ELECTRONIC SURVEYING : ACCURACY OF ELECTRONIC POSITIONING SYSTEMS (*)

by Henry W. BIGELOW
Deputy Head, Geodetic Branch
U. S. Naval Oceanographic Office

SYNOPSIS

There are many different electronic systems that can be used for positioning or navigation, on land, at sea, or in the air. Each system differs from the others because each was designed to meet a specific need. There is no one system that meets all needs. The systems may be classified by the kind of line of position generated : hyperbolic, concentric circles, radials, or composite.

The accuracy of position obtained by any one of these systems is a function of two independent terms : repeatability and predictability. Repeatability is the measure of the reliability with which the system permits the user to return to a given spot on the surface of the earth using the electronic lines of position generated by the system. This is what most people mean by " accuracy ". It includes the random errors of the system and the effect of the spacing of the lines of position and the angle at which they intersect. The errors of repeatability depend on known quantities. A rapid and reasonably accurate analysis can be made of any proposed installation. The methods of analysis for the different kinds of systems are developed.

The predictability of an electronic positioning system is the measure of the reliability with which the system can define the location of a given point in terms of geographic rather than electronic coordinates. If there is no distortion of the electronic lattice as it is actually propagated, then the position, corrected for errors of repeatability, is the true location. There is always some distortion. The factors causing the distortion are : the conductivity of the surface over which the signal propagates and the refractivity of the atmosphere through which it propagates. These factors are considered

(*) Article reprinted from the *Journal of the Surveying and Mapping Division, Proceedings of the American Society of Civil Engineers*, October 1963, pages 37 to 76.

in detail, but evaluation is difficult. At present (1963), they can only be derived empirically. A formula for the error of predictability is derived, but it is of little practical use in the present state of the art.

Additional study is needed to determine which factors govern the variations in propagation over different kinds of surfaces and the boundaries between them. Knowledge is needed concerning the manner in which the variations in the meteorological conditions of the atmosphere affect the propagation velocity. Techniques must be developed to determine the variation with time and the variation in space of the elements that are found to affect wave propagation. Techniques also are required for the adequate calibration of systems, particularly those with long range.

INTRODUCTION

The basic navigation instrument for surface navigation has been and still is the time-honored hand-held sextant. With it, the navigator fixes his position from the stars. A competent navigator can usually fix his position within a few miles, weather permitting. In the past few years, a number of electronic systems have been developed to assist the navigator. These systems could aid navigation or be used for precise positioning.

An electronic aid to navigation permits the navigator to fix his position more frequently and without regard to cloud cover and other weather conditions. However, it does not fix his position with an accuracy greater than he would have obtained from good sextant fixes. A precise positioning system will give a position with a few ship lengths anywhere within its service area.

The distinction between navigation and positioning is important and should be kept in mind. A positioning system can be used for normal navigation, but a navigation system cannot be used for precision work.

Most electronic positioning depends on an accurate measurement of the time required for a radio signal to travel from the transmitter to the receiver. The signal may be thought of as a wave front which is propagated in all directions. That part remaining beneath the ionosphere is called the groundwave, and it may : (1) take a direct path from the transmitter to the receiver; (2) be refracted in the troposphere; or (3) follow along the surface. For purposes of precise positioning, only the groundwave can be used at present (1963).

That part making use of ionospheric reflection or refraction to provide a path between transmitter and receiver is called the skywave. Because the ionosphere is not a stable layer, corrections for skywave paths, although they can be computed, are approximations only. Thus, skywaves are useful only for navigation and communication. When ionospheric variations can be predicted with accuracy and reliable corrections computed, skywaves will be as valuable to precise positioning as they now are to navigation. For

TABLE 1

Electronic Positioning Systems

System (1)	Range (2) (**)	Accuracy (3)	Frequency (4)	Signal Type (5)	Users (6)	Shore Sites (7)	Power (8)	Notes (9)
Hyperbolic Systems								
Loran-A	800 nm	1/4-5 nm	2 Mc/s	Pulsed	Multi	3	1 kw	Navigation only
Loran-B	25 nm	45-300 ft	2 Mc/s	Pulsed	Multi	3	1 kw	Experimental
Loran-C	1 200 nm	50-1 200 ft	100 kc/s	Pulsed	Multi	3	80 kw	Phase comparison
Lorac-A	200 nm	15-400 ft	2 Mc/s	CW	Multi	3	500 watts	
Lorac-B	300 nm	15-400 ft	2 Mc/s	CW	Multi	4	500 watts	
Decca Nav.	250 nm	1/4-2 nm	100 kc/s	CW	Multi	3	600 watts	Navigation only
Decca Survey	200 nm	25-300 ft	100 kc/s	CW	Multi	3	600 watts	
Omega	4 500 nm	1-10 nm	10 kc/s	Pulsed	Multi	3	100 kw	Experimental
Hi-Fix	40 nm	25-150 ft	2 Mc/s	CW	Multi	3	10 watts	Portable stations
Ranging Systems								
Shoran	30 nm	30-50 ft	300 Mc/s	Pulsed	4-6	2	12 kw	Line of sight
EPI	400 nm	195 ft-1/4 nm		Pulsed	1 (*)	2	12 kw	
Lambda	425 nm	25-250 ft	100 kc/s	CW	1 (*)	2	600 watts	2 Range Decca with Lane Identification
DM Raydist	200 nm	12-100 ft	1.6 & 3.2 Mc/s	CW	2	2	100 watts also 10 watt set	
Hydrodist	25 nm	12-100 ft	3 000 Mc/s	Pulsed	1	2		Line of sight
Hi-Fix	30 nm	12-100 ft	2 Mc/s	CW	1 (*)	2	10 watts	Portable stations
Azimuthal Systems								
Consol	700 nm	6-24 nm	250-400 kc/s	CW	Multi	2	3 kw	Navigation only
MPFS	30 nm		9 375 Mc/s	CW	Multi	3	7 kw	Experimental
Composite Systems								
Raydist	75 nm	15-150 ft	1.6&3.22 Mc/s	CW	2	2	10 & 100	Range-Hyperbolic

(*) Two ships, by time-sharing.
(**) One nautical mile equals 6076.115 ft or 1852.000 m.

the time being, they are more of a hindrance than a help in long-range precise positioning.

There are a number of different electronic systems that can be used for positioning or navigation. Each system differs from the others because each was designed to meet a specific need. Consequently, there is no one system that answers all needs.

Electronic positioning systems may be classified into four general groups : hyperbolic systems, ranging systems, azimuthal systems, and composite systems. Each has common characteristics. At the same time, those of one group differ significantly from those of another group. The distinguishing characteristic is the kind of lattice (network of lines of position) generated by the transmitters.

In a hyperbolic system, the lines of position are hyperbolae. In a ranging system, they are concentric circles. In an azimuthal system, they are radials. In a composite system, two different kinds of lines of position are generated.

Table 1 presents some of the systems now in use (1963) and lists some of their more important characteristics. The range given in the table is for the groundwave under normal conditions. Two accuracy figures are given. The first is for the best part of the service area under excellent conditions. The second is for the range quoted in column 2 under normal conditions.

ACCURACY

Accuracy of position, as determined by an electronic system, is a function of two independent terms called repeatability and predictability. Repeatability is the measure of the ability of the system to permit the user to return to a specific point on the surface of the earth. The specific point is defined only in terms of the electronic lattice peculiar to that system. Predictability is the measure of the ability of the system to define the location of that point. The definition is in terms of geographic rather than electronic coördinates.

Lattices are usually plotted on a chart or put into a computer program. The intersection of two observed electronic lines of position is then plotted on the chart or fed into the computer program. However, it is a fallacy to think that the intersection of the observed lines of position is actually at the plotted or computed geographic position. The positions would agree only if the propagated lattice was not distorted from the computed lattice or if all lattice distortions were predicted and the necessary corrections made.

What is herein called repeatability is what most people mean when they talk about the " accuracy " of a system. The term repeatability includes the random errors found in any system : instrumental errors, operator errors, ephemeral propagation anomalies, etc. Because they are strictly random, careful calibration can establish their limits, and they may be allowed for. Repeatability also includes the effect of the net geometry. Note that the " accuracy " figures of table 1 are, in fact, the repeatability of the given system.

NET GEOMETRY

A specific installation of an electronic positioning system may be called either a net, a rate, a chain, or a triad. Every net will have its own particular geometry. This geometry depends on the orientation and length of the base-lines connecting the master and the slave stations of the net. Net geometry

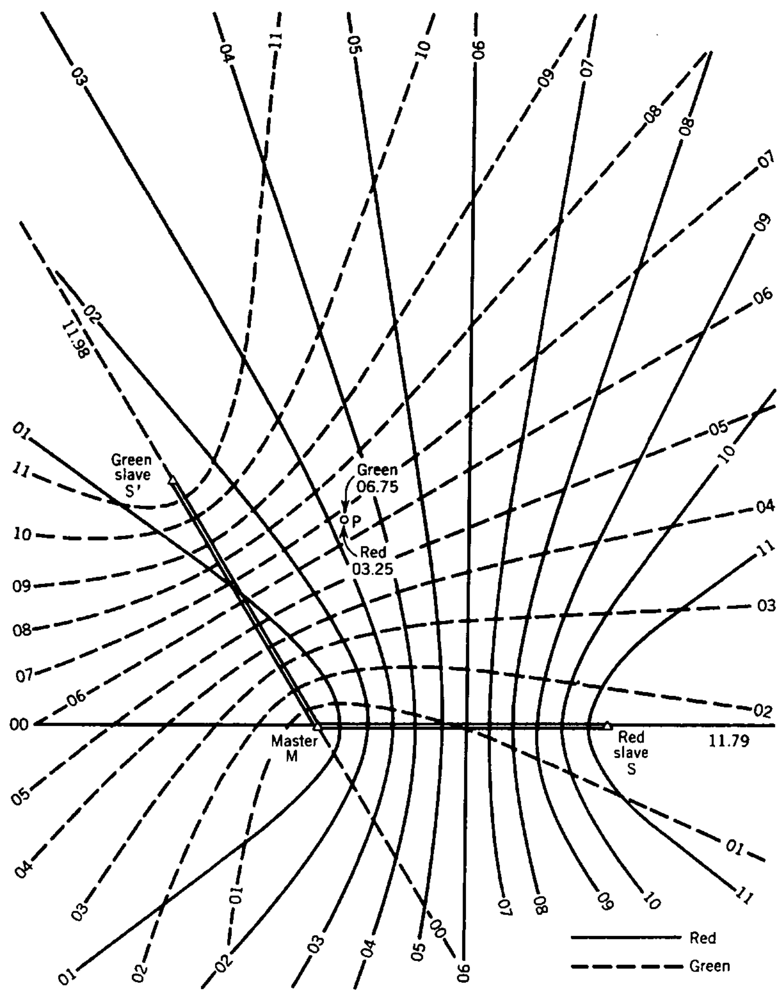

FIG. 1. — Hyperbolic triad.

may be considered in two respects, lane width and the angle of intersection of the lines of position.

The lane width is that distance represented by a unit of the electronic lattice. In a phase comparison (or CW) hyperbolic system (see table 1), a lane is one-half a wave-length of the frequency used. This concept has been extended to other systems. On the baseline of a hyperbolic system one lane has a definite value. For a 2 Mc per sec phase comparison (CW) system, the lane width on the baseline is approximately 75 metres (m). For a 300 kc per sec system, it is approximately 500 m. The " lane width " of a pulsed system is one microsecond, equivalent to approximately 150 m. The " lane width " of a pulse system is not frequency dependent.

If it is assumed that the hyperbolic triad in fig. 1, showing the lines of position generated by three stations, is using a frequency of 300 kc per sec, then the spacing of the lines along the red baseline between M and S will be 500 m. As the user departs from the baseline, a lane becomes wider because of the divergence of the hyperbolae. At point P, one red lane, between line of position Red 03 and Red 04, is still only one-half a wave-length, but the distance it represents is nearly twice that on the baseline.

This expansion factor is an important consideration in the geometry of hyperbolic systems. In ranging systems, similar to the one in fig. 2, which shows the lines of position generated by two shore stations, the lane

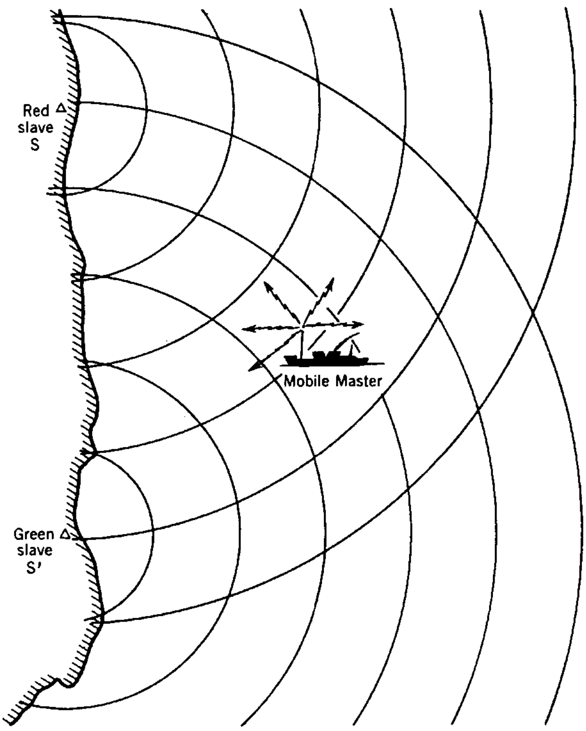

Fig. 2. — Ranging pair.

width is constant throughout the net. In azimuthal systems, the angular resolution of the receiver takes the place of the lane width.

Fix Strength

The strength of a fix is measured by the angle of intersection of the lines of position. Those lines of position intersecting with the smaller angle between 60° and 90° give strong fixes. As a rule of thumb, a net should not be used in an area where the smaller angle of intersection is less than 15°. Where the angle of intersection is between 15° and 30° the fix is considered weak, between 30° and 60° it is considered good.

Usability

The area over which a given net can be used is determined by : (1) the geometry of the net; (2) the radiated power; (3) the signal-to-noise ratio; and (4) the characteristics of the transmission path. As a general rule, the signals can be heard at distances well beyond the range at which they can be used for accurate positioning.

HYPERBOLIC SYSTEMS

A hyperbola is defined as the locus of a point moving so that the difference between the distances measured from two fixed points to the moving point remains a constant. A hyperbolic positioning system is designed so that the mobile user measures the difference in transmission time between signals from two shore stations. This is only one line of position. To obtain a fix, two lines of position must be generated. In fig. 1, point P is defined as the intersection of the two hyperbolic lines of position. Position line Red 03.25 is generated by the master station (M) and the red slave station (S), and position line Green 06.75 is generated by M and the green slave station (S′).

There are three types of hyperbolic systems : the pulsed, the CW (phase comparison), and the combined systems. In the pulsed (time difference) system, the master station (M) transmits a coded series of pulses at short intervals. The receipt of this master pulse at the slave stations (S and S′) triggers their respective transmissions. Fixed time delays are established so that the mobile user always receives the master pulse first and can identify the signals arriving from S and S′ respectively. The receiver measures the difference in time of receipt of M and S, and M and S′; removes the fixed time delays; and displays the resulting time differences, which define two hyperbolic lines of position. These time differences are expressed in microseconds. Loran-A (LOng RAnge Navigation) is such a system.

In the CW (phase comparison) systems, the operation is more complex. These systems emit a continuous wave signal and use different frequencies to distinguish between M minus S, and M minus S′. While the actual operation is quite complex, it follows much the same sequence as the time-difference system. The distinction between them is that the phases of the signals transmitted from M and S are compared, rather than the travel times. The phase comparison gives a precise measure of the fractional part of a wave-length, or lane, but no indication of the total number of whole lanes existing. Some auxiliary means of keeping track of this number of whole lanes is therefore essential. Decca and Lorac (LOng Range ACcuracy) are typical phase comparison systems.

The combined system uses a combination of time-difference and phase comparison techniques. An approximate position is obtained by the difference in arrival time of the pulsed signal, and this is refined by a comparison of the phase of the signal within the pulse. Lorac-C is a combined system. There is no limit to the number of users of any of these hyperbolic systems. All the user needs is the proper kind of receiver and the correct charts, tables, or programmed computer.

RANGING SYSTEMS

A ranging system is comprised of a mobile master and two shore stations (fig. 2). The master interrogates the shore stations; these stations respond; and the round trip time is measured at the master. There are two basic types, those using a pulsed signal and those using a continuous wave signal.

Shoran (SHOrt RAnge Navigation) is a pulsed system consisting of one or more mobile interrogators and two fixed responder beacons. The system is line of sight, because it uses frequencies in the vicinity of 230 Mc. The output reading is given in statute miles. With average station elevation, the normal range of surface transmission is 30 miles under ideal conditions; if the slave stations have sufficient elevation, the range may be extended to 40 miles.

The EPI system (Electronic Position Indicator) was developed to provide a positioning system with the repeatability of Shoran and with greater range. The frequency was lowered to the 2-Mc band and, under ideal conditions, the system can be used 400 miles from the coast. EPI is essentially a single-user system, but with time-sharing, two interrogators can share the same pair of responder beacons. Although EPI is excellent for small-scale development and exploratory work, it cannot be used satisfactorily for large-scale development. A scale of one mile to the inch is the practical limit.

Lambda and Two Range Decca (2RD) are continuous wave systems, using the same frequencies and developed directly from the standard hyperbolic Decca Survey System. The master station is on the mobile user, and the two shore stations are the two slaves. Slave stations spaced 400 miles apart have been used with adequate repeatability at similar distances offshore. However, at these distances, the signal strength is badly degraded. These ranging systems are generally single-user systems. Lambda and 2RD require careful calibration because ship-board installation places the master receiver within the induction field of the master transmitter.

Hydrodist is an adaptation of the Tellurometer used ashore for precise distance measurements. The Hydrodist is a pulsed system using the 3 000 Mc per sec band. It consists of two shipborne master units. Each master unit has its own responder beacon located at a known point ashore. It is the only ranging or hyperbolic positioning system having more than one master unit. Two operators are needed in the vessel but only a caretaker is required at each shore station.

It should be noted that any precise distance measuring equipment could, in theory, be used in this manner. Conversely, a precise positioning system, such as Hiran (HIgh accuracy shoRAN), can be used for distance measurement. Lines over 500 miles long have been measured over water by Hiran. The precision of these measurements is greater than that of any other method by which the distance can be checked.

AZIMUTHAL SYSTEMS

The azimuthal systems differ from those previously examined in that the mobile user measures direction rather than distance. Radio direction finder and radio compass stations are azimuthal systems for navigation. The only other systems now in use are Consol, Consolan, and MPFS.

Consol, now used mainly in Europe, is a British version of Sonne, a German World War II development. Consolan is a slightly modified American version of Consol. The user needs only a standard broadcast receiver to determine his position. Two rotating patterns, one of dots and one of dashes, permit the mobile user to determine his bearing from the station. Two or more stations are required for a fix. It has only navigational accuracy.

The Canadian Microwave Position Fixing System (MPFS) is a line of sight device consisting of three microwave transmitters at known locations ashore, and a mobile receiver that measures the included angles between the radius vectors to the three shore stations. From these angles the position of the user can be plotted or computed as a solution of the " three-point-problem ".

COMPOSITE SYSTEMS

Most of the composite systems have been developed to meet a particular need. For example, an island platform may not be extensive enough to support a full hyperbolic triad with adequate baselines. However, one hyperbolic line of position can be generated and crossed with a ranging system line of position. The Raydist system can be used in this manner. One shore station generates a circular (ranging) line of position and also acts as master for a hyperbolic line of position. The other station is only the slave for the hyperbolic pair. Several make-shift systems have been used, such as one Lorac pair and a single Shoran or other ranging system station. The principal reason for these systems is an attempt to improve system geometry.

REPEATABILITY

The repeatability of any electronic positioning system defines the limits within which the user may return to, or repeat, a position whose lattice coordinates are known. This repeatability is a function of a number of random errors and the net geometry. The importance and magnitude of the random sources of error vary with the system.

Error of Repeatability

The several sources of error could be determined separately and independent corrections could be applied. However, it is more convenient to compute a single term, d_r, the root-mean-square error of repeatability which includes the effects of both random errors and net geometry. In fig. 3, point P is the apparent location of a fix, somewhere in the service area of any positioning system. The lines of position intersect with an angle B, t_1 and t_2 are the respective displacements of the two lines of position.

The actual position may lie anywhere within a circle with centre at P and a radius of d_r, with a probability of 68 %. The root-mean-square error is usually defined as the square root of the sum of the squares of the major and minor semi-axes of the error ellipse. It has been assumed herein that the major semi-axis of the error ellipse is the radius of the error circle. This assumption, while not entirely correct, simplifies the derivation.

The value of d_t, the position error of one observation, can be derived in terms of t_1, t_2, and B. In fig. 3, t_1 and t_2 are the displacement of lines of position 1 and 2, respectively. B is the angle of intersection of these lines.

$$d_t^2 = \frac{t_1^2}{\sin^2 B} + \frac{t_2^2}{\sin^2 B} + \frac{2\ t_1 t_2 \cos B}{\sin^2 B} \tag{1}$$

By definition

$$d_r^2 = \frac{\Sigma\, d_t^2}{n} \tag{2}$$

in which n is the number of observations. If p_1 and p_2 are the standard deviation, or RMS error, of each line of position, and R is a correlation factor which is one if every radius vector is common to at least two pairs

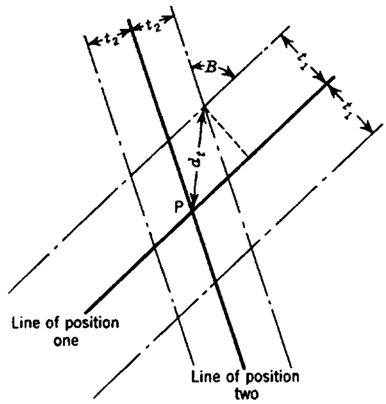

FIG. 3. — Derivation of the error of position.

and zero if none are used in more than one pair, then

$$p^2 = \frac{\Sigma\, t^2}{n} \tag{3}$$

and

$$R = \frac{\Sigma\, t_1 t_2}{n\, p_1 p_2} \tag{4}$$

Substituting in eq. 1

$$d_r = \frac{1}{\sin B} \sqrt{p_1^2 + p_2^2 + 2\, p_1 p_2\, R \cos B} \tag{5}$$

Eq. 5 has been derived assuming a normal distribution of random errors. That is, 68 % of the observations may be expected to be equal to or less than the computed value of d_r and 32 % may be expected to exceed that value. This is known as a one sigma distribution (*). If higher probabilities are desired, the appropriate value of sigma may be found in table 2. This

TABLE 2
Probability of a Given Multiple of Sigma

Probability, in percentage	Sigma
50	0.67
68	1.0
87	1.5
95	2.0
98.8	2.5
99.7	3.0
99.95	3.5
99.994	4.0
99.99932	4.5
99.99994	5.0

(*) *Applied General Statistics*, by F. E. CROXTON and D. J. COWDON, Prentice-Hall, Inc., Englewood Cliffs, N.J., 2nd edition, 1955.

value is used as a multiplier in eq. 5. For example, if a 95 % probability of the repeatability being equal to or less than the computed d_r is desired, the term outside the radical of eq. 5 would read :

$$\frac{2}{\sin B}$$

HYPERBOLIC SYSTEMS

Along the baseline connecting the master and slave stations of a hyperbolic pair, the uncertainty, *p*, can be given a definite, constant value based on the ability of the system to define a position. This value has the form

$$p = k\,E\,s \qquad (6)$$

in which *k* is a factor to convert time to distance, *E* denotes the expansion factor, which is one on the baseline, and *s* represents the standard deviation of the fix reading.

The expansion factor has been derived by SITTERLY (*). He shows that a hyperbola bisects the angle between the radius vectors from P to M and S in fig. 1. The angle of intersection between hyperbolae produced by the two pairs of a triad will be $B = \varphi_1 + \varphi_2$. It is also stated that

$$E = \frac{1}{\sin \varphi} \qquad (7)$$

in which φ is one-half the angle between the radius vectors. Substituting in eq. 6

$$p = \frac{k\,s}{\sin \varphi} \qquad (8)$$

If *s* is in microseconds, and *V*, the propagation velocity, is assumed to be 299 690 km per sec, then

$$k = \frac{V}{2} \times 10^{-6} = 149.845 \text{ m} \qquad (9)$$

TABLE 3

Value for k for Various Distance Units

Distance Units	*s* in microseconds	*s* in lane counts
Metres	194.845	149 845/*f*
Feet	491.617 320	491 617.320/*f*
Nautical Miles (1 852 m)	0.080 910	80.910/*f*
Statute Miles (1 609.3 m)	0.093 109	93.109/*f*

(*) *Demonstration Concerning the Geometry of Loran Lines*, by B. W. SITTERLY, Appendix C in *Loran* by PIERCE, MCKENZIE and WOODWARD, McGraw-Hill Book Co., Inc., New York, N.Y., 1948.

and p, in eq. 8, will be in metres. If p is desired in any other units, the correct value for V in those units must be used. If s is given in lanes, or lane count, the frequency f in kc per sec of the shore station pair is included. Table 3 lists the values of k for various units.

RELIABILITY DIAGRAM

As an approximation, the term d_r can be defined as a function of sin B and sin φ. The following graphical method can be used to determine the effective coverage of a hyperbolic triad.

This method requires the determination of the area within which the expansion factor is equal to or less than a given specific value, and areas within which the angle of intersection is equal to or greater than certain specific values. Significant angles of intersection are 30° for the best working area, and 15° for the outside limits of usable area. The expansion factor should not exceed six.

A suitable geometrical proposition can be stated as follows : the locus of points at which a line subtends a given angle is the arc of a circle constructed by using that line as a chord. The centre is located at a dis-

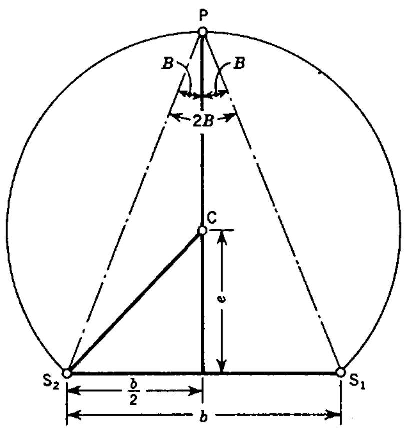

FIG. 4. — Center of arc of constant angle.

tance e along the perpendicular bisector of the chord (fig. 4). It can be shown that hyperbolic lines of position intersect at an angle one-half that subtended by the radius vectors from point P to the two slave stations. Therefore, a circular arc corresponding to a constant angle of 2B can be constructed on the line connecting the slave stations. This circle encloses an area within which the angle of intersection will be equal to or greater than B. In fig. 4, the equation

$$e = \frac{b}{2 \tan 2B} \tag{10}$$

will yield the location of the centre of the required circle along the perpendicular bisector of the line connecting the stations. If angle B is 30°, e equals 0.288 675 (b) in which b is the length of the line S_1S_2. Because the curve of constant φ defines the locus of points of constant expansion factor ($1/\sin \varphi$), the same method can be used.

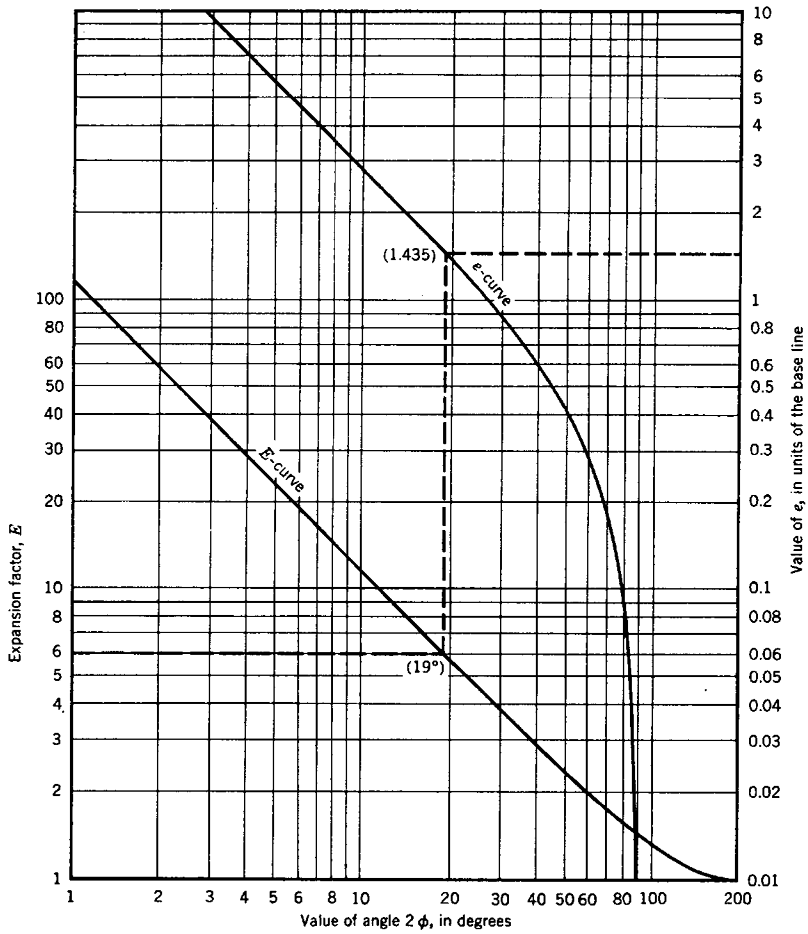

FIG. 5

Fig. 5 has been prepared to yield the value (without computation) of e either for any angle of intersection or for any expansion factor. If it is required to find the contour of an expansion factor of six, this value of E intersects the E-curve at approximately 19° (2φ). Follow this line up to the e-curve, which it intersects at 1.435. Multiply the length of the MS baseline by 1.435 and lay off this product along the perpendicular bisector of MS, above the baseline, locating point A on fig. 6, a typical reliability diagram for a hyperbolic triad. The cross-hatched area is that of the highest reliability; the expansion factor of both pairs is less than six and the angle of intersection is between 30° and 90°. The points labeled A, B, C, D, are the centres from which the several arcs were drawn. Normally they would not be shown. Using the distance AM as a radius, draw a circle with A as the centre. This arc will be the contour for the expansion factor of six. Multiply distance MS′ by 1.435 and lay this distance off on the MS′ perpendicular bisector as was done previously, thus locating point B. Using the distance MB as a radius draw the circle with B as the centre.

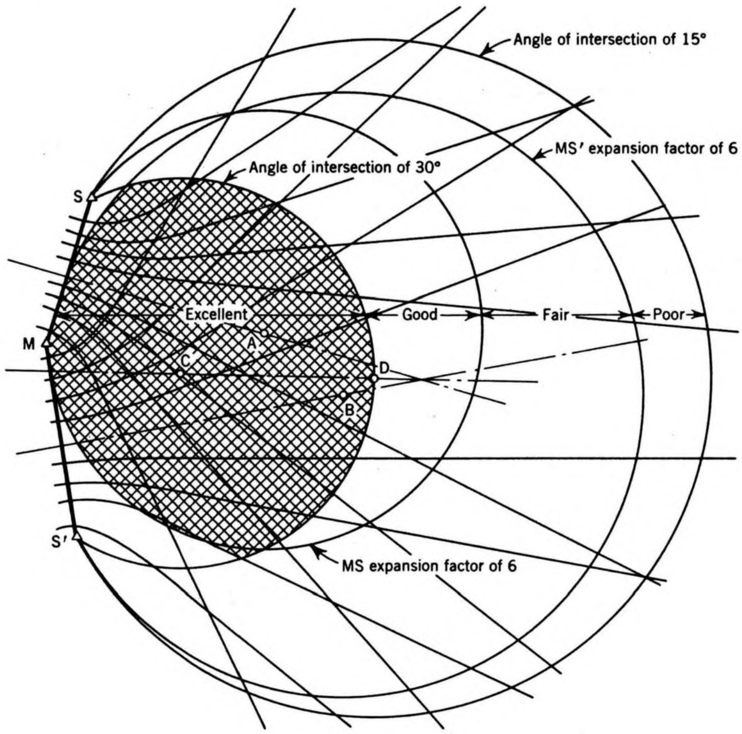

FIG. 6. — Typical reliability diagram, hyperbolic triad.

For the centre of the circle with a constant angle of intersection of 30°, the angle $2\,\varphi = 60°$ intersects the e-curve at 0.289. Lay off point C on the perpendicular bisector of the line joining the slave stations SS′ at a distance above the line equal to 0.289 times distance SS′. With the distance SC as a radius and C as a centre draw the circle. Other expansion factors and angles of intersection can be constructed in the same manner.

Repeatability Contours

A more rigorous method may be developed that is quantitative as well as qualitative. By substituting eq. 8 into eq. 5

$$d_r = \frac{k\,s}{\sin B}\sqrt{\frac{1}{\sin^2\varphi_1} + \frac{1}{\sin^2\varphi_2} + \frac{2\,R\cos B}{\sin\varphi_1\,\sin\varphi_2}} \tag{11}$$

In a triad (three-station net) one line of position is common to both pairs, so that a value of $R = 0.33$ may be assumed.

From eq. 11 a family of curves can be constructed to display convenient values of d_r in metres, in terms of the two angles, $2\ \varphi_1$ and $2\ \varphi_2$. Eq. 11 has been programmed for computer solution, and fig. 7 is derived from this

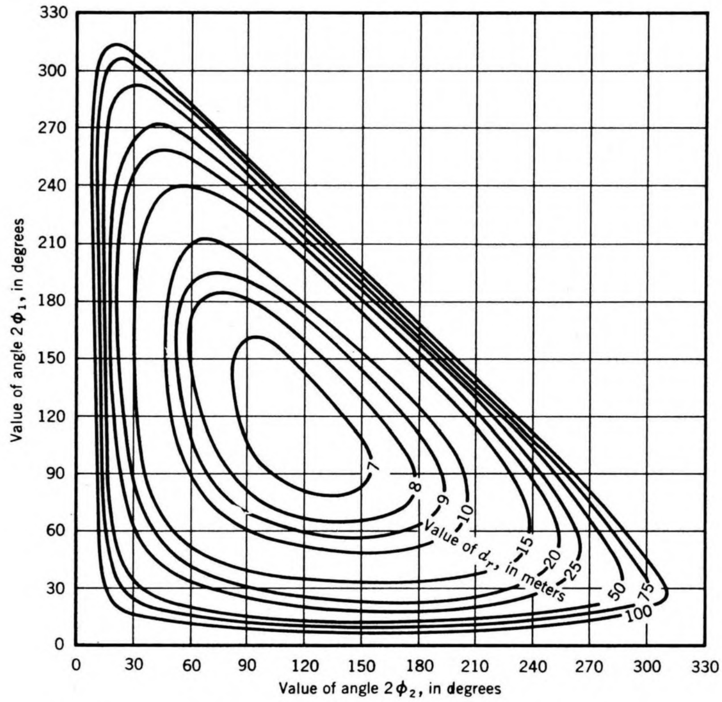

Fig. 7. — Solution of d_r in terms of angles $2\ \varphi_1$ and $2\ \varphi_2$.

solution. A standard deviation of 0.025 lanes, a frequency of 2 Mc per sec and a probability of 95 % were used in the computation. The value of k, from table 3, is 74.9225 m. If 150° is set on one side of a three-arm protractor, and 30° on the other, $2\ \varphi_1$ and $2\ \varphi_2$, respectively, the plotted position will be a point on the 15-metre d_r contour. Additional points on the contour can be plotted using other angles defined by the 15-metre curve.

These data can now be used to plot a reliability diagram showing the repeatability of the net. The pairs of angles defining successive points along a given d_r curve are selected from the curves. Using a three-arm protractor, these angles are then plotted on a conveniently scaled chart. The points thus determined will give the location of that curve for the specific conditions of triad location and baseline length as shown in fig. 8. These repeatability contours for a hyperbolic triad are in metres based on a standard deviation of 0.25 microseconds and 95 % probability. If the signal-to-noise radio is very high during the night, the standard deviation might be 0.50. In this case, the probable error of position at a point might be 400 m at night and 200 m during the day, the value of contours being multiplied by the ratio of the two deviations (0.50/0.25 = 2). Note that the outer contours have been deformed, because the signal cannot be received with reliability beyond a given distance from the farthest station.

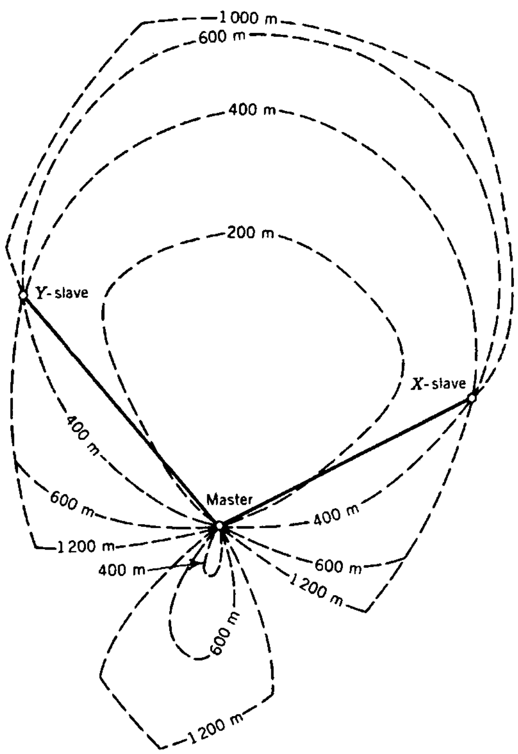

FIG. 8. — Repeatability contours for a hyperbolic triad.

Ranging Systems

In ranging systems, the lines of position are concentric circles. Except at close ranges, short segments of these lines of position can be considered straight lines. In a ranging system, the spacing between successive circles is uniform over the entire area, so that E equals one. The standard deviation, s, is assumed to be the same for both stations. The two stations are independent, so that no lines are shared in common, and $R = 0$. The angle of intersection of the lines of position is equal to the angle between the radius vectors at point P, so that $2\varphi = 2B$, and p is defined by sk. Substituting into eq. 5

$$d_r = \frac{ks\sqrt{2}}{\sin 2B} \tag{12}$$

Because d_r is defined by the angle 2B, eq. 10 will yield the distance from the line joining the two shore stations to the centre of a circle of constant angle of intersection.

In the most reliable part of the area, where the angle of intersection is 90°, eq. 12 reduces to

$$d_r = ks\sqrt{2} \qquad (13)$$

If

$$g = 2\,ks\sqrt{2} = 2\,(is) \qquad (14)$$

in which the probability is 95 % and (i) is given in table 4, the

TABLE 4

Value of i for Various Distance Units

Distance Units	sigma in microseconds	sigma in lane counts
Metres	211.910 799	211 910.799/f
Feet	695.245 214	695 245.214/f
Nautical Miles	0.114 423	114.423/f
Statute Miles	0.131 675	131.675/f

substitution of eq. 14 in eq. 12 results in

$$d_r = \frac{g}{\sin 2B} \qquad (15)$$

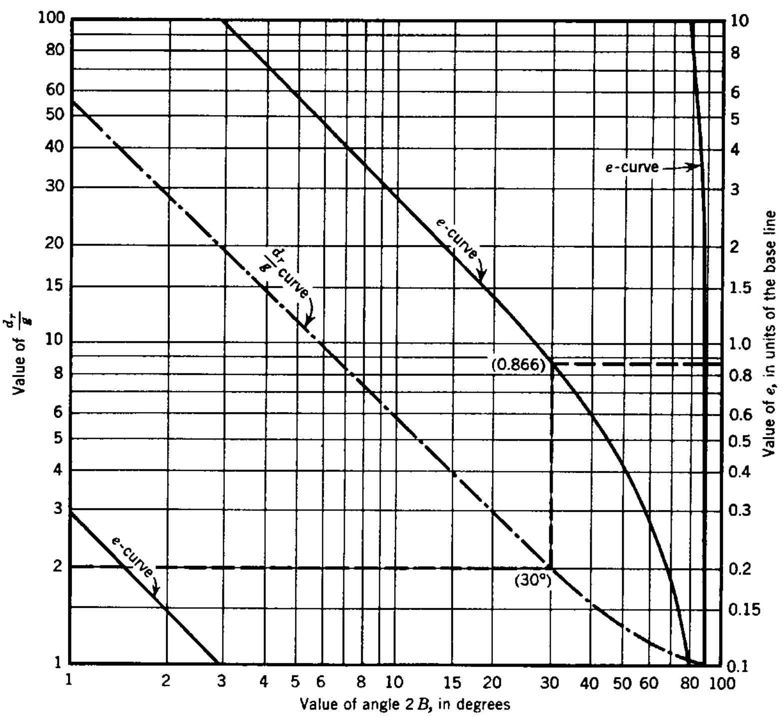

FIG. 9

and

$$\frac{d_r}{g} = \frac{1}{\sin 2B} \tag{16}$$

The contour for a given d_r will then be the circle for an angle 2B. Fig. 9 provides a convenient means of obtaining the value of e in terms of a unit baseline. It should be noted that the e-curve extends upward to the left and downward to the right. The small extension in the lower left has values of e between 10 and 30; for the extension on the right, the values given are divided by 100. Neither of these extensions are very important. For example, if $g = 100$ ft, the contour for $d_r = 200$ ft is defined by $d_r/g = 2$. Enter fig. 9 at $d_r/g = 2$, which intersects the d_r/g curve at 30°. Follow the 30° line to the e-curve, which it intersects at $e = 0.866$. For a specific pair, where $b = 317\,382$ m, $(e)\,(b)$ will be 274 853 m.

TABLE 5

Centre for Given d_r Contour

d_r Contour	d_r/g	e	$(e)\,(b)$, in metres
100 ft	1.00	0	—
200 ft	2.0	0.866	274 853
300 ft	3.0	1.415	449 096
400 ft	4.0	1.936	614 452
15 m	1.0	0	—
20 m	1.33	0.441	139 966
30 m	2.0	0.866	274 853
40 m	2.67	1.237	392 602
50 m	3.33	1.591	504 955
60 m	4.0	1.936	614 452

Other contours are defined in table 5. To plot the contours of constant d_r, first plot the location of the two shore stations at a convenient scale. Erect a perpendicular bisector to the line joining them. From fig. 9, determine the distance e for the desired d_r contour and lay off the quantity e times the distance S_1S_2 (e times b) along the perpendicular bisector. This should be done both above and below the line. Draw circles with these points as centres and the distance from the centre to the shore stations as a radius.

Fig. 10 shows such a diagram for an actual case, with a d_r of 200 ft maximum, and a useful range of 200 nautical miles. Note that the outer contour is deformed by the effect of the distance from the shore stations. It is the degradation in the standard deviation that limits the useful range. That is, the net is limited not by net geometry but by the ability of the system to provide reliable signals at long range. If the signals could be received reliably at distances up to 800 miles, and the shore stations spaced 600 miles instead of 200 miles (380 400 m), the service area would be tremendous.

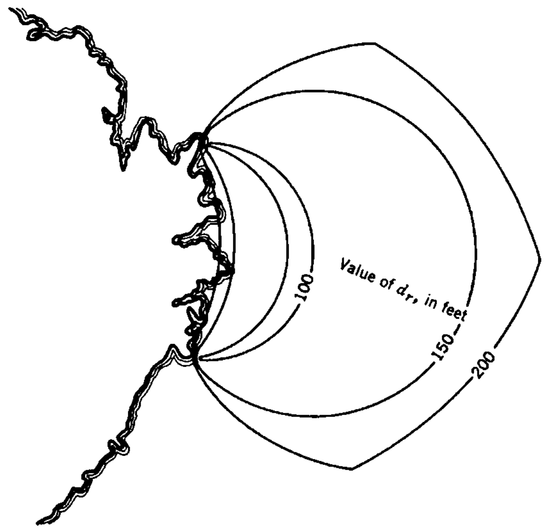

FIG. 10. — Coverage diagram, ranging pair.

Azimuthal Systems

In an azimuthal system such as Consol, the lines of position are the radials from the shore stations, the repeatability is a function of the angular resolution of the equipment (the aperture angle), and the angle of intersection of the rays. The functions used for the ranging system also can be used for the azimuthal.

The two shore stations are independent, therefore R is zero. The errors of position depend on (1) the distance, h, along the radial, (2) the angle of intersection, 2B, of the radials, and (3) the aperture angle, a, measured in degrees. Because one radian equals 57.29578 degrees, p will have the form

$$p = \frac{h\,a}{57.296} \tag{17}$$

It may be assumed that the reliability of both beacons is the same, and the same in all directions. Substituting in eq. 5 for a 95 % probability yields

$$d_r = \frac{0.0349\ a}{\sin\ 2B}\sqrt{h_1^2 + h_2^2} \tag{18}$$

When

$$h_1 = h_2 = h = \frac{b}{2\ \sin\ B} \tag{19}$$

is derived from fig. 4, substitution results in

$$d_r = \frac{0.02468\ (a)(b)}{(\sin\ 2B)\ (\sin\ B)} = \frac{j}{\sin\ 2B} \tag{20}$$

in which j is equal to 0.025 (a) $(b)/\sin$ B. This is in a form that can be used with fig. 9. A reliability diagram can be prepared as for the ranging system.

The MPF System produces the two angles required for position by a solution of the three-point problem. This may be done either graphically or by computation. The mathematics involved in the computation are quite involved, therefore it is more satisfactory to solve the problem by graphical plot. It has been determined that a three-point fix may have an error of position of ± 10 yards at 1/10 000 plotting scale. The error is proportionately greater at smaller scales (*) (**).

Composite Systems

In a composite system the repeatability of position is the resultant of the repeatabilities of the two different lines of position.

CONDUCTIVITY

When both the transmitting and receiving antennae are close to the ground, the direct- and ground-reflected components of the groundwave tend to cancel out. The resulting field intensity is principally that of the surface wave. The groundwave is affected by : (1) the electrical conductivity of the surface over which it propagates; (2) the dielectric constant of that surface; (3) the index of refraction of the air through which it propagates; and (4) the lapse rate of the index of refraction with altitude. The surface wave is able to follow the curvature of the earth. It is not confined to the earth's surface, however, but extends upward a considerable height.

The surface wave propagating over the surface of the earth is vertically polarized. Any horizontal component of the electrical field in contact with the ground would be short circuited by the earth. The groundwave induces charges in the earth that travel with the wave and so constitute a current. In carrying this induced current, the earth behaves like a leaky capacitance, and can be represented by a resistance (or conductance) shunted by a capacitative reactance. The characteristics of the earth as a conductor can therefore be described in terms of conductivity and dielectric constant.

Because no surface is a perfect conductor or a perfect ground, losses retard the grounded edge of any given wave front. This causes the wave front to tilt in the direction of travel so that successive wave fronts have a forward inclination. Poor conducting surfaces cause high loss and greater tilt. Table 6 shows the variation in angle of tilt from the vertical for

(*) *Resection by Intersection*, by Erwin SCHMID, Journal of the Coast and Geodetic Survey, No. 6, August 1955.
(**) *The Three Point Problem*, by Lansing G. SIMMONS, Journal of the Coast and Geodetic Survey, No. 6, August 1955.

frequencies from 20 kc to 20 Mc, propagated over sea water and over dry ground. As frequency increases, the angle of tilt increases. This tilting of the electric vector of an electromagnetic wave is not to be confused with the bending of a wave, or defraction, which is a phenomenon associated with a wave front striking the edge of a solid object, the greatest bending taking place at the lowest frequencies.

TABLE 6

Angle of Tilt versus Frequency

Frequency	Angle of Tilt over	
	Sea Water	Dry Ground
20 ks per sec	0° 02.'5	4° 18′
200 ks per sec	0° 08′	13° 30′
2 Mc per sec	0° 25′	32° 12′
20 Mc per sec	1° 23′	35° 00′

THE ELECTRICAL CHARACTERISTICS OF THE EARTH

The electrical characteristics of the ground are determined by the nature of the soil, its moisture content, its temperature, and the geological structure of the ground. The effective depth of penetration of the waves for given ground conditions is a function of the frequency of the signal. The absorption of energy by vegetation, buildings, and other objects on the surface, while slight, cannot be neglected.

The frequency of the groundwave determines the particular component of the wave that will prevail along any given signal path. At frequencies below 10 Mc per sec, the conductivity of the terrain determines the transmission characteristics of the surface wave. At frequencies between 10 Mc per sec and 30 Mc per sec, the conductivity becomes less important and the dielectric constants of the terrain determine the surface wave transmission characteristics. The signal is strongest for higher dielectric constants and lower frequencies. At frequencies greater than 30 Mc per sec the losses suffered by the surface wave become excessive. Transmission at high frequencies, therefore, is usually possible only by means of the direct wave or, approximately, line of sight. The permeability of the ground relative to free space can be regarded as unity, and has no effect on propagation problems.

It has been established by numerous measurements that the conductivity and dielectric constant of the soil vary with the nature of the soil. However, it seems probable that this variation may be due, not so much to the chemical composition of the soil, as to its ability to absorb and retain moisture. The normal conductivity of loam is on the order of 10^{-2} mhos per m. However, when dried, loam will have a conductivity as low as 10^{-4} mhos per m. This is of the same order as the conductivity of granite.

The moisture content appears to be the major factor in determining the electrical characteristics of the ground. Below a depth of approximately one metre the moisture content of a particular soil, at a particular site, seems to remain more or less constant throughout the year. There are variations, but they tend to fluctuate about a mean value. The moisture content for a particular soil may, however, vary considerably from place to place, depending on the drainage conditions.

Laboratory measurements of the electrical characteristics of different solid samples have been made. They show that the temperature coefficient of conductivity is on the order of 2 % per degree centigrade. The temperature coefficient of the dielectric constant is negligible. At the freezing point there is a large and rapid change in both factors. However, temperature variations decrease rapidly with depth. Therefore, this change at the freezing point is important at high frequencies, where the penetration of the waves is small; and in arctic and subarctic areas, where the ground is frozen to a considerable depth.

The ground involved in overland propagation is not usually homogeneous so that the effective conductivity is determined by several different types of soil. It is, therefore, of great importance to have a complete knowledge of the general geologic structure of the region of interest. The effective conductivity is determined, not by the nature of the surface soils alone, but rather of the nature of the soil profile. The extent of this profile is determined by the depth to which there are ground currents of appreciable magnitude. The underlying strata form a part of the medium through which the waves propagate. These strata have an indirect effect by determining the height of the water table, and hence the amount of moisture in the ground.

The depth of penetration is defined as that depth at which the wave has been attenuated to 37 % of its surface value. For frequencies between

TABLE 7

Depth of Penetration of Waves into Ground

	GROUND TYPE		
Conductivity	5	10^{-2}	10^{-3}
Dielectric Constant	81	10	5
FREQUENCY	PENETRATION, IN METRES		
10 kc per sec	2	50	150
100 kc per sec	0.67	15	50
3 Mc per sec	0.2	5	17
10 Mc per sec	—	2	9

10 kc per sec and 10 Mc per sec the penetration is given in table 7 (*). Because of the relatively deep penetration of the earth at these frequencies, conductivity is not particularly sensitive to conditions at the actual surface of the ground, such as recent rainfall. The conductivity of typical varieties of soil conditions are shown in table 8.

TABLE 8

Typical Ground Constants (**)

Type of Terrain	Dielectric Constant	Conductivity (mhos/metre)
Pastoral, low hills, rich soil, typical of Dallas, Texas and Lincoln, Neb., areas	20	3×10^{-2}
Pastoral, low hills, rich soil, typical of Ohio and Illinois	14	10^{-2}
Pastoral, medium hills and forestation, heavy clay soil, typical of Central Virginia	13	4×10^{-3}
Pastoral, medium hills and forestation, typical of Maryland, Penna., and New York except for mountainous territory and sea coasts.	13	6×10^{-3}
Flat country, marshy, heavily wooded, typical of Louisiana near the Mississippi River..	12	7.5×10^{-3}
Rocky soil, steep hills, typical of New England	14	2×10^{-3}
Sandy, dry, flat, typical of coastal country ..	10	2×10^{-3}
City, industrial area — Average attenuation ..	5	10^{-3}
City, industrial area — Maximum attenuation.	3	10^{-4}
Sea Water varies with Salinity and Temperature:		
At salinity of 35 ‰ and temp. 20°C(68°F)	81	4.78
At salinity of 20 ‰ and temp. 10°C(50°F)	81	2.29
Fresh Water varies with Temperature at 20 °C		4×10^{-2}
10 °C	80	3×10^{-2}
0 °C		2×10^{-2}

Of the electrical characteristics (conductivity, dielectric constant, and permeability), only conductivity is of importance for transmission over sea water at frequencies below microwave. The conductivity of water varies with salinity and temperature, as shown in fig. 11, in which the temperature

(*) *Report 139*, International Radio Consultative Committee (CCIR), IX Plenary Assembly, Vol. III, Los Angeles, Calif., p. 272.

(**) Data from : TERMAN, F.R., *Radio Engineer's Handbook*, p. 709, McGraw-Hill Book Co., N.Y., with addition from HO Sp 11 (1956) and other sources.

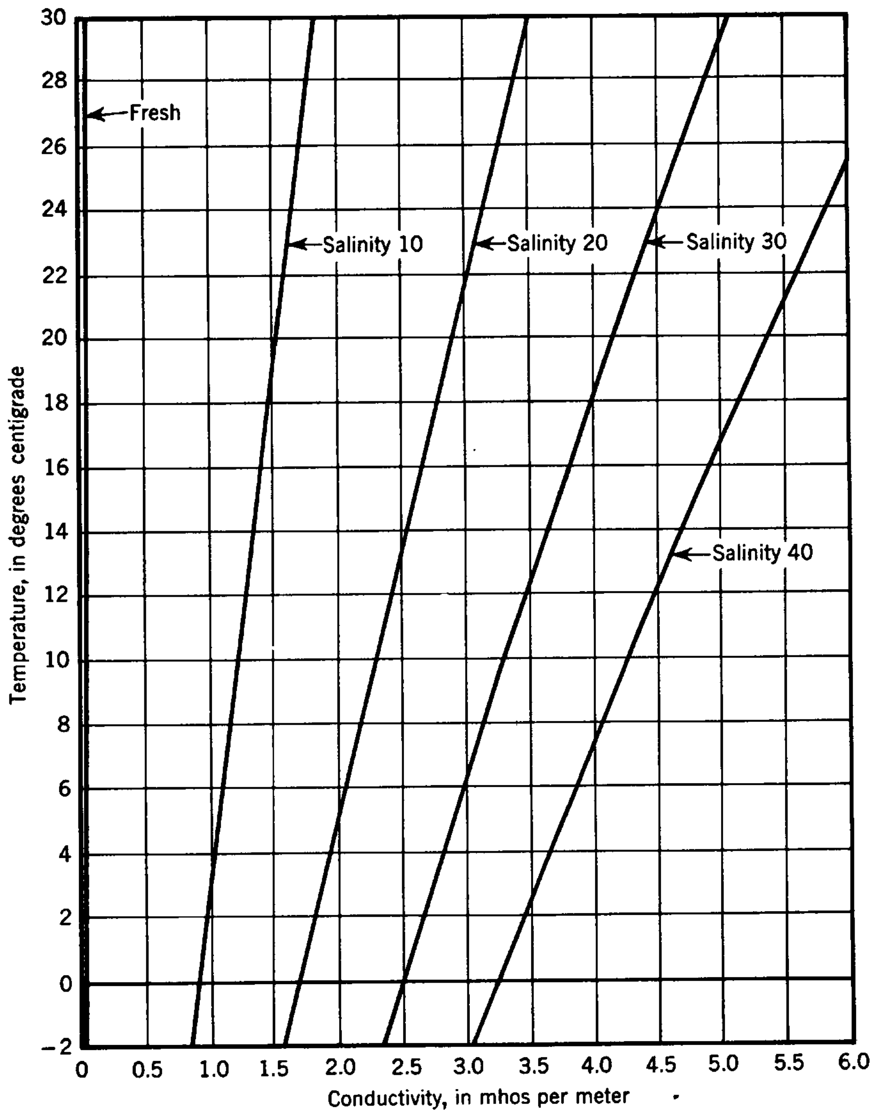

FIG. 11. — Conductivity of water.

is in degrees centigrade and salinity in parts per thousand. There is a tendency, clearly indicated in the literature, to assume that the conductivity of sea water is a constant, with a value of approximately five mhos per m. Although the conductivity of water is not a constant, conditions along a water path tend to be more homogeneous than along a land path. At its worst, the conductivity of water is far better than the conductivity over the best land path. For consistent results minimum land paths are highly desirable.

THE SECONDARY FACTOR

It has been found convenient to express the total electromagnetic field propagated by a groundwave as the product of two factors, the primary field, or free-space field, expressed in volts per metre, and the secondary factor, a dimensionless number. This secondary factor takes into account the disturbing influence of the earth, primarily its conductivity.

In considering the propagation of a groundwave over an actual path, it must be assumed that it will pass over a series of sections each with a different conductivity, but homogeneous within the section.

In the vicinity of the transmitter, but beyond the range of the induction field, the radiation initially obeys an inverse-distance law of field strength. For a spherical earth, the law of attenuation with distance becomes an exponential type, with an attenuation coefficient that is dependent on the earth constants.

Consider a non-homogeneous earth in which, at a certain distance from the transmitter, there is a boundary where the wave crosses from one kind of earth to another. It may be assumed that the new section extends an indefinite distance. At a sufficient distance beyond the boundary, the propagation becomes characteristic of the earth constants of the new section. In other words, the type of transmission becomes the same as if the new section had extended back to the transmitter.

At the boundary itself, a disturbance in transmissions exists because of the change in propagation from one characteristic type to another. It is assumed that the disturbance at the boundary must extend in some degree toward the transmitter so that the wave must be somewhat modified before it reaches the boundary. It has been supposed that this effect will be small enough to be negligible. There is now reason to doubt the validity of this assumption.

The problem of the land-sea boundary has been given some study, but this problem needs more field investigation. The secondary factor may be expected to vary from hour to hour with fluctuations in tidal conditions if the land-sea boundary is a relatively flat foreshore such as a mudflat or sandy beach. If the boundary is steep, such as a bluff or cliff, the secondary factor may be expected to remain relatively stable.

The secondary factor is defined in some detail by J. R. JOHLER, W. J. KELLER and L. C. WALTERS (*). Fig. 12 is derived from them, and other sources, and presents the effect of conductivity on the phase of the secondary factor at various distances from the transmitter for selected conductivities and a frequency of 100 kc per sec. Table 9 shows the slope of the curves beyond 500 nautical miles.

(*) *Phase of the Low Radio-Frequency Ground Wave*, by J. R. JOHLER, W. J. KELLER, and L. C. WALTERS, U.S. National Bureau of Standards Circular 573, Washington, D.C., June 1956.

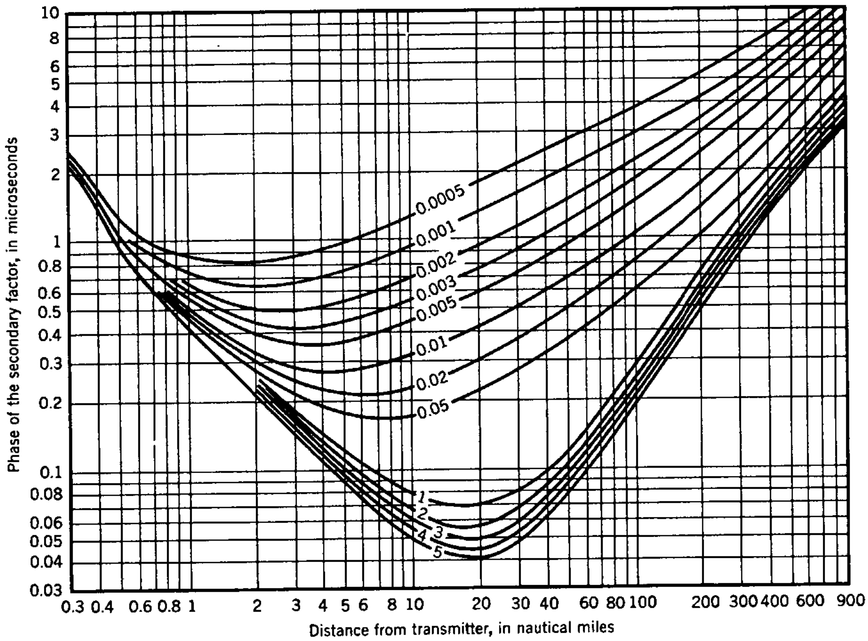

FIG. 12. — Phase of secondary factor for selected conductivities.

TABLE 9

Slope of the Curves beyond 500 Nautical Miles

Conductivity	Microseconds per nm
0.0005	0.0107
0.001	0.0105
0.002	0.0102
0.005	0.0086
0.01	0.0081
0.02	0.0069
0.05	0.0051
1.0	0.0048
2.0	0.0045
3.0	0.0044
4.0	0.0043
5.0	0.00419

It can be shown that the phase of the total wave is the sum of the phase of the primary wave and the phase of the secondary factor. The time in microseconds required by the total field to reach a given point is the sum of the time for the primary wave and that for the secondary wave.

In most computations of electronic lattice the secondary factor used is based on the assumption that all propagation is over sea water with a conductivity of 5 mhos per m. This is almost never true and therefore there will be a secondary phase correction which we may call Δt_c.

This correction may be determined from fig. 12 by the modified Millington Method.

SECONDARY PHASE CORRECTION

The field strength of a signal received from a distant source must be the same, no matter which end of the line is the transmitter. For a homogeneous path this is true. For composite paths this is not always the case.

Millington's Method is based on the following premise : the phase distortion due to a composite land-sea path is the arithmetic average of the phase distortion found in the forward and reverse direction.

The phase distortion effect is cumulative. Therefore, a composite path may be formed using the various curves for conductivity from fig. 12. The parts of the curve corresponding to the distance from the source for each section of the path are used.

A diagrammatic representation of a hypothetical path is shown in fig. 13. The corresponding secondary phase correction is computed for this path in table 10.

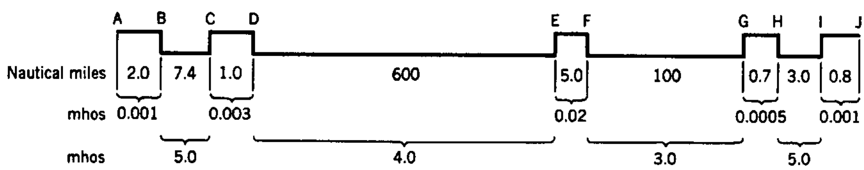

FIG. 13. — Diagram of the transmission path.

The phase distortion caused by the terrain of the several sections of the path is computed on the successive lines of table 10. The length of the section and the conductivity are given in fig. 13. The transmitter is first assumed to be at A. The phase distortion is obtained from fig. 12, using the curve for the indicated conductivity and the distances from the transmitter to the two ends of the section. The phase distortion, t_1, is for the distance from the transmitter to the section limit listed in column 1. The phase distortion, t_2, is for the distance from the transmitter to the section limit in column 2.

In the first section, the distance A to B is 2.0 miles and the conductivity is 0.001 mhos per m. Because A is the transmitter, t_1 is zero. From fig. 12, using the curve for conductivity 0.001 and a distance of 2.0 miles, t_2 is found to be 0.63 microsecond. The phase distortion for the section is $t_2 - t_1$ or 0.63 microsecond.

TABLE 10

Computation of the Secondary Phase Correction by the Modified Millington Method

Section from 1	Section to 2	Length of Section, in nm	Distance from Transmitter, in nm	Conductivity in section, in mhos per m	t_2	t_1	t_2-t_1
A	B	2.0	2.0	0.001	0.63	0.0	0.63
B	C	7.4	9.4	5.0	0.05	0.22	— 0.17
C	D	1.0	10.4	0.003	0.56	0.54	0.02
D	E	600.0	610.4	4.0	2.17	0.05	2.12
E	F	5.0	615.4	0.02	4.00	3.96	0.04
F	G	100.0	715.4	3.0	2.75	2.31	0.44
G	H	0.7	716.1	0.0005	10.83	10.82	0.01
H	I	3.0	719.1	5.0	2.12	2.11	0.01
I	J	0.8	719.9	0.001	8.41	8.40	0.01
					t (forward)		3.09
J	I	0.8	0.8	0.001	0.82	0.0	0.82
I	H	3.0	3.8	5.0	0.11	0.54	— 0.43
H	G	0.7	4.5	0.0005	0.95	0.91	0.04
G	F	100.0	104.5	3.0	0.28	0.12	0.16
F	E	5.0	109.5	0.02	0.85	0.80	0.05
E	D	600.0	709.5	4.0	2.60	0.24	2.36
D	C	1.0	710.5	0.003	7.51	7.50	0.01
C	B	7.4	717.9	5.0	2.50	2.47	0.03
B	A	2.0	719.9	0.001	9.50	9.48	0.02
					t (back) ..		3.06
					sum		6.15
Total Phase Distortion, t					mean		3.08
Secondary Phase Factor for all seawater path, 719.9 nm, 5.0 mhos per m, t_s							2.52
Secondary Phase Correction, Δt_c							0.56

In the second section, the distance from B to C is 7.4 miles and the conductivity is 5.0 mhos per m. The distance from the transmitter to B is 2.0 miles, and to C is 9.4 miles. Using the curve for 5.0 on fig. 12, t_1 is found to be 0.22 and t_2 is found to be 0.05. The difference is — 0.17 for the section.

This procedure is continued until the phase distortion has been computed for each section. The sum of these values is t_f, the phase distortion in the forward direction. The transmitter is then assumed to be at J. A new series of computations is then made in the same manner. The sum of these phase distortions is t_r, the phase distortion in the reverse direction.

The total phase distortion, t_c, is the mean of t_f and t_r, or 3.08 microseconds. However, in computing the tables and charts, the assumption is made that there is a homogeneous seawater path, with a conductivity of 5.0 mhos per m. This quantity, t_s, is obtained from the curve for conductivity 5.0 of fig. 12, and the total distance, 719.9 miles. The value is 2.52. The secondary phase correction, Δt_c, is the difference between the total phase distortion and the distortion for the homogeneous seawater path. In this case, 3.08 minus 2.52, or 0.56 microseconds. This is the secondary phase correction for this one point in the service area. The same computation must be made to obtain the secondary phase correction for each position of the receiver in the service area.

REFRACTIVITY

At low frequencies the surface wave is the most important element in radio propagation. As the frequency increases, the attenuation of the surface wave increases, until in the vicinity of 30 Mc per sec the surface wave virtually disappears. As the surface wave diminishes, the tropospheric wave increases in importance. As the conductivity becomes less important, the refractivity becomes more important.

While the entire radiated wave is influenced to some degree by the radio refractive index of the atmosphere through which it propagates, that part of the wave that remains within the troposphere is most strongly affected. The atmosphere causes a downward curvature of the horizontally launched radio waves. Under some meteorological conditions the radio energy may be confined to thin layers near the earth's surface, which act as ducts. As a result, abnormally high field strengths may be observed well beyond the normal radio horizon. At other times a transition layer between differing air masses will act as a reflector of radio energy. In addition to these gross effects, the atmosphere is always more or less turbulent with the result that radio energy is scattered out of the normal radiation pattern.

The radio refractive index of air, n, is a function of atmospheric pressure, temperature, and humidity, combining in one parameter three of the normal meteorological elements used to define climate. The lapse rate of the index of refraction with altitude is of minor importance, except in connection with high frequencies, aerial navigation, and skywave propagation.

The relation of the refractive index and free space velocity of propagation to the effective velocity of propagation, through the atmosphere over a specific path, is shown by the following expression

$$V_e = \frac{C}{n} \tag{21}$$

in which V_e is the effective velocity of propagation, C denotes the free space velocity (approximately 299 792.5 km per sec), and n represents the effective index of refraction of the atmosphere near the surface.

THE RADIO REFRACTIVE INDEX OF AIR

Near the surface of the earth, the refractive index is a number on the order of 1.0003. Since the variation is seldom more than a few parts in 10^{-4} it is more convenient to use the refractivity, N, of the atmosphere, in which

$$N = (n - 1)\, 10^6 \qquad (22)$$

E. K. SMITH and S. WEINTRAUB (*) have shown that N may be determined within 0.05 % for reasonable ranges of the several terms by the expression

$$N = \left(\frac{77.6}{T}\right)\left(A - 4810\, \frac{e_s R}{T}\right) \qquad (23)$$

in which T is the temperature in degrees Kelvin, A describes the total atmospheric pressure in millibars, R denotes the relative humidity, in percentage, expressed as a decimal and e_s refers to the saturation vapor pressure at temperature T.

At any given point, the meteorological factors may be observed and hence the effective velocity may be computed at that point. Fig. 14 shows the effective velocity of propagation of radio waves for various conditions of temperature and humidity, at an atmospheric pressure of 1 013 millibars.

The requirement is not for the conditions at a point, but for those along a line. If conditions are relatively homogeneous, the factors may be observed at the two ends of the line, and a mean value used. This will probably be adequate for short periods, of short distances, but these meteorological factors change from place to place and from time to time. Homogeneous conditions cannot be assumed for any length of time, or for any great distance, even over the oceans. What is required is not data for points, but conditions over areas. This calls for synoptic studies. But even synoptic studies of the meteorological conditions are based on data at relatively few points in an area. Such studies assumed more or less uniform variations between observations both in time and in space

B.R. BEAN, J.D. HORN and A.M. OZANICH (**) have made a study of the world-wide distribution of sea-level refractivity, N_0, based on world weather records. They have prepared charts for the range and distribution of N_0.

(*) *The Constants in the Equation for Atmospheric Refractive Index at Radio Frequencies*, by E. K. SMITH and S. WEINTRAUB, Proceedings, IRE, Vol. 41, No. 8, August 1953.

(**) *Climatic Charts and Data of the Radio Refractive Index for the U.S. and the World*, National Bureau of Standards Monograph 222, Washington, D.C., November 1960.

These data, however, are statistical generalities. Such charts will not give the actual refractivity at any given time or place. They give the average conditions; the effect of atmospheric turbulence must be allowed for

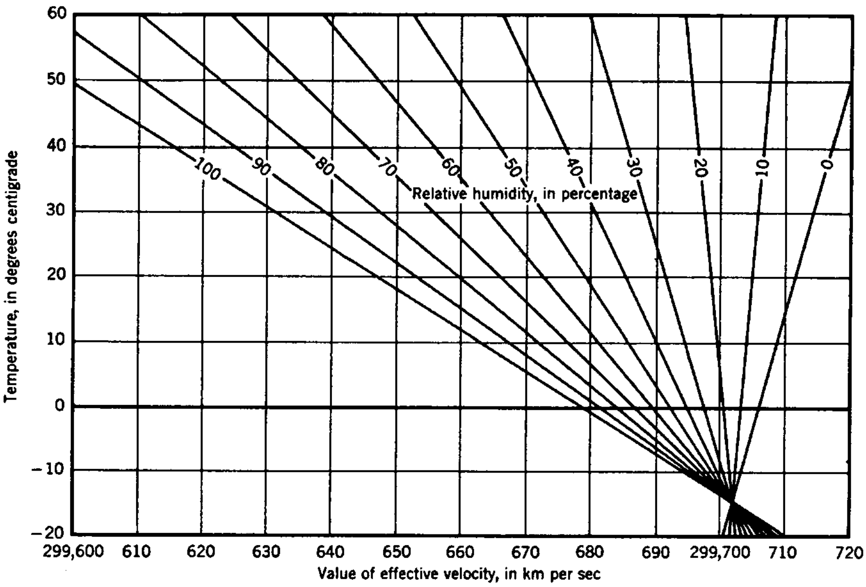

FIG. 14. — Effective velocity of propagation.

Perhaps the meteorological effect over the service area should be measured by some function of the refractivity, rather than by the refractivity itself at any one point, or even the mean of a few observations.

CORRECTION FOR EFFECTIVE VELOCITY

If V_c is the nominal velocity used to compute an electronic lattice, then the effective velocity, V_e, over a given transmission path will be

$$V_e = V_c + \Delta V_N \tag{24}$$

in which ΔV_N is the modification to the nominal velocity due to the actual refractivity of the atmosphere through which the wave is being propagated. For any given V_c there will be a nominal refractivity, N_c. If the conditions along the transmission path are different from those assumed, then N_c will be modified by an amount,

$$\Delta N = \Delta n \cdot 10^6 \tag{25}$$

The effective velocity will be

$$V_e = V_c + \left(\frac{-V_c \Delta N}{N_c + \Delta N + 10^6} \right) = \frac{V_c N_c + V_c 10^6}{N_c + \Delta N + 10^6} \tag{26}$$

V_c and N_c are known quantities, because they are the nominal velocity and refractivity assumed to fit the general conditions. ΔN is the change in the refractivity caused by the difference between conditions assumed to exist and those actually existing along the transmission path. For relatively short distances, conditions can be assumed to be homogeneous, and ΔN will be zero. For long distances, homogeneous conditions cannot be assumed to exist. A knowledge of the conditions actually existing along the transmission path is required.

PREDICTABILITY

The predictability of any electronic positioning system is the ability of that system to define the location of a given point on the surface of the earth. This definition is in terms of geographic coordinates, rather than in terms of the electronic lattice developed by the system. Once the random and systematic errors affecting the system have been taken into account, it is possible to return to a given point on the surface of the earth within reasonable limits. This is the repeatibility of that system, and it is in terms of the electronic lattice of that system.

The assumption is usually made that the lattice as computed and laid down on a chart (or built into a computer program) is a true representation of the actual lattice being propagated by the system. That is, when the user arrives at a given intersection of lines or position, he has arrived at the latitude and longitude shown on the chart of indicated by the computer. This is true only if there is no distortion of the lattice. If the effective velocity of propagation is accurately known, then it will be possible to compute (or predict) the real geographic coordinates of a given point in the system. The factors that modify the free space velocity to produce the effective velocity, and thus distort the lattice, are the refractivity of the atmosphere and the conductivity of the surface.

The velocity of propagation of waves in the electromagnetic spectrum in a vacuum has been given considerable study. A value of 299 792.5 ± 0.4 km per sec has been used in this study. It is the value recommended by the International Union of Geodesy and Geophysics in 1957.

It has been stated that the predictability is a function of the refractivity and the conductivity. Because these functions are difficult to evaluate, at this time (1963) they are derived empirically.

THE ERROR OF PREDICTABILITY

The error of predictability can be defined in the following terms. At point P in the service area of any system, it can be shown that the observed lines of position give distances measured by

$$D = t\ V_c \qquad (27)$$

in which t is the time interval measured in microseconds, or in lanes divided by frequency; and V_c denotes the assumed velocity of propagation used to compute the system.

If V_c is modified by the refractivity of the atmosphere the distance will be

$$D + d_N = t\ (V_c + \Delta\ V_N) \qquad (28)$$

and

$$d_N = t\ \Delta\ V_N \qquad (29)$$

in which d_N is the error in the line of position at point P due to the refractivity of the atmosphere, and $\Delta\ V_N$ represents the change in the effective velocity due to the refractivity, having due regard for sign.

It has been shown that the transmission time is subject to a correction for the effect of overland transmission, t_c , so that the distance will be

$$D + d_L = (t + \Delta t_c)\ V_c \qquad (30)$$

and

$$d_L = V_c\ \Delta t_c \qquad (31)$$

The error of predictability in any one line of position will be

$$d_N + d_L = t\ \Delta\ V_N + V_c\ \Delta\ t_c \qquad (32)$$

An expression for the error of position in terms of the displacement of the lines of position and the angle of intersection was developed as eq. 5. If it is assumed that the correlation factor is zero, the error of predictability will be

$$d_p = \frac{1}{\sin\ B} \cdot \sqrt{d_{p1}^2 + d_{p2}^2} \qquad (33)$$

in which d_{p1} and d_{p2} are the errors of the two lines of position, respectively, and

$$d_p = \frac{1}{\sin\ B} \sqrt{(t_1\ \Delta\ V_{N1} + V_c\ \Delta\ t_{c1})^2 + (t_2\ \Delta\ V_{N2} + V_c\ \Delta\ t_{c2})^2} \qquad (34)$$

Eq. 34, although it has no immediate practical value, is of academic interest. The terms can only be evaluated empirically. There is not enough known about what happens in the atmosphere through which the wave is propagated. Nor is there exact information on how the wave propagates over different kinds of surfaces.

At present (1963), there is only one way to determine the predictability of an electronic system. This is to take a receiver to known positions within the service area. The difference between the computed lattice coordinates and the observed lattice coordinates (corrected for errors of repeatability) is the predictability at that point. If observations are made at a sufficient number of known points, a map of corrections can be prepared. Two problems exist : how many points are a " sufficient number " (10 ? 100 ? 1 000 ?) and how are the " known points " to be established ? On land this latter problem is merely the extension of geodetic control to the " sufficient number " of " known points ". This is time consuming, but feasible. However, at sea, where the long range systems are most useful, it is a real problem.

If the monitor is positioned by visual means, this limits the area sampled to the immediate vicinity of shore-based control. Here propagation anomalies are known to exist. If the monitor is positioned by another electronic system, how is that system calibrated ? How can it be assured that some factor will not influence both systems ? The calibration of a low frequency system by a high frequency system seems to be the safest. This is because one is relatively insensitive to refractivity and the other to conductivity.

With long range systems there is another problem that is seldom encountered with short range systems. For the predictability of a system to be meaningful, all of the shore stations must be positioned on the same geodetic datum. If, for example, a long range net is established with three transmitters on three islands. If the positions are each referred to an independent astronomical determination, there is trouble. The spread between the three independent astronomical observations of position can be on the order of plus or minus a mile. The spread will be several hundred yards at a minimum. Any ambiguity of position of the transmitters is magnified when projected into the service area. In such a case predictability is not a dream, it is a nightmare.

CONCLUSIONS

Before World War II the possibility of precise positioning out of sight of land was not even considered. Presently, more-or-less pinpoint positioning is available several hundred miles from land. World-wide coverage with no more than navigational accuracy is feasible. The problem is primarily one of economy. It is necessary to obtain and install the system or systems which will provide the best coverage with the fewest units.

Repeatability on the order of a few feet and almost perfect predictability are possible using frequencies above 30 Mc per sec. However, at these frequencies the useful range is limited to approximately line-of-sight. For the most part, these short-range systems are ranging systems. The principal exception is the MPFS, which is azimuthal. Hi-Fix is a short-range system which uses the intermediate frequency (2 Mc per sec) band. It can be operated in either a hyperbolic or ranging mode.

Repeatability and predictability adequate for precise positioning can be obtained at medium ranges with both hyperbolic and ranging systems. There are two frequency bands that can be used for positioning at range greater than line-of-sight : below 300 kc per sec and above approximately 2 Mc per sec. Between 300 kc per sec and 2 Mc per sec daylight transmission suffers serious absorption. The intermediate band between 2 Mc per sec and 30 Mc per sec is more efficient. However, skywave contamination renders transmission at this frequency unreliable at ranges greater than 200 miles. Lorac uses the intermediate frequency band, while Decca uses the low frequency band.

With intermediate frequencies, baseline lengths for precise work are restricted to approximately 200 miles by the necessity for maintaining net stability. The requirements of net stability, net geometry, and power limit the intermediate frequency systems to a maximum range of approximately 400 miles.

Low frequency systems are better adapted for long range requirements than are the intermediate ones. While the low frequency systems are not as sensitive to skywave contamination as are the intermediate ones, the problem still exists and becomes more serious with increasing range. The repeatability of the low frequency systems can be determined with some precision.

The predictability of the electronic positioning system is too often a matter of guess work. For the short range, high frequency system, it is not a problem. For the medium range, intermediate frequency system, the problem is there, but is not significant. For the long range, low frequency system the problem is significant. At present (1963) the only practical method of calibration is to determine the position of a receiver in the service area by some other system. This other means may be visual control, which demands that it sample only a small area close to the available control. At sea, the check must be made in close proximity to the shore line where propagation anomalies are known to exist. If the other means is another electronic system, then propagation anomalies may exist that affect both systems, and which will therefore remain undetected.

Until the predictability of a long range system can be clearly defined, the real accuracy of that system is in doubt. It is evident that more work is needed to determine which factors govern the variations in propagation over different kinds of surfaces and the boundaries between them. Knowledge is needed as to how the variations in the meteorological conditions of the atmosphere affect the propagation velocity. Techniques need to be developed to determine the variation with time and the variation in space of the elements that are found to affect the propagation. Techniques are also required for the adequate calibration of systems, particularly those with long range.