When the boundaiy conditions are applied, equation (115) defines the value of the current I in the wave-antenna in terms of the impressed electromagnetic field specified by V' and f_w'. By equation (101) the resultant voltages at the ends of the antenna are:

$$V''(0) = V'(0) + V(0), \tag{116}$$

$$V''(s) = V'(s) + V(s). \tag{117}$$

To this point, the solution of the wave-antenna problem has been in a rigidly analytic form. While it is possible to determine completely the received current by following through this method of solution, the problem can be greatly simplified and a physical picture of the problem gained by a synthetic process.

The synthetic method of attack consists of replacing the impressed field by a set of electromotive forces identically equivalent to the impressed field in the sense that it produces the same currents and charges.[31]

The proposed set of electromotive forces is as follows:

A. A distributed longitudinal electromotive force f_w' per unit length in the wire, i.e., an electromotive force $f_w'dx$ in each element of length dx;

B. A distributed vertical electromotive force, V', in the superposed shunt admittance Y between the wire and ground, i.e., an electromotive force V' in each elemental admittance path Ydx;

C. In each end of the wire, $x = 0$ and $x = s$, localized series electromotive forces, equal respectively to minus and plus the impressed voltages at those points; i.e., equal to $- V'(0)$ and $+ V'(s)$ respectively.

The electromotive force of A is suggested by (107), that of B by (109) and (110), that of C by the terminal conditions expressed in (116) and (117). In the case of a wave-antenna constructed to maintain high insulation resistance, the conductance portion of the superposed admittance Y can be made negligibly small. Under this condition, the susceptance part of this admittance can be combined with the linear capacitance of the wire to alter the propagation constants (K and γ) of the antenna and the voltages induced in the superposed shunt admittances neglected.

By reference to Fig. 17, the impressed field may be identically defined at each point along the antenna.

The longitudinal electromotive force in each element of the wire is

$$\begin{aligned} f_w' dx &= F'(x, \theta) \cos \theta \, dx, \\ f_w' dx &= F'(0) \epsilon^{-\gamma' x \cos \theta} \cos \theta \, dx. \end{aligned} \tag{118}$$

The impressed voltage at the point x along the antenna is

$$\begin{aligned} V'(x) &= h \cdot E'(x, \theta), \\ V'(x) &= h \cdot E'(0) \epsilon^{-\gamma' x \cos \theta}. \end{aligned} \tag{119}$$

In (118) and (119) $F'(0)$ and $E'(0)$ represent the horizontal and vertical components respectively of the impressed electric field at the end of the antenna $x = 0$, and h represents the height of the antenna above ground. For the purpose of this discussion, it will be assumed that F' and E' are not dependent upon θ. The current produced at the receiving end s by the horizontal component of the impressed field is given by

$$I_{F'\theta} = \int_0^s \frac{F'(0) \epsilon^{-\gamma' x \cos \theta} \cos \theta \, dx}{2K} \epsilon^{-\gamma(s-x)}, \tag{120}$$

from which

$$I_{F'\theta} = \frac{sF'(0) \cos \theta}{2K} \frac{\epsilon^{(\gamma - \gamma' \cos \theta)s} - 1}{(\gamma - \gamma' \cos \theta)s} \epsilon^{-\gamma s}. \tag{121}$$

The current produced at the receiving end s by the vertical component of the impressed field is evaluated as follows:

$$I_{E'\theta} = \frac{V'(s)}{2K} - \frac{V'(0)}{2K} \epsilon^{-\gamma s} \tag{122}$$

and by combination of (119) and (122)

$$I_{E'\theta} = \frac{hE'(0)}{2K} \left[\epsilon^{(\gamma - \gamma' \cos \theta)s} - 1 \right] e^{-\gamma s}. \tag{123}$$

Zenneck's theory of wave propagation [32] has been developed by Breizig [33] to show that the horizontal and vertical components of the impressed field are related by the expression

$$-\frac{F'}{E'} = \epsilon^{j\delta} \tan T. \tag{124}$$

[32] J. Zenneck, "Ueber die Fortpflanzung ebener electromagnetischer Wellen längs einer ebenen Leiterfläche und ihre Beziehung zur drahtlosen Telegraphie," *Ann. der Phys.*, *23*, 846; June, 1907.

[33] Franz Breizig, "Theoretische Telegraphie," Braunschweig, 1924. 2d ed., pp. 482–487.

The total current produced at the receiving end s by the impressed field is

$$I_\theta = I_{F'\theta} + I_{E'\theta} \tag{125}$$

and by application of (124) the constituents of the total current are

$$I_{F'\theta} = \frac{S\lambda' F'}{2K} \cos\theta \frac{1 - \epsilon^{-[\alpha S\lambda' + j2\pi S(m - \cos\theta)]}}{\alpha S\lambda' + j2\pi S(m - \cos\theta)} \epsilon^{-j2\pi S\cos\theta}, \tag{126}$$

$$I_{E'\theta} = -\frac{S\lambda' F'}{2K} \frac{h}{S\lambda'} \frac{1}{\epsilon^{j\delta}\tan T}(1 - \epsilon^{-[\alpha S\lambda' + j2\pi S(m - \cos\theta)]})\epsilon^{-j2\pi S\cos\theta}. \tag{127}$$

In (125), (126), and (127), the symbols have the following meanings

Symbol	Definition	Unit
I_θ	The total current produced at the receiving end of the antenna s by an impressed field propagated at an angle θ from the axis of the antenna.	amperes
$I_{F'\theta}$	The portion of I_θ produced by the horizontal component of the impressed field.	amperes
$I_{E'\theta}$	The portion of I_θ produced by the vertical component of the impressed field.	amperes
F'	The horizontal component of the impressed field. (Positive direction in the direction of propagation along the ground.)	volts per kilometer
E'	The vertical component of the impressed field. (Positive direction downward.)	volts per kilometer
δ	Phase angle between the horizontal and vertical components of the impressed electric field.	radians
T	"Quasi-tilt angle" of the impressed electric field.	radians
K	The characteristic impedance of the wave-antenna.	ohms
γ	The propagation constant of the wave-antenna.	
α	The real part of the propagation constant of the wave-antenna or the attenuation constant.	napiers per kilometer
β	The imaginary part of the propagation constant of the wave-antenna or the phase constant.	radians per kilometer
γ'	The propagation constant of the space waves.	
α'	The real part of the propagation constant of the space waves (assumed equal to zero).	napiers per kilometer
β'	The imaginary part of the propagation constant of the space waves.	radians per kilometer
s	The length of the wave-antenna.	kilometers
h	The height of the wave-antenna above ground.	kilometers
$S = s/\lambda'$	The length of the wave-antenna.	space wave-lengths
$\lambda' = 2\pi/\beta'$	The wave-length of the space waves.	kilometers
$V = 2\pi f/\beta$	Apparent velocity of propagation of waves along the wave-antenna.	kilometers per second
V'	The velocity of propagation of the space waves ($= 3 \times 10^5$ km per second).	kilometers per second

V/V'	Velocity ratio.	numeric
$m \equiv V'/V = \beta/\beta'$	Reciprocal of the velocity ratio.	numeric
$j = \sqrt{-1}$		
θ	The angle between the axis of the wave-antenna and the direction of propagation of space waves measured in a clockwise direction.	
R.D.R. $= \dfrac{I_\theta}{I_0}$	Relative directional receptivity.	numeric

APPENDIX 2

Antenna Arrays

The directional discrimination yielded by a single antenna can be increased by utilizing several such antennas in an array.[17] In Fig. 18,

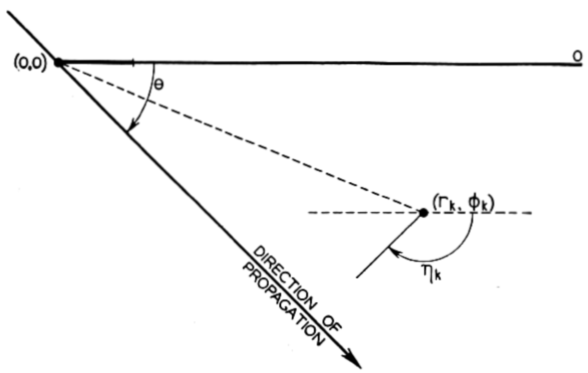


Fig. 18

a general array of n antennas is indicated, of which only the first and the k'th are portrayed.

Each antenna in the array is completely specified by the coordinates of the initial end of the antenna, the angle between the zero axis of the coordinate system and the axis of the antenna, and the current delivered at the receiving end of the antenna for a given electric field impressed on the antenna at each angle of incidence with the antenna. Literally, the first and the k'th antennas are specified as follows:

	First Antenna	*k'th Antenna*
Coordinates of initial end of antenna...	(0,0)	(r_k, ϕ_k)
Direction of antenna..................	0	η_k
Current delivered by antenna for a constant electric field propagated in the direction θ	$I_{\theta 0}$	$I_{\theta k}$

For the purpose of this discussion, it is sufficiently accurate to assume that the propagation of space waves over the area covered by the array only involves phase retardation, i.e.,

$$\gamma' = j\beta'. \tag{201}$$

The output of the k'th antenna is transmitted through a linear transducer having a transfer constant P_k to a common point where it is combined with the outputs of the other antennas of the array. The current from the k'th antenna at the point of combination is therefore

$$J_{k\theta} = I_{\theta k}\epsilon^{-j\beta'[r_k/V']\cos(\theta-\phi_k)}\epsilon^{-P_k}, \tag{202}$$

where

$$\theta_k = \theta - \eta_k \tag{203}$$

and

$$\beta' = \frac{2\pi V'}{\lambda'}. \tag{204}$$

The total current received from the n antennas of the array is equal to the sum of the currents received from the individual antennas, or

$$J_\theta = \sum_{k=1}^{k=n} I_{\theta k}\epsilon^{-j[2\pi r_k/\lambda']\cos(\theta-\phi_k)}\epsilon^{-P_k}. \tag{205}$$

Equation (205) gives the total current received from *any* array of antennas for any direction of wave propagation in a horizontal plane. This general expression is not adapted to ready determination of directional characteristics of antenna systems, but it may be simplified by placing the following restrictions on the individual antennas forming the array and their space relations in the array:

(1) The antennas are all alike. This restriction may be defined by the expression:

$$I_{\theta k} = I_{\theta(k+1)}.$$

(2) The axes of the antennas are parallel, as defined by the expression

$$\eta_k = 0 \text{ or } \pi.$$

(3) The initial ends of the antennas are equally spaced along straight lines in each subgroup and the subgroups are equally spaced along straight lines. All of the subgroups are identical.

The general antenna array conforming to these restrictions is shown in Fig. 19. In this figure, there are q groups of antennas equally spaced by the distance a along a line 90 deg. from the zero axis. In each of these q groups of antennas, there are p antennas, divided into two series, those for which $\eta = 0$ being numbered 1, 3, $\cdots$, $(2l-1)$, $\cdots$ $(p-1)$ and those for which $\eta = \pi$ being numbered 2, 4, $\cdots$, $2l$, $\cdots$, p, the initial ends of the second series being removed by a distance s from the initial ends of the first series, along the axes of the antennas of the first series.

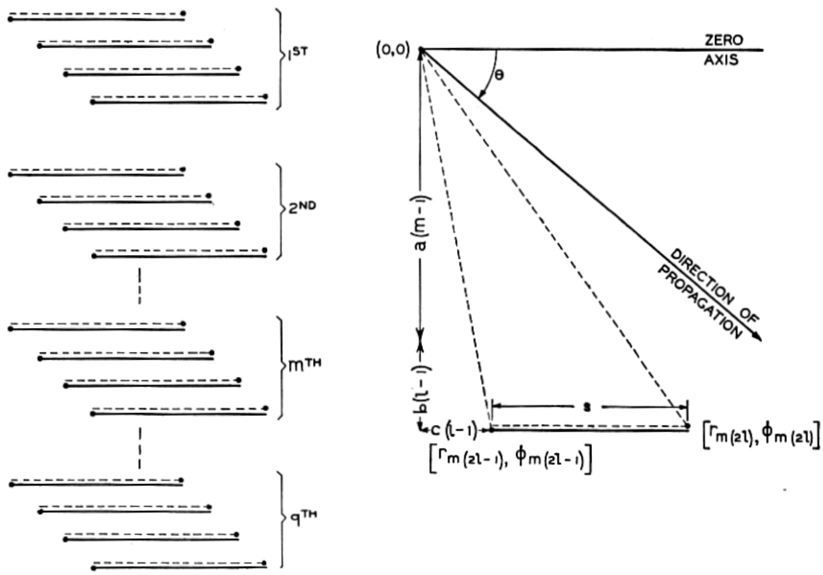


Fig. 19

Equation (205) applied to this general array gives for the total current received from the array

$$J_\theta = I_\theta \sum_{m=1}^{m=q} \sum_{2l=2}^{2l=p} \epsilon^{-j\frac{2\pi r_{m(2l-1)}}{\lambda'}\cos[\theta-\phi_{m(2l-1)}]}\epsilon^{-P_{m(2l-1)}} + I_{\theta-\pi} \sum_{m=1}^{m=q} \sum_{2l=2}^{2l=p} \epsilon^{-j\frac{2\pi r_{m(2l)}}{\lambda'}\cos[\theta-\phi_{m(2l)}]}\epsilon^{-P_{m(2l)}}, \quad (206)$$

where

$$r_{m(2l-1)} = \sqrt{[a(m-1)+b(l-1)]^2 + [c(l-1)]^2}, \quad (207)$$

$$r_{m(2l)} = \sqrt{[a(m-1)+b(l-1)]^2 + [c(l-1)+s]^2}, \quad (208)$$

$$\phi_{m(2l-1)} = \tan^{-1}\left[\frac{a(m-1)+b(l-1)}{c(l-1)}\right], \quad (209)$$

$$\phi_{m(2l)} = \tan^{-1}\left[\frac{a(m-1)+b(l-1)}{c(l-1)+s}\right]. \quad (210)$$

In a double summation, the result is independent of the order in which the summations are taken. If then we write

$$u_\theta = \sum_{2l=2(m=1)}^{2l=p} \epsilon^{-j\frac{2\pi r_{m(2l-1)}}{\lambda'}\cos[\theta-\phi_{m(2l-1)}]}\epsilon^{-P_{m(2l-1)}}, \tag{211}$$

$$v_\theta = \sum_{m=1(2l=2)}^{m=q} \epsilon^{-j\frac{2\pi r_{m(2l-1)}}{\lambda'}\cos[\theta-\phi_{m(2l-1)}]}\epsilon^{-P_{m(2l-1)}}, \tag{212}$$

$$w_\theta = \sum_{2l=2(m=1)}^{2l=p} \epsilon^{-j\frac{2\pi r_{m(2l)}}{\lambda'}\cos[\theta-\phi_{m(2l)}]}\epsilon^{-P_{m(2l)}}, \tag{213}$$

$$y_\theta = \sum_{m=1(2l=2)}^{m=q} \epsilon^{-j\frac{2\pi r_{m(2l)}}{\lambda'}\cos[\theta-\phi_{m(2l)}]}\epsilon^{-P_{m(2l)}}. \tag{214}$$

The expression for the total current may be written

$$J_\theta = I_\theta u_\theta v_\theta + I_{\theta-\pi} w_\theta y_\theta. \tag{215}$$

If the transducers in the circuits from each antenna of a pair are so related that

$$P_{m(2l)} - P_{m(2l-1)} = P_c, \tag{216}$$

the expression for the total current becomes

$$J_\theta = u_\theta v_\theta [I_\theta + I_{\theta-\pi}\epsilon^{-j[2\pi s/\lambda']\cos\theta}\epsilon^{-P_c}]\epsilon^{-P_1}. \tag{217}$$

The directional diagram in terms of relative directional receptivity is

$$RDR = \frac{J_\theta}{J_0} = \frac{u_\theta}{u_0} \times \frac{v_\theta}{v_0} \times \left[\frac{I_\theta + I_{\theta-\pi}\epsilon^{-j[2\pi s/\lambda']\cos\theta}\epsilon^{-P_c}}{I_0 + I_{-\pi}\epsilon^{-j[2\pi s/\lambda']}\epsilon^{-P_c}}\right]. \tag{218}$$

Since there has been no assumption to this point of the character of I_θ, the significance of the coefficients u_θ and v_θ may be determined by assuming

(1) $I_\theta = I_0$, which is the directional characteristic of a vertical antenna

(2) $s = 0$

(3) $P_c = \infty$.

Consideration of (218) in light of (211) and (212) under these conditions leads to the conclusion that

$$\frac{u_\theta}{u_0} \quad \text{and} \quad \frac{v_\theta}{v_0}$$

are the relative directional receptivities of two arrays of vertical antennas placed at the initial ends of the antennas comprising the desired array. If, then, we designate the relationship between antennas indicated by the expression

$$J_c = [I_\theta + I_{\theta-\pi}\epsilon^{-j[2\pi s/\lambda']\cos\theta}\epsilon^{-P_c}] \tag{219}$$

as *compensation* [14] and recognize that this expression gives the directional characteristic of a compensated antenna, we may formulate the rule that the directional characteristic of an array of similar parallel unit antennas is equal to the product of the directional characteristic of the unit antenna and the directional characteristic of an array of unit vertical antennas placed at the initial ends of the unit antennas forming the array, the product being taken point for point as the angle of incidence increases. The relative directional receptivity of each fundamental array of vertical antennas is termed the array factor, so that similarly, the relative directional receptivity of an array of similar parallel unit antennas is given by the product of the relative directional receptivity of the unit antenna and the array factor. This method may be extended to the solution of a complicated array such as that shown in Fig. 19, by determining the relative directional receptivity for groups of unit antennas, then determining the array factor for these groups taken as unit antennas. Expressed literally for a complex array of this type:

$$RDR_{\text{array}} = [A_1 \times A_2 \times \cdots \times A_n]RDR_{\text{unit antenna}}, \tag{220}$$

where $A_1 \cdots, A_n$ are the array factors for the fundamental groups into which the complete array may be divided.

APPENDIX 3

Wave Tilt and Ground Conductivity

In Zenneck's [32, 33] exposition of the relation between the horizontal and vertical components of a plane electric wave propagated along a horizontal surface between two media, it is demonstrated that these two constituents of the wave in the upper medium (1) are related by the expression

$$-\frac{F'}{E'} = \epsilon^{j\delta} \tan T = \sqrt{\frac{\dfrac{9 \times 10^{11}}{\rho_1} + j\dfrac{1}{4\pi}\omega k_1}{\dfrac{9 \times 10^{11}}{\rho_2} + j\dfrac{1}{4\pi}\omega k_2}}, \tag{301}$$

where

Symbol	Definition	Unit
F'	The horizontal component of the electric wave in medium 1. (Positive direction in the direction of propagation along the interface.)	volts per kilometer
E'	The vertical component of the electric wave in medium 1. (Positive direction downward.)	volts per kilometer
ρ_1	Specific resistivity of medium 1.	ohms per centimeter cube
ρ_2	Specific resistivity of medium 2.	ohms per centimeter cube
k_1	Dielectric constant of medium 1 and equal to unity for a vacuum.	numeric
k_2	Dielectric constant of medium 2.	numeric
f	Frequency	cycles per second
ω	$2\pi f$	

Our primary interest is in the case where the first medium is air, and the second medium is the earth beneath an antenna system. In this case the constants of the media may be given the values:

$$
\begin{aligned}
\rho_1 &= \infty && \text{(air)},\\
\rho_2 &= \rho && \text{(earth)},\\
k_1 &= 1 && \text{(air)},\\
k_2 &= k && \text{(earth)}.
\end{aligned}
$$

Substituting these values into the general equation (301)

$$
-\frac{F'}{E'} = \epsilon^{j\delta} \tan T = \frac{1}{\sqrt{k}} \left[\frac{\left(\dfrac{fk\rho}{18 \times 10^{11}} \right)^2}{1 + \left(\dfrac{fk\rho}{18 \times 10^{11}} \right)^2} \right]^{\frac{1}{4}} \epsilon^{j[1/2 \tan^{-1}(18 \times 10^{11}/fk\rho)]}. \quad (302)
$$

At this point it is desirable to indicate the significance of the term "quasi-tilt angle" as applied to T. It is seen that ($\tan T$) is the absolute magnitude of the ratio of the horizontal and vertical components of the electric field. In the case that the time phase between the two components of the field is zero (i.e., $\delta = 0$), T would represent the angle of forward inclination of the propagated wave front. In general, δ is unequal to zero and hence the angle of inclination of the major axis of the ellipse traced by the electric vector is less than T, but it still remains convenient to express the ratio of the magnitudes of the two components of the field as the tangent of an angle. This angle cannot be called the wave tilt, however, but the term "quasi-tilt angle" may safely be applied to it.

The ground constants may be determined from measurement of the "quasi-tilt angle" as the following development shows:

Equation (302) may be written as two equations

$$\tan T = \frac{1}{\sqrt{k}}\left[\frac{\left(\dfrac{fk\rho}{18 \times 10^{11}}\right)^2}{1 + \left(\dfrac{fk\rho}{18 \times 10^{11}}\right)^2}\right]^{\frac{1}{4}}, \tag{303}$$

$$\delta = \frac{1}{2}\tan^{-1}\left(\frac{18 \times 10^{11}}{fk\rho}\right). \tag{304}$$

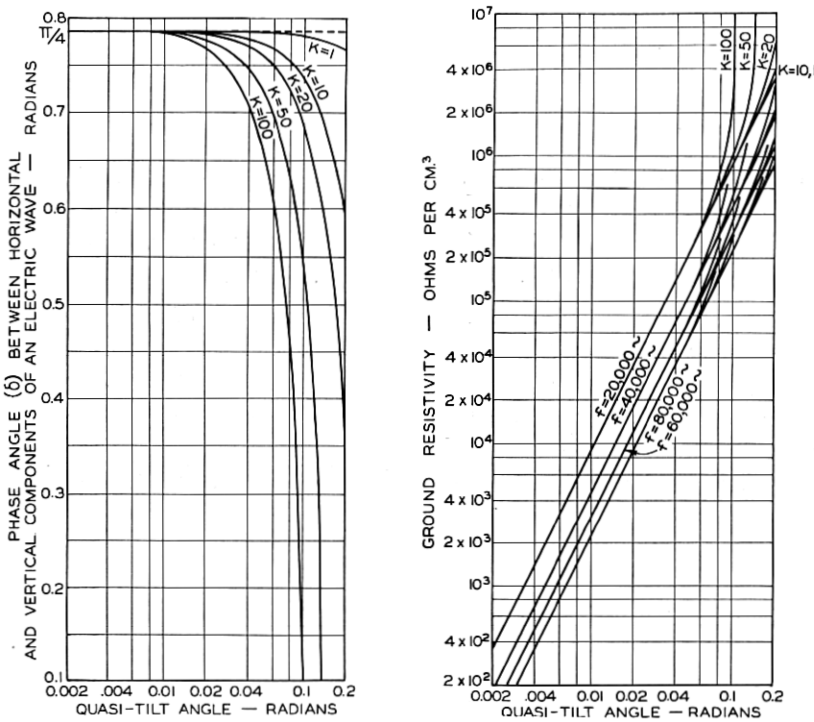


Fig. 20—Relation between quasi-tilt angle, ground resistivity, and phase angle between horizontal and vertical components of an electric wave. (By Zenneck's formula.)

Solving equations (303) and (304) as simultaneous equations for δ in terms of k and T and for ρ in terms of f, k, and T yields

$$\delta = \frac{1}{2}\cos^{-1}(k \tan^2 T), \tag{305}$$

$$\rho = \frac{18 \times 10^{11}}{f}\,\frac{\tan^2 T}{\sqrt{1 - k^2 \tan^4 T}}. \tag{306}$$

These two expressions have been evaluated for the extreme range of values of k that would be met in practice (k between 1 and 100) and for values of T between 0.002 and 0.2 radian and are plotted in Fig. 20. The figures for dielectric constant given by Fleming [34] show that for earth, the maximum value of k to be expected is below 20. It is evident, therefore, that δ is negligibly different from $\pi/4$ for values of T below 0.05 radian in the vicinity of an antenna which is constructed over land. Also Fig. 20 shows that the specific resistivity is practically independent of k for the same range of T. Fortunately, the measured values of T lie within these limits, so that the time phase difference between the horizontal and vertical components of an electric wave, and the ground resistivity may be evaluated with but slight error from measurements of the quasi-tilt angle.

APPENDIX 4

Probability of Voltages Greater than any Specified Value Resulting from the Simultaneous Reception of Several Radio-Telegraph Stations in a Restricted Frequency Range

In order to determine the required load capacity of vacuum tubes for a radio receiver, it is necessary to obtain some estimate of the voltages from interfering signals which may occur at the input of the radio receiver and during how much of the time certain specified voltages are exceeded.

If we assume that there are N telegraph stations within a restricted frequency range, that each station contributes equal unit voltage at the receiver, and that the probability of the key being closed at any one station is constant, then the probability that exactly n stations have their keys depressed at the same time is

$$P_n = \frac{N!}{n!(N-n)!} K^n (1-K)^{(N-n)}, \tag{401}$$

where K is the fraction of the total time that each station has its key depressed.

In order to determine the probability that n stations will produce a voltage equal to or greater than any specified value x we have followed Rayleigh's problem of random phases as explained in Volume 6 of his "Scientific Papers," page 618. While the conditions are not all satisfied it can be shown that they are approximately satisfied for

[34] J. A. Fleming, "Principles of Electric Wave Telegraphy and Telephony," Longmans, Green and Co., 1916. 3d edition, p. 800.

the great majority of possible combinations and for small time intervals. The formula of Rayleigh gives the probability that the resultant of n vectors lies within an arbitrary interval $(r - dr/2,\ r + dr/2)$. Since we will assume sinusoidal voltages in the actual problem under consideration we require the probability that the projection of the resultant on the real axis is greater than a given value of x. This can be calculated by changing the polar coordinates of Rayleigh's formula to rectangular coordinates and integrating with respect to y from $-\infty$ to $+\infty$ and then with respect to x from x to $+\infty$.
The integrated formula then becomes
Probability of a voltage greater than $x = P_x$

$$P_x = A_1\Gamma\left(\frac{1}{2}, \frac{x^2}{n}\right) + A_2\Gamma\left(\frac{3}{2}, \frac{x^2}{n}\right) + A_3\Gamma\left(\frac{5}{2}, \frac{x^2}{n}\right) + A_4\Gamma\left(\frac{7}{2}, \frac{x^2}{n}\right) + A_5\Gamma\left(\frac{9}{2}, \frac{x^2}{n}\right), \tag{402}$$

where

$$A_1 = \frac{1}{2\sqrt{\pi}}\left(1 - \frac{3}{16n} - \frac{5}{24n^2} + \frac{105}{16.32n^2}\right),$$

$$A_2 = \frac{1}{2n\sqrt{\pi}}\left(\frac{3}{4} - \frac{25}{64n}\right),$$

$$A_3 = \frac{1}{2n\sqrt{\pi}}\left(-\frac{1}{4} + \frac{155}{192n}\right),$$

$$A_4 = \frac{1}{2n\sqrt{\pi}}\left(-\frac{47}{144n}\right),$$

$$A_5 = \frac{1}{2n\sqrt{\pi}}\left(\frac{1}{32n}\right)$$

and

$$\Gamma(p, u\sqrt{p}) = \Gamma(p)[1 - I(u, p - 1)],$$

in which

$$p = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \frac{7}{2}, \frac{9}{2} \text{ and } u\sqrt{p} = \frac{x^2}{n}.$$

Having found u, the I functions of $(u, p - 1)$ can be obtained from Pearson's "Tables of the Incomplete Γ-Functions." $\Gamma(p)$ for the values of p given above is found to be

$$\sqrt{\pi},\ \frac{1}{2}\sqrt{\pi},\ \frac{3}{2}\cdot\frac{1}{2}\cdot\sqrt{\pi},\ \frac{5}{2}\cdot\frac{3}{2}\cdot\frac{1}{2}\cdot\sqrt{\pi},\ \frac{7}{2}\cdot\frac{5}{2}\cdot\frac{3}{2}\cdot\frac{1}{2}\cdot\sqrt{\pi}.$$

The probability of exactly n stations being on at the same time multiplied by the probability that exactly n stations will give a voltage equal to or greater than x equals the probability of obtaining a voltage equal to or greater than x from just n stations.

Hence the summation from $n = 1$ to $n = N - 1$ of these probabilities for a given value of x will give the probability of obtaining a

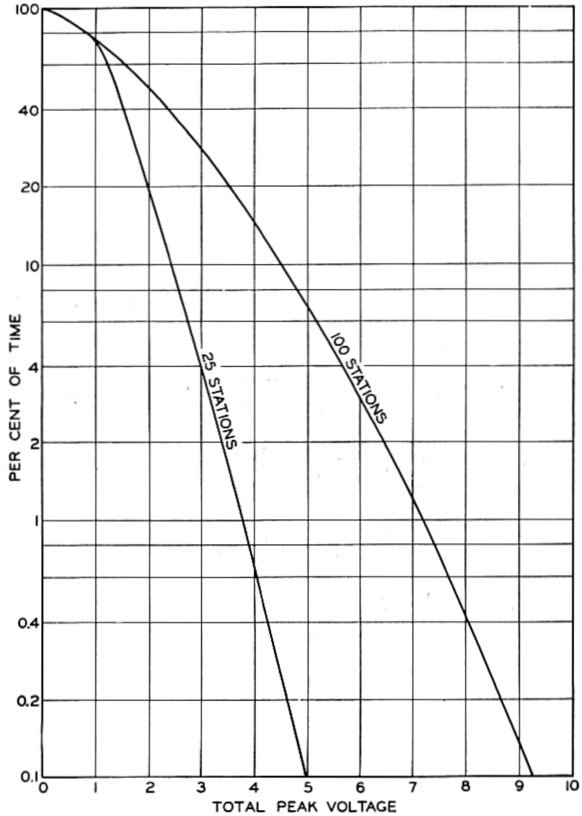


Fig. 21—Voltages resulting from several unit voltages each applied 15 per cent of the time and with random phase and frequency

voltage equal to or greater than x from all of the N stations in the restricted frequency range considered or

$$P_{xN} = \sum_{n=1}^{n=N-1} P_n P_x. \tag{403}$$

In a vacuum tube large negative voltages are equally as important as large positive voltages in producing distortion. Equation (403) has been derived for positive values greater than x but negative values greater than $-x$ are equally probable and therefore the fraction of the time that the absolute value of voltage is equal to or greater than x is $2P_{xN}$ or

$$P_{|x|N} = 2P_{xN}. \tag{404}$$

Specific cases which approximate the existing conditions of long-wave transatlantic reception have been calculated from equation (404) and are shown in Fig. 21. These curves are based on the following assumptions:

1. That the number of stations lying in the restricted frequency range is $N = 100$ and $N = 25$.
2. That each station contributes unit peak voltage to the input of the radio receiver.
3. That each station has its key depressed 15 per cent of the total time during any day. $K = 0.15$.
4. That transmissions from all stations are random.

Considerable valuable assistance in the preparation of this appendix has been obtained from Dr. F. H. Murray of this department and the authors wish to express their appreciation of this aid.