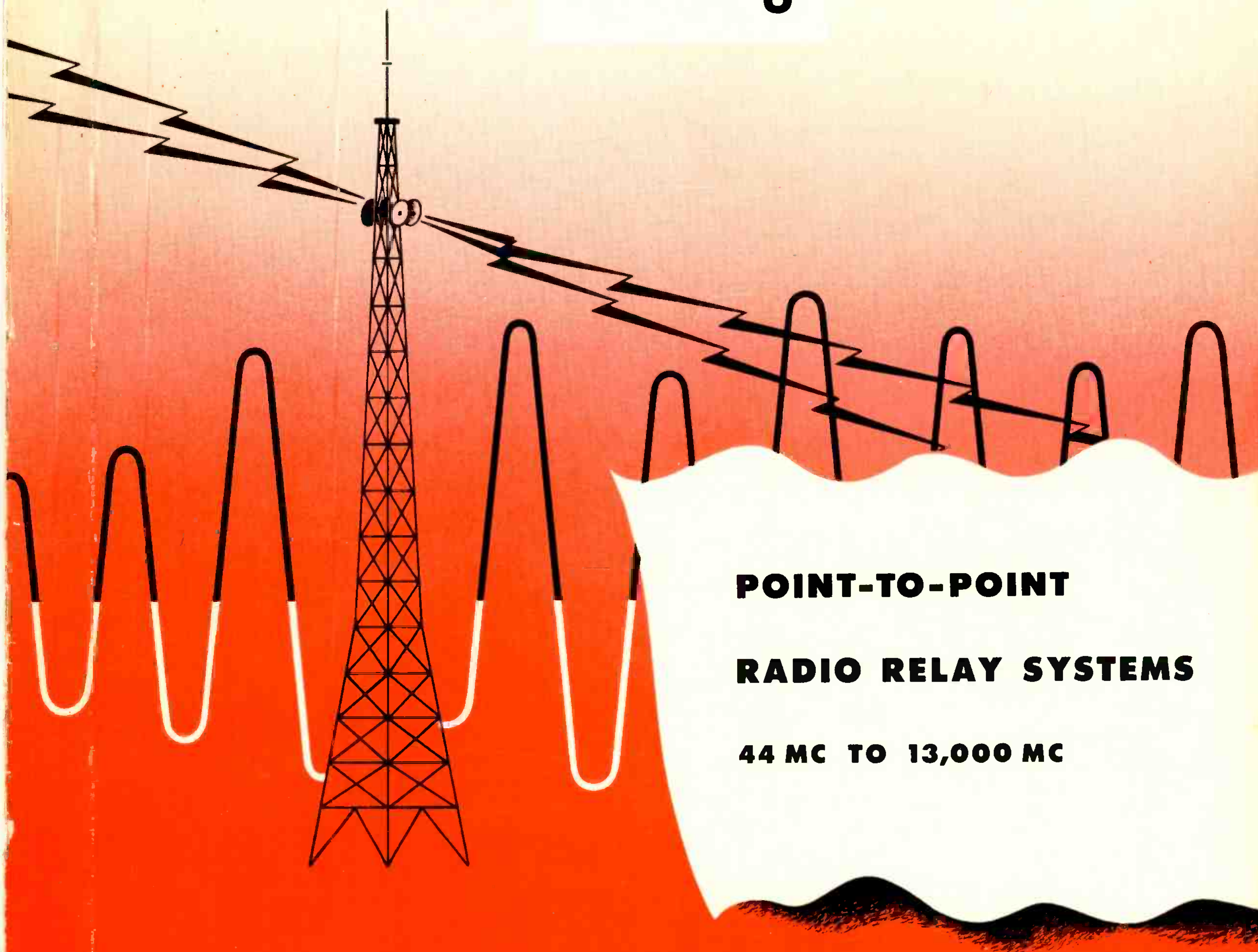
RCA electronic training series



POINT-TO-POINT RADIO RELAY SYSTEMS

44 MC TO 13,000 MC

Published by Government Service Department • RCA Service Company, Inc., Camden, N. J.
A Radio Corporation of America Subsidiary

POINT-TO-POINT

RADIO RELAY SYSTEMS

44 MC to 13,000 MC

PREPARED BY

RCA SERVICE COMPANY, INC.

A Radio Corporation of America Subsidiary

CAMDEN, N. J.

for

AIR FORCE CAMBRIDGE RESEARCH CENTER
AIR RESEARCH AND DEVELOPMENT COMMAND

Reprinted by permission
Air Force Cambridge Research Center

30 September 1954

POINT-TO-POINT RADIO RELAY SYSTEMS

TABLE OF CONTENTS

TABLE OF CONTENTS (cont'd)

TABLE OF CONTENTS (cont'd)

TABLE OF CONTENTS (cont'd)

LIST OF TABLES AND ILLUSTRATIONS

Figure *Page*

LIST OF TABLES AND ILLUSTRATIONS (cont'd)

LIST OF TABLES AND ILLUSTRATIONS (cont'd)

LIST OF TABLES AND ILLUSTRATIONS (cont'd)

FOREWORD

This publication is intended to present, in a simplified fashion, the many problems encountered when operating radio-relay systems. The material is limited to the frequency range extending from 44 to 13,000 megacycles, since the majority of radio-relay systems are designed to operate within this range.

The first four chapters deal exclusively with anomalous factors such as wave propagation, fading, and weather effects. Although the effect of these factors on the successful operation of radio-relay systems may be calculated for any given set of conditions, there is no easy way to predict, in advance, the amount by which such factors will vary. Therefore, it is necessary to engineer radio-relay systems to operate over a wide range of varying conditions.

A great portion of the basic material in this publication, then, is devoted to the problems of equipment siting, system reliability, and interference elimination. However, since the characteristics of each individual system component (transmitters, receivers, antennas, etc.) determine the overall operation of the system, these components are covered under separate chapter headings.

Additional information also presented includes (1) computations of wave propagation based on a series of nomograms (2) general ionospheric transmission problems, and (3) a section on testing and calibration, complete with a list of present standard test equipment.

This text is intended to cover only the basic problems of radio-relay system operation, to give a ready method for computing such systems, and to provide additional information where necessary.

CHAPTER 1

BASIC WAVE PROPAGATION

INTRODUCTION

The movement of a wave washing up on a beach or the disturbance caused by a stone when dropped into still water is easy to understand. In turn, a leaf floating on the pond where the stone was dropped will not be disturbed except for a slight bobbing motion. This shows that the wave travels through the water without moving very much of the water itself. Such movement of visible waves through a substance like water, or invisible waves through the atmosphere, is called wave propagation. The many kinds of invisible waves include sound, heat, and electromagnetic waves.

In this chapter we shall deal with electromagnetic wave propagation of frequencies ranging from 44 through 13,000 megacycles per second. It is primarily a chapter of definitions intended for use throughout the remainder of the book. A good knowledge of wave propagation is an important aid towards understanding the construction and use of radio-relay communications equipment.

MECHANICS OF WAVE PROPAGATION

WAVES IN GENERAL. Probably everyone has thrown a stone into quiet water and watched ever-widening circular waves spreading over the surface from the point where the stone entered the surface. The water does not move as a whole, but rather, only the surface pattern. In order to set up such a motion, the individual particles of water must move in transmitting the wave, but they move over rather short paths.

Waves can also be produced by vibrations other than those of material particles. For example, the current in an antenna circuit of a radio station produces an alternating field in the region around it. As the current oscillates, this field continually builds up and collapses, and in so doing creates electromagnetic waves which are propagated outward from the antenna. However, these waves are not transmitted by the motion of air particles, but by changes in the electric and magnetic conditions of space.

Wave motion in general can be classified as either longitudinal or transverse. A longitudinal wave is one in which the oscillating motion vibrates forward and backward, parallel to the direction in which the wave is propagated. Sound waves produced by a radio loud speaker as it vibrates in and out are longitudinal. The oscillating motion of a transverse wave, however, is always at right angles to the direction of propagation. Water waves are a good example of transverse waves, since the wave crests and troughs move up and down while the wave pattern as a whole moves outward from the source. Electromagnetic waves are also propagated as transverse waves, because the stress produced by such waves in space is perpendicular to the line of propagation. The two types of waves are illustrated diagrammatically in figure 1-1.

CHARACTERISTICS OF WAVE MOTION. In order to interpret the characteristics of wave motion, four basic wave measurements must be known and are listed as follows:

1. *The period of oscillation* which is the time necessary to complete one cycle,
2. *The frequency*, or the number of cycles completed in one second,

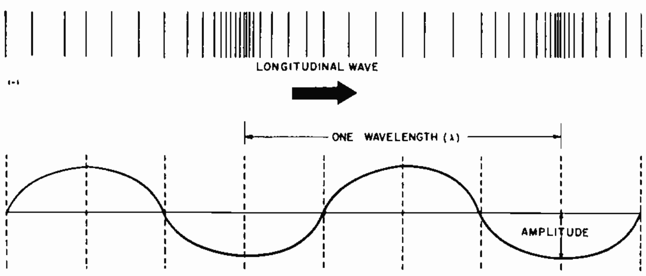


Figure 1-1. Types of Waves

3. *The amplitude of oscillation* which is the maximum displacement of the wave from its undisturbed or null position, and

4. *The wave length* which is the distance, measured along the direction of propagation, from any point on a wave to a corresponding point on the preceding or following wave (one cycle). The Greek letter lambda (λ) is used to signify wavelength.

Two waves moving simultaneously along the same line of propagation produce characteristics which are the direct result of their combination, although they advance independently. For the case where the two waves are oscillating at the same frequency, the characteristics are as follows:

1. An *in phase relationship* occurs when they continually pass through corresponding points of their paths at the same time. Otherwise, they are *out of phase.*

2. *Wave reinforcement* is produced when such waves are in phase. The combined action of both waves can be pictured by adding the coordinates of the component waves, point by point, as illustrated in figure 1-2A.

3. *Phase opposition* occurs when the two waves reach their maximum amplitudes in opposite directions at the same time.

4. *Destructive interference* is produced when such waves are in phase opposition, and further, if they have equal amplitudes, the result is a complete annulment, as illustrated in figure 1-2B. Destructive interference also occurs, although to a lesser degree, when the waves are only partially in phase opposition, or if they do not have equal amplitudes.

When two waves of a slightly different frequency are propagated in the same direction, they combine to produce a type of pulsating interference. The resultant amplitude is alternately large and small, producing pulses or throbs which are known as beats.

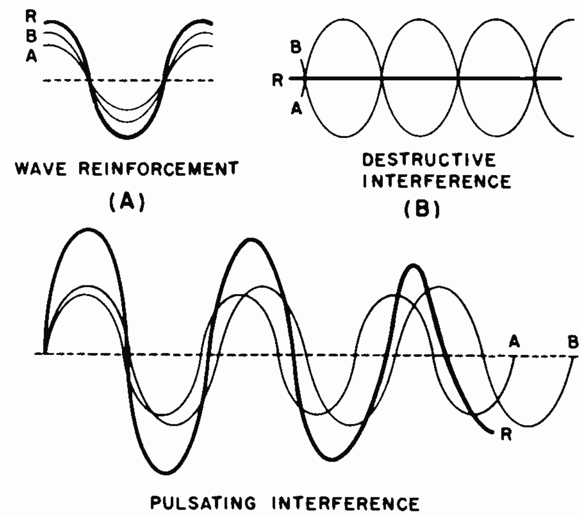


Figure 1-2. Interference of Waves

ELECTROMAGNETIC WAVES. It is customary to speak of radio waves traveling through the atmosphere or along a transmission line. Such waves are not simple in form like sound or heat waves but complex in nature. To start with the foundation underlying electromagnetic phenomena, we have the field theory as developed by James Clerk Maxwell. This theory states that an electromagnetic wave contains two fields: a moving electric field which creates a magnetic field, and a moving magnetic field which creates an electric field. The created field, at any instant, is in phase in time with its parent field, but is perpendicular to it in space. These laws hold true whether a conductor is present or not.

When an electric current flows in a single wire conductor, such as a simple antenna, a magnetic field, H, is set up around the antenna. Separated positive and negative charges also appear on the antenna, causing an electric field, E, to be set up. The magnetic lines of force form circles about the conductor, while the electric lines of force are normal to the conductor as illustrated in figure 1-3A. In this case, the sinusoidal variation in charge magnitude lags the sinusoidal variation in current by one-quarter cycle (90 degrees) and, therefore, the two fields must also be out of phase by 90 degrees. Thus, in spite of the fact that the two fields are perpendicular, they do not constitute the radiated electromagnetic field referred to above. This energy surrounding the conductor is called the *induction field* which, since it is a static field, cannot be detached from the antenna. Such energy travels along the surface, or skin, of the antenna, and its amplitude varies inversely as the square of the distance from the antenna. Consequently, its effect is local.

Now consider only the electric field set up about the antenna. As the polarity of the incoming voltage changes, the charges producing the electric field are constantly moving from one end of the antenna to the other. One end of the antenna may be at one instant, positive; an instant later, uncharged; and still later, negative. The antenna is then uncharged again and the whole cycle repeats itself.

In figure 1-3B, flux lines are drawn between positive and negative charges. An instant later, in figure 1-3C, the antenna is nearly discharged and the charges approach each other, bringing with them the two ends of the flux lines. When the charges merge, they seem to disappear, and most of the induction field flux does disappear. However, some flux is repelled by other lines nearer the antenna and, as shown in figure 1-3D, they become closed fields. Therefore, a closed electric field has been created without an associated electric charge.

At some instant after the independent field has been formed, the antenna charges again in the opposite direction, producing lines of force that repel the recently formed field. The repelling field is of the proper polarity to do this, as illustrated in figure 1-3E, and the radiated field is forced away from the antenna with a velocity equaling that of light.

As previously stated, a moving electric field generates a perpendicular magnetic field in phase with it. Since the independent electric field described above is moving, it therefore generates a magnetic field in accordance with this principle. The result is a *radiated electromagnetic field,* or wave, which may be propagated to great distances.

By similar reasoning, the magnetic lines of force may also become detached from the antenna and, as they move away, generate a perpendicular in-phase electric field. The result is also a radiated electromagnetic field, or wave. Therefore, the electromagnetic wave radiated from an antenna is made up of two components: the electric-generated field and the magnetic-generated field. The two fields, identical in composition but 90 degrees out of phase in time, will add vectorially and produce a single sinusoidally varying radiated field.

In figure 1-4, a segment of an electromagnetic wave is shown in three dimensions. Here, the electric field is represented by the vector E, and the magnetic field is represented by the vector H. The direction of propagation is along the vector Z. The direction of the electric field is called the direction of polarization of the wave. In the case of the electromagnetic wave in the figure, the electric field lines are in the vertical plane, and, therefore, the wave is *vertically polarized.* If the electric field lines were horizontal, then the wave would be *horizontally polarized.* As an electromagnetic wave moves forward, the electric field may rotate, thereby producing a circularly polarized field. Should little or no rotation exist, the field is *linearly polarized.*

The transfer of electromagnetic energy in a medium depends on certain electromagnetic properties of the medium as well as on the similar properties of the bounding media. Thus wave transfer in the atmosphere will depend in varying degree on the properties of the earth over which the transmission takes place. These properties are defined by the three parameters listed below.

1. The dielectric constant (κ) or the capacity of electrostatic energy a medium will store. A dielectric is a nonconducting material and, therefore, an insulator. Air, rubber, glass and mica are all good dielectrics. The dielectric constant for a vacuum is expressed by the symbol κ_0, and is equal to 8.854×10^{-12} farads per meter.

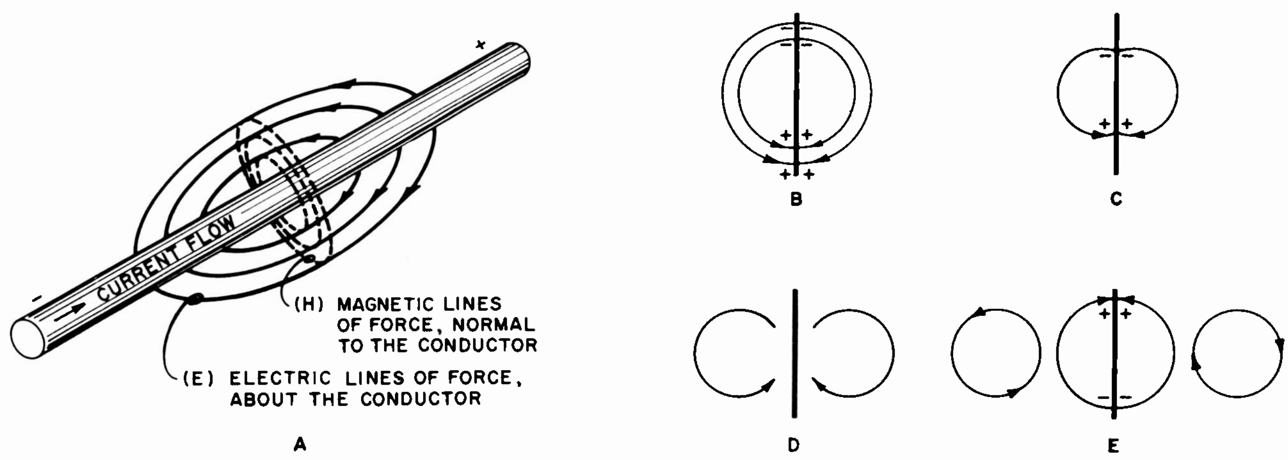
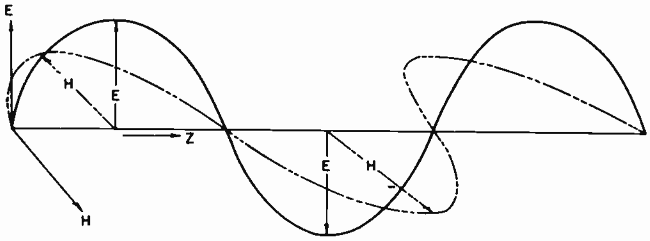


Figure 1-3. Induction Field



E - ELECTRIC FIELD
H - MAGNETIC FIELD
Z - DIRECTION OF PROPAGATION

Figure 1-4. Electromagnetic Wave

2. The permeability (μ) or the measure of the superiority of a material over a vacuum as a path for magnetic lines of force. Ferromagnetic materials such as iron, steel, nickel, cobalt, and magnetic alloys all have high permeabilities. On the other hand, diamagnetic substances such as copper, brass, and bismuth have permeabilities comparable to that of free space. The value of μ for a vacuum is equal to 1.257×10^{-6} Henrys per meter.

3. The conductivity (σ) or the measure of the ability of a material to pass electricity. All pure metals are conductors with some having more conductivity than others. Conductivity is the reciprocal of resistivity and is measured in mhos.

While σ is zero in a vacuum, κ and μ never are; in fact, the velocity (ν) of an electromagnetic wave in any medium is given by

$$\nu = \frac{1}{\sqrt{\mu\kappa}} \text{ meters per second} \qquad (1\text{-}1)$$

where

κ = The dielectric constant of the material through which the wave is being propagated (MKS Units).

μ = The permeability of the material (MKS Units).

Since the values of μ and κ are essentially the same for air as for a vacuum, it will be found that the velocity of electromagnetic waves through both materials is approximately 3×10^8 meters per second, or 186,000 miles per second. This value may be used for all electromagnetic waves through either air or vacuum without appreciable error.

WAVE FRONTS. The simplest antenna, or radiator, is a theoretical point source commonly called an *isotropic radiator*. In unobstructed space, such a source radiates equally well in all directions, and electromagnetic energy is propagated fundamentally with velocity of 186,000 miles per second. At the end of 1/1000 of a second, the field energy will be present a distance of 186 miles from the source. Similarly, at the end of one second, the wave would be the surface of a sphere with a radius of 186,000 miles. In each case, the spherical wave forms a front surface, or wave front, with the isotropic antenna located at the center of the sphere. According to Christian Huygens, a wave front may be considered to consist of an infinite number of isotropic radiators, each one sending out wavelets always away from the source. The collective effect of such action constitutes a new wave front as illustrated in figure 1-5.

The illustration shows an isotropic antenna from which is radiated the primary wave front, arc AG. Along this wave front, points A through G may be considered as an infinite number of isotropic radiators, each oscillating in phase. These radiators send out little wavelets of their own, such as shown coming from point E on the wave front, which combine with the wavelets sent out by all the other points on the primary wave front to form a new wave front at some distance further along the direction of propagation.

As such a wave front advances, a small section of its face would appear to be a plane, and both the electric field intensity (E), and the magnetic field intensity (H) would be vectors perpendicular to each other and located in the plane wave front as illustrated in figure 1-6. The direction of wave propagation would be perpendicular to the wave front.

REFLECTION AND REFRACTION. By studying the changes in a wave front which occur as a wave advances through a medium of one density, it is possible to predict the effects that will be produced when a wave encounters a medium of a different density which either reflects, refracts, or absorbs energy.

The behavior of a wave upon striking a reflecting surface may be determined by an adaptation of Huygens' construction. In figure 1-7, an electromagnetic wave is shown being reflected from the ground somewhere between a transmitting and receiving antenna. A close-up drawing of the wave front at the point of reflection shows wave front AB impinging upon the surface of the ground, through which it cannot penetrate. If the earth's surface had been absent, the wave would have advanced without change in direction and, in a certain time interval, would have reached the position A'B. However, the presence of the earth's surface causes a change in the direction of the wave front as illustrated by the heavy

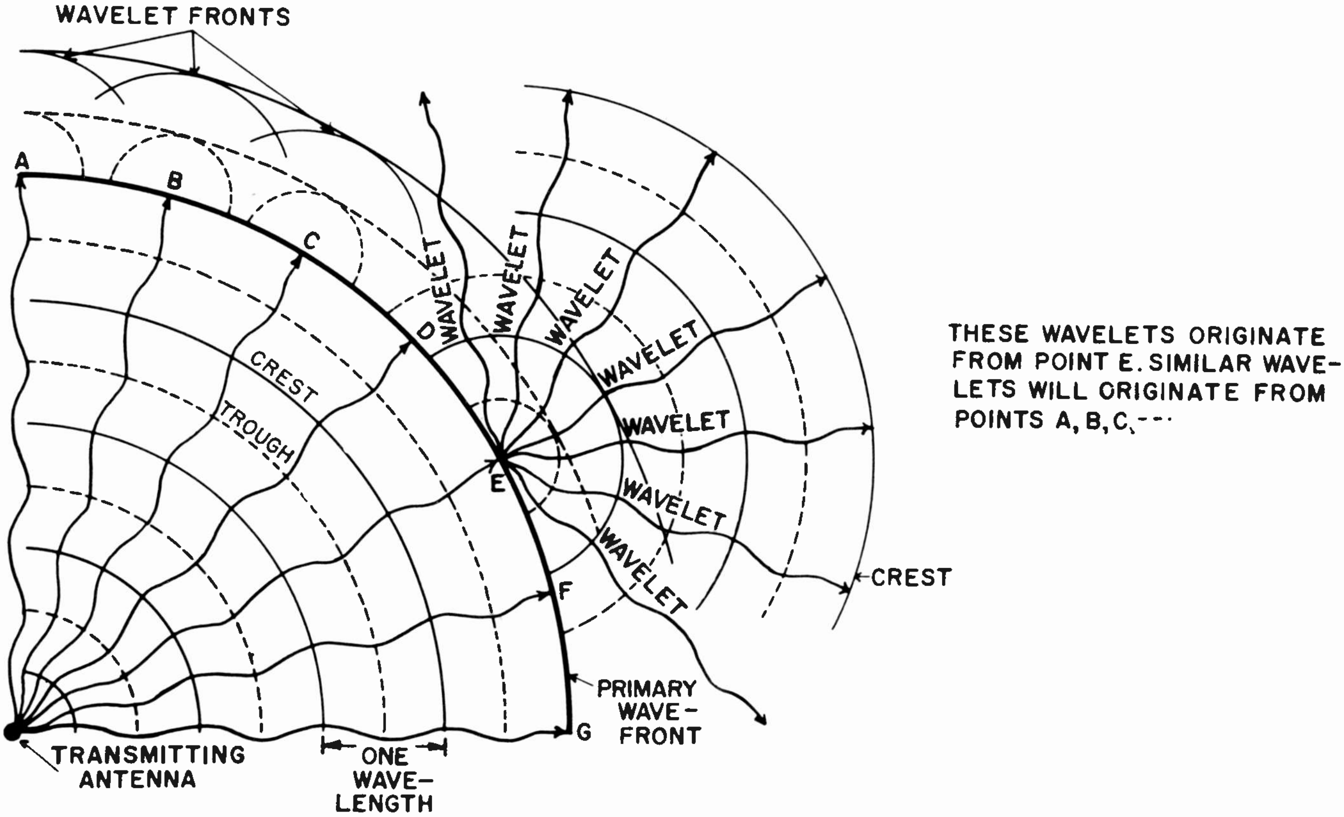
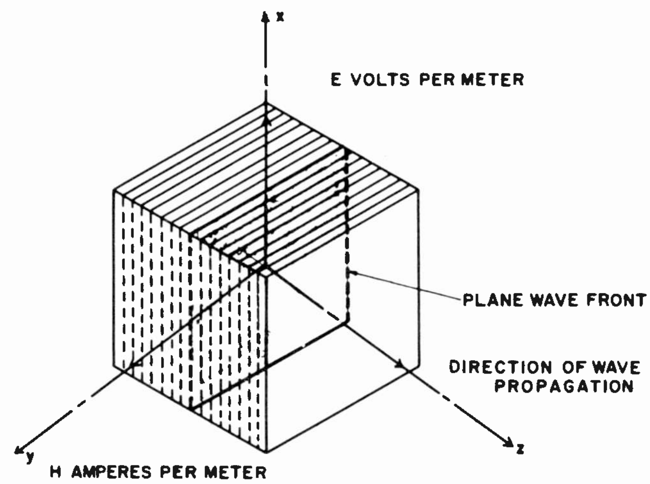


Figure 1-5. Huygen's Wavelets

line AOB. The line OB represents the incident wave front, and the line AO represents the reflected wave front. The angle i, which is the angle of incidence, and the angle r, which is the angle of reflection, are equal and lie in the same plane.

Inasmuch as we deal with the transfer of electromagnetic energy in the case of wave propagation, it is convenient to express the energy relations of the incident and reflected waves by a ratio called the coefficient of reflection. The *coefficient of reflection* (ρ) is defined as the square root of a power ratio, and is found by dividing the reflected energy per second leaving a reflecting surface by the energy per second incident to the same surface. If the two energies are of equal magnitude ($\rho = 1$), we have perfect reflection. If the reflected energy is smaller than the incident energy, the difference is either dissipated at the reflecting surface, or partially dissipated at the surface and partially passed through the surface as a refracted ray. For example, when an electromagnetic wave impinges upon a cloud, most of the energy will be transmitted through the cloud. However, due to various particles in the cloud, a portion of the wave will be returned by reflection at the surface of the cloud and another portion will be absorbed within the cloud itself. When the wave energy is absorbed, it is converted into heat. The part of the wave that is passed through the cloud will be refracted, or changed in direction, if the electromagnetic properties of the cloud differ from those of the surrounding air. In fact, when a radio wave encounters any medium whose electromagnetic properties differ from those of the previous medium,



ONE CUBIC METER OF THE GENERAL WAVE FRONT

Figure 1-6. Electric & Magnetic Vectors

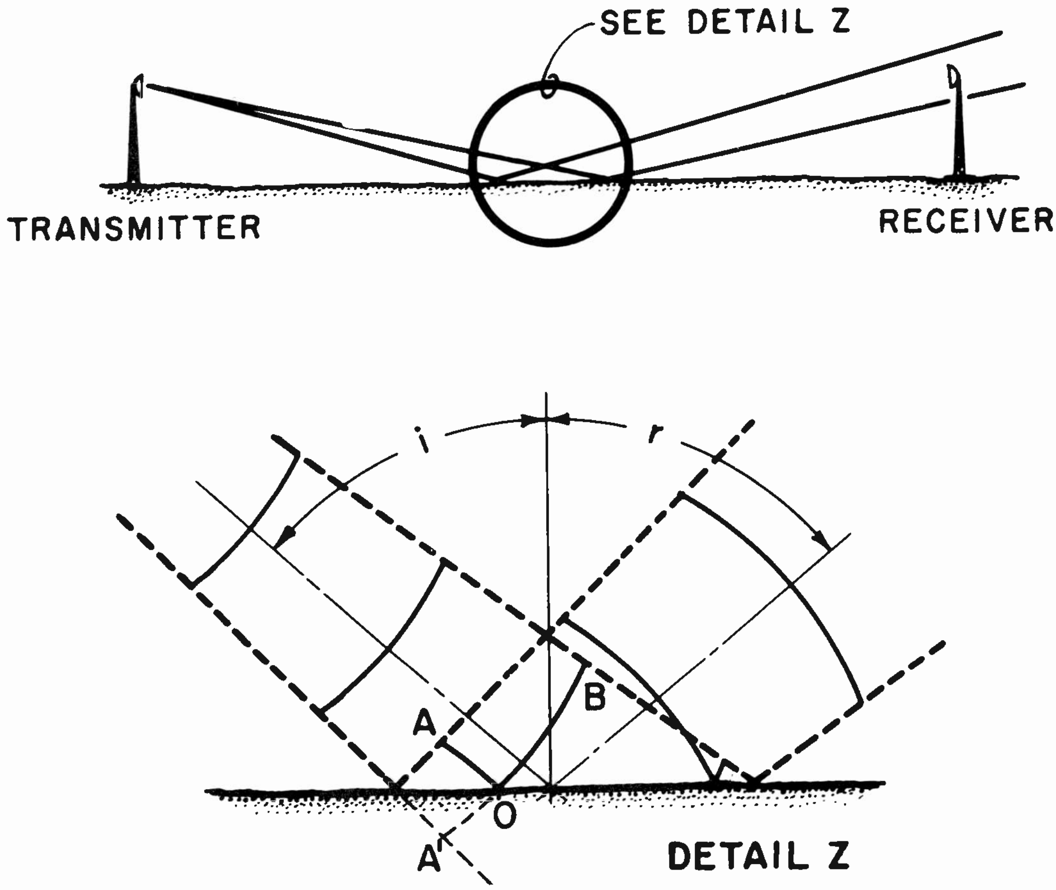


Figure 1-7. Reflected Wave

reflection and refraction take place simultaneously.

In figure 1-8 an electromagnetic wave is illustrated being refracted through a mass of air which is *denser than* the air surrounding it. For the sake of simplicity, all the action is shown taking place at the interface between the air mass and the surrounding atmosphere. Actually, however, the refraction occurs in gradual steps, since air masses have no sharply defined interface. Since the mass of air is denser than the surrounding air, the wave will be slowed down as it enters and therefore bent. If the air mass had not been present, point B on the wave front AOB would have advanced to B′ at a uniform rate of velocity. However, since the wave velocity is less in the air mass than in the surrounding air, the line OB represents the new direction on the wave front, and the angle r′ which it makes with the normal is called the angle of refraction.

If the density of the air mass was *less than* the surrounding air, the wave velocity would have increased and the wave refracted upward. Therefore, a wave is refracted toward the normal when its velocity is reduced, and away from the normal when its velocity is increased. Consequently, the law of refraction states than an incident wave, traveling obliquely from one medium to another, will undergo a change in direction if the velocity of the wave in one medium is different from that in the other. It was discovered by Willebrod Snell, Dutch astronomer and mathematician, that the ratio of the sine of the angle of incidence to the sine of the angle of refraction is the same as the ratio of the respective wave velocities in these media, and is a constant for two particular media. Expressed as a formula, Snell's law becomes

$$\eta = \frac{\text{Sin i}}{\text{Sin r}'} = \frac{\nu_1}{\nu_2} \tag{1-2}$$

where η is called the refractive index of the second medium relative to the first. The absolute refractive index of a substance is its index with respect to a vacuum, and is practically the same value as the index with respect to air. Thus, if the velocity of a wave is ν in a particular substance, ν_o in a vacuum, and ν_a in air, the refractive index of the substance is given by the formula

$$\eta = \frac{\nu_o}{\nu} = \text{approximately } \frac{\nu_a}{\nu} \tag{1-3}$$

where the symbol η represents the index of refraction.

It is the change in the index of refraction that determines the path of electromagnetic waves through the atmosphere, as will be discussed later in this chapter.

DIFFRACTION. The amount of r-f energy in the form of electromagnetic waves that will travel from a transmitting antenna to a receiving antenna is determined by the path over which the waves must go. It was often thought that high frequency waves traveled according to geometric optics (geometric line of sight) or slightly below this line if atmospheric refraction existed. However, experiments have shown that, at frequencies below approximately 100 mc, waves may reach much farther beyond the horizon than refraction can account for.

Wave reflections do not originate at one point only, but, according to Huygens' wavelet principle, from the entire surface of an obstacle which is in the path of electromagnetic waves. Wavelets, therefore, will re-radiate in all directions from a multitude of elementary radiation centers at the earth's horizon as they receive incident wave energy. Electromagnetic diffraction is, therefore, the bending of waves as they graze the earth's surface or any other intervening obstacle, and is caused by Huygens' wavelets. Diffraction of electromagnetic waves into the geometrical shadow region behind a mountain ridge is illustrated in figure 1-9.

FRESNEL WAVE INTERFERENCE. As mentioned previously, each wavefront progressing from a transmitting source to a receiving point consists of an infinite number of secondary sources. Therefore, even in the simple case of transmitting energy from point T to point R in free space (figure 1-10) there are an infinite number of paths to consider, each path originating from a secondary source on the progressing wavefront. In figure 1-10, the wavefront described by the arc AG is a particular part of the beam of energy that is being sent from a transmitting antenna. The points A through G on the wavefront are designated as secondary sources, otherwise known as Huygens' radiation centers. The radii of the circles whose diameters are described by the point AG, BF,

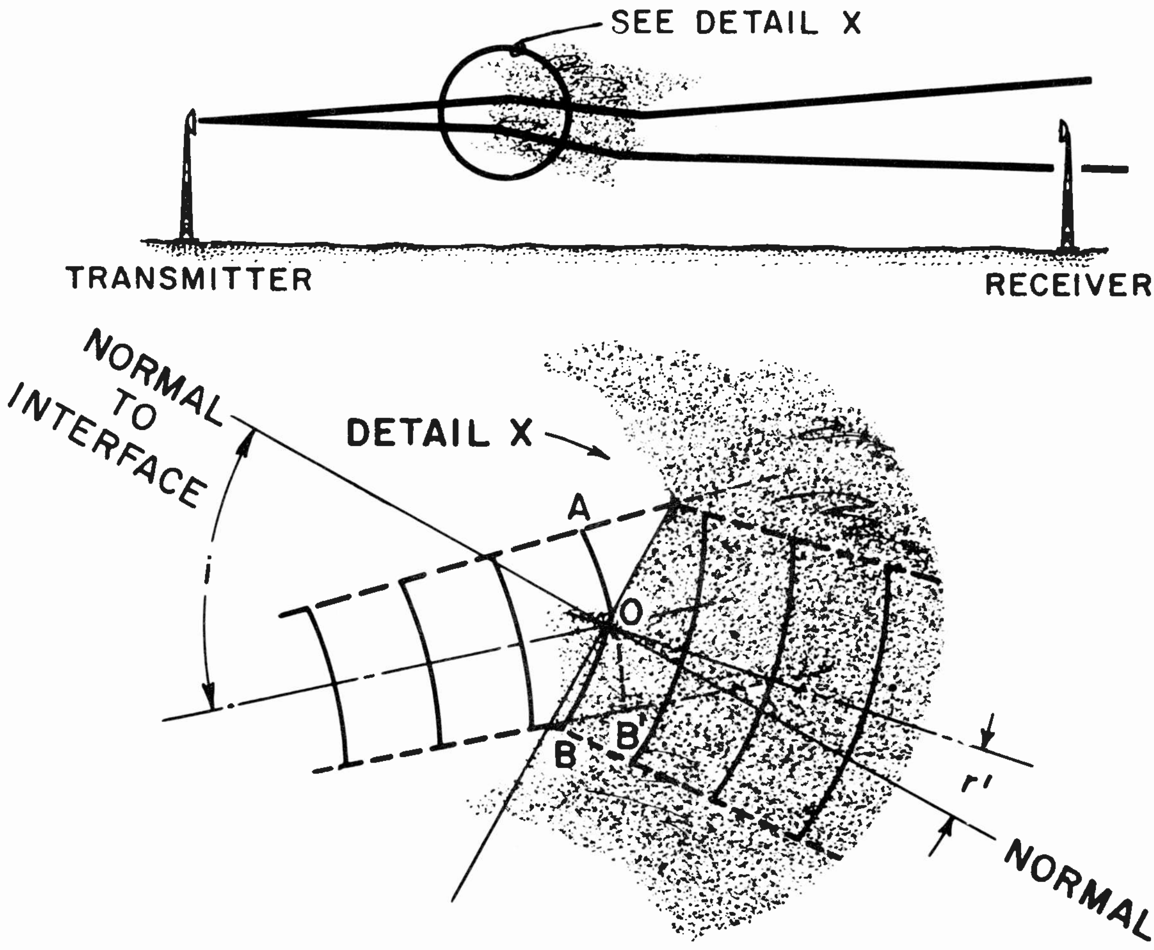


Figure 1-8. Refracted Wave

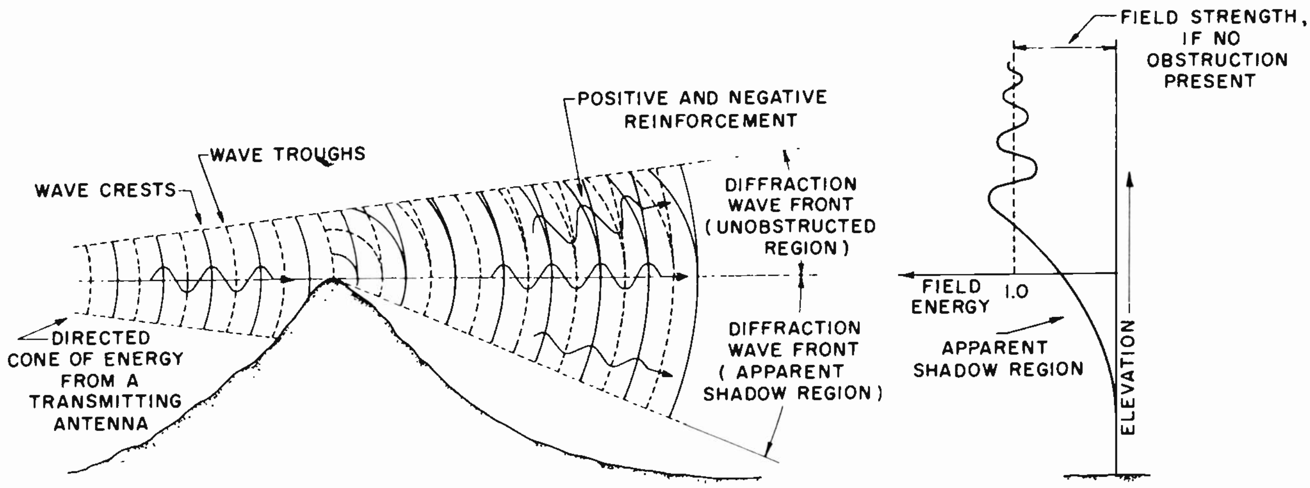


Figure 1-9. Diffraction

and CE are so chosen that the total path length from T to R via each circle is n $\lambda/2$ greater than the shortest path TDR, where n is an integer. Therefore, the distance ER is longer by a half wavelength than the distance DR, and the distance FR is longer than the distance ER by a half wavelength. The circular regions formed by these radii are called *Fresnel zones,* and are not equal, but decrease in energy in proportion to the distance from the central or first Fresnel zone. The field contributed by each zone at point R is alternately positive and negative, as determined by the difference in path lengths of the secondary waves. The total field received at point R is, therefore, the vector sum of all the incident fields. For example, consider the wave arriving at point R from the secondary source C on the wavefront. Since the path length CR is one-half wavelength longer than the direct path DR, the wave arriving from point C will be out of phase with the direct wave and its contribution to the total received field will be negative. Conversely, a wave arriving at point R from point B on the wavefront will be in phase with the direct wave and its contribution to the total received field will be positive, since the path BR is longer by a full wavelength than the path DR. Therefore, Fresnel zones of odd numbers are zones whose contribution to the total received field is positive, while Fresnel zones of even numbers are zones whose contribution to the total received field is negative. Actually, there are an unlimited number of Fresnel zones on a wavefront, but, for simplicity, the number is limited to three in the diagram.

The region described as the first Fresnel zone accounts for approximately one-quarter of the total received field energy, and, if possible, should be clear of all path obstructions if the received signal is to have a maximum signal strength. It must also be remembered that the area of the first Fresnel zone depends upon both the transmitting and receiving antennas and upon the operating frequency. The radius of the first Fresnel zone at any point (P) along the transmission path may be found by the expression

$$R = 13.15\sqrt{\lambda \frac{D_1 \times D_2}{D_3}} \qquad (1\text{-}4)$$

where

- R = The first Fresnel zone radius in feet at any point P.
- D_1 = The distance in miles from the transmitting antenna to any point P.
- D_2 = The distance in miles from the receiving antenna to any point P.
- D_3 = $D_1 + D_2$.
- λ = Wavelength in centimeters.

Optimum Fresnel zone clearance actually occurs when the edge of an obstruction is located a slight amount within the first Fresnel zone. This slight penetration into the first Fresnel zone causes wave reinforcement to occur since the reflected wave is one-half wavelength longer than the direct wave and also reversed in phase by 180° at the reflection point. The reflected wave is therefore in phase with the direct wave, thus increasing the effective signal strength at the receiving antenna. Optimum clearance is illustrated in figure 1-11A. When there is more than optimum clearance, such as illustrated in figure 1-11B, multipath fading may occur, caused by impinging waves from the 2nd or 4th Fresnel zones reflecting from the surface of the obstacle. Less than optimum clearance, as illustrated in figure 1-11C, may cause a possible shadow loss of power, depending upon the degree of obstruction, due to multiple re-

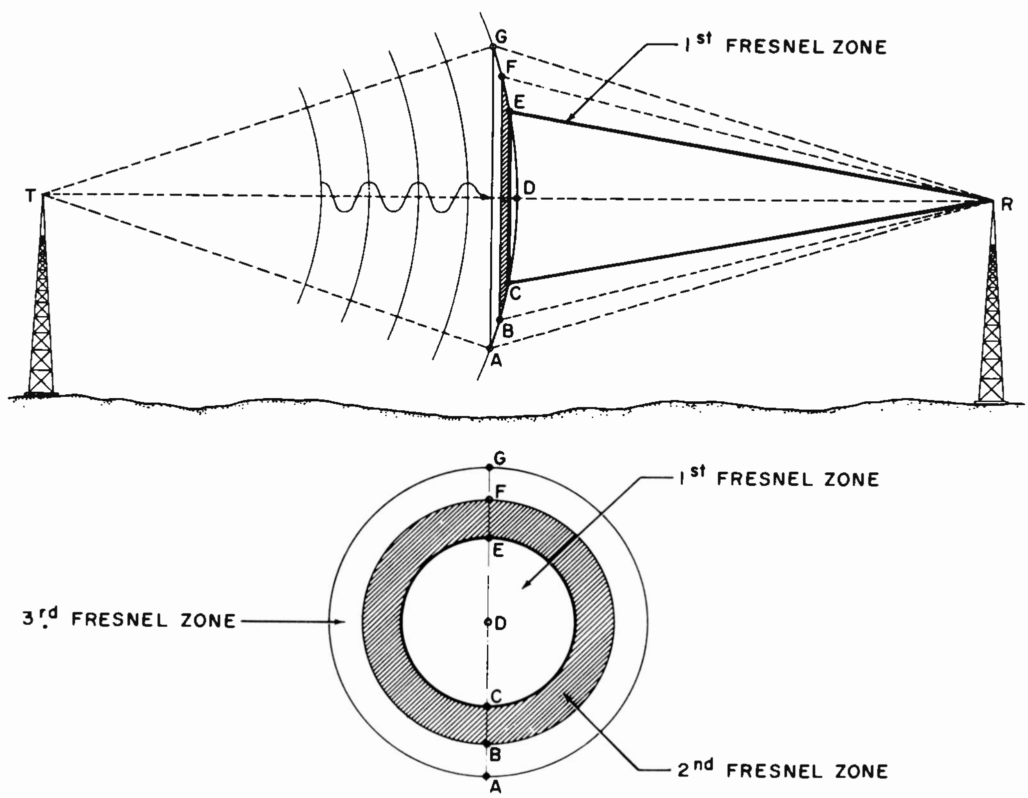


Figure 1-10. Fresnel Zone

flections from the obstructed side of the first Fresnel zone.

EARTH REFLECTION ZONES. When a radio wave is incident upon the earth's surface it is not actually reflected from a point on the surface, but from a sizeable area. This reflection area may be large enough to include several Fresnel zones, or it may be in the form of a ridge or peak including only a part of the first Fresnel zone. For the case where the wave is incident upon a plane surface, the resulting Fresnel zones formed on the reflecting surface take the form illustrated in figure 1-12A. The elliptical zones formed on the reflecting surface are similar to those which would be formed on an oblique plane placed between a transmitting source and a receiver in free space as shown in figure 1-12B. Therefore, the earth-reflected Fresnel zones are simply a projected image of the free-space Fresnel zones at the plane of reflection and may be determined by the same geometry as that used for the free-space Fresnel zones.

The significance of the ground reflected Fresnel zones is similar to that mentioned for the free-space zones. However, as radio waves are reflected from the earth's surface, they are generally changed in phase depending upon the wave polarization and the angle of incidence. For horizontally polarized waves at the range of frequencies encountered in this book, such waves reflected from the earth's surface are shifted in phase by about 180°, effectively changing the electrical path length by a half wavelength. On the other hand, for vertically polarized waves, a considerable variation in the phase angle will be found to exist for different angles of incidence and reflection coefficients, and will vary between 0° and 180° lag depending upon ground conditions. For horizontally polarized waves, therefore, (and in some cases for vertically polarized waves) if the area of the reflecting surface is large enough to include the total area of any odd-numbered Fresnel zones, the resulting wave reflections will arrive out of phase at the receiving antenna with the direct wave and cause destructive interference.

The radiation pattern of an antenna near a reflecting surface such as the ground differs from the free-space pattern primarily because of the existence of ground reflections. Since the direct path and the reflected path will not be of the same physical length, and since there will be a phase change upon reflection, the two waves may arrive at the receiving antenna with any phase relationship. This phase relationship of the two waves will cause either an increase or decrease in the signal strength at the

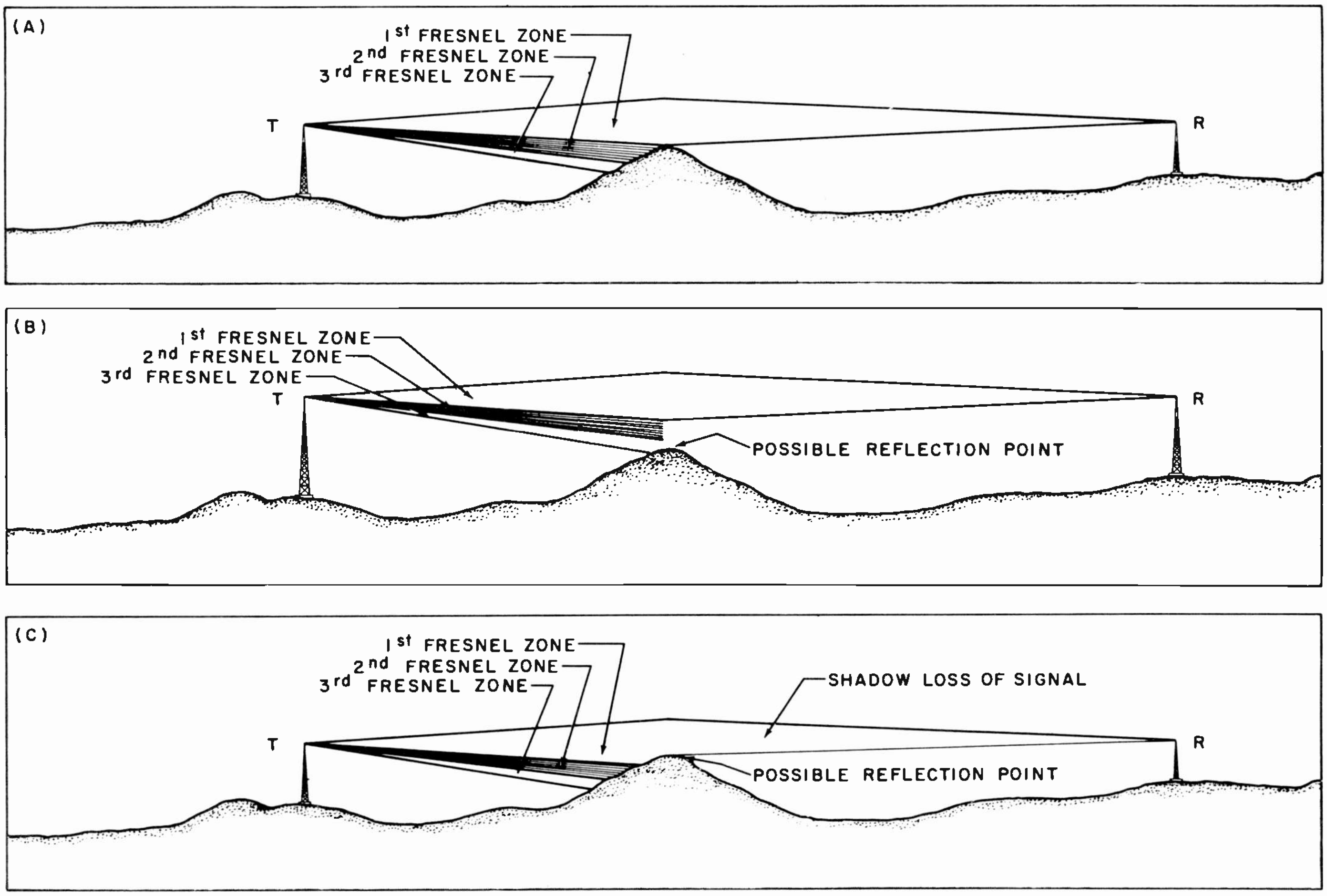


Figure 1-11. Fresnel Zone Clearances

receiver and will produce the effect of distinct lobes and nulls in the radiation pattern since the two rays add vectorially at the receiver.

PROPAGATION CHARACTERISTICS

FREE-SPACE TRANSMISSION. According to the IRE, free-space transmission is electromagnetic radiation over a straight line path in a vacuum or in an ideal atmosphere, sufficiently removed from all objects that might affect the wave in any way. In this case, only the direct wave propagated from a transmitting antenna is effective at the receiving antenna, and the electric field intensity present at any point in space is commonly called the free-space field strength, E_0. The field strength obtained in free space is a convenient standard for reference. Signal fluctuations of high-frequency waves as propagated through the lower atmosphere may, on occasion, rise above the free-space value, but the average field strength of such fluctuations will always be somewhere below the free-space value.

Free-space field strength depends only upon the amount of power transmitted and the distance over which the wave is propagated, and may be found by the formula

$$P_r = \frac{G_1 G_2 \lambda^2 P_t}{(4\pi)^2 d^2} \qquad (1\text{-}5)$$

where

P_r = Power in watts at the terminals of the receiving antenna

P_t = Power in watts at the terminals of the transmitting antenna

G_1, G_2 = Power gains of the transmitting and receiving antennas, respectively

d = Distance between antennas in meters

λ = Wavelength in meters

In the above formula, the transmitting and receiving antennas are assumed to be properly oriented for maximum transfer, and the receiving antenna is assumed to be properly matched to its receiver.

In order to determine the direction of maximum radiation or maximum response, a radiation pattern may be used. A *radiation pattern* of a transmitting antenna is a diagram indicating the intensity, in all directions, of the radiated field as it would occur under actual operating conditions. In the case of a receiving antenna, it indicates the response of the antenna to a signal of uniform intensity arriving from all directions. Such patterns may be plotted in terms of field strength in volts per meter or in terms of power in watts per square meter. Figure 1-13 illustrates a typical power radiation pattern for a linear antenna whose physical length equals approximately one-half wavelength of the frequency used. This type of pattern gives the relative directivities based on a percentage of power, but not the absolute magnitude of the field.

The radiation patterns of a transmitting antenna and a receiving antenna with the same physical characteristics are the same. Therefore, for example, at a given distance, with a transmitting antenna height of 100 feet, and a receiving antenna height of 50 feet, the received field would be the same even though the transmitting and receiving units were interchanged. This fact is called the *law of reciprocity.*

CONDUCTIVE AND DISSIPATIVE WAVE MEDIA. In a conducting wave medium such as copper wire, we have found that electromagnetic field energy is absorbed because of the internal resistance of the wire itself. The energy that is absorbed is converted into heat. When the medium surrounding an electromagnetic current-carrying conductor is not a perfect dielectric but exhibits a certain conductivity, heat losses will also occur in the medium. Therefore, we have, in addition to the induction field and the propagated field about a conductor, radiated heat generated in the electromagnetic field.

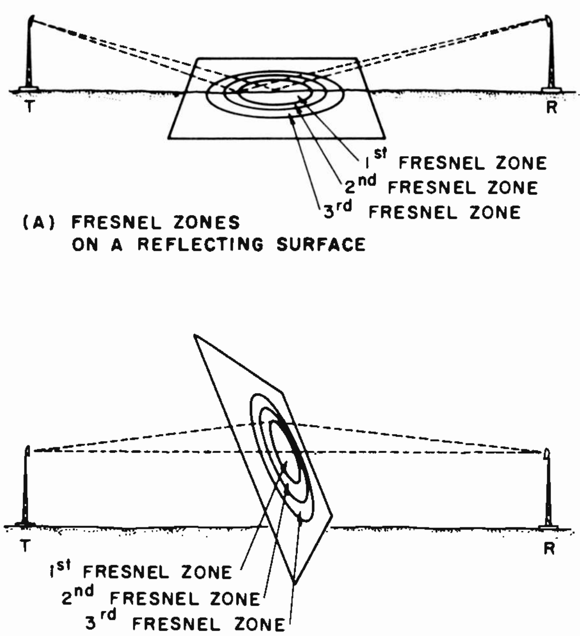


Figure 1-12. Earth-Reflected Fresnel Zone

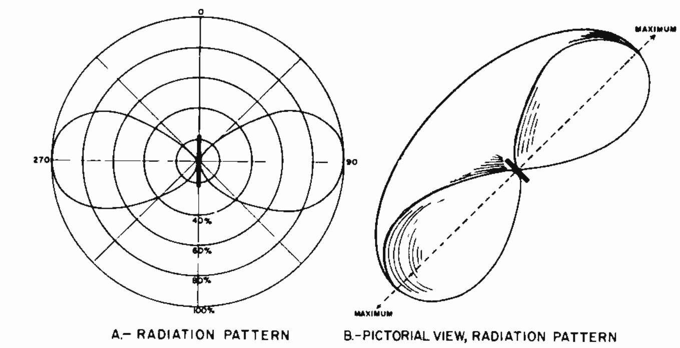


Figure 1-13. Radiation Pattern

A medium that is not a perfect dielectric is termed a *dissipative medium.* As an electromagnetic wave is propagated through a dissipative medium, heat losses will occur, and account for the attenuation of the wave's energy. In the earth's atmosphere, at frequencies above approximately 10,000 mc, precipitation and water vapor are two factors that can cause significant attenuation of energy by absorption.

CHARACTERISTIC IMPEDANCE. The value of the total opposition that transmission media offer to the flow of electromagnetic field energy is called the *characteristic impedance* (Z_o) and is measured in ohms. The characteristic impedance of free space is represented by the formula

$$Z_o = \sqrt{\mu_o/\kappa_o} = 377 \text{ ohms} \tag{1-6}$$

where

μ_o = the permeability of a vacuum

κ_o = the dielectric constant of a vacuum

The characteristic impedance of free space also expresses the ratio E/H of the associated electric and magnetic fields. Since

$$Z_o = \frac{E}{H} = 377 \text{ ohms} \tag{1-7}$$

where

E = electric field

H = magnetic field

for electromagnetic waves in unobstructed normal atmosphere, the electric field in volts per meter at any space point is numerically 377 times larger than the numerical value of the associated magnetic field in amperes per meter at the same space point.

Q-VALUE. The figure of merit (Q value), which must be a large numerical value for an efficient wave medium, is given by the ratio

$$Q = \frac{\omega\kappa}{\sigma} \qquad (1\text{-}8)$$

where

ω = $2\pi f$ with f in cycles or $6.28 \times 10^6 f$ with f in megacycles

κ = the dielectric constant of the medium

σ = the conductivity of the medium

Since the conductivity of dry air is zero, it will be seen from the formula that the Q-value of the atmosphere is infinite, and therefore waves of any frequency may be propagated. The numerator ($\omega\kappa$) is a measure of the *displacement current* or the current at right angles to the direction of propagation expressing the rate of expansion and contraction at which the field energy changes. The denominator (σ) is a measure of the *conduction current* or the current in *watts*, moving in the direction of propagation, that may or may not be absorbed because of the heat losses. The conduction current may be measured in amperes per square meter by using the formula

$$\text{Conduction current (J)} = \text{E amp/m}^2. \qquad (1\text{-}9)$$

In general wave media, conduction and displacement currents exist by virtue of an electromagnetic field. For a pure dielectric medium only displacement currents are present, but when a wave strikes the earth's surface or other obstacle, such obstacles exhibit the characteristics of a dissipative medium. *To what extent the obstacle is a dielectric or a reflector depends upon the operating frequency.*

PROPAGATION CONSTANT. In free-space as well as in all other media, there is a propagation characteristic which defines both the wavelength and the dissipation of electromagnetic field energy. This characteristic is called the *propagation constant,* $\Gamma = \alpha + j\beta$, and expresses what happens because of energy transfer. The part of the propagation constant which defines the wavelength is known as the *phase constant* (β) and is by no means fixed, even though the frequency may be perfectly stabilized. The wavelength of any given frequency will vary, depending upon the media through which the wave is propagated. The part of the propagation constant that defines the dissipation is generally known as the *attenuation constant* (α) and will explain the losses that account for the amplitude decay of both the electric and magnetic field vectors. The propagation constant is complex in nature and is included in this chapter only for the purpose of definition.

THE POYNTING THEOREM. We have established some of the components that offer opposition to the flow of electromagnetic field energy in various media. Now we must establish the rate of flow of the energy through these media. The *Poynting Theorem* states that the rate of flow of electromagnetic field energy is at any instant proportional to the magnitude of the resultant vector product of the electric and magnetic field intensities.

A *vector* is a quantity having both magnitude and direction. A current, for instance, is a vector, since it has a magnitude expressed in so many amperes and a definite direction of flow. The electric field intensity (E) is a vector, since it has a magnitude of so many volts per meter in a certain space direction, just as the magnetic field intensity (H) is a vector, since it has a magnitude of so many amperes per meter in a certain space direction. Volts, degrees centigrade, or energy in watts are known as *scalar magnitudes*, since they are completely characterized by magnitude alone.

The vector product of the Poynting Theorem is termed the *Poynting vector*, whose magnitude in watts per square meter gives the amount of total field power transmitted. Since it is a vector, it must also give the direction in which the power flow occurs. In figure 1-14 is illustrated the effective electric field (E) in volts per meter propagated along the vertical (x) direction (vertical polarization) and the associated effective magnetic field (H) in amperes per meter propagated along the horizontal (y) direction. The Poynting power flow, p = EH watts per square meter, is then along the direction of propagation (z). The quantity p, therefore, represents the absolute amount of the Poynting vector and is the amount of energy which crosses an area of one square meter each second, perpendicular to the direction of power flow. The formula

$$P = E^2/Z_o = E^2/377 \qquad (1\text{-}10)$$

represents the energy density per square meter at any point in space between transmitting and receiving antennas. Similarly, the formula

$$P = 1/2\ (\mu H^2 + \lambda E^2) \qquad (1\text{-}11)$$

represents the energy density per *cubic* meter of field energy.

ATMOSPHERIC TRANSMISSION. The electromagnetic wave intercepted by a receiving antenna may have been propagated in the form of three different waves, or any combination thereof. They

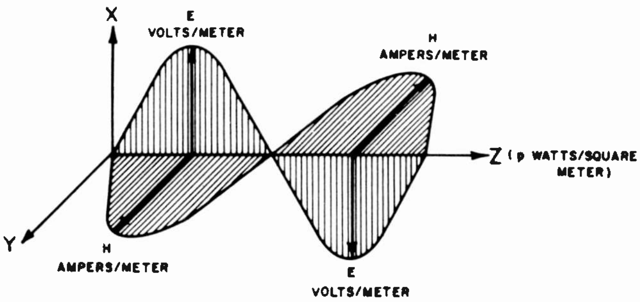


Figure 1-14. Poynting Power Flow

are: the *ground wave*, which is propagated over the earth; the *ionospheric wave* or sky wave, which is propagated by way of the ionosphere; and the *tropospheric wave* which is propagated by reflection from a place of abrupt change in the dielectric constant or its gradient in the troposphere.

The ground wave includes all components of a radio wave over the earth except ionospheric and tropospheric waves. Better transmission is nearly always obtained over water than over land, and over soil of high conductivity rather than soil of poor conductivity. The ground wave may also be refracted because of variations in the dielectric constant (which determines the index of refraction) of the troposphere including the condition known as a surface duct. At frequencies below about 3 mc, ground wave propagation up to 200 miles may be obtained because of the diffraction of the wave below the radio horizon. At higher frequencies, diffraction effects become less prevalent than refraction effects in regard to the propagation of the ground wave beyond the earth's horizon.

The ionospheric wave may be employed for communication over greater distances since it can be reflected from the ionosphere, thus returning to the earth at points remote from the transmitting source. Generally the higher the frequency, the less the wave is reflected, until at frequencies above 100 mc only an extremely small portion of the signal is ever reflected from the ionosphere. Ionospheric return for such frequencies does not occur often and when it does it is erratic in nature (see chapter 15). The frequencies above 100 mc are, however, affected by what is known as the troposphere.

THE TROPOSPHERE. The troposphere extends upward from the earth's surface to about 6 miles, and contains almost all the atmospheric conditions and changes known as "weather". Ionized layers of air occur rarely in the upper part of the troposphere, and not at all in the lower half. For this reason, waves returned from the troposphere are almost always the result of changes in the atmospheric characteristics, rather than reflections from ionized layers. Where reflections do occur from these ionized layers, there is little effect on propagation since the reflecting ability of the layers is very low. Another source of reflected wave from the troposphere is an air mass differing considerably in its dielectric constant from the surrounding air. Electromagnetic waves reflected from such a mass are called tropospheric waves, and may seriously interfere with radio propagation. Waves traveling through the lower part of the troposphere (ground waves) on the other hand are not reflected, but curved. The curvature is caused by the gradual change in the refractive index of air that accompanies an increase in elevation.

The refractive index of air depends on the temperature, atmospheric pressure, and water vapor pressure of the air, and ordinarily decreases uniformly with elevation. As stated above, this results in a curving of the radio beam, usually downward. This curving follows the law of refraction which states that a beam passing from a lighter to a denser medium is refracted toward the normal, or perpendicular to the surface between the two media. In figure 1-15A, this is illustrated between the transmitting antenna and the midpoint. At the midpoint, however, the beam commences to travel from the denser to the lighter medium. The beam continues to bend downward, since, according to the law of refraction a beam traveling from a denser to a lighter medium is refracted away from the normal. This condition is shown in the figure between the midpoint and the receiving antenna, and the whole path, therefore, is curved downward.

In figure 1-15A, both the earth's surface and the beam are curved. Any calculations based on such curves may be simplified by multiplying the earth's radius by some constant that will cause the radio beam to appear as a straight line (figure 1-15B).

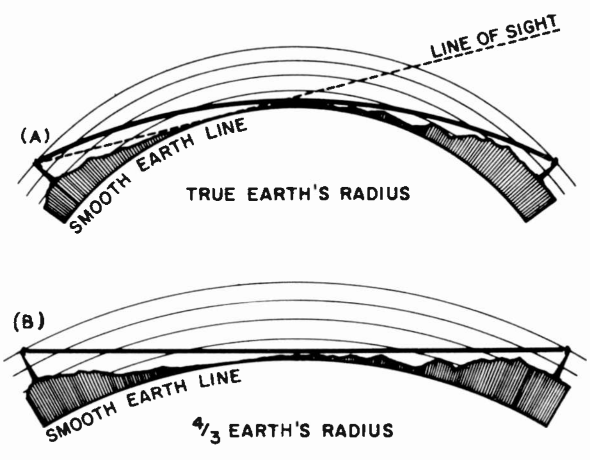


Figure 1-15. Refraction through Standard Atmosphere

This constant, "k", is dependent upon the rate of change of the refractive index of air with altitude, which in turn is dependent upon weather conditions. The symbol k may vary between 0.8 and 3.0, with a value of 4/3 generally in agreement with conditions normally found to exist in the atmosphere. The actual path over which the radio waves are propagated is called the radio path, and the radio horizon is the point at which such a path is tangent to the earth. Radio horizons and path attenuation values based on 4/3 times the true earth's radius are found to be in agreement with values determined by field strength measurement, and most of the formulas and nomograms used in this book are based on the 4/3 earth radius, rather than true earth radius.

CHAPTER 2

ATTENUATION

INTRODUCTION

ATTENUATION may be defined as a decrease in the intensity of energy, such as the energy of electrc-magnetic waves. There are a great many factors which cause attenuation. In the case of radio waves propagated through the atmosphere, the fundamental loss is the spreading of energy.

When installing radio-relay stations, the distance between stations is of great importance. The attenuation of the transmitted signal will largely determine the maximum range obtainable for reliable reception of the signal. Attenuation will occur in the transmission lines from the transmitter to the transmitting antenna and from the receiving antenna to the receiver, but most of the signal attenuation will be found in the radio path between antennas. Figure 2-1 illustrates the attenuation of electromagnetic field energy from the transmitter output circuit to the receiver input circuit.

PATH ATTENUATION

FREE SPACE. The light radiated from an ordinary light bulb is a good example of path attenuation. A 60-watt bulb may supply sufficient light for a small room but not for a large room. In both cases, an equal amount of light is supplied (60 watts), but in the larger room the light is distributed over so much more space that the amount of light per unit volume becomes relatively small.

A point source (isotropic antenna) in free space theoretically radiates electromagnetic field energy equally well in all directions in much the same way as a light bulb radiates light. The wave so radiated spreads out in a spherical form and is called a spherical wavefront. As the spherical wavefront expands in free space, no energy is actually lost, but only spread over a larger spherical surface. Although there is no decrease in the total amount of radiated energy, there is a decrease in the amount of energy *per square meter* of the wavefront. This decrease in energy between transmitting and receiving antenna terminals is called *path attenuation* and is generally expressed in decibels. The *decibel* expresses the magnitude of a change in signal level, and is equal to the common logarithm of the ratio of power level transmitted over power level received. The general attenuation formula

$$\text{Attenuation in db} = 10 \log_{10} \frac{P_t}{P_r} \qquad (2\text{-}1)$$

where

P_t = Power in watts at the terminals of the transmitting antenna

P_r = Power in watts at the receiving antenna terminals

may be used for any situation, regardless of frequency, to determine path attenuation between transmitting and receiving antennas when the effective power at both antennas is known.

ATMOSPHERIC. When an electromagnetic wave is propagated through the atmosphere with a certain Poynting power flow (p) watts per square meter, the available power at the output terminals of a receiving antenna (P_r) is dependent upon the antenna's effective area. The formula

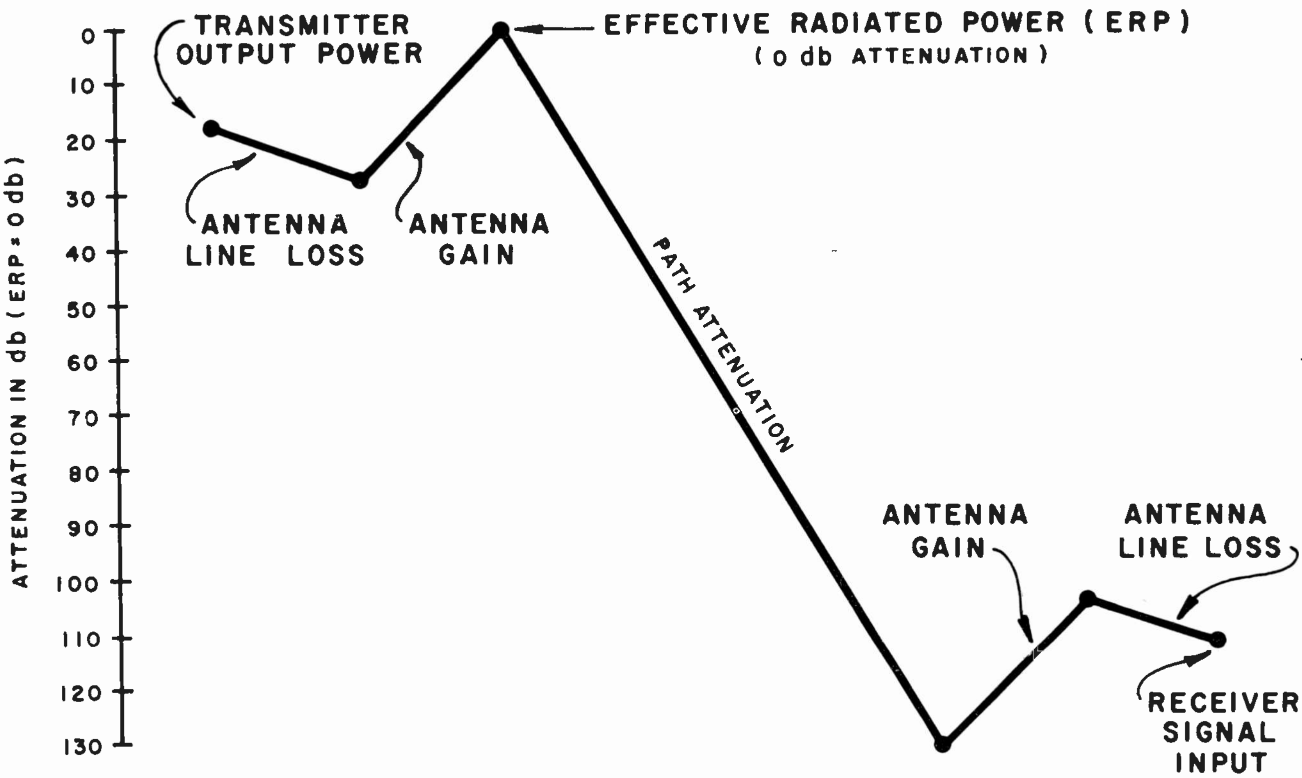


Figure 2-1. Relative Attenuation

$$P_r = p \times (A_r) \quad (2\text{-}2)$$

where

A_r = The effective area of the receiving antenna in square meters may be used to determine the available power, or *antenna gain*, if the Poynting power flow is known. The value A_r is expressed in terms of the wavelength of the received wave.

In order to determine the Poynting power flow (p) at any given distance from the transmitting antenna, the following formula may be used:

$$\text{Poynting power flow (p)} = \frac{P_t A_t}{d^2 \lambda^2} \text{ watts/meter}^2 \quad (2\text{-}3)$$

where

A_t = The effective area of the transmitting antenna expressed in square meters

d = Any given distance from the transmitting antenna expressed in meters; i.e., between transmitting and receiving antennas

λ = Wavelength in meters

By combining formulas (2-2) and (2-3) and solving for P_t/P_r, we find the *general attenuation formula* between transmitting and receiving antennas to be

$$\frac{P_t}{P_r} = \frac{d^2 \lambda^2}{A_t A_r} \quad (2\text{-}4)$$

or

$$\text{Attenuation in db} = 10 \log_{10} \frac{P_t}{P_r} = 10 \log_{10} \frac{d^2 \lambda^2}{A_t A_r} \quad (2\text{-}5)$$

where all units are consistent.

Since an isotropic radiator is the basic, theoretical form of antenna, it is used as a relative standard for all other antennas. The effectiveness of an antenna's directivity, such as that used with radio-relay stations, as compared with an isotropic antenna, is called *antenna gain*, or power gain. A half-wave dipole is quite often used as a standard for practical work instead of the isotropic antenna, but for theoretical work the isotropic radiator is more convenient. The actual power directivity of a half-wave dipole and the equivalent power directivity for an isotropic radiator with the same energy content are shown in figure 2-2. The effective area of an isotropic radiator is given by

$$A_e = \frac{\lambda^2}{4\pi} \quad (2\text{-}6)$$

and the gain of such an antenna is taken as unity. Therefore, the effective area of any antenna may be expressed by

$$A = \frac{G\lambda^2}{4\pi} \cong 0.08G\lambda^2 \quad (2\text{-}7)$$

where

G = Power-gain ratio of the antenna over an isotropic antenna.

By assuming that, in radio-relay stations, the same type and size of antennas are used for both transmitting and receiving, we can simplify formula (2-5) by substituting formula (2-7) in the place of A_r and A_t. The formula

Attenuation in db = (2-8)

$$10 \log_{10} \frac{P_t}{P_r} = 10 \log_{10} \frac{4.1 \times 10^{12} d^2}{G^2 \lambda^2}$$

is the simplified version of formula (2-5) where

d = The distance between antennas in miles

G = The power-gain ratio of the antenna over an isotropic antenna

λ = Wavelength in centimeters

Formula (2-8) is intended for use when expressing the attenuation between antennas under free space conditions, and may be found in nomogram form in the appendix. (See nomogram A.)

The power intercepted by a receiving antenna is in direct relation to the path attenuation between transmitting and receiving antennas. Assuming again that in radio-relay stations identical antennas are used for both transmitting and receiving, we can use formula (2-1) to determine the power in watts occurring at the output terminals of the receiving antenna.

$$\text{Attenuation } (\alpha) \text{ in db} = 10 \log_{10} \frac{P_t}{P_r} \qquad (2\text{-}9A)$$

or

$$\log_{10} P_r = \log_{10} P_t = \alpha/10 \qquad (2\text{-}9B)$$

where

α = Attenuation in db as derived from formula (2-8).

Formula (2-9B) may be used to express received power under any conditions if the path attenuation is known. If the value of (α) is taken from formula (2-8) then formula (2-9B) is used to express power under free space conditions only. Free space conditions are nearly realized when the antennas used are placed high enough to give first Fresnel zone clearance, except with frequencies above approximately 8000 mc when the transmission is through rain.

Rain scatters radiation from a passing electromagnetic wave. In addition, if the particles of precipitation are comprised of a dissipative medium, they will absorb energy from the wave and convert it into heat. Both phenomena have little effect on waves below the frequency of 8000 mc; but as the frequency increases, attenuation, due to scattering and absorption, becomes more and more noticeable. For frequencies above 10,000 mc both scattering and absorption are quite pronounced and impose a serious limitation on wave propagation through rain. However, at times, radiation may be scattered far beyond the radio horizon making it possible to receive signals which would otherwise be unintelligible.

When water, such as in fog or clouds, forms in droplets, each drop acts as a small particle of absorbing and scattering material. If the drop size is small in relation to the wavelength of the impinging wave, the absorption is independent of drop size and dependent only upon the total water content per unit volume. Since the concentration of water in fog or clouds is extremely small, except possibly in the case of very heavy sea fogs, attenuation by these media is negligible.

If the drop size increases as compared to the wavelength, absorption will increase at a fairly rapid rate. When the concentration of the rain is large enough, scattering will also result in an appreciable rate of attenuation.

Since it is extremely difficult to measure drop size in a typical rainstorm, attenuation by rainfall is generally expressed in terms of rate of precipitation per unit of time. The following empirical formula is a close approximation of attenuation by rain

$$\text{Attenuation in db} = \frac{10\, d\, R}{\lambda^2} \qquad (2\text{-}10)$$

where

R = Rate of rainfall in inches per hour

d = Path length in miles

λ = Wavelength in centimeters

The above formula may be found in nomogram form in the appendix. (See nomogram B.) Attenuation by

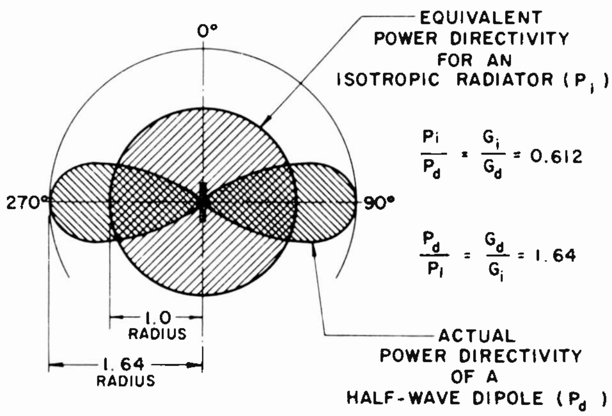


Figure 2-2. Relative Antenna Gain

rain is additional to the free-space path attenuation as given in formula (2-8). It must be remembered that there are many irregularities in a typical rainstorm and any attenuation calculations for path lengths much over three miles will be of an approximate value only.

For frozen precipitation, the absorption loss is considerably smaller than for a liquid. However, if the particles, such as hailstones or snow, are sufficiently large, the scattering loss will again result in an appreciable rate of attenuation.

FIELD STRENGTH. Field strength refers to the quantity of the electric field (E) at any given point and is measured in terms of volts per meter. One volt per meter is equivalent to a potential of one volt induced in an antenna wire one meter long. Since the unit, one volt per meter, is too large for most practical considerations, millivolts per meter and microvolts per meter are used more frequently. For example, if 50 microvolts are induced in a conductor one meter long when a radiated wave cuts across it, the field strength is 50 microvolts per meter or 50 μv/m.

The variation of field strength around an antenna system can be shown graphically by polar diagrams, which are simply circular charts which resemble the face of a compass. The antenna is taken as the center of the chart and the circumference of the chart is laid off in angular degrees. Computed or measured values of field strength then may be plotted radially in a manner that shows both magnitude and direction for a given distance from the antenna. (See figure 2-3). Field-strength values of this sort may be measured directly in the field by the use of a standard field-strength meter. Field strengths in the vertical plane are plotted on a semi-circular polar chart and are referred to as vertical polar diagrams.

The expected free-space field strength (E_o) in volts per meter, at a distance (d) in meters from a transmitting antenna, is given by

$$E_o = \sqrt{30\, G_t P_t}\, /d \qquad (2\text{-}11)$$

where

G_t = Power-gain ratio of the transmitting antenna over an isotropic antenna

P_t = Transmitter power in watts

d = Distance in meters

Formula (2-11) is simply a straight line relationship between the effective radiated power (G_tP_t) of a transmitting system, and the distance traveled by the wave. It is often convenient to determine the expected field strength of a definite antenna by using the value of one watt for (P_t). The answers obtained by this procedure may then be plotted graphically resulting in *field-strength below one watt.* A quick reference is then available for determining the expected strength below one watt by multiplying the field strength below one watt by the square root of the desired transmitter power ($\sqrt{P_t}$).

At frequencies above 300 mc, where the operating wavelength is considerably less than one meter, the received field is generally measured in terms of power rather than in field strength. The conversion from volts per meter to watts per square meter under free space conditions is given by the formula

$$P = \frac{E^2}{377} \qquad (2\text{-}12)$$

where

P = Field strength in watts per square meter

E = Field strength in volts per meter

By substituting in the above formula the free-space value of E as given in formula (2-11), the following field strength equation is obtained:

$$P = 0.08 G_t P_t / d^2 \text{ watts/meter}^2 \qquad (2\text{-}13)$$

where all units are the same as given for formula (2-11).

OBSTACLE GAIN. In VHF and UHF transmission, the first requirement has always been a clear path between the transmitting and receiving antennas, with approximately first Fresnel zone clearance above obstructions as a second, and less important, consideration. While there has been a definite minimum clearance acceptable for relay work, very few paths have been rejected for reason of excessive clearance, and any land mass intercepting the radio beam more than about 20 percent has been considered reason for discarding a site in favor of another with a clear radio path. Occasionally in the past, field strength measurements made with a land mass intervening between transmitting and receiving antennas have shown less attenuation than indicated by calculations, but these values were generally credited to exceptional atmospheric conditions, and ignored. There was no attempt made to calculate the effects of a large mass obstructing the signal, since it was generally believed that the attenuation would be far too great to give an intelligible signal at the receiving antenna.

Now it has been found that a large mass of land intercepting the signal path may actually improve the field strength over the value found for a similar path above smooth earth. This possibility was noted

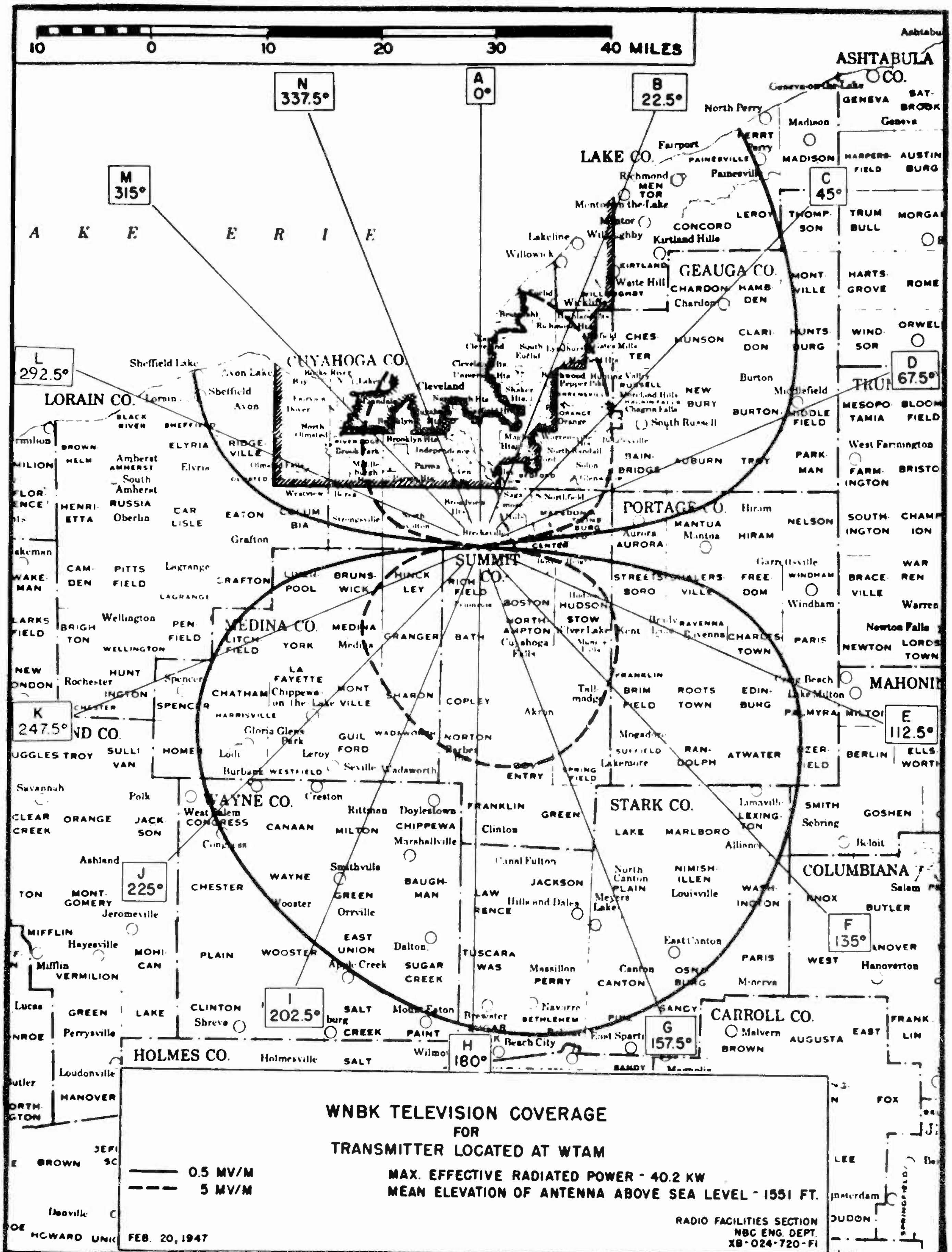


Figure 2-3. Polar Diagram of Field Strength

in articles by Schelling, Burrows, and Ferrell, in 1933, and by Bullington (Proc. I.R.E. Oct. 1947), although no installations based on this principle have been made until quite recently. According to an article in the Proc. I.R.E. for Sept. 1953 entitled "Large Reductions of VHF Transmission Loss and Fading by the Presence of a Mountain Obstacle in Beyond-Line-of-Sight Paths", by Dickson, Egli, Herbstreit, and Wickizer, higher field strength may be obtained, as well as better fade conditions. This article describes results found on a 38 mc, 160 mile path between Yakutat and Gustavous, Alaska. The transmitting and receiving antennas are approximately 200 feet high, with Mt. Fairweather intercepting the path as a single knife-edge, 8,775 feet high. Such a path, by all previous beliefs, would guarantee a signal completely lost in the noise. Instead of finding an extremely low field strength, the attenuation was found to be 134 db, as compared with 207 db attenuation for the same path over smooth earth. This amounts to an increase of 73 db, which is referred to as the "obstacle gain" of the intervening mountain. Recorded observations made over a considerable period of time showed a remarkable absence of fading.

This increase in signal strength, and almost complete absence of fades, is most easily explained by referring to the diffraction of light over a knife-edge. The radio signal is also diffracted by a knife-edge interception, and may also be reflected from the earth's surface. Figure 2-4 shows a source of an electromagnetic wave at T, and the point where the wave is received R. H represents the effective height of the intercepting obstacle, which is not necessarily the true height. The effective height is higher, and is determined by the ray paths from points T and R as they follow a grazing path over the mountain. Points P′ and P″ are reflection points on the earth's surface. As shown in the figure, there are four paths the radio beam takes in traveling between the two antennas: (1) direct-direct (T-P-R), (2) reflected-direct (T-P_1-P-R), (3) reflected-reflected (T-P_1-P-P_2-R), and (4) direct-reflected (T-P-P_2-R). It might appear that the reflected paths in mountainous country would be unimportant, due to irregularities of the ground. Actually, reflection will be surprisingly high, since the angle of incidence is almost at the grazing point. With such a high angle of incidence, the irregularities of the surface have far less effect than if the angle were more nearly at right angles to the surface. It is obvious from the figure that these paths will vary in length, and therefore the different rays arriving by the four paths will vary in their phase relationship at the receiving antenna. The four-path diffraction

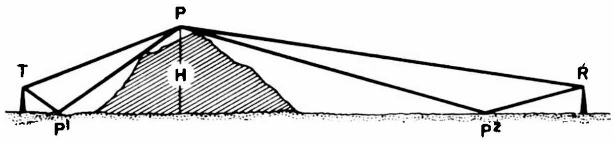


Figure 2-4. Four-Ray Path

problem of optics applies in this case, and may be used to determine the phase relationship. Any change in path lengths or operating frequency will alter the phase relationship of the four waves arriving at the receiving antenna. For any set condition of antenna location and obstacle height, a frequency may be determined that will allow all four rays to arrive at the receiving antenna in phase, or nearly in phase. Naturally, some may arrive many wavelengths behind, but still in phase, and no appreciable effect on the audio signal will be noted. A graph of attenuation against varying depth of penetration of the beam by the knife-edge will show the possible range of attenuation. (Figure 2-5.) The attenuation will be found to vary from near infinity to almost free space value, as the phase relationship of the arriving waves varies from opposing to aiding phase relationship.

This discussion of the effect of large obstacles on radio propagation is included primarily for its interest. At present, methods of using large obstacles to improve reception have not been definitely applied, and it is therefore best to use methods that have been tried and proved. There can be little doubt but that the useful effects of obstacles will be of value in the radio relay field, but considerable work remains before it can be satisfactorily applied.

FREQUENCY CHARACTERISTICS

For all practical purposes it is convenient to classify radio waves into bands of frequencies within which propagation effects are similar. However, any such classification must be of a general nature only, since changes in propagation characteristics with frequency are not sharply defined. Thus, when an upper or lower limit of frequency is designated for a certain propagation effect, it does not mean that such an effect stops at those limits, but rather that it becomes negligible beyond such limits.

30 MC TO 300 MC. This range of frequencies is part of the oldest known frequency band of the entire radio spectrum, since both Hertz and Marconi conducted their famous experiments in the region from 30- to 3,000-mc.

As a rule, ionospheric propagation is negligible at

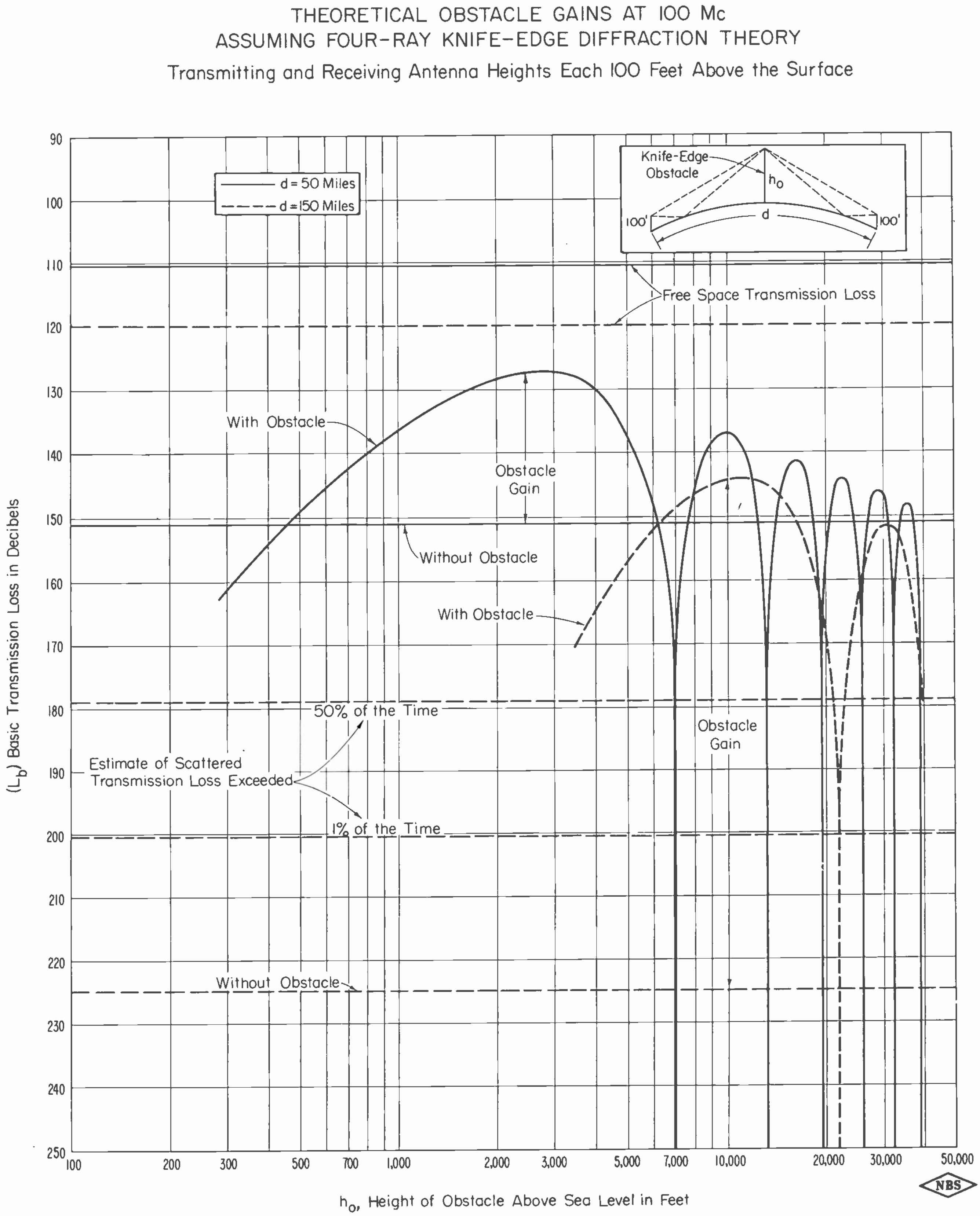


Figure 2-5. Attenuation with Obstacle

these frequencies and cannot be relied upon for propagation to great distances. However, irregular ionospheric reflections are possible (due to sporadic E variations) and can cause a signal in the lower part of this band to be propagated several hundred miles. Such propagation, by way of sporadic E reflections, decreases rapidly with frequency and becomes of little importance above 100 mc.

The most common method of propagation at these frequencies is, therefore, through the atmosphere along the surface of the earth. Over a smooth terrain, power radiated from an elevated antenna varies inversely as the square of the distance, and under normal operating conditions ground reflections are particularly important and depend upon the coefficient of reflection of the reflecting surface. For certain antenna heights the reflected wave will cancel the direct wave, while for other antenna heights wave reinforcement occurs. Within the horizon, interference between the direct and reflected waves will produce a series of maximum and minimum field strengths which are a function of the earth's dielectric constant and conductivity.

Since the density of the earth's atmosphere and also its index of refraction normally decrease with altitude, radio waves are propagated along a curved path by refraction. This is often represented as a straight-line propagation path occurring over an earth whose radius is 4/3 the true earth's radius. In effect, refraction tends to increase the range of a transmitted wave slightly beyond the geometric horizon, such a distance being called the radio horizon. In this frequency range, the radio horizon is governed not only by refraction, but by diffraction as well.

Diffraction plays a major role in this frequency band, depending upon the path over which the waves are propagated. The effect of obstacle gain caused by the diffraction of radio waves over a mountain ridge was first discovered in this frequency range. Diffraction effects are primarily important in regard to irregular terrain where it makes possible the reception of signals in the geometric shadow region of a hill or other intervening object. Generally, the higher the frequency, the less the diffraction effects.

Atmospheric noise at these frequencies is fairly low and decreases with increasing frequency. Such noise is usually caused by local electrical storms only. However, one of the more important sources of noise in this range is man-made noise such as that from ignition systems, diathermy machines, and X-ray equipment. Receiver noise also begins to become prevalent here, although considerable improvement in circuit design has made possible its reduction.

300 MC TO 3,000 MC. Almost all the energy transmitted at these frequencies is propagated through the earth's atmosphere along a curved path. The refracted path may again be assumed as a straight-line path over an earth's radius of 4/3 the true radius.

Ground reflections are still present at these frequencies and can cause multipath fading due to destructive interference, although such reflections become of less importance at the higher end of the band. However, a second type of multipath fading can occur when parts of the wave are refracted through other higher layers of the atmosphere and become bent sufficiently to return and combine with the wave received over a lower and more direct path.

At frequencies above 1,000 mc, attenuation of the transmitted signal by trees or other vegetation can range anywhere from 12 to 46 db per mile, although if the antennas are located to give first Fresnel zone clearance *above the trees* such attenuation becomes negligible.

Atmospheric noise in this frequency band is extremely low, as well as man-made noise, with the exception of ignition pulses which can become serious at times. Receiver noise is somewhat greater at these frequencies and increases with increasing frequency, thus calling for special receiving tubes and circuit design.

3,000 MC TO 13,000 MC. At these frequencies, transmission is strictly limited to line of sight distances based on 4/3 true earth's radius. Very little wave reflection will radiate from the earth at these frequencies. Instead, the earth will act as if it were made up of an infinite number of small mirrors, each reflecting the incident wave in a different direction. This is sometimes called diffuse reflection. In addition, incident radiation will also be absorbed by the earth's vegetation. Consequently, the amount of reflected energy from the earth's surface is small, and very little wave interference will occur from this source. However, multipath fading caused by tropospheric reflections or the multiple refraction of several propagation paths through the atmosphere is important throughout the entire band. There is also a tendency for buildings and other man-made objects to cast sharp shadows at these frequencies and, if the surface of such objects is smooth, even to reflect the waves in a new direction.

Rain scattering and absorption can cause a serious loss of radiated power at the higher end of this frequency range. If the drop size is comparable to the wavelength of the propagated wave, a very substantial portion of the transmitted energy will be reradiated from the raindrop in a wide range of directions. This phenomenon, known as scattering, has an attenuating effect on radio waves somewhat like that of diffuse earth reflections. However, not all the energy incident upon a raindrop is re-radiated, but instead, it is virtually trapped or absorbed and converted into heat. If the drop size is comparatively small in relation to the wavelength, such losses are dependent only upon the volume of water in suspension and therefore are generally negligible.

The use of sharply beamed waves to overcome the losses due to atmospheric attenuation in this band conserves enough power to enable a few watts of directed power to be as effective at a distant receiver as many kilowatts of undirected power.

Receiver noise at these frequencies has a great effect on the practical range of a receiving system. Special techniques, such as the use of crystal mixers, have been developed to minimize this noise by converting the received signal to a lower frequency before amplification. Such noise depends both upon the operating temperature and the receiver bandwidth and increases directly with both.

CHAPTER 3

FADING

INTRODUCTION

The term "fading" as used in this chapter is defined as the variation of radio field strength caused by changes in the transmission medium with time. There are two general sources of fades, inverse beam-bending and multipath interference, with multipath interference by far the more common of the two.

TYPES OF FADES

FADING DUE TO INVERSE BEAM-BENDING. In this type of fade, only the direct beam between the transmitting and receiving antennas need be considered, since the fading is caused by changes in the refractive index of the air through which this beam passes. Ordinarily, the beam follows a slightly curved path between the transmitting and receiving antennas, with the center of the beam striking the receiving antenna under standard atmospheric conditions. Any change in the refractive index of air through which the beam passes will result in a change in the beam path. Such changes will cause the beam to pass above or below the antenna, and a weaker received signal will result. The name, inverse beam-bending, is given because the bending is most likely to occur in an upward, or inverse, direction (figure 3-1A). It is possible, in extreme conditions, for the beam to be bent upward to such a degree that the signal becomes lost in the noise. Extreme fades of this type rarely take place, however, since a considerable change in the index of refraction must occur, and when they do occur, they are usually slow in nature.

MULTIPATH FADING. The beam directed at the receiving antenna is not a narrow pencil of energy, but spreads out in the form of a cone, with its apex at the transmitting antenna. Under normal conditions the central part of the cone strikes the receiving antenna, with part of the beam passing above and part below the receiving antenna. Should the upper part of this cone enter a region of the atmosphere with an unusually small index of refraction, it would be refracted downward toward some part of the earth's surface. Such a refracted beam might easily strike the receiving antenna and cause fading, the degree of the fade depending upon the signal strength of the refracted beam, and the phase relationship of the refracted and direct beam (figure 3-1B).

The comparative strength of the two signals will be determined more by the part of the cone refracted downward than by the difference in path length. For paths of fifty miles or less, the difference in path lengths will be comparatively small. The phase relationship between the two beams, however, may easily vary from 0 degrees to 180 degrees, depending upon the path length and the transmitted frequency. Should the path of the refracted beam, for example, be one wavelength longer than the direct beam, the signals would arrive in phase, and the resultant field strength would be approximately twice the expected field strength. A difference in path length of one-half wavelength would result in the arriving signals opposing each other, and deep fading would result. It should be noted that the refracted beam might arrive many wavelengths behind the direct beam, and still be in phase, or any odd multiple of half-wavelengths, and still be 180 degrees out of phase with the direct beam. In either case, the modulating signal would be unaffected, except by fades, because of the tremendous length of the modulating signal at audio frequencies. Since the difference in paths is measured in wavelengths, rather than miles, such fades would be expected more often at shorter wavelengths than longer ones; and so it is found to be in practice.

Fading caused by radio signals reflected from some part of the earth's surface are similar to fades due to refracted signals, although the operation is a little different. The received signal strength will depend not only on the initial strength of the transmitted beam, but on the amount of energy reflected from the earth and the phase lag introduced at the reflection point. As the angle of reflection in radio relay work is small, seldom exceeding 2 degrees, the surface will be smooth for practically all wavelengths over approximately ten centimeters. (Under 3,000 mc.) At frequencies above 3,000 mc, earth reflections generally become diffuse in nature, and their effect upon the received field gradually becomes negligible.

HEIGHT–FREQUENCY DEPENDENCE. At frequencies below 30 mc, antenna height above the earth is important, since the radiated signal is affected by the presence of the earth. Antenna heights under one-half wavelength cause the radio signal to be deflected downward to such an extent that the signal is rapidly attenuated. By increasing antenna height to one-half wavelength, or higher, the transmitted wave will then travel outward, instead of deflecting downward. Obviously, since the wavelength height of the antenna, rather than its physical height, is the measure of elevation, antenna height may be increased by increasing the operating frequency, or by actually raising the antenna. The earth's surface will affect the signal to a lesser extent as the height (or frequency) is increased, until full first Fresnel zone clearance is obtained. At such time, the signal will radiate with no shaping of the wave by the earth's surface. Beamed signals such as are used for microwave relay work may therefore be considered as originating in free space so long as first

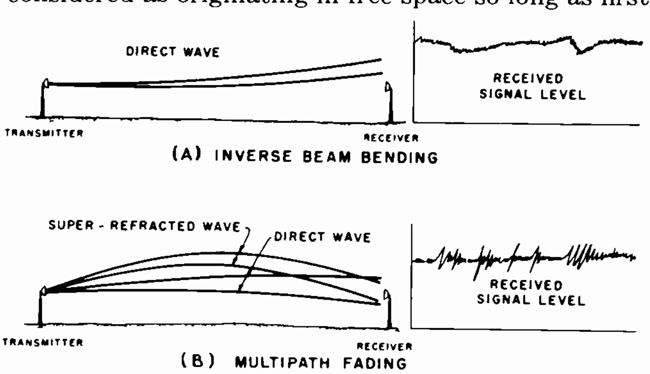


Figure 3-1. Single & Multipath Fading

Fresnel zone clearance is obtained over the full path length. Fading may still take place, however, since some part of the beam may be reflected from the earth with sufficient intensity to affect the received signal. Such conditions are particularly true where the path is over a good reflector, such as sea water. For example, a beamed signal from Los Angeles to Catalina Island was first tried by using extremely high antennas, in order to obtain line of sight transmission. Reflections from the surface of the water caused almost completely unintelligible signals, however. Had it been feasible to raise the antennas higher, a point would eventually be reached where free space conditions would prevail, and such reflections cease to cause trouble.

ATMOSPHERIC EFFECTS

M-PROFILE. An M-profile is a graphical cross-section of the atmosphere showing the change of the index of refraction with altitude, and is drawn by plotting (M) as abscissa, and (H) as ordinate. Under standard atmospheric conditions, the value of M increases by 3.8 per hundred feet, and the standard M-profile will therefore be a straight line, having a slope of 3.8 per hundred feet. Field strength at the receiving antenna, calculated on the basis of a standard M-profile will be the normal field strength to be expected, that is, with neither fades, nor abnormally strong signal. The field strength value based on the standard M-profile agrees with values calculated by use of effective earth's radius equal to 4/3 the actual radius.

The standard M-profile is dependent upon homogeneous atmosphere, the most commonly found condition, but weather changes can produce variations in the M-profile, and in the transmission of radio signals. (See Chapter IV.) In figure 3-2A is shown a standard profile; figure 3-2B shows the effect of an abnormally high surface temperature or water vapor content. Such a condition will result in changing the curvature of the radio signal so that it will approach a straight line (based on true earth radius) or in curving the beam away from the earth. Such a curve is called inverse beam-bending, and may occur if the area is sufficiently deep, or sub-standard.

Figure 3-2C illustrates a condition that is just the opposite of that shown in 3-2B, both in the curve of the M-profile, and in the effect on the radio signal. A sharp rise in temperature with increasing height, or a sharp decrease in water vapor content, or both will produce such a condition. This phenomenon produces a marked bending of the wave path and is known as super-refraction. Since the radio signal will tend to stay within the elevation limits shown as A and B, the condition is frequently referred to as ducting, or trapping. Where the M value is decreasing with elevation, at elevations close to the earth's surface, a surface duct will result, and the effect on the radio signal will depend on the degree of change in the M value with elevation, and the height of the duct. A similar condition, known as an elevated duct, is shown in 3-2D. As in 3-2C, the boundaries of the duct are indicated by A and B. Such elevated ducts have been found near San Diego, Cal., while surface ducts are often found over a large body of water, such as the Caribbean Sea, where a surface wind is prevalent.

Trapping of a radio signal may not be relied on for transmission other than for short periods of time, since the conditions producing the duct are subject to change. The trapping found in the Caribbean Sea, for example, is predominantly a daytime phenomena.

Observations have shown that a duct tends to confine electromagnetic waves with an action much like that of a waveguide. Field strength measurements taken inside a duct are considerably higher than those taken outside the duct. Therefore, in order for a duct to be used to maximum advantage, both the transmitting and receiving antennas should be located within the duct. Another factor to consider in trapping is the transmitted frequency. A duct, like a waveguide, is frequency sensitive and will confine certain frequencies much better than others. As a general rule, the higher the duct, the lower the frequency that may be used. The strength of the trapped signal will also vary according to the nature of the duct boundaries. If the rate of change of the M-profile is considerably less than standard, the refraction at the boundary of the duct will be greater than in the event of a slight change, and a greater part of the signal will remain within the duct.

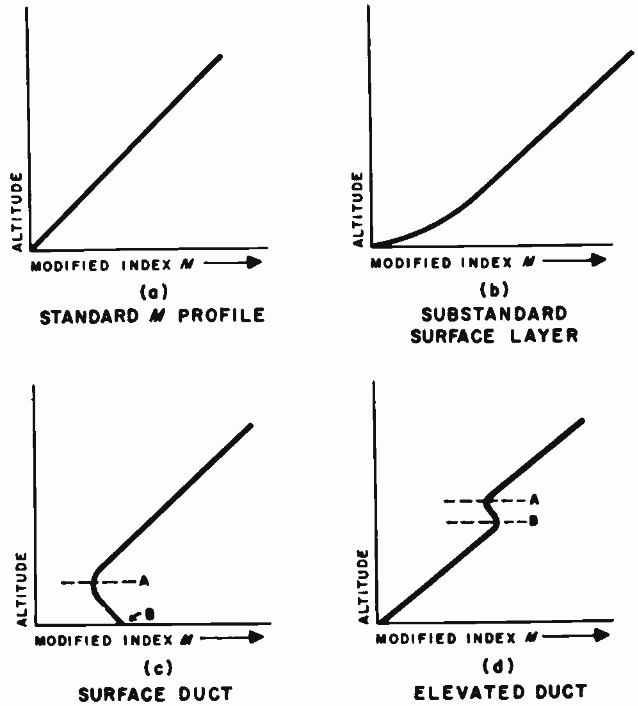


Figure 3-2. Typical M Profiles

DIURNAL SIGNAL VARIATIONS. Assuming there is no degree of obstruction or object penetration of the transmission path (first Fresnel zone clearance), steady transmission occurs when the air along the path is well mixed, as on sunny afternoons, or in stormy, windy weather. The M-profile is then nearly standard except where irregular temperature and water vapor pressure conditions near the ground cause a slight increase in the M value at that elevation. Signal fading occurs in the presence of irregularities in the M-profile. A marked change in signal level invariably accompanies these irregularities over non-optical paths because of multipath fading, but for optical paths the change in signal level is more *diurnal* in nature, indicating the fading to be of the single-path variety. In general, a *diurnal signal* changes from a steady condition at midday to a type characterized by pronounced fading at night. It is usually observed that on non-optical paths the *average* signal level rises when fading occurs, as in over-water paths, but on optical paths it may either rise or fall for considerable periods. Diurnal signal variations are considerably more pronounced in summer than in winter because of the high water-vapor content of the air in summer, which increases the effectiveness of temperature gradients in producing strong M gradients.

Since changes in the M-gradient occur during or after the passage of temperature fronts or high wind conditions, it is possible to compare the passage of such fronts with the change in signal level and arrive at an approximate relationship between the two. Here again, any hard and fast statement is dangerous, since exceptions occur due to seasonal effects, or combinations of weather changes, but the general effects will be as given below.

Warm fronts are usually associated with lower signal strength, whether or not the signal prior to the arrival of the front was normal or high. This may be accepted as a fairly safe rule for most seasons, although in winter time, the warm front may produce a higher signal strength, rather than a low one. A warm front with a high moisture content will occasionally result in much stronger signals during winter. For example, television broadcasts originating from Philadelphia, Pa., channel 3, have been received at Riverhead, L. I., far beyond the normal range of the frequency used.

Cold fronts are usually accompanied or followed by an increase in signal strength. This increase may continue for several hours after the passage of the front.

High winds produce a thorough mixing of the atmosphere, and hence a normal signal will result. An exception would occur if the high winds should be accompanied by an air mass that was either exceptionally dry, or exceptionally moist.

Stormy conditions are often accompanied by a variable signal. The diurnal trend also is more noticeable during stormy conditions.

TYPES OF SIGNALS

FADING VS M-PROFILE. Four different combinations of fades have been determined, as shown on records of field strength taken over a number of years. These types of fades are the result of a direct combination of inverse beam-bending and multipath fading, and are illustrated in figure 3-3.

Figure 3-3A shows the relationship between the type of fade encountered with a substandard M-profile. Such fades occur quite rapidly, with signal strength below normal. Occasionally slower variations of signal strength are superimposed on the fading shown in the figure. This type of fading is known as scintillation and is fairly common in summer, although almost non-existent in winter.

With M-profiles that are standard, or nearly so, the average signal level is steady within about ± 5 db. Scintillation is usually present with an amplitude varying between 5 and 10 db, although the upper limit may approach 25 db. This is the kind of fading most frequently encountered throughout the year, and is often referred to as a standard signal (figure 3-3B).

When shallow surface M-inversions, with their accompanying ducts, occur, the signal level is above standard and increases with duct height. The signal never rises above the free-space value, and only minor, short-time fades occur. This is defined as partial trapping, and is shown in figure 3-3C. A partial-trapping condition is quite stable, with very little fading. This condition is quite common in winter, when it may last for several days at a time.

With deep M-inversions at the surface, the signal level is near the free-space value. The signal is unsteady, with *roller fading* of amplitudes as large as 40 db and for periods ranging from a few minutes to an hour. As the M-inversion deepens, the amplitude of fading increases, and the period of fading decreases. Roller fading is characterized by broad flat maxima and sharp deep minima, and represents the most violent fluctuations possible. This type of signal, called the strong-trapping signal, (figure 3-3D), occurs in the summer and fall and then only on rare occasions. These periods of extreme fades may exist from one to five hours, and usually occur between midnight and sunrise.

While these types of fading are distinct, one from another, changes from one type to another may take place quite rapidly. Such changes may affect either the average level of the signal, or the type of signal, and may occur at the same time. The magnitudes of these changes generally increase with frequency.

FREQUENCY DEPENDENCY. Signal types have been found to vary, not only according to changes in the M-profile, but according to frequency as well. In cases where a grazing path exists, substandard M-profiles will be accompanied by a scintillating signal at all frequencies considered in this book. The variations due to scintillations are irregular and rapid, with periods of about 30 seconds duration, and amplitudes of at least 15 db. The decreases from a standard signal are approximately equal at frequencies between about 3,000 mc to 13,000 mc. As the clearance for the radio path is lessened by an obstacle, the signal strength becomes lower, provided there is no counteracting effect, as noted in obstacle gain (Chapter II).

During periods of standard or nearly standard atmospheric conditions, the signal strength at all frequencies will be approximately standard as well. No unusual signal strength due to ducts, or multipath conditions will occur, since these conditions themselves are dependent upon unusual weather conditions. For any given path, for example, the signal will be approximately as calculated, since such calculations would take into account standard refraction, reflection from the earth's surface, or any other unusual physical condition. Where a superstandard condition occurs, as indicated by an M-inversion at some layer of the atmosphere, different frequencies will be affected differently. A shallow surface duct, as might occur over water, will have little effect on a signal at 3,000 mc, even though both antennas are within the duct. A signal of 13,000 mc transmitted through a shallow duct may increase in strength, due to the duct. If the duct increases in height, the signal strength of the 13,000 mc signal will increase, although at first there will be no change in the lower frequency signal. As the duct size is increased to 50 feet, the field strength of both the 3,000 and 13,000 mc signals will rise above the value to be expected

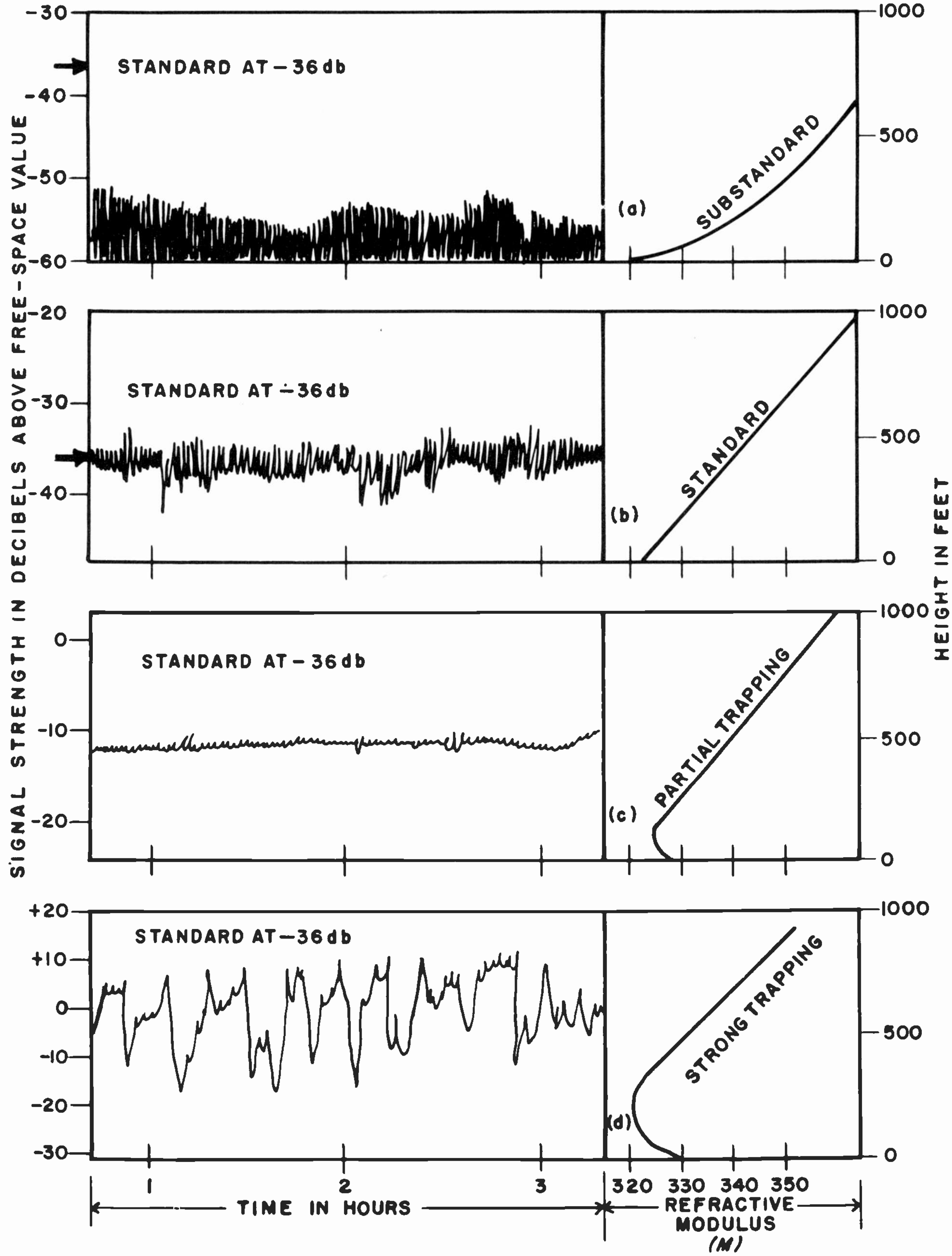


Figure 3-3. Types of Signals

without the duct, and the higher frequency will approach free space value. As the duct height approaches about 70 feet, roller fading will occur at the higher frequency, while the 3,000 mc signal will continue to increase in strength. With duct heights of greater value, the 13,000 mc signal will decrease in strength, with increased effects of multipath fading, and roller fading will begin to occur at the 3,000 mc level.

The discussion on signal strength and fading is mentioned as occurring at the two levels; actually, the intermediate frequencies are also affected, with the effect varying between the two extremes mentioned. If the duct height is too low for any given frequency, it will have little effect; if too high, roller fading will take place, succeeded by a reduction in field strength, as the duct height is further increased. Should the elevation at which the M-inversion occurs be sufficiently high for the wavelength under consideration, no ducting will occur, but a multipath fade, caused by refraction of the signal at the layer may take place. At any given distance within the radio horizon, the period and amplitude of the fades on the upper and lower frequencies are nearly equal. As shown in figure 3-4, the fading, however, is far from synchronous. Usually the higher the frequency, the shorter the fading period, with the amplitude of fade increasing with increasing frequency.

EFFECTIVE EARTH'S RADIUS. Experimental data on fading at frequencies above 30 mc can be correlated reasonably well with the concept of an effective earth's radius if it be assumed that such a value is rarely less than 0.8 times, or more than 3.0 times, the true earth's radius. At higher frequencies the received signal at distances beyond the radio horizon is seldom less than would be expected for an earth's radius of 0.8 true radius, and may be equal to or above the free-space value. The theory of trapping would indicate that the received signal may be considerably higher than indicated by free-space transmission. Instantaneous peaks of 18 to 20 db above the free-space value have been reported, but the average signal is greater than 3 or 4 db above the free-space value for only a small percentage of the total time. This means that the fading range at 3,000 mc over a path ten miles beyond the radio horizon may be as much as 60 or 70 db.

The percentage of time that the signal is either extremely high or extremely low depends on meteorological conditions as mentioned, which, in turn, are functions of geography and season of the year, as well as daily weather conditions. Within the radio horizon, the received power at frequencies above

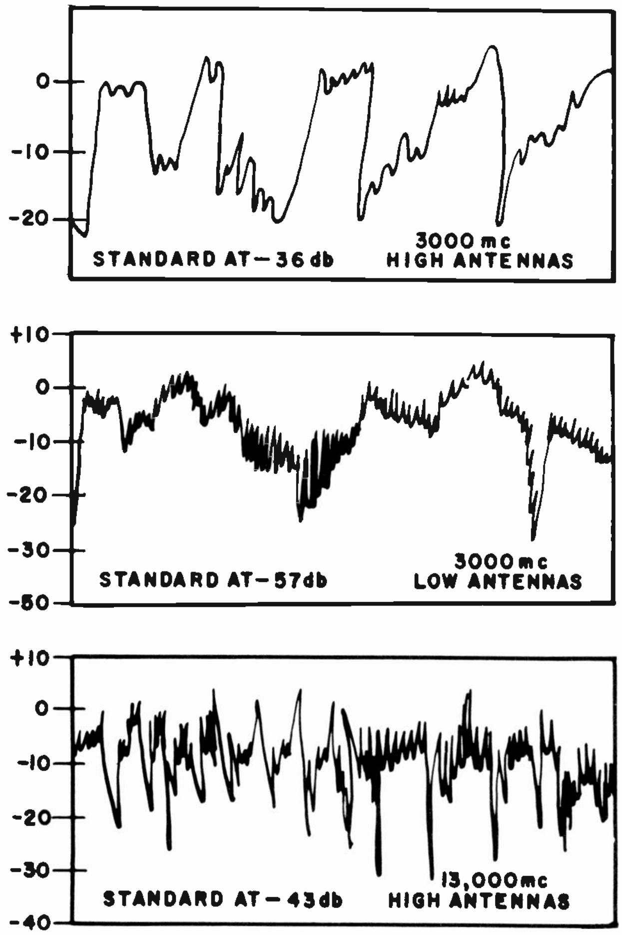


Figure 3-4. Fading Comparison

3,000 mc may vary from several db above, to 15 db or more below the free-space value, even over a good optical path. A good optical path is defined as one with full first Fresnel zone clearance as discussed in Chapter 5. The fading appears to be more severe over sea water than over land, but whether this difference results primarily from a difference in atmospheric conditions or from a difference in the reflecting properties of sea water has not been clearly established. An outline of fading causes plotted against frequency is illustrated in figure 3-5, although the frequency divisions shown are general, and such causes may extend far beyond the ranges given.

DIVERSITY RECEPTION

SPACE DIVERSITY. Multipath fading may be minimized by the use of a practice called space-diversity reception. Two or more antennas are spaced several wavelengths apart, generally in a vertical direction. It has been observed that multipath fading will not occur in synchronism at both antennas, although a diurnal change in signal intensity will.

Sufficient output is almost always available from one of the antennas to provide a useful signal, and only the strongest signal is applied to the receiver amplifier. The use of two antennas at different heights provides a means of compensating, to a certain extent, for changes in the electrical path differences between direct and reflected rays by the selection of the stronger signal. However, fading due to inverse beam-bending is generally not affected by space diversity.

The antennas should be separated far enough to give an approximate half wavelength between the geometrical path differences of the two signals. An approximation of the spacing is given by

$$\text{Spacing in feet} = 43.4\lambda d/h_t \qquad (3\text{-}5)$$

where

λ = Wavelength in centimeters

d = Path length in miles

h_t = Height in feet of the transmitting antenna above a plane tangent to the earth at point of reflection.

FREQUENCY DIVERSITY. Since the period of fading decreases as the frequency of a signal increases, another method of diversity reception has been utilized to some extent with varying degrees of success. This method is commonly known as frequency diversity, and must be used simultaneously at both the transmitter and receiver.

Frequency diversity is, as the name implies, simply the use of two transmitters and two receivers, each pair tuned to a different frequency. If the fading period at a precise frequency extends a definite length of time, the same signal on a higher frequency will still be received. Likewise, if the amplitude of fade tends to be too great, the signal at a lower frequency will be steadier. Since this method doubles the amount of equipment necessary, it is used only on the most expensive systems, such as the trans-Atlantic network where high reliability is a necessity.

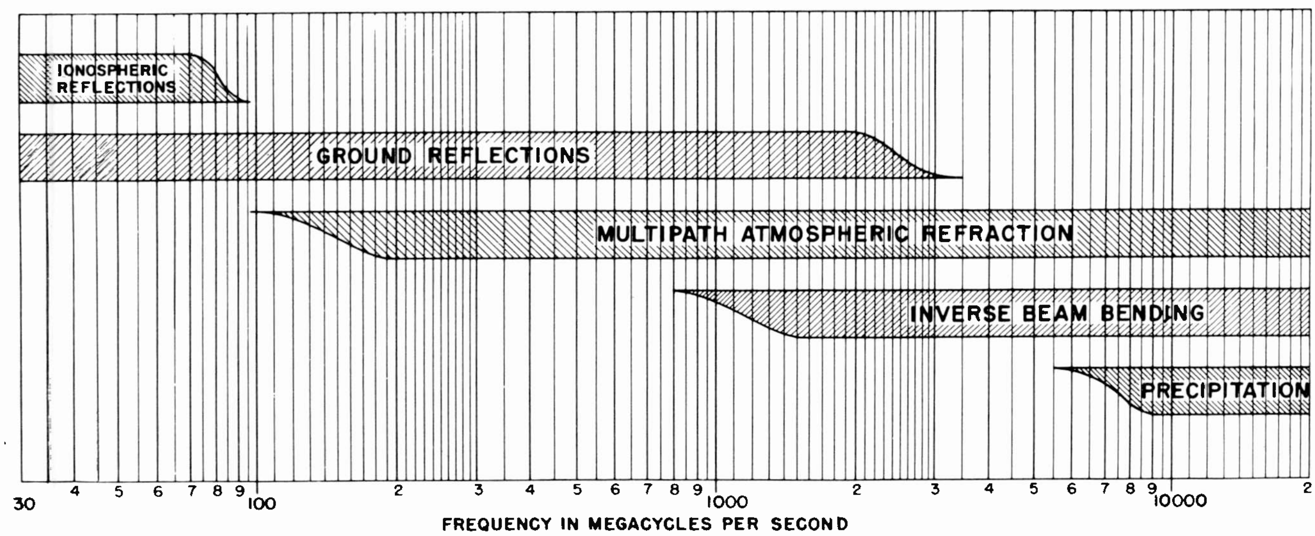


Figure 3-5. Fading Causes by Frequency

CHAPTER 4

WEATHER VERSUS PROPAGATION

INTRODUCTION

Weather is one of the many factors affecting electromagnetic wave propagation. The ways in which weather may affect wave propagation are many, and it is the purpose of this chapter to consider the various phenomena known as weather, and to show their relationship to radio propagation. Wind, air temperature, and water content of the air can combine in many ways, with different combinations causing radio signals to be heard many hundreds of miles beyond their ordinary range, or attenuated to a point where the signals may not be picked up over a normally satisfactory path. Unfortunately, no hard and fast rules may be given concerning the effects of weather on radio transmission, since the variables of weather are extremely complex, and subject to frequent change. Any discussion on the effects of weather on radio must be limited to general terms, therefore, and will be so treated here. As an illustration, rain is considered to be a negligible factor in the attenuation of radio signals at frequencies below 6,000 mc. Rain may cause excessive attenuation of signals below 6,000 mc, in the event of a cloudburst, or a rain with exceptionally large drops. Normal rainfall, even though it covers the whole radio path, will cause so little attenuation that it may be safely ignored. At frequencies above 6,000 mc, normal rain may still be ignored, but the chances of attenuation by rainfall become greater, and increase as the frequency increases.

WEATHER VS PROPAGATION

ATMOSPHERIC CONDITIONS. As stated in Chapter 3, wave propagation in free space is subject to path attenuation only. The total amount of power traveling outward from any antenna in free space remains undiminished with distance, but the amount of power falling on a unit area will decrease as the wave spreads outward from the antenna. Neither absorption nor refraction can take place in free-space propagation, only path attenuation due to the distance between transmitting and receiving antennas. Free-space propagation, of course, represents an impossible ideal, since the earth's atmosphere will affect wave propagation in a number of ways with the result that the signal strength may fall far below a satisfactory value, or possibly reach a value greater than free space. A consideration of the atmosphere and its make-up will be of help in understanding the effects of weather on radio propagation.

The three basic elements of weather are atmospheric pressure, temperature, and vapor pressure. These elements are not uniform with respect to elevation above the earth, but tend to decrease uniformly with increase in elevation. Their variations have been determined from meteorological soundings taken in the lower part of the troposphere, over a period of years, and standard gradients have been established. These gradients are used to establish standard atmospheric conditions. Pressure and temperature normally decrease with increase in elevation; temperature, for example, decreasing approximately 5.4° F. per 1,000 feet. Relative humidity increases with elevation, and may be determined from wet and dry bulb temperature readings applied to relative humidity charts. In formulas concerning the refractive modulus, vapor pressure is used, and may be determined from charts designed to show the relationship between wet and dry bulb readings, atmospheric pressure and vapor pressure.

Standard atmospheric conditions are based on homogeneous air, which is defined as air that is uniform in composition, and having uniformly varying temperature caused solely by uniform decrease in pressure with increase in elevation. Elevation has no effect on the water vapor content, although both the dew point and relative humidity will change with elevation due to change in pressure and temperature. It should be noted that temperature, pressure, and vapor pressure at the earth's surface are not specified for standard atmospheric conditions, nor is the elevation of the surface considered. The important factors concerning standard atmosphere are uniform vertical change of pressure and temperature, with pressure, temperature, and vapor pressure constant at any given elevation.

HOMOGENEOUS AIR

Homogeneous air is the result of thoroughly mixing the air, which may be done in a number of ways. One way in which the air becomes mixed is by means of convection currents. Open areas of the earth's surface may be heated by the sun, producing warmer, and lighter, air above the area. This mass of air rises due to its comparative lightness until an elevation is reached where the pressure of the mass of air is the same as the surrounding air. As the mass of air moves upward, it is replaced by cooler air from the surrounding terrain. The result is a thorough mixing of the air. Another source of homogeneous air is storms. Surprisingly enough, the winds producing storms also produce homogeneous air, which is one reason why relay systems are superior to wired lines. Storms that might damage a wired line contain enough turbulence to produce homogeneous air, with consequently good operating conditions for radio transmission. A third source of homogeneous air is cool air moving over warm earth. The air mass becomes heated from the earth, with upward air currents resulting in a thorough mixing of the air. In discussing homogeneous air, it should be pointed out that the homogeneity is produced only in the lower troposphere. Similarly, only a small part of the earth's surface may be covered by homogeneous air, but this can still be great enough to be important in radio propagation.

Where standard atmospheric conditions do not exist, the normal change in vapor pressure and temperature with elevation may be considerably varied. One cause of such a change is moving masses of air. Throughout the earth's atmosphere, a continual exchange of air takes place between the equatorial regions and the poles. Large masses of polar air sweep downward in a south-easterly direction in the northern hemisphere, (north-easterly in the southern hemisphere), while masses of tropic air move towards the poles, resulting in a continuous, but irregular, exchange of air in intermediate areas. Polar air masses are characterized by low temperature, moderately low vapor pressure, and high atmospheric pressure. Tropic air masses usually originate over salt water near the equator, and hence have high temperature, high vapor pressure, and low atmospheric pressure.

AIR MASSES

POLAR. While these air masses themselves may be homogeneous, they differ considerably from the air in the regions entered. The polar air mass will be denser than the existing air mass, and colder. The denser air will tend to displace the lighter, warmer air, and will tend to underlay the lighter air. As a result, the lighter air will rise, and a temperature increase with elevation (temperature inversion) will exist. This temperature inversion, however, will not be as great as might be supposed. As the warm air mass is pushed upward, the pressure will decrease, and the temperature will therefore also decrease. The vapor pressure will increase, so that, while a temperature inversion will exist, the inversion will be less than if the displaced mass retained its original pressure and temperature. The temperature from the earth's surface will decrease upward through the polar air mass at approximately the standard rate of 5.4° F. per thousand feet, to the boundary between the two masses.

In the boundary between the two masses, the temperature will increase with elevation, since the displaced air will be warmer than the polar mass. From the lower boundary of the displaced mass upward, the temperature will again decrease with elevation, and at an approximately normal rate. As may be seen from this explanation, the boundary between the air masses is not sharply defined, but possesses thickness. While the thickness may vary, there will still be a uniform change in temperature through the boundary between the two air masses.

TROPIC. Tropic air masses moving into an area of colder air will overlay the colder air, so that similar conditions to those found when a cold air mass enters an area will prevail. Regardless of which air mass is moving, the warm air will tend to overlay the cold air, and a temperature inversion will exist. It may occasionally happen that the warm air is denser than the existing air, and will therefore push the existing air mass upward. When this happens, the decrease in temperature at the boundary between the two masses will be greater than standard, and thus the temperature will decrease with elevation at a greater than normal rate.

SUBSIDENCE. There are other air mass movements than those already mentioned. Air may move in a circular direction which is not necessarily accompanied by a forward movement. A counterclockwise movement, called cyclonic in the northern hemisphere, occurs as a low pressure area, and is characterized by surrounding air at low levels flowing into the area and upward through the center of the area. Such an air mass movement has little effect on radio propagation, although the thunderstorms which may accompany such a movement will produce interference.

Another stationary circular movement is clockwise or anti-cyclonic, and is characterized by high pressure, downward flow through the center of the area, and outward flow at low levels. Where the area involved is large, and the sinking mass of air covers a large area, the movement is called subsidence. As the air flows, or sinks, downward, the pressure of the mass increases with a resultant increase in temperature. As the sinking movement continues, the temperature and pressure continue to rise until the mass reaches a level where the pressure is equal to the pressure of the surrounding air. At this point, the downward movement ceases, for the air is balanced, as it were. The temperature is greater than that of the surrounding air, because of the temperature increase produced by the increased pressure. Thus, a temperature inversion exists due to subsidence. The condition produced by subsidence tends to be short-lived, because of the sinking layer spreading outward as it reaches a level where pressure is balanced, and downward movement ceases. As the spreading movement continues, the layer becomes increasingly thinner, eventually losing its identity altogether. While the sinking process is under way, however, and for some time after, a temperature inversion at the layer will exist due to warming by pressure increase. Where subsidence exists, the vapor pressure may or may not increase with atmospheric pressure. While it would be extremely difficult for the radio operator to know whether to expect the vapor pressure to increase or decrease, vapor change can materially affect transmission characteristics of the air (see M-characteristics, this chapter). The transmission characteristics may be determined, where necessary, by means of meteorological soundings of the atmosphere.

DIURNAL CHANGE. The air above the earth is continually subject to change by either heating or cooling from the earth's surface. This effect is, of course, tempered by winds and the horizontal movement of air masses, and may be negligible in comparison to effects caused by winds. In still air, however, cooling or heating of the atmosphere by the surface of the earth can have appreciable effects on the refractive modulus, and hence on wave propagation.

Heating of the atmosphere from the earth's surface occurs almost exclusively in day time, when the sun's heat is partially reflected from the surface. This heating results in upward convection currents of warm air. As the warm air rises, it displaces heavier, cooler air, which is in turn warmed by the earth's surface. This results in a thorough mixing of the atmosphere with a standard temperature gradient, and homogeneous air will result. Such heating is effective above 1,000 feet elevation.

Just as heating from the surface is a daytime phenomenon, surface cooling is a nocturnal phenomenon. The heat absorbed by the earth during the day is dissipated during the early evening hours, the time of dissipation depending on the temperature and type of soil. Since no heat reaches the earth, further cooling will cause the earth's temperature to drop below that of the air, and produce a temperature inversion near the surface of the earth. The cool air will be heavier than the overlaying warm air, and consequently, no convection currents can exist. Any interchange of heat between the layers of air will be by conduction only. With no vertical air movement, no mixing can take place, and the air becomes stratified, with cool air underlying the warm air. This increase of temperature with elevation, or temperature inversion, may or may not produce an M-inversion (see Chapter III), depending on the humidity of the air, and the moisture content of the earth. If the air is not saturated, and the ground is dry, nocturnal cooling will produce a higher humidity of the cooling layer, and an M-inversion will result. If the ground is wet, or even moist, the problem becomes extremely complicated, and it is impossible to predict the M-profile, although it may be plotted from actual values of temperature and vapor pressure at different elevations. Since the ground is usually moist during winter, and dry during summer, nocturnal cooling generally takes place during the warmer months, and its effects are frequently observed on 24-hour signal strength charts. Nocturnal cooling is sufficiently common during the summer and can be expected more than half of the summer nights.

PRECIPITATION ATTENUATION

INTRODUCTION. Calculating the effect of weather on radio propagation would be comparatively simple if there were neither water nor water vapor in the atmosphere. The refractive modulus would depend only on the temperature gradient, and a homogeneous atmosphere would probably exist most of the time. However, some form of water (vapor, liquid, or solid), is always present in the atmosphere, even in arid regions, and must be considered in all microwave calculations.

The discussion up to this point has considered only the effect on the refractive modulus of the atmosphere due to temperature and vapor pressure. Water in the atmosphere, whether as vapor or liquid, will also affect radio transmission by attenuation. Water can attenuate an electromagnetic wave by absorption and by scattering. These effects are so slight at frequencies below 6,000 mc, that attenuation by these causes at such frequencies may be safely ignored.

Attenuation by absorption is caused by the water molecule acting as a dipole. This dipole absorbs the electric component of the wave, dissipating the absorbed energy in the form of heat. Absorption due to the molecule acting as a dipole may occur with other gases, but such absorption affects radio signals at frequencies beyond the range of this book.

RAIN. As previously stated, attenuation due to raindrops is greater than the attenuation due to other forms of water. Attenuation may be caused by absorption, whereby the raindrop, acting as a lossy dielectric, absorbs power from the electromagnetic wave and dissipates the power as heat, or by scattering. Scattering by raindrops will cause more attenuation than absorption at the frequencies considered in this book, with the effects of scattering becoming quite pronounced at the higher frequencies. This scattering is similar to the scattering of light by minute particles in the atmosphere, as determined by Lord Rayleigh in his experiments on light waves. The scattering of the electromagnetic wave is dependent upon the size of the raindrops, their density in the atmosphere, and the wavelength of the electromagnetic wave. Large drops produce greater attenuation than small ones, with the attenuation increasing as the diameter of the drops approaches the wavelengths of the radio wave.

One of the difficulties in attempting to determine the attenuation by scattering is caused by the variation in drop size. There is no uniformity of drop size in any rainfall; the droplets vary in diameter from less than one millimeter to five millimeters or more. As a general rule, the heaviest rate of rainfall is accompanied by the greatest drop size, and, therefore, the greatest attenuation. This tendency toward increasing drop size with increasing rate of precipitation is shown in table 4-1, compiled from a study of nine different rains by A. C. Best.* The table shows the approximate number of raindrops per hour, of different sizes, as found in nine different rainfalls. As may be seen from the table, there is no exact relationship between drop size and rate of precipitation, only a tendency toward larger drops with heavier rains. Column "E", in particular, shows this lack of relationship. While the rate of rainfall is 2½ times that of column "D", the smallest drops are still predominant. Table 4-2 shows the relation between attenuation and precipitation rates. This table was prepared from data in table 4-1, and represents the approximate attenuation to be expected. Notice that the column "B" is omitted in this table, since this rain is composed almost completely of droplets 1 millimeter or less in diameter.

Another difficulty in calculating the expected attenuation by rain over a radio path is the variation of the rate of rainfall at different points along the path. It is highly improbable that a path length of 30 to 50 miles will ever have a uniform rate of rainfall, and probable that somewhere along the path there will be no rain at all.

In calculations involving attenuation due to rain, it probably will be simplest to obtain rates of rainfall to be expected in the area under consideration. This information is readily available from the weather bureau for most parts of the world, and may be applied to table 4-2, or some similar source, to obtain the approximate expected attenuation. Nomogram B, "Attenuation by Rainfall", (appendix I) may be used, and will give the maximum attenuation that is to be expected. This nomogram is based on a uniform rate of precipitation along the total path length. If the maximum rate of rainfall is used in conjunction with the nomogram, the computed attenuation will be at a maximum. A figure of 75% of the attenuation, as given by this nomogram will usually be a safe value for any path length over 25 miles.

FOG. Fog may be considered another form of rain, as far as attenuation is concerned. Fog remains suspended

*_Water in the Atmosphere_, Part I of the Interim Report of the Ultra Short Wave Panel Working Committee, July 18, 1944.

TABLE 4-1: NUMBER OF RAINDROPS (PER M^3) FOR VARIOUS RAINS

	Distribution A	B	C	D	E	F	G	H	I
D, cm	*p,mm/hr* 2.46	3.6	4.0	6.0	15.2	18.7	22.6	34.3	43.1
0.05	28.5	476	752	61.4	3.33	.....	323	245	47.6
0.10	71.8	512	30.8	25.6	59.7	12.8	134	108	333
0.15	31	27	11.4	14	21.5	9.52	66	68.4	95.2
0.20	3.13	22	31.2	15.6	7.2	23.4	46.1	21.6	31.2
0.25	2.76	...	.....	4.0	0.96	0	28.3	21.5	0
0.30	.....	...	.....	7.2	0	25.3	10.2	17.6	0
0.35	.....	...	.....	....	3.83	0	3.35	0	0
0.40	.....	...	.....	....	4.48	5.75	2.3	0	0
0.45	.....	...	.....	....	0	.....	.....	11.3	22.5
0.50	.....	...	.....	....	2.71	.....	.....	.....	.....
Liquid water g/m^3	0.130	0.439	0.217	0.242	0.521	0.673	0.930	1.25	1.55

After A.C. Best

TABLE 4-2: RELATION BETWEEN ATTENUATION (DB/KM^3) AND PRECIPITATION

mm/hr	*λ, cm* 1.25	3	5	8	10	15	20	30	50	75	100	*Distribution*
2.46	1.93×10^{-1}	4.92×10^{-2}	4.24×10^{-3}	1.23×10^{-3}	7.34×10^{-4}	2.80×10^{-4}	1.52×10^{-4}	6.49×10^{-5}	2.33×10^{-5}	1.03×10^{-5}	5.85×10^{-6}	A
4.0	3.18×10^{-1}	8.63×10^{-2}	7.11×10^{-3}	2.04×10^{-3}	1.19×10^{-3}	4.69×10^{-4}	2.53×10^{-4}	1.08×10^{-4}	3.88×10^{-5}	1.72×10^{-5}	9.75×10^{-6}	C
6.0	6.15×10^{-1}	1.92×10^{-1}	1.25×10^{-2}	3.02×10^{-3}	1.67×10^{-3}	5.84×10^{-4}	3.02×10^{-4}	1.25×10^{-4}	4.34×10^{-5}	1.93×10^{-5}	1.09×10^{-5}	D
15.2	2.12	6.13×10^{-1}	5.91×10^{-2}	1.17×10^{-2}	5.68×10^{-3}	1.69×10^{-3}	7.85×10^{-4}	2.95×10^{-4}	9.23×10^{-5}	4.15×10^{-5}	2.35×10^{-5}	E
18.7	2.37	8.01×10^{-1}	5.13×10^{-2}	1.10×10^{-2}	6.46×10^{-3}	1.85×10^{-3}	9.09×10^{-4}	3.60×10^{-4}	1.20×10^{-4}	5.36×10^{-5}	3.03×10^{-5}	F
22.6	2.40	7.28×10^{-1}	5.29×10^{-2}	1.21×10^{-2}	6.96×10^{-3}	2.27×10^{-3}	1.17×10^{-3}	4.81×10^{-4}	1.66×10^{-4}	7.41×10^{-5}	4.19×10^{-5}	G
34.3	4.51	1.28	1.12×10^{-1}	2.32×10^{-2}	1.17×10^{-2}	3.64×10^{-3}	1.75×10^{-3}	6.83×10^{-4}	2.24×10^{-4}	9.95×10^{-5}	5.63×10^{-5}	H
43.1	6.17	1.64	1.65×10^{-1}	3.33×10^{-2}	1.62×10^{-2}	4.96×10^{-3}	2.29×10^{-3}	8.71×10^{-4}	2.78×10^{-4}	1.23×10^{-4}	6.98×10^{-5}	I

After A.C. Best

in the atmosphere, and the attenuation to be expected is determined by the quantity of water per unit volume, and the size of the droplets. Attenuation due to fog is of minor importance at wavelengths longer than 1.5 cm, since the droplets are exceedingly small, and the density rarely exceeds 1.0 gm/m^3, usually being less than 0.5 gm/m^3. Fog can cause serious attenuation by absorption, but only in the millimeter range.

SNOW. Scattering due to snow is difficult to compute, owing to the irregularities of the flakes. Attenuation from smooth, spherical drops, such as raindrops or hail, may be calculated quite easily from Rayleigh's formula for scattering:

$$\alpha_s = \lambda^2 \left(\frac{2\pi r_r}{\lambda}\right)^6 \qquad (4\text{-}1)$$

where

α_s = the attenuation due to scattering

λ = the wavelength

r_r = the radius of the drop.

Where the scattering object is snow, the parameter r_r is harder to determine, and should be an average dimension of the snowflake. While information on the attenuating effect of snow is limited, it is probable that attenuation from snow is less than from rain falling at an equal rate. This is borne out by the fact that the density of rain is eight times the density of snow. As a result, a rainfall of 1 inch per hour, for example, would have far more water per cubic meter of atmosphere than an equal snowfall, even with the lower terminal velocity of snow.

HAIL. Attenuation by hail is determined by the size of the stones, and their density. Attenuation of electromagnetic waves by scattering due to hailstones is considerably less than by rain, since ice has a lower index of refraction.

SLEET AND GLAZE. Sleet is defined meteorologically as very small pellets of ice, and as such has little effect on the electromagnetic wave in the frequency limits of this book. Glaze, defined meteorologically as rain that freezes on contact with any object, may be safely treated as rain having equal drop size.

ATMOSPHERIC CHARACTERISTICS

REFRACTIVE INDEX. The refractive index of the air is determined by its pressure, temperature and water vapor pressure, or relative humidity. Since these qualities vary with elevation, the refractive index also varies with elevation, so that any electromagnetic wave is continually traveling through a medium of constantly varying refractive index. In free space, or a uniform atmosphere, the electromagnetic wave would travel in a straight line. Thus we have an electromagnetic wave traveling in a straight line through the atmosphere which follows the curvature of the earth. Since the refractive index is gradually changing along the beam path, the beam becomes curved because of this changing index of refraction. This bending action takes place at all radio frequencies and at light frequencies as well, although to a lesser degree.

The change of the refractive index of air with increase in elevation causes the radio signal to be curved downward, and as a result the radio horizon extends beyond the optical horizon. The radio horizon compares with the optical horizon in that it is the point on the earth's surface where the transmitted beam is tangential to the earth's surface. The 4/3 earth's radius figure often used in microwave calculations is the direct result of refraction of the radio beam by the standard atmosphere.

REFRACTIVE MODULUS. In discussing the effect of the curvature in the atmosphere of a radio wave, only a standard atmosphere has been considered. When non-standard conditions exist, the radio beam will be refracted to a greater or lesser degree, depending upon the atmospheric changes that have taken place. Since these changes affect the refractive index at some elevation in the atmosphere, an allowance must be made for the decreasing temperature and pressure to be found as the elevation is increased. This is done by use of the formula

$$M = \frac{77.6}{T}\left(p + \frac{4810e}{T}\right) + 3.8H \qquad (4\text{-}2)$$

where

M = refractive modulus

T = temperature in degrees Kelvin

p = pressure in millibars

e = water vapor pressure in millibars

H = elevation above the earth in hundreds of feet

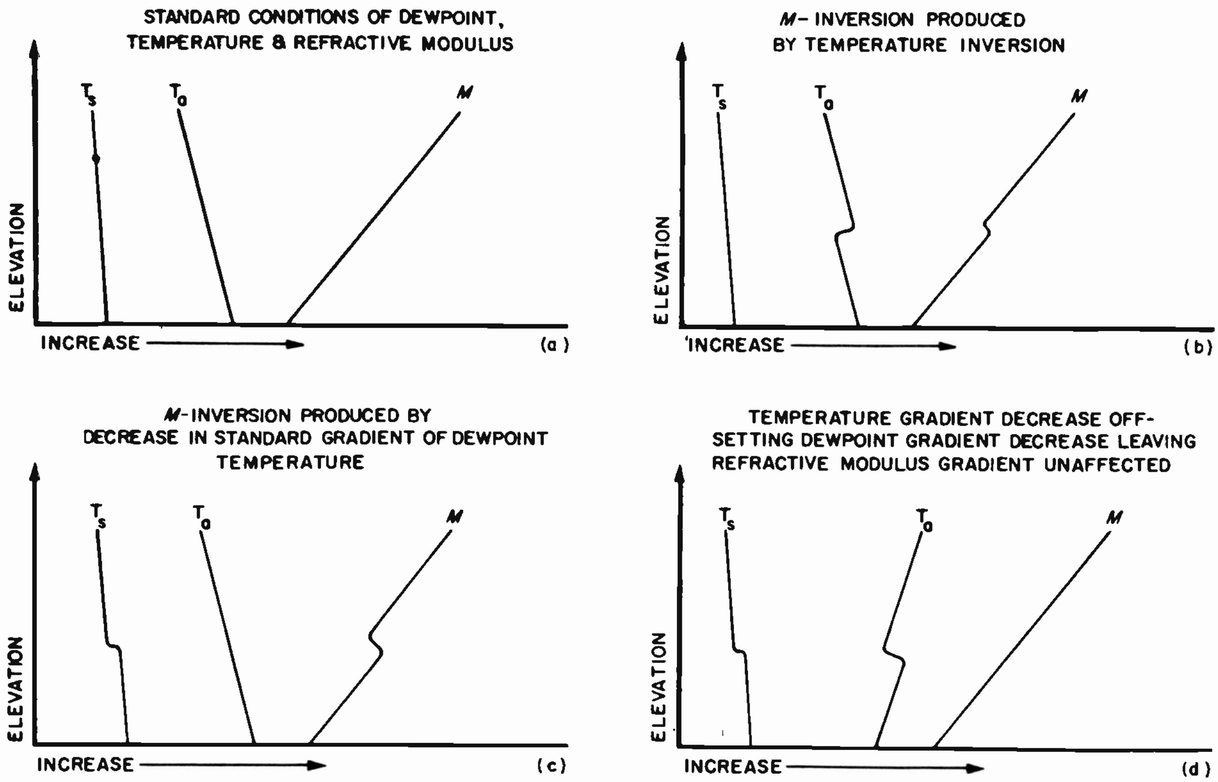
The value 3.8H is introduced to account for the earth's curvature.

Formula 4-2 gives the refractive modulus (M) of the atmosphere at any elevation where meteorological soundings have been taken. The values used for plotting the M-profile are rarely taken at elevations above 3,000 feet, since the radio signal will usually remain below this elevation, if it is to be received within a few hundred miles. (Ionospheric reflection of signals is an exception, and represents a different type of radio transmission.) While these values are taken at some specific point on the earth, they may be safely assumed to apply for the general area around that point, since atmospheric conditions do not change abruptly. The area will vary considerably in size, depending upon weather changes taking place at the time the tropospheric soundings are taken.

M-PROFILE. The M-profile is a straight line, with a slope of 3.8 per hundred feet, under standard atmospheric conditions. Where the plotted values of meteorological soundings do not fall in a straight line, atmospheric variations will be noted, and the extent of these variations will be indicated by the degree that the M-profile deviates from a straight line. The value of the refractive modulus at any one point is less important than the difference in the refractive modulus between two points that bound an abnormal condition. Where an abnormal condition exists, the differences in values for M are noted, and referred to as the M-gradient, usually expressed in terms of the difference per hundred feet. Where the M-gradient is less than standard, refraction toward the earth is greater than normal, while an above standard M-gradient is accompanied by less refraction, so that the curvature of the signal will be less than normal.

A below standard M-gradient, or M-inversion, may vary in three ways; namely, in elevation, in depth, and in intensity. By intensity is meant the amount the M-gradient varies from the standard.

Figure 4-1 shows four different conditions of the M-profile, together with the associated dewpoint temperature (T_s), and atmospheric temperature (T_a). Standard conditions of T_s, T_a, and M are shown in figure 4-1A. Notice that the slope of T_s and T_a are



T_s = DEWPOINT, WET BULB TEMPERATURE
T_a = TEMPERATURE ABSOLUTE, DEGREES KELVIN
M = REFRACTIVE MODULUS

Figure 4-1. Relation between Dewpoint, Temperature and M-Profile

such that at some elevation the dew point and atmospheric temperature will be the same, with completely saturated air resulting.

Figure 4-1B shows a temperature inversion, such as would be produced by a cold front or subsidence. The M-inversion occurs at the same elevation, and the degree of the M-inversion will be determined by the degree of the temperature inversion. Figure 4-1C shows an M-inversion produced by an increase in dewpoint temperature, with the degree of the M-inversion determined by the extent of the change from the normal dewpoint temperature. In figure 4-1D, a possible, although highly improbable, condition is shown whereby the increase in the M-gradient caused by an abnormal temperature decrease is offset exactly by a decrease in the M-gradient produced by a decrease in T_s. This will show approximately what to expect from the M-gradient, since abnormal changes in either T_s or T_a are almost always accompanied by changes in T_a or T_s.

CHAPTER 5

EQUIPMENT SITING

INTRODUCTION

Siting might very well be called the practical application of radio relay theory. All the various factors mentioned elsewhere in this book must be taken into account in order to lay out a radio relay system. It is entirely possible, of course, to set up a relay system by installing the repeaters on hills, without regard to any of the factors mentioned in this book. Such a system would probably work, too, but the chances of satisfactory and efficient operation are remote. Consideration of the many factors involved in radio relay work will produce a satisfactory relay system, using equipment that may be available. It is the purpose of this chapter to point out the problems encountered in siting, and explain the remedies commonly used in overcoming them. Chapter 14 gives a problem in siting, worked out according to the information contained in this chapter, so no attempt is made to give a definite illustration here.

SURVEY PROCEDURES

In locating a radio relay system, especially where maximum reception time is necessary, considerable attention to the location of relay stations is required, particularly in rolling country, or in areas where scattering or reflection may introduce problems.

While a straight line of relay stations, like a string of telephone poles, may appear desirable, there are two good reasons why the stations will usually take a zigzag course. One reason is that the terrain seldom lends itself to a straight line of repeater stations without having some hops unnecessarily short and others too long for satisfactory reception. Also, hills are rarely spaced so that the repeater stations may be aligned, and still have adequate clearance. A second reason for zigzagging stations is to prevent signal skipping. Since atmospheric conditions can produce vertical beam-bending, a signal beamed to an adjacent station may be picked up by another repeater farther away, assuming that the transmitting station and the other two repeaters are in a straight line, and use the same frequency. This produces interference and cross talk, since the skipping signal will more than likely be out of phase with the signal intentionally beamed at the receiving repeater.

For every rule, there are exceptions, and zigzagging is not invariably necessary. It is possible and practical to align repeater stations and still have reliable reception. Alternate polarization of antennas, a high front-to-back ratio for the antennas, and alternate frequencies may be used, either separately or in combination, to give satisfactory reception.

METHODS OF SITING. Siting may be done by a number of methods, or by a combination of methods. Contour maps, aerial photographs and aerial surveys may be used satisfactorily. Generally speaking, the sites are tentatively chosen from contour maps, and a field trip or aerial survey made to verify the choice. While the aerial surveys are generally made in airplanes, helicopters fitted with microwave transmitting and receiving equipment have also been found useful, especially where the terrain is too rough for vehicular travel.

POINTS OF CONSIDERATION. Some factors to be considered in locating the relay stations are: power supply, accessibility, type of ground on which the station is to be built, obstructions to be removed before building, and necessary antenna height in order to clear obstructions. Many of these factors may be estimated from contour maps, profiles made from the contour maps, and aerial photographs, thereby saving a great deal of time in the field.

Terminal stations for military use are usually located near the source of the messages to be transmitted, unless the terrain is such as to require remote installation of the terminal transmitter. A good illustration of such a condition is a military encampment located in a valley, or behind a hill. If the terminal station is located away from the source of the messages, a wired link may be installed to the transmitter, or the messages may be transmitted by a low powered transmitter beamed to the terminal station. The latter method is recommended where the cost of running the transmission line is greater than the cost of a radio transmitting link. Such a system is frequently used in commercial installations where the studio of an FM station, for example, is located in a large city, and the transmitting antenna placed on a hill nearby.

USE OF CONTOUR MAPS. Contour maps are valuable in laying out sites for the relay stations, since both elevation and distance between the sites may be readily determined. Sites may be tentatively chosen on the basis of information taken from the contour maps, and profiles of the terrain beneath the radio beams may then be drawn to determine the clearance above hills or other obstacles. In using these profiles to determine path clearance, both the curvature of the earth and the refraction of the radio beam in the troposphere must be considered. If the profile is drawn on rectangular coordinate paper, or graph paper, corrections must be made for the earth's curvature, and refraction of the beam in the air. 4/3 earth's radius paper, when available, is recommended, since corrections for curvature and refraction are taken into account in the paper itself (See Appendix I). After plotting the profile, a straight line may be drawn between the antennas at the repeater sites, and any obstacles to the radio beam will become apparent. The attenuation caused by the obstacle may be determined from the nomograms in the appendix, and an alternate path chosen if excessive attenuation is found. Some knowledge of the accessibility of the sites may be acquired from the contour maps, but field inspections should be made in order to determine the amount of work required to clear a road to the site for inspection and maintenance.

FIELD INSPECTION. Since siting, as done on contour maps, is easily changed before the relay stations are built, it is advisable to locate several possible sites for alternate location of each relay station. Some of these alternates will be discarded on the basis of information obtained from the profiles, while the most desirable can best be determined from information acquired in the field.

The field trip should be for the purpose of locating accessible roads, determining whether the ground is suitable for the erection of the antenna and building for the equipment, and the amount of clearing needed to prepare the site. Field parties equipped with binoculars may operate in pairs, in order to check the actual line between the two points. This is especially important where the path clearance may be subject to question. Where portable beam-antenna equipment is available, a further check on signal attenuation may be made by the two parties between adjacent sites.

SITING BY RADIO ALTIMETER. A second method of siting repeater stations is by means of an airplane equipped with an aneroid barometer (pressure altimeter) and/or a radio altimeter. The plane may be used to provide profile maps between tentative relay sites. These maps can be made from a recording altimeter or from frequent readings taken along the path with an indicating altimeter. The difference in readings between the aneroid altimeter and radio altimeter can be used to determine the elevations of the terrain along the beam path, and the values plotted on 4/3 earth's radius paper, as previously mentioned. In using these data, just as with profiles developed from contour maps, the curvature of the earth and average refraction will be taken into account. Earth's curvature, and atmospheric refraction, will frequently elevate some obstruction several hundred feet higher than it appeared on a plane-earth, or rectangular coordinate profile, especially where a long hop occurs. Photographs made from the plane would be of use in determining roads, power lines, or other details that might affect the final choice. An inspection of the site is still desirable, however, since some problem concerning the use of the site might arise that could not be definitely solved either in the office or by an aerial survey.

SITING BY HELICOPTER. The newest and quickest method of siting is by use of helicopters. Tentative sites may be chosen from contour maps, or actually worked out in the field by observers in the helicopters. Even field trips may be eliminated, since the observers can easily examine the sites from the helicopter, or descend to the ground by means of a rope ladder for closer inspection. Antenna heights may be worked out to determine the minimum antenna height required; even the road to be built to the site, if one is needed, may be located from the helicopter.

The equipment required for siting by helicopter is as follows: a compass for triangulation, binoculars, two-way radio communication between the helicopters, a beamed microwave transmitter situated on one helicopter, and a receiver on the other, an aneroid altimeter or plumb line, preferably both.

The transmitter is equipped with a parabolic antenna on each side of the helicopter, each mounted to turn more than 180 degrees horizontally. This is necessary, since the helicopter must head into the wind to hover, and unless a full 360-degree coverage is available, the transmitted beam may not be aimed properly at the receiving helicopter. The transmitted beam should be fairly wide, since precise alignment on the receiving helicopter is difficult, and the vertical beam should be at least 90 degrees wide, to facilitate picking up the signal in rolling or hilly terrain.

The antennas mounted on the receiving helicopter must also be arranged for a full 360-degree horizontal coverage, but a much smaller beam angle may be used. Received signal strength may be estimated aurally by an audio note modulating the carrier, or visually, by means of an oscilloscope.

Observations are carried out as follows: Tentative sites are located by triangulation on known landmarks, and the helicopters hover above the sites to be inspected. The transmitted signal is beamed toward the receiving helicopter, and the signal strength noted. The helicopters may then descend, until the signal is attenuated by some obstacle, and the minimum antenna height noted, either by radio altimeter or measured plumb line. This will complete all necessary observations regarding signal attenuation between the two sites, and the transmitting helicopter may then proceed to the site beyond the one occupied by the receiving helicopter for another set of observations. By this method, the sites may be found adequate or inadequate, and a satisfactory alternate site quickly established if necessary.

SITING CONSIDERATIONS. In the ensuing discussion on siting considerations, an attempt is made to review the various phases of the siting problem. The relative weight of these considerations is not fixed, but will vary according to the conditions of the area where the repeater is to be located. Certainly, if only one location is available for a repeater station, that location will be used, whether or not there is a convenient power line and road nearby. By the same token, a repeater system operating on storage batteries charged from a power line could continue to operate for several hours after a power failure, and such a type of system would have less need for standby equipment than a station powered directly from the power line. In other words, these recommendations are intended for general guidance, and must not be considered as hard and fast rules.

The most important requisite to be considered in choosing the repeater site is accessibility. Occasionally a microwave repeater may be located near a fire-tower, or some other form of observation post. Where this can be done, some sort of road to the site will already exist, and usually a source of power as well. Many roads built to reach fire-towers are not designed for year round use, and will require additional work if they must be used in all seasons.

Where no road exists to a proposed site, a field trip will determine the amount of work necessary to build a road. For rolling terrain and a mild climate, road construction will be rather easy, while in rocky, hilly terrain the problem of road construction may be difficult enough to eliminate a site otherwise entirely satisfactory. Another consideration is the damaging effect of weather on roads. In hilly country, threats of landslides and washouts could increase the initial cost, as well as the cost of maintenance. In flat, marshy country a site might be discarded because of the expense of the fill, and the additional work necessary in maintaining such a road.

The quality of the access road will depend considerably on the amount of use it will receive. If the receiver must be powered by a local generator, frequent trips with fuel will be necessary, and the road must be built to withstand this extra use. The conditions governing the road construction, then, will depend on climate, terrain, and the amount of use the road will receive. In order that the repeater stations may be kept operating satisfactorily, these factors must be taken into account in building the road.

It frequently happens that the final choice between the two sites is determined by the power supply available. The reliability of the repeater, and therefore of the whole system, is dependent upon the source of power. If a power line is available at the site, which is highly unlikely, it should be examined carefully to determine whether it is capable of carrying the additional load of the repeater. The power should be regulated, both in frequency and voltage, and the wires large enough to carry the extra load without voltage variation due to line loss. Quite often a power line may be run to a desirable site from a nearby power line without excessive expense. In any case, a site requiring a local power supply, such as a motor generator, should be avoided if possible, because of the additional equipment maintenance required.

The advantage of a radio relay system lies in its freedom from breakdown by weather, but if the power supply should fail due to an ice or windstorm, the advantage is lost. A standby power plant installed in the repeater station to take over in the event of power failure by the regular source will overcome this handicap.

SUMMARY. The profiles will give most of the information required concerning the terrain beneath the beam path, but field trips are important for a final check on conditions, and to disclose any possibility that might give trouble, but is not otherwise shown on the profile or contour map. Contour maps rarely give tree heights, and in many instances some man-made object, such as a power line, or building, may have been erected after the map was drawn. Seasonal problems may also be present, yet not show up on contour maps. A bare hill, for example, may be a cornfield in certain seasons, and thus present a different problem in attenuation than it would if used as a pasture, or hay field. A flat area may appear on a contour map, and apparently be no problem, but inspection by a field trip might show that during a rainy season it would fill with water, and perhaps cause a reflected signal to appear at a receiving antenna. These field trips, then, serve two useful purposes, to verify calculations made from maps and aerial photographs, as well as to discover any source of trouble that might not appear on the maps. Needless to say, a trained field crew could find such sources of trouble easier than an untrained crew, and find them much sooner.

SYSTEM RELIABILITY

GOVERNING FACTORS. Like the automobile salesman, expounding on the various gadgets of the new car and forgetting the engine, it is easy to overlook the main purpose of the relay system. Just how well is it going to work, once it's installed? Having satisfied all of the many requirements in selecting sites for the repeater stations, one question still remains; how much of the time will the signal be intelligible? This may be figured out with reasonable accuracy from available data, part of which is found in the training manual (TM) on the equipment, and part from Federal Communications Commission (FCC) charts.

There are many factors to be considered in estimating the reliability of a relay system, other than failures of the power supply or equipment. Transmitter output, distance between stations, and receiver sensitivity are some of the characteristics that will affect the reliability of the system. Other factors, such as reflected or refracted signals, fades due to ionospheric storms, and even seasonal changes in the terrain can affect the received signal strength, and thereby affect the reliability. A formula for taking these factors into consideration would be very complicated, and as a result, many of these variables are eliminated in careful choice of the repeater site, while others can be included in a "cushion", or allowance for fades.

One way in which high reliability can be obtained is to decrease the distance between repeaters. More powerful transmitters may be used, or larger parabolic reflectors for the antennas, with a correspondingly larger gain. Such solutions, however, are expensive, and tend to become unsound economically. Repeater spacing of 30 to 50 miles is generally found to be satisfactory for frequencies below 10,000 mc, provided good path clearance (first Fresnel zone or better) is available.

So far in this discussion, everything appears to indicate satisfactory performance. It still remains to determine just how satisfactory it will be. It is not difficult to estimate the signal strength arriving at the receiver of the repeater station. Notice that this is the signal at the receiver, not at the antenna. Such a signal will be affected by the receiver antenna gain, and the transmission line loss between the antenna and the receiver. In order for the signal to be intelligible it must be stronger than the receiver noise level. The manufacturers of receiving equipment determine a signal level for their equipment which depends upon the type of modulation used, and the receiver design. This level is called the *threshold level*, and represents the lowest signal level that will be intelligible. Although the signal may be just barely above this level, satisfactory transmission will result, but with no allowance for fades. Since fades will lower the signal strength at the receiver, it is necessary to have a much stronger signal arrive under normal conditions, in order that an intelligible signal may be received under expected fade conditions. The difference between the field strength of the level normally received, and the threshold level, is called the fade margin. The field strength to be expected for any repeater site may be calculated within ± 5 db, using information on the equipment to be installed, and data taken on the individual links of the relay system.

ACCEPTABLE RELIABILITY. At this point, we can discuss the reliability to be expected from the system. The reliability is determined by the percent of the time that the signal strength at the receiver is above a predetermined level. This level may be found by applying the various factors concerning the repeater link to the nomograms given in the appendix, and is the signal strength to be expected 90 percent of the time. It should be emphasized that this level is independent of the receiver characteristics and fade conditions to be expected in the area where the link is to be installed. Should the signal strength be equal to the threshold level of the receiver, the outage time would be 10 percent, obviously very unsatisfactory operation. No relay system would be planned to operate at the receiver threshold level, however, since a fade of any magnitude would immediately result in an outage. By designing the links so that the received signal is higher than the threshold level of the receiver, a cushion to absorb fades is available, and as a result, higher reliability may be obtained. Assume, for example, that the received field strength at the receiver is -10 dbm, and the receiver threshold level is -50 dbm. The difference, 40 db, is the fade margin for the link. On the basis of previous experience in radio, it has been found that a 40 db fade may be expected not over 0.01 percent of the time, so that the 40 db fade margin gives a reliability of 99.99 percent. It is possible to have still higher reliability, but the cost of additional equipment necessary would hardly warrant the slight increase to be obtained. Nomogram C in the appendix gives the relationship between reliability and fade margin, so that the relay system may be engineered for whatever reliability is desired. In most instances, the path attenuation will be the greatest source of attenuation, so that where a lower reliability is acceptable, the repeater stations may be spaced farther apart.

In calculating the reliability to be expected from a relay system, the simplest, and safest, method is to assume that outages do not occur simultaneously. For example, if a repeater system has 10 links, with an efficiency of 99% for each link, it is safe to estimate that each link will be out 1% of the time. Since the outages are assumed to occur at different times, the total outage time for the ten links would then be 10%, so that the system as a whole would have a reliability of 90%. As the number of links increase, the likelihood of the outages occurring simultaneously also increases, so such a figure represents the blackest possible condition.

CHAPTER 6

TRANSMISSION LINES

INTRODUCTION

A transmission line may be likened to a belt system, in that it is merely a means of transmitting power from its source to the place where it may be used. The energy to be transmitted may be power to operate a factory, or to supply a city. It may be a telephone conversation between homes in a city, or a radio program transmitted from the studio to the transmitter. Each type of use has different problems in transmission, and only the problems concerned with radio relay systems will be considered here.

BASIC THEORY

ELECTRIC AND MAGNETIC FIELDS. While there are many different types of transmission lines used in radio work, all may be considered modifications of a two-wire line. Such lines, carrying direct current, present no difficulties other than the loss of power due to resistance in the wires. If the wires are sufficiently large, the power loss in the line will be negligible, and no particular problem will exist. There is a problem, however, when the rate of current flow through the line is changed. Consider a long pair of wires connecting a battery to a lamp (figure 6-1). Before the switch is closed there will be no current flow in the wires, nor will there be any magnetic field surrounding the wires. At the instant the switch is closed, current will start to flow outward, and a number of changes will take place. As shown in the figure, the current flow has started toward the lamp. The wires become charged just as a capacitor becomes charged, and an electric field will appear between the wires, exactly as in the dielectric between capacitor plates. This will be a moving electric field, since the current is flowing along the wire and establishing the field as it flows along. Another type of field will also exist. Each wire will be surrounded by a magnetic field, with the field being strongest between the two wires since each wire contributes to the field in the space between them.

Now, as the current flows outward along the wire, the two fields move outward with it, the magnetic field circling the wire, and the electric field normal to the wire. These fields will continue to move until the wire is completely charged, that is, until the current has reached the lamp, which will then light. At this time the flow of current is uniform along the line, the magnetic and electric fields are established and are stable. The fields will continue to exist as long as current flows through the wires, and will remain unchanged until some change takes place in the rate of current flow.

After the stable condition is reached, let the switch be considered as open. Current will cease to flow from the battery, but for an extremely brief interval the lamp will continue to light. The current in the wire will continue to flow due to the electric and magnetic fields surrounding the wires. These fields which were originally created by the current flow do not represent expended energy, but rather, stored energy. Since the source which established and maintained these fields has been removed, the fields will collapse. In collapsing, the fields restore to the line the energy originally used in creating the fields. Therefore, some part of the power to operate the lamp may be considered as being stored in the space surrounding the wires, and this power is referred to

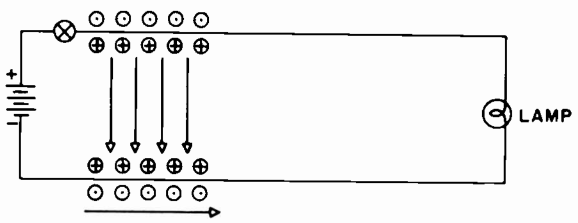


Figure 6-1. Field Created by Current Flow

as an induction field, or the field produced by the current flow (see Chapter I).

CAPACITANCE AND INDUCTANCE. According to elementary theory, a magnetic field is produced by current flowing through an inductance, and an electric field by current flowing into a capacitance. Since these fields exist in the region surrounding the transmission line, obviously some capacitance and inductance must exist in the lines. The inductance is present, since the wires act as a coil of one turn, and the capacitance is also present due to the two wires acting as individual plates of a capacitor. For direct current work, or alternating current power (60 cycles), the capacitance and inductance of the line are negligible except in the case of very long a-c lines. At radio frequencies, particularly at the higher frequencies, capacitance and inductance in transmission lines can be a serious problem.

In many problems of circuitry, the evenly distributed capacitances and inductances may be lumped, that is, considered as being concentrated in one or two places. In a transmission line, lumped constants would mean that instead of having the inductance and capacitance distributed evenly, they would be located at different points along the line; with equivalent values substituted. While this would simplify calculations, serious errors in the impedance of the line would result, and unsatisfactory calculations would be obtained. In explanations, however, lumped constants are frequently used, in order to explain the results of the line impedance, figure 6-2.

In the illustrations mentioned, where the direct current flows along the wire toward the lamp, certain conditions were described. These conditions were produced by the movement of the electrons along the wire, and took place only during the rise and decay of the current. In an alternating current flow, the value of the current is changing continuously, and thus an accompanying change is taking place in the induction field.

Figure 6-2 shows lumped constants for a short length of transmission line, which could be considered as continuing indefinitely. Alternating current flowing through the line will create both electric and magnetic fields, just as were created in the example given for the d-c line, but with the difference that the fields will be continually changing in value and direction. In this figure, two values of resistance are added, one in series with the line, and one across the line for each group. The series resistance represents the resistance of the line itself, the resistance placed across the line representing the conductance of the dielectric surrounding the wires. The series resistance may be kept low by the use of suitable wires, and the conductance is generally negligible if a satisfactory dielectric, such as air, can be used. The capacitance as shown represents the capacitance between the two wires, although another also exists, not shown. Each wire has capacitance with itself.

In order to explain the capacitance of a wire with itself, the length of the wire must be taken into account. This length may be measured either in terms of physical length, such as feet, or miles, or it may be measured in terms of its electrical length. One wavelength is the unit of electrical length, and it may be 3,100 miles for 60-cycle current, or 118 inches for 100 mc current. In both cases, the wire length is one wavelength for the current carried. Thus, in discussing the self-capacitance of a wire, the wire length and wavelength of the current must be considered. If the wire is long electrically, that is, in terms of wavelength, there may be several points along the wire where the voltage is momentarily at a maximum or minimum. Since this produces two or more points of unlike potential, a capacitance will exist between these two points. It also follows that such capacitance not only exists between the two points of maximum and minimum potential, but also between any two points where a difference of potential exists. Therefore, the wire may have self-capacitance.

The value of the self-capacitance is determined by the frequency. In the case of 60-cycle power, the points of maximum and minimum potential will be so far apart as to render the self-capacitance of the wire negligible. As the frequency increases, however, the physical length of the wire in terms of wavelength decreases, until at UHF self-capacitance becomes an appreciable part of the transmission line problem.

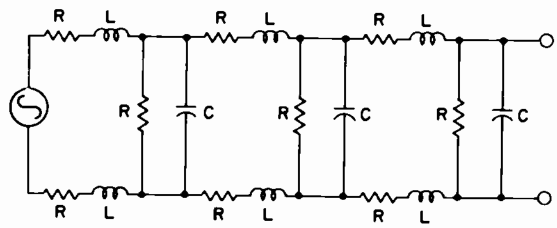


Figure 6-2. Lumped Constants

The plates of the capacitor, in other words, are much closer together at UHF than at long wave frequencies, with the result that the capacitance will be correspondingly greater.

One other source of capacitance may exist, and that is the capacitance between the wires and ground. This source is negligible in coaxial cables or waveguides, and is less important than the capacitance between the wires themselves.

LINE TERMINATIONS. The length of a transmission line, and the termination or load fed by the line, has a considerable effect on the wave conducted by the line. If a line of infinite length could be built, the transmitted wave would be carried onward forever, (assuming that there was no loss of power due to radiation, leakage between the conductors, or resistance in the wires). The amplitude of the voltage would be determined by the potential applied to the transmission line, and the loss in the line between the point of interest and the source. The important thing to consider in an infinite line is the fact that there is only one wave traveling along the line, and that the wave travels outward from the source. An infinite line, having only one wave, is known as a non-resonant line. Any other line of finite length that is terminated in such a way as to have no reflected wave is also called a non-resonant line.

The conditions where the line is not infinite, but of some definite measurable value, will affect the wave movement, and the manner in which the line is terminated will also affect the wave. For example, consider the transmission line as being any whole number of wavelengths long, and terminating in an open circuit. When the transmitted, or incident, wave reaches the end of the transmission line, the open circuit acts like a mirror, reflecting the incident wave back on itself. This reflected wave is actually a continuation of the incident wave, so that two traveling waves exist on the line. Since these waves both appear upon the transmission line at the same time, a resultant wave will appear, which can be measured both for amplitude and wavelength. This resultant wave is called a *standing wave*, since its nodes and maxima remain always at the same point along the transmission line. Since the incident and reflected waves travel at the same speed, but in opposite directions, the amplitude of the standing wave at any point is continually changing, but the nodes and maxima do not travel along the transmission line as they do in the case of the incident and reflected wave.

Continuing the discussion of the open-circuited line, the next consideration is the value of the standing wave voltage at some particular point along the line. This value will be maximum at the open end of the transmission line, since at this point the impedance will be very high. From this point back towards the source, the voltage will decrease, reaching a minimum ¼ wavelength from the end, then increasing to a maximum ½ wavelength from the end, and so on back to the source. In the event that the transmission line terminates in a short-circuit, rather than an open circuit, the voltage at the termination would be zero, since the extremely low resistance of the short circuit would cause the voltage to drop, just as in any d-c circuit.

The amount of variation of the rms voltage developed on a transmission line because of wave reflections is called the voltage standing-wave ratio, VSWR. Such a quantity is defined as the ratio of the maximum rms voltage to the minimum rms voltage on the line. Transmission lines whose terminations have the same magnitude of reflection coefficient (ρ) will also have equal values of VSWR. For a maximum transfer of energy, it is essential that the VSWR of a transmission line be as low as possible, approaching unity for well-matched systems. A method for measuring the rms standing-wave voltage is given in chapter 16.

So far, other than to mention the fact that a transmission line possesses resistance, inductance, and capacitance, the effect of line impedance has been ignored. It may be considered at this point, however, to illustrate the current flow along the transmission line. The value of the current at any point is determined by the impedance at that point, which in turn is determined by the type of termination of the line, just as in the case of the line voltage. An open-circuited line, as stated above, has a maximum voltage at its end. The standing current wave will be zero at this point, since the open end acts as an extremely high resistance and the current flow at this point will therefore be negligible. At each half wavelength, then, measured back from the open end of the transmission line, the voltage standing wave will be a maximum, and the current standing wave will be minimum. The intermediate values will vary sinusoidally, according to the impedance of the line, and the impedance may be determined at any point along the line by measuring the voltage and current at that point.

Figure 6-3A shows a transmission line one wavelength long, open-circuited at the end. The source is shown at the left, with the impedance indicated above the line. At the outer end, the impedance, consisting of resistance, is very high as shown in the impedance curve. One-quarter wavelength back from the open end, the resistance is low, as shown. As a result, the

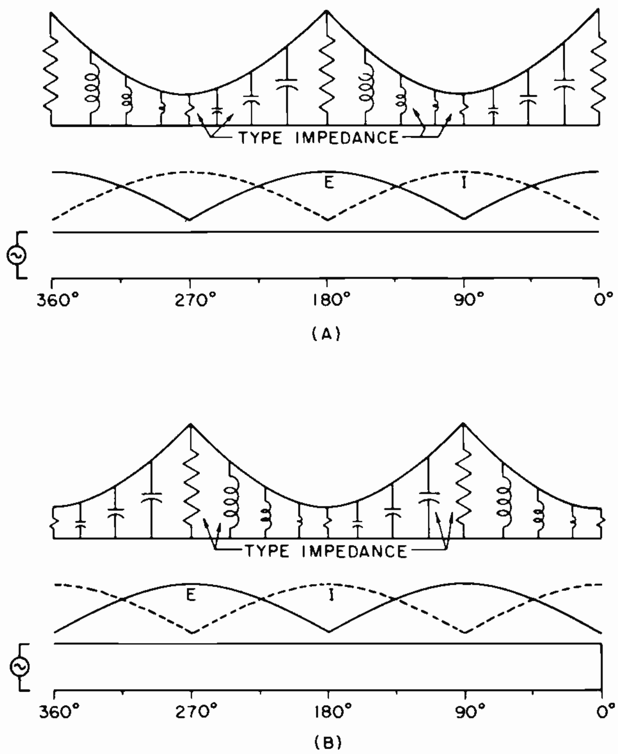


Figure 6-3. Open and Shorted Transmission Lines

current at this point is high, and the voltage is low. It should be noted that the impedance curve does not reach the base line at any point. This is caused by line losses due to dielectric leakage between the conductors, and resistance in the wires of the transmission line. The voltage and current, also, never quite reach the zero point because of these same losses. At each point along the line, the value of the current, voltage and impedance is as indicated, with the type of impedance also indicated. The diagram does not present a true picture of the resistance and reactance of the line, which is evenly distributed along the line, but it does show the effective impedance.

The actual length of a UHF transmission line is rarely as short as a single wavelength, usually being many wavelengths long, plus some fractional part of a wavelength. It is this fraction, and only this fraction, that need be considered in discussing the length of the transmission line. Since the standing wave is reflected back from the end of the transmission line, the transmitter or source of power feeds a line that is in effect only that fraction of a wavelength. Referring to figure 6-3A, a source is shown one wavelength from the open end, but not connected. If the source were connected at that point, it would be directly below the large resistance shown on the impedance curve. In other words, the source would be feeding a load consisting of pure, high resistance. As a result, the current would be low, since the high resistance would limit the flow of current, and the voltage would be high. Now assume that the source is connected at the 270° point, and that the transmission line to the left of that point does not exist. The source will now be feeding on equivalent load of low, pure resistance. As a result, the current will be high, and the voltage low, just as shown on the voltage and current curves. For any length of line, then, the load fed by the source will be of the type indicated on the impedance curve at that point, and the current and voltage relations will also be as shown. If the transmission line is terminated in a short circuit, as shown in figure 6-3B, the impedance, voltage and current relations will be as shown. It will be noted that the short-circuited line appears to the source as an open-circuited line, ¼ wavelength longer than the open-circuited line.

A discussion based on infinite, open-ended, and short-circuited lines may seem odd, since transmission lines are usually considered as being of measurable length, and carrying a radio signal to a load. Actually, the lines may be used for much more than the transmission of power. The infinite line, for example, has no standing wave, since there is no point of reflection to return the incident wave. As a result, all the power from the source continues outward. Such an arrangement is desirable in a transmission line feeding an antenna, since all the power from the transmitter should be fed into the antenna. If the impedance of the antenna matches that of the line, then the line acts exactly as though it were infinite, with no reflections. Open-ended and short-circuited lines may be used as filters to remove unwanted harmonics. "Metallic insulators", transformers and test equipment (chapter 16) also are based on transmission lines.

CHARACTERISTIC IMPEDANCE. Every transmission line, whether two-wire, coaxial, or waveguide, has a certain characteristic impedance. This impedance is constant for each type of line, and is determined by the size of the conductors, their spacing, and the dielectric. In the case of coaxial lines using beads as spacers, the type of the material in the beads will also affect the characteristic impedance. While there are no conductors in a waveguide in the usual sense of the word, waveguides still possess characteristic impedance, just as in the case of any other type of transmission line.

In the beginning of this chapter it was shown that when a battery was connected to an uncharged pair of long wires, the voltage between the conductors does not jump instantly to a maximum at all points of the line. Instead, a steep voltage wavefront travels at a definite speed towards the far end of the line. If the inductance in henrys per meter and the capacitance in farads per meter distributed along a line are represented by L and C, respectively, the wavefront travels at a velocity of

$$v = \frac{1}{\sqrt{LC}} \text{ m/sec} \tag{6-1}$$

and reaches the lamp after a time interval depending on the distance between the battery and the lamp. Generally, the time interval is extremely short as the speed of propagation along transmission lines is of the same order of magnitude as the speed of light.

The voltage wavefront (V) is always accompanied by a similar current wavefront (I) both traveling in phase, with the ratio of voltage to current being a constant at any point on the line. This constant is called the *characteristic impedance* Z_0, and in a lossless line is given by the relations

$$Z_0 = \frac{V}{I} = \sqrt{\frac{L}{C}} \tag{6-2}$$

Of fundamental importance is the fact that a wave of *any shape* is propagated without change in shape or magnitude along a uniform infinite line at a constant velocity since the characteristic impedance ratio $\sqrt{L/C}$ remains constant. The characteristic impedance of the line must be matched at both the transmitter and antenna (or receiver and antenna) for maximum efficiency. Theoretically, such matching would produce an equivalent infinite line, with no standing wave; actually, there will always be a standing wave, due to objects near the line or antenna that may affect the characteristic impedance.

LOSSES AND ATTENUATION. In the foregoing discussion, the characteristics of transmission lines are based upon the assumption that the lines are lossless. Actually, a small amount of power is always dissipated along a transmission line, such losses including the following:

a. Conductor loss, or power lost as heat because of current flow through the resistance of the conductors.

b. Dielectric loss, or power loss because of currents flowing in imperfect insulating material between conductors, usually stated in terms of the conductance of the dielectric.

c. Losses caused by radiation and by interaction of the fields with adjacent objects and circuits. These losses are eliminated in the case where the fields surrounding the conductors are confined by a conducting shield, or in the case where coaxial lines are used.

Where such losses occur, the wave is propagated in the same way as on a lossless line, in phase and at a velocity as given by formula (6-1) with an amplitude ratio equal to the characteristic impedance of the line. However, as the wave progresses along the line, its power is continually decreased by the line losses. Likewise, the amplitudes of the voltage and current also decrease or attenuate as the distance from the source increases. Thus, the effect of a transmission line is to reduce the power of the traveling wave in a definite ratio.

If the input power is designated by P_1, and the load power by P_2, the attenuation of the line may be specified by the ratio P_2/P_1, or more conveniently by the expression

$$\alpha_s = \frac{1}{2} \ln \frac{P_1}{P_2} \tag{6-3}$$

where

α_s = Attenuation in nepers of a line of length s, or α = nepers per unit length.

ln = Napierian logarithm ($\log_2$)

The neper, like the decibel, is a logarithmic measure of a power ratio and is equal to 0.1151 times the number expressing the same attenuation in decibels. Given in decibels, formula (6-3) for the attenuation of a line of length "s" becomes

$$\alpha_{db} = 10 \log_{10} \frac{P_1}{P_2} \tag{6-4}$$

At the same time that the amplitudes of the voltage and current decrease with distance, the phase angles also change in proportion to the distance. The combined effects of attenuation and propagation velocity can therefore be expressed as a constant, called the propagation constant (Γ) given by the equation

$$\Gamma = \alpha + j\beta \tag{6-5}$$

The propagation constant is made up of a real (resistive) term called the attenuation constant (α), and an imaginary (reactive) term known as the phase constant (β). The attenuation constant (α) is generally measured in nepers per meter while the phase constant (β) is equal to 2π times the reciprocal of the wavelength measured in meters.

Line losses also affect the characteristic impedance of a transmission line, and for most cases encountered in high-frequency practice Z_0 may be expressed by the formula

$$Z_0 = \sqrt{(R + j\omega L)/(G + j\omega C)} \text{ ohms} \quad (6\text{-}6)$$

where

L = Line inductance in henrys per unit length

C = Line capacitance in farads per unit length

R = Line resistance in ohms per unit length

G = Dielectric conductance in mhos per unit length.

The value of the attenuation constant now becomes

$$\alpha \cong \frac{R}{2Z_0} \text{ nepers/meter} \quad (6\text{-}7)$$

and the value of the phase constant will be given by

$$\beta \cong CZ_0 \text{ radians/wavelength} \quad (6\text{-}8)$$

where the values given by the formulas are used to determine various transmission line characteristics.

It was stated earlier that the propagation constant also accounts for the velocity of propagation of a wave along a transmission line. Electro-magnetic waves traveling along a transmission line have a definitely slower velocity, and a correspondingly shorter wavelength than that obtained in free space. The following expression may be used to determine wave velocity in any line, and consequently the wavelength.

$$\nu = \frac{1}{\beta} \text{ meters/sec} \quad (6\text{-}9)$$

The difference between the velocity of propagation in a transmission line and the velocity of light is usually expressed as a ratio for any particular instance, and is generally called the velocity constant (k), or

$$k = \frac{\nu}{3 \times 10^8} = \frac{1}{3 \times 10^8 \beta} \quad (6\text{-}10)$$

where k is always less than unity except in the case where waveguides are used. For example, the velocity constant of the RG8/U cable is 0.659 or 65.9%. The velocity of a wave in the RG8/U cable therefore becomes $0.659 \times 3 \times 10^8$ or 1.97×10^8 meters per second, and the wavelength in the cable becomes 0.659×600 or 395.4 centimeters, 600 cm being the free-space wavelength.

PRACTICAL TRANSMISSION LINES

Transmission lines may be generally classified into five general groups, as follows:

1. Two-Wire Line
2. Four-Wire Line
3. Shielded Pair
4. Coaxial or Concentric Cable
5. Waveguide

TWO-WIRE LINES. The oldest and simplest type of transmission line is the familiar two-wire line, such as that mentioned in the early part of this chapter. A two-wire transmission line consists of two parallel wires or tubes uniformly spaced by means of insulators or spreaders, with a spacing between the wires smaller than the operating wavelength.

Since the characteristic impedance of any transmission line is a function of its distributed inductance and capacitance, an increase in the spacing of the wires increases L and decreases C, thus increasing Z_0. Similarly, a reduction in the diameter of the wires also increases the characteristic impedance, for it is equivalent to decreasing the size of the plates in a capacitor in order to obtain lower capacitance. Changes in the dielectric material between the two wires also changes the characteristic impedance of the line due to a change in the distributed capacitance and conductance.

For non-loaded open wire lines where the dielectric material between the lines is air, the value for Z_0 may be found by the formula

$$Z_0 = 276 \log_{10} \frac{2D}{d} \text{ ohms} \quad (6\text{-}11)$$

where

D = Distance between the centers of the wires

d = Diameter of the wires

This formula is sufficiently accurate at high frequencies where the value of Z_0 is practically pure resistance, and may be found in nomogram form in appendix I (nomogram D).

Another type of two-wire line is the twisted pair, consisting of two insulated wires twisted to form a flexible line without the use of spacers. Such a line is not generally used at frequencies higher than about 30 mc due to extremely high radiation losses and leakage losses between conductors.

FOUR-WIRE LINES. One type of four-wire line resembles a pair of two-wire lines with the four wires uniformly spaced. Such lines have the advantage of reduced radiation and less mutual effect on nearby lines when diagonally opposite wires are connected in parallel. An improvement on the uniformly spaced system is the spiral-four line, where four wires instead of two are twisted together to form the transmission line. Spiral-four lines are used mainly to carry audio frequencies, and, like the parallel four-wire line, diagonally opposite wires are connected in parallel.

SHIELDED PAIR. The shielded pair consists of two parallel conductors separated from each other by a dielectric material. The conductors are contained in a copper-braid sheath which, in turn, is covered with a flexible coating of rubber to protect the line from moisture or friction. Thus, the shielded pair closely resembles the sheathed cable used for electric power wiring.

The shielded-pair transmission line possesses an outstanding advantage in having its two conductors balanced to ground; that is, the capacitance between each conductor and ground is uniform along the entire length of the line. Since the wires are completely shielded, losses due to radiation are eliminated as well as interaction with adjacent objects and circuits.

Radiation from an unshielded line may also be prevented provided the current flow in each conductor sets up a field equal and opposite to the field set up by its paired conductor. This condition may be obtained if the line is sufficiently removed from any conducting objects, and if the spacing between the wires is small. If a conducting object is near the transmission line, a certain amount of capacitance exists between the object and the wires of the line. Unless this capacitance is identical for both wires, the current flow will be greater in the wire nearest the conducting object, resulting in unequal current flow in the two conductors. This unequal current flow will produce incomplete cancellation of the electromagnetic fields resulting in radiation.

COAXIAL CABLE. The concentric or coaxial cable has several important advantages which make it quite practical for reliable operation at high frequencies. The first advantage is the fact that radiation loss is eliminated by surrounding one wire by the second which is at ground potential. Thus, the conductor to ground capacitance remains uniform at all points. Another advantage of a coaxial cable is that the line is free from noise pickup as radiated from external sources.

Two types of coaxial cables are in use, flexible and rigid. In the flexible cables, dielectric materials such as synthetic rubber or polymeric resin compounds are used to separate the inner and outer conductors. Such cables have the advantage of easy handling and installing, although they have a greater attenuation than rigid coaxial cables because of a finite value of dielectric conductance.

Rigid coaxial lines use a solid outer tube, rather than the braid used for flexible lines, and the inner conductor is held in place by uniformly spaced insulators. Since moisture increases the dielectric constant of air, the outer joints of a rigid coaxial line are usually sealed, and the dielectric within the cable carefully dried. Occasionally, coaxial cables are pressurized with either dried air or nitrogen in order to prevent moisture-laden air from entering the cable.

The spacers used in the rigid coaxial cable may be either porcelain beads, regularly placed along the inner conductor; pins, or rods, through the conductor; or a plastic belt wound spirally around the inner conductor. While the ideal coaxial cable would have an air dielectric throughout, such a cable would not have uniform spacing between the two conductors. Any spacing device, such as the beads surrounding the inner conductors are necessary evils. Beads, for example, tend to reflect the signal, thus altering the characteristic impedance of the line. This may be corrected by using undercut beads; that is, beads whose inner diameter is less than the diameter of the inner conductor. This requires a stricture of the inner conductor at the point where the bead is attached. Cable using the spiral spacer mentioned has a lower attenuation than the other types, and is also sturdier, being capable of withstanding a blow that might shatter a bead or rod.

As with all other transmission lines, the characteristic impedance of a coaxial line varies with its distributed inductance and capacitance, and therefore with the size and diameters of the conductors. For air-filled coaxial lines, the characteristic impedance (Z_0) may be found by the formula

$$Z_0 = 138 \log_{10} \frac{b}{a} \qquad (6\text{-}12)$$

where b is the *inner* diameter of the *outer* conductor and a is the *outer* diameter of the *inner* conductor. (See Nomogram E in Appendix I.) It will be found, in most cases, that the value for Z_0 ranges from 40 to 80 ohms for coaxial lines.

The total attenuation imposed by a coaxial line is made up of two components, namely the resistance of the conductors themselves and by losses in the dielectric separating the conductors. The characteristic impedance of the conductors, for a given line, is 72 ohms when the ratio of b/a is equal to 3.59, such a value being used in most cases. For the particular case where the above ratio is satisfied, the attenuation contributed by solid copper conductors may be found by the formula

$$\alpha_R = \frac{0.187\sqrt{\kappa_r}}{b\sqrt{\lambda}} \text{ db per meter} \qquad (6\text{-}13)$$

and the attenuation contributed by the dielectric is

$$\alpha_G = \frac{2730\sqrt{\kappa_r} + \tan\delta}{\lambda} \text{ db per meter} \qquad (6\text{-}14)$$

where

- b = Radius of the outer conductor
- λ = Free-space wavelength
- δ = $\sigma/\omega\kappa$, where σ is the conductivity and κ is the dielectric constant of the separating material. ($\delta \cong 0.003$ for polyethylene)
- κ_r = The relative dielectric constant of the separating material (referred to a vacuum). (κ_r = 2.25 for polyethylene)

The total attenuation of a coaxial line is simply the sum of $\alpha_R + \alpha_G$, and is given in decibels per meter. For metals other than copper, the attenuation given by formula 6-13 should be multiplied by the following factors: 2.0 for brass, 1.8 for dural, 1.3 for aluminum, and 0.58 for silver.

In a coaxial cable, the electric lines of force are *normal* to the surface of the inner conductor. The electric field is made continuous by the presence of the two conductors so that the lines of force in the space between the wires are everywhere perpendicular to the direction of propagation. Conversely, the magnetic lines of force form closed paths around the *inner* conductor and are also perpendicular to the direction of propagation. Such a field configuration is called *transverse electromagnetic* (TEM), since both fields are everywhere normal to the direction of propagation. The TEM wave is the basic form of field configuration in all transmission lines, with the exception of waveguides, and is commonly called the principal *mode of operation* (figure 6-4A). In addition, however, there may exist higher modes of operation in a coaxial cable (figure 6-4B) if the mean diameter of the cable exceeds the free-space wavelength, although such waves are highly attenuated and not commonly used.

In a waveguide, there is only one conductor and, therefore, no closed physical path through which the field energy may be returned. Here, the lines of force must form continuous loops within the guide itself, and since the electric and magnetic fields are always perpendicular to each other, one of the fields must travel part of the time in the *direction of propagation*. Thus, both fields cannot be transverse to the direction of propagation at the same time and a TEM mode is impossible to achieve.

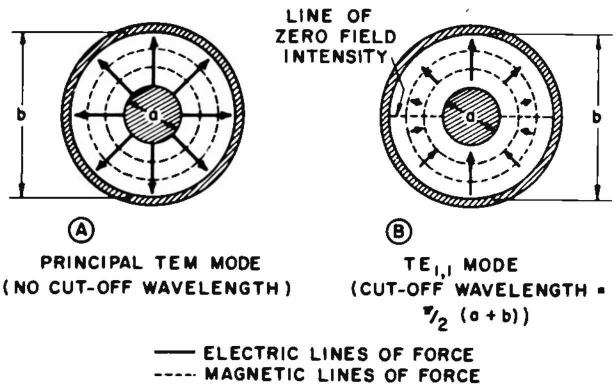


Figure 6-4. Coaxial Modes of Operation

WAVEGUIDES

GENERAL. A waveguide may be defined as a structure consisting of either a conductor or a dielectric, or both, in which the boundaries containing the guide are continuous. Waveguides are similar in action to wire transmission lines in that they are used to transmit r-f energy from one place to another. However, here the likeness to wire lines stops, as a waveguide consists of only one hollow metal conductor or pipe, which is capable of guiding electromagnetic waves through its interior. Two types of metal waveguides commonly used have either a rectangular or circular cross-section. The dielectric waveguide, or rod, is not used extensively because its losses are great, and because the electromagnetic field cannot be wholly contained within the rod. Other types of waveguides include elliptical, corrugated, and flexible, all of which are special forms and will not be discussed here.

One reason for using a hollow waveguide is that it has lower loss than either open-wire lines or coaxial lines in the frequency ranges for which it is practical. Generally, waveguides are used at frequencies ranging from 3000 mc and up, while wire lines are commonly used below 3000 mc. A second reason for using a waveguide is that it is capable of transmitting more power than a coaxial line of the same size. The maximum power transmitted on a coaxial line is given by the formula

$$P = V^2/Z_0 \qquad (6\text{-}15)$$

where V is the voltage and Z_0 the characteristic impedance of the line. Therefore, the power transmitted on a given line can be increased only by raising the voltage, which, if raised too high, will cause the dielectric to break down. The length of the voltage break-down path in a coaxial line is the distance from the inside surface of the outer conductor to the outside surface of the inner conductor. This distance is less than half the voltage break-down distance in a round waveguide of the same size, since when operating in the principal mode, the maximum voltage in a round waveguide appears across its diameter.

Summarized, the *advantages* of a waveguide transmission line are as follows:

1. Complete shielding, eliminating radiation loss.

2. No dielectric loss, since there are no insulating beads within the guide.

3. Conductor loss due to the internal resistance of the guide is less than that of a coaxial line of the same size, operated at the same frequency, since only one conductor is used.

4. Greater power-handling capacity than for a coaxial line of the same size.

5. Simpler construction than that of a coaxial line.

There are, of course, important *disadvantages* to consider when using waveguides. First, the minimum size of waveguide that can be used to transmit a certain frequency is proportional to the wavelength at that frequency. The proportionality depends on the shape of the guide and the manner in which electromagnetic fields are set up within it. In all cases, however, there is a minimum frequency, called the *cut-off frequency,* below which no other frequencies can be transmitted. As a result of this characteristic, a rectangular waveguide would have to be wider than two inches to transmit a 3,000 mc signal, 20 inches wide for a 300 mc signal, and 17 feet wide for a 30 mc signal.

Second, the installation and operation of a waveguide transmission system are somewhat more difficult than for other types of lines. The radius of bends in the line must be greater than one wavelength to avoid excessive attenuation. In addition, if the guide is dented, or if solder beads are permitted to run inside when joints are made, the attenuation of the line will be greatly increased and the voltage breakdown point will be reduced.

MODES OF OPERATION. Solutions of Maxwell's equations, subject to boundary conditions imposed by perfectly conducting metal walls, show that any given guide can propagate an infinite number of different types of electromagnetic waves. Each type or mode has its own individual electric and magnetic field configuration and is designated by that configuration. Each mode of operation has a critical frequency below which it will not be propagated through the guide. The critical frequency or cut-off frequency (f_t) for that mode and the corresponding free-space wavelength or cut-off wavelength (λ_c) are related by the formula

$$\lambda_c = \frac{c}{f_c} \tag{6-16}$$

where c is the free-space velocity of light. In a given size waveguide, the mode with the lowest cut-off frequency is called the principal mode for that guide, and will be the only mode propagated if the frequency is *greater* than the principal cut-off frequency but *less than* the cut-off frequency for the second lowest mode. It is a noticeable characteristic of waveguides that the cut-off wavelength of the principal mode is comparable to the size of the guide. For example, in a round air-filled waveguide, the cut-off wavelength of the principal mode is equal to 1.71 times the diameter, and in rectangular air-filled guides λ_c is equal to twice the large cross-sectional dimension.

In general, two types of modes are possible. The configurations of one type are called *transverse electric* (TE) since their electric fields are everywhere perpendicular to the direction of propagation. Occasionally TE waves are referred to as H-plane waves since part of the magnetic field is parallel to the direction of propagation The second type of mode configurations are called *transverse magnetic* (TM) since their magnetic fields are everywhere perpendicular to the direction of propagation. TM waves may also be called E-plane waves because they have electric field components parallel to the direction of propagation.

The operating wavelength (λ_g) in a waveguide for any mode may be found by the expression

$$\lambda_g = \frac{\lambda}{\sqrt{\kappa_r}\sqrt{1 - (\lambda/\lambda'_c)^2}} \tag{6-17}$$

where

λ = The free-space wavelength

λ_c = Cut-off wavelength of a particular mode of operation

κ_r = The relative dielectric constant of the material with which the guide is filled. Since most guides are air-filled, κ_r may be taken as unity.

The guide wavelength (λ_g) for all gas-filled waveguides is always larger than the corresponding free-space wavelength. This longer wavelength is caused by the wave apparently traveling faster than light, although this seems to violate the fundamental law of physics which states there can be no real velocity greater than that of light. Actually, the rate of energy flow in a guide, called *group velocity* (ν_g), is always somewhat less than that of light and approaches zero as the free-space wavelength approaches the cut-off value. However, superimposed on the traveling electromagnetic waves is a variation of field intensity which appears to move at a much higher velocity. This superimposed velocity of a change in field intensity is called the *phase velocity* (ν_g), and is related to the group velocity by the expression

$$\nu_g \times \nu_p = c^2 \qquad (6\text{-}18)$$

where c is the speed of light in free space. Therefore, since the point of maximum field intensity, *but not the actual energy*, moves down the waveguide at the phase velocity, the apparent wavelength in a guide is always greater than the wavelength in free space. Typical wavelengths found in practice usually run from 1½ to 2 times the wavelength outside the guide.

Since the different modes have correspondingly different wavelengths in a guide, it is generally impossible to make a guide which will be matched for more than one mode. Also at any discontinuity, such as a junction or sharp corner, other modes will tend to be excited and, therefore, it is desirable to choose a guide which will propagate only the lowest mode. If this is done, then any mode beyond cut-off will die out within a very short distance of its source. In general, higher modes are used only because of their special polarization or attenuation features.

RECTANGULAR GUIDES. The lowest mode, or principal mode, in rectangular guides is the $TE_{1,0}$ mode. An instantaneous picture of its field is given by figure 6-5A. The electric field can be represented by electric lines of force which extend from the bottom to the top of the guide. The intensity of the field varies sinusoidally along the "A", or long dimension

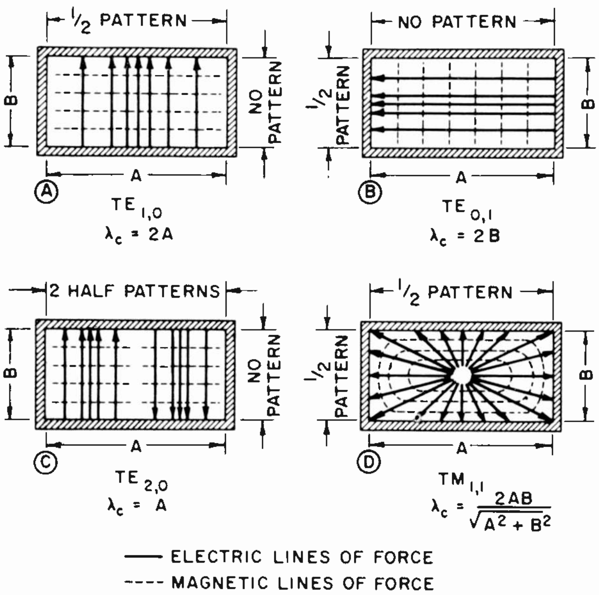


Figure 6-5. Rectangular Waveguide Modes

of the guide, while along the "B", or short dimension, it is uniform. The magnetic field is pictured by means of the magnetic lines of force, which are always perpendicular to the electric lines of force. If the electric field is known, the magnetic field configuration can be determined and, for the $TE_{1,0}$ mode, the magnetic lines form plane loops which are perpendicular to the electric lines.

The cut-off wavelength in an air-filled waveguide for the principal $TE_{1,0}$ mode is given by the expression

$$\lambda_c = 2A \qquad (6\text{-}19)$$

where A is the long dimension of the guide, and must be at least one-half the free-space wavelength in order to propagate the $TE_{1,0}$ mode. In a rectangular waveguide which is filled with a dielectric other than air

$$\lambda_c = 2A\sqrt{\kappa_r} \qquad (6\text{-}20)$$

where κ_r is the relative dielectric constant of the material within the guide.

The various modes of any guide are designated by the subscripts m and n, where m and n are whole numbers. In a rectangular guide, the first subscript m of a $TE_{m,n}$ or a $TM_{m,n}$ mode refers to the number of *half-cycle* variations in the fields along the "A" dimension. The second subscript n refers to the number of *half-cycle* variations along the "B" dimension. For example, the $TE_{0,1}$ mode, as shown in figure 6-5B, is just like a $TE_{1,0}$ mode, which has been rotated through 90°, but it has the subscript "0,1" because the fields are uniform along the "A" dimension and have a half-cycle variation along the "B" dimension.

The cut-off wavelength for the $TE_{0,1}$ mode may be found by the formula

$$\lambda_c = 2B \qquad (6\text{-}21)$$

where B is the short dimension of the guide. The next higher mode in a rectangular guide is the $TE_{2,0}$ mode, shown in figure 6-5C. It is best described as two $TE_{1,0}$ modes side by side and 180° out of phase in time. The cut-off wavelength for the $TE_{2,0}$ mode is

$$\lambda_c = A \qquad (6\text{-}22)$$

Thus, the above data on cut-off wavelengths may be used as a basis to determine the size of rectangular guide which will propagate only the $TE_{1,0}$ mode. The dimensions "A" and "B" of the guide must be such that

$$2A > \lambda > A \qquad (6\text{-}23)$$
$$\lambda > 2B$$

where λ is the free-space wavelength of the frequency to be propagated. The inequality ($2A > \lambda$) allows the $TE_{1,0}$ mode to be propagated; the inequality ($\lambda > A$) prevents the $TE_{2,0}$ and higher modes; while the inequality ($\lambda > 2B$) prevents the $TE_{0,1}$ and similar modes. The cut-off wavelength for any $TE_{m,n}$ or $TM_{m,n}$ mode in a rectangular guide is given by

$$\lambda_c = \sqrt{\left(\frac{m^2}{A}\right) + \left(\frac{n^2}{B}\right)} \qquad (6\text{-}24)$$

For any dielectric other than air, multiply the above expression by $\sqrt{\kappa_r}$.

In ordinary transmission-line practice, characteristic impedance may be defined in terms of voltage and current, power and current, or in terms of power and voltage. All three, under normal conditions, give identical results. For waveguides, however, they do not do so although their numerical differences are not large. For an air-filled guide the characteristic impedance, based on power and voltage, follows the relationship

$$Z_0 = \frac{754B\ \lambda_g}{A\ \lambda}\ \text{ohms} \qquad (6\text{-}25)$$

It may be seen from the formula that, depending upon the guide dimensions, Z_0 may be given almost any desired value. However, representative values of Z_0 for most waveguides now in general use lie between 500 and 1,000 ohms, although some have been used with values of Z_0 less than 100 ohms.

Another matter of prime importance in the practical use of waveguides is the amount of attenuation present within the guide. Such attenuation may be divided into two causes: (1) the attenuation due to losses in the conducting walls of the guide, and (2) the attenuation due to the shunt conductivity of the dielectric filling the guide. For air dielectric, where the shunt conductivity may be neglected, the attenuation developed in a rectangular copper guide is found by

$$\alpha = \frac{0.104B^{-1} + 0.52\ \lambda^2A^{-3}}{\sqrt{\lambda}\qquad\sqrt{1 - 0.25\ \lambda^2A^{-3}}} \qquad (6\text{-}26)$$

CIRCULAR WAVEGUIDES. The principal mode of a circular waveguide is the $TE_{1,1}$ mode, which is the round waveguide equivalent of the rectangular $TE_{1,0}$ mode. The $TE_{1,1}$ mode is shown in cross section in figure 6-6A. The cut-off wavelength of such a mode may be found by the formula

$$\lambda_c = 1.71d \qquad (6\text{-}27)$$

where d is the inside diameter of the guide. However, cut-off wavelengths for the higher modes in circular guides cannot, as a rule, be expressed in easily remembered terms since they involve roots of Bessel functions.

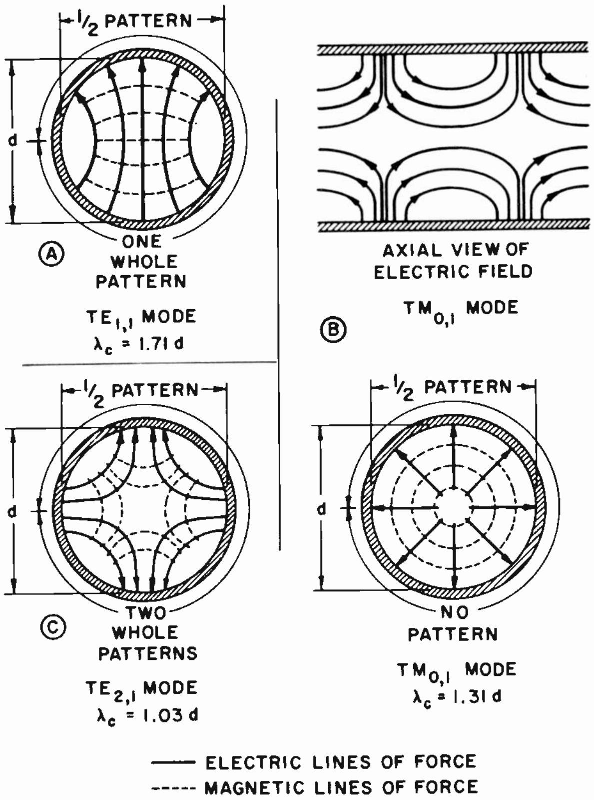


Figure 6-6. Circular Waveguide Modes

In a circular guide, the first subscript m of a $TE_{m,n}$ or $TM_{m,n}$ mode indicates the number of *whole* (or full-wave) patterns in the fields encountered around the circumference of the guide. The second subscript n indicates the number of *half-wave* patterns that exist along a diameter of the guide. Thus, the $TE_{m,n}$ mode, for example, will have m angular modes (diameters) and (n-1) radial modes (circles).

The next highest mode of propagation which occurs in circular waveguides is the $TM_{0,1}$ mode as illustrated in figure 6-6B. The magnetic lines form concentric circles and lie in planes perpendicular to the direction of propagation. Since the magnetic lines are everywhere perpendicular to the guide axis, the surface currents must everywhere flow parallel to the guide axis. Likewise, an axial view of the electric field would show that each line of force emerges perpendicularly from the guide wall, turns and travels parallel to the guide axis, and then bends back into the guide. The cutoff wavelength for the $TM_{0,1}$ mode in a round guide may be found by the expression

$$\lambda_c = 1.31d \qquad (6\text{-}28)$$

Similarly, the third mode of a circular waveguide is the $TE_{2,1}$ mode illustrated in figure 6-6C. The cut-off wavelength for this mode is

$$\lambda_c = 1.03d \qquad (6\text{-}29)$$

Therefore, to choose the correct size of circular guide which will propagate only the $TE_{1,1}$ mode, for example, the guide diameter should be such that

$$1.71d > \lambda > 1.31d \qquad (6\text{-}30)$$

where λ is the free-space wavelength of the operating frequency. Since circular waveguides are very seldom used in present-day equipments, and the mathematics involving α and Z_0 is beyond the scope of this book, no attempt will be made to go further into circular waveguide theory.

TRANSMISSION LINE USES

The theoretical development of transmission lines and their use as a means of transmitting r-f power from one point to another have been discussed previously. However, transmission line sections are commonly used for many other purposes. For example, r-f lines are used as:

1. Frequency control elements.
2. Filters.
3. Phase shifters.
4. Delay lines.
5. Impedance matching sections.
6. Attenuators.

FREQUENCY-CONTROL ELEMENTS. In the frequency range from about 200 mc to about 1,000 mc parallel line sections are commonly used to control the frequency of oscillators and amplifiers. The frequency stability of an oscillator is a function of the Q of its tuned circuits. The Q, or figure of merit, of a tuned circuit is equal to the inductive reactance at resonance divided by the resistance and as such is a function of frequency. Q may also be thought of as a constant times the ratio of the energy stored in a circuit to that dissipated in the resistance of the circuit. For a given inside diameter of the outer conductor in a coaxial line, the Q is a maximum when the ratio of the diameters is equal to 3.59. This ratio corresponds to a characteristic impedance of 76 ohms so that the use of nominal 72 ohm coaxial line will approximate the conditions of maximum Q. The larger the inside diameter of the outer conductor, while maintaining the same diameter ratio, the larger will be the Q of the line. As discussed in the development of transmission lines, a finite length transmission line which is terminated in other than its characteristic impedance will have standing waves and will reflect an impedance back to the source. The value and type of impedance, resistive, capacitive or inductive, will depend on the type of termination and the length of the line. A line one-quarter wavelength long terminated in a short circuit will reflect a very high resistive impedance to the source and so act as a parallel resonant circuit. The same line terminated in an open circuit will reflect a low resistive impedance and so appear as a series resonant circuit.

Tuning these lines may be accomplished by several methods as illustrated in figure 6-7A. An adjustable shorting strap can be used to change the effective length of the line if very low resistance connections are made so as not to lower the Q of the circuit. A second method of changing the effective line length is the use of a capacitor as a shorting bar to change the resonant frequency of the line. Parallel plate capacitors with air dielectric are usually used for this

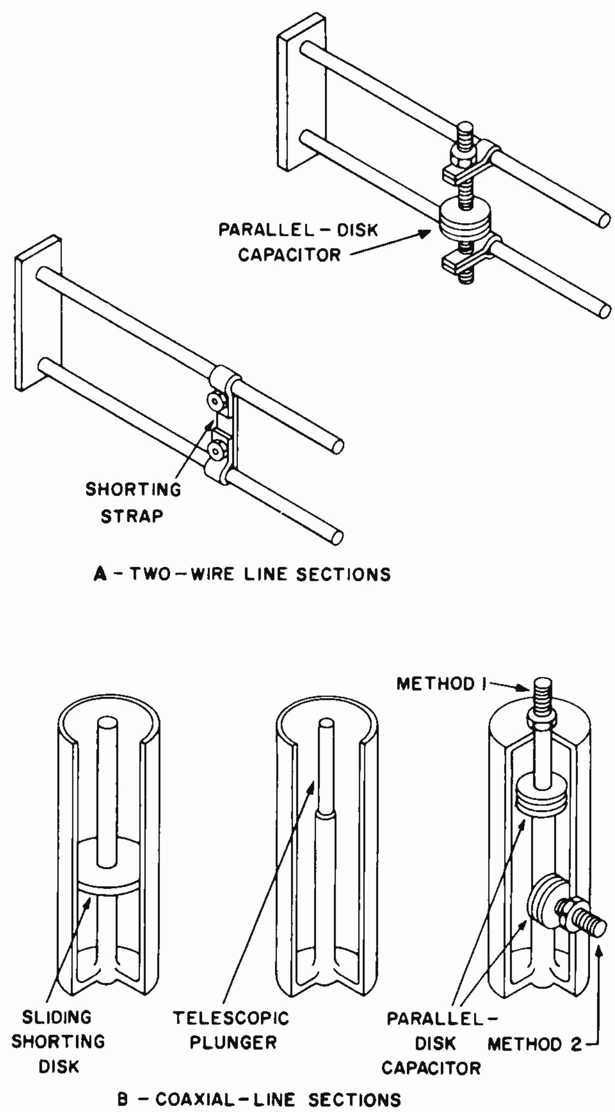


Figure 6-7. Common Tuning Methods

purpose, but unless carefully designed and adjusted may unbalance the line.

Coaxial lines are also used at frequencies ranging to about 3,000 mc as tuned circuits. Although they are harder to tune (see figure 6-7B) they have very little, if any, radiation loss and the possibility of stray coupling is reduced, allowing a more compact design to be used. Increasing the diameter of the line will increase the Q of the circuit, but the diameter should be kept smaller than that required to propagate any mode other than the TEM mode, or oscillation at other than the desired frequency may result.

Harmonic response is controlled easily in a quarter-wave tank circuit, since it responds only to the odd harmonics. The third harmonic can be practically eliminated by loading with a small value of capacitive reactance or by tapping onto the tank at the third-harmonic voltage minimum. Although quarter-wave and half-wave line sections are most commonly used, the statements made here about the basic quarter-wave sections also are true for sections composed of any odd multiple of a quarter wave. Statements about half-wave sections hold true for any multiple of a half wave.

By closing off both ends of a length of coaxial line with conducting material, standing waves will be set up if any energy is introduced. When a transmission line completely encloses the energy in the line, it is called a cavity resonator. If the cavity has a center conductor so that the TEM mode can exist, the cavity will be resonant at any frequency for which the length is a multiple of a half-wave length. Additional resonant frequencies will exist due to TE and TM modes of resonance. The number of modes other than TEM with a wavelength longer than the free-space wavelength in a given coaxial cavity can be approximated by

$$N = 4.4 \frac{V}{\lambda^3} \qquad (6\text{-}31)$$

where N is the number of modes, V is the volume in the same units as λ^3, and λ is any wavelength. By having the diameter of the coaxial cable small enough to prevent the occurrence of any mode other than the TEM mode near the operating frequency, a stable frequency control element with a very high Q is obtained.

Coaxial cavities can be tuned in a number of ways. The most commonly used method is a plunger which changes the length of the cavity. Another method used for fine tuning is the insertion of a tuning slug which may be either capacitive or inductive depending on its position with respect to the electric and magnetic lines of force.

Cavity resonators can also be produced by closing off both ends of a waveguide. Modes are identified by the addition of a third index to the mode designation for waveguides, the third number being the number of half waves along the axis. The principal mode in a right circular cylindrical cavity is the $TM_{0,1,0}$ mode in which the axial length has little control over frequency. Higher mode cavities can be tuned by varying the axial length, which is usually the frequency critical dimension.

Cavity resonators may be made in almost any shape in which a dielectric is completely enclosed by a conductor. Figure 6-8 illustrates some of the forms which may be used, and gives their approximate Q.

Energy may be fed into or removed from a cavity by an inductive loop which couples the magnetic lines, by a capacitive probe or antenna which couples to the electric lines, or directly by a waveguide oriented so that the magnetic lines in the waveguide are parallel to the magnetic lines in the cavity.

FILTERS. Sections of coaxial line are often used as filters to remove even harmonics. A quarter wave-length shorted section in parallel with the line will appear as a very high impedance at the fundamental frequency and all odd harmonics, but will appear as a very low impedance at the even harmonics. Elimination of other frequencies requires the use of a half-wave shorted line at the interfering frequency plus an additional section to cancel out the reactance introduced to the desired frequency. For example, suppose the desired frequency is 400 mc and the interfering frequency is 250 mc. A shorted half-wave section at 250 mc is placed across the line to short out the 250 mc signal. This introduces a capacitive reactance at 400 mc and requires the use of an inductive section to cancel the reactance out in order to prevent attenuation at the desired frequency. (See Nomogram F, Appendix I.)

Waveguide sections can be used similarly to filter out undesired frequencies either as a shunt filter or by placing a section of the proper size in the line to act as a high-pass filter. Low-pass filter action can be obtained by shunting the transmission line with a section of waveguide with the proper cut-off wavelength and terminating it with a resistive load so that no reflection of energy takes place. All frequencies above the cut-off frequency will be partially absorbed and all frequencies below the cut-off frequency will be passed with little attenuation.

It is often desirable to be able to use the same antenna for simultaneous reception and transmission. This can be accomplished by the use of antenna duplexers and different frequencies for reception and

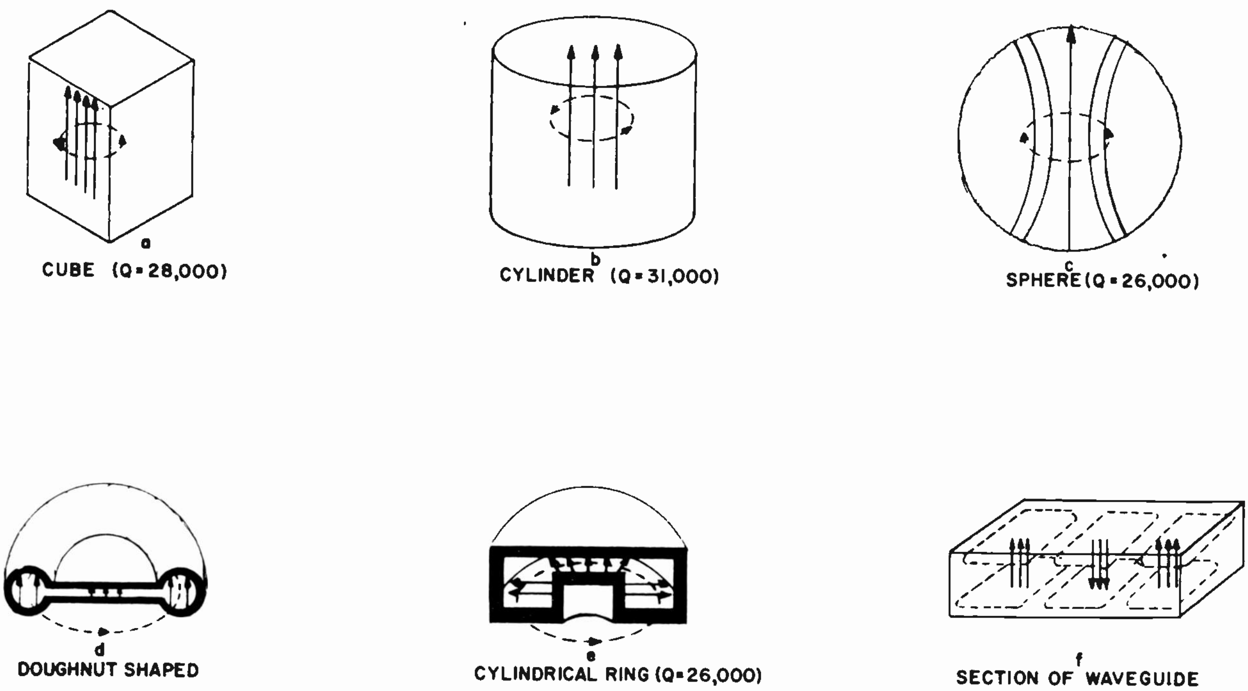


Figure 6-8. Types of Cavities

transmission. Frequency sensitive sections of transmission line separate the signals. The transmission line to the transmitter should have a high mismatch at the received frequency so as to avoid a loss of power in the transmitter, and the receiver transmission line should have a high mismatch at the transmitter frequency to prevent power loss in the receiver and possible damage to the receiver.

Since a shorted quarter-wave-long transmission line has a very high impedance, it is often used as a "metallic insulator". Mechanically and electrically the best coaxial lines for centimeter wavelengths are the so-called stub-supported lines which are used to a great extent. The stub is in parallel with the line and, if it is made exactly an electrical quarter-wavelength long, it will give no reflection at that frequency. Since a quarter-wave stub may be thought of as a loaded resonator, it will be perfectly reflectionless at only one wavelength. At longer wavelengths, the stub is less than a quarter-wave and therefore shunts the line with an inductance, and at shorter wavelengths the stub is more than a quarter-wave and so shunts the line with a capacitance. To compensate for this effect, it is possible to place two stubs in the line an odd number of quarter-waves apart so that their respective reflections cancel.

PHASE SHIFTERS AND DELAY LINES. Since there is a progressive phase shift of 360° per wavelength along a matched transmission line, a short length of line can be used to obtain any desired amount of phase shift between two points in a circuit. For example, a line one-eighth wavelength long will have a phase shift of 45° between input and output. By lengthening the line a time delay may be obtained with the amount of delay depending on the length of the line. To obtain any large time delay would require a very long and bulky line. To avoid this, artificial transmission lines are built using lumped capacitance and inductance in which delays of several hundred microseconds may be obtained.

IMPEDANCE MATCHING SECTIONS. Quarter-wavelength transmission lines are often used for impedance matching. By shorting one end of the line standing waves are set up causing the line impedance to vary from zero at the shorted end to a very high impedance at the open end. Any impedance can be obtained by tapping on to the line at the point where the V/I ratio equals the desired impedance.

If the quarter-wave line is not shorted, but has a load impedance of Z_r across the load end of the section the input impedance will be

$$Z_s = \frac{(Z_0)^2}{Z_r} \qquad (6\text{-}32)$$

where Z_s is the input impedance and Z_0 is the characteristic impedance of the line. If Z_s and Z_r are both known, the required Z_0 to match the two can be determined from

$$Z_0 = \sqrt{Z_s Z_r} \qquad (6\text{-}33)$$

For example, it is desired to match a 600-ohm input impedance to a 72-ohm output impedance. The required line would have a characteristic impedance of 208 ohms. (See Nomogram G in Appendix I.)

Usually in matching applications, the characteristic impedance of a transmission line is of great importance. However, by the use of half-wave or integral multiples of a half-wave section, the impedance of the line becomes unimportant. The impedance of a device connected to one end of a half-wave line will be reflected unchanged to the other end regardless of the characteristic impedance of the line although there will be a 180° phase shift per half-wave section of line.

When connecting a device that is balanced to ground to one which is not, some means of balance converter must be used. A typical example is the connection of a dipole antenna which is usually balanced to ground to a coaxial line which is unbalanced. One device that may be used is known as a bazooka line balance converter (figure 6-9). This is used when the impedance of the devices already match and no impedance transformation is desired. It consists of a quarter-wavelength shield placed around the end of the coaxial line. The balance transformation can take place in either direction.

Another type of line balance converter is shown in figure 6-10. This functions as a balance converter as well as an impedance transformer with a ratio of 1:4. Since the U-shaped line is a half-wave long, there is a 180° phase shift around it with both ends of the line at a high impedance to ground. If, for instance, 50 volts is applied to the converter, 50 volts will appear at each end with respect to ground, but the ends will be of opposite polarity to each other. Doubling the voltage while maintaining the same power corresponds to an impedance transformation ratio of 1:4. Like the bazooka converter, this converter can be used for transformation in either direction.

Although waveguides have a definite characteristic impedance, as yet no exact method exists for finding it. Approximations may be made from current, voltage, and power ratios, but these all give different values. Fortunately, it is not necessary to know the characteristic impedance to match two sections of waveguide. All that is required is to insert a tapered section several wavelengths long between the two. Transformations of shape as well as size can be made in this way without reflection provided the taper is gradual.

Coaxial line to waveguide transformation usually consists of a right angle junction with the center conductor of the coaxial line extending into the waveguide approximately a quarter-wavelength and functioning as an antenna. The end plate of the waveguide is about a quarter guide wavelength from the antenna, so that it reflects radiation in phase with that radiated directly down the guide.

In actual practice neither of these lengths is a quarter wavelength, but must be determined by experiment. We may consider the end-plate distance as varying the amount of coupling between the coaxial line and the waveguide through the varying phase of directly radiated and end-plate-reflected fields. This produces a match of the characteristic impedance of the guide to that of the coaxial line except for the reactive component of the transition. This reactance may be matched out by adjustment of the antenna length.

ATTENUATORS. Attenuators are of two general types. The first class operates on the principle of a waveguide below cut-off. The second class operates by introducing dissipative elements to absorb energy. These will be referred to simply as the cut-off type and the dissipative type.

The cut-off type is commonly used as a variable attenuator. The calibration of such an attenuator may be computed theoretically, and may be made substantially independent of frequency. The theory underlying this type of attenuator is as follows:

Propagation through a waveguide takes place with

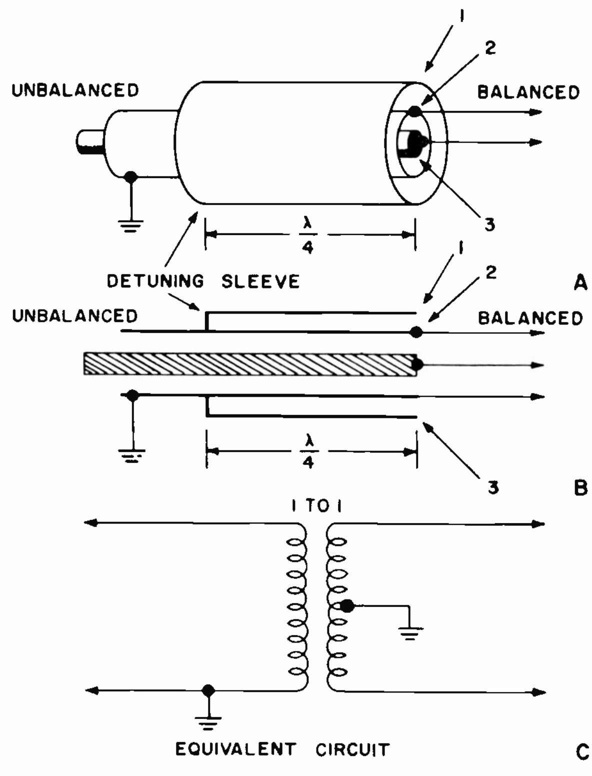


Figure 6-9. Bazooka Line Balance Converter

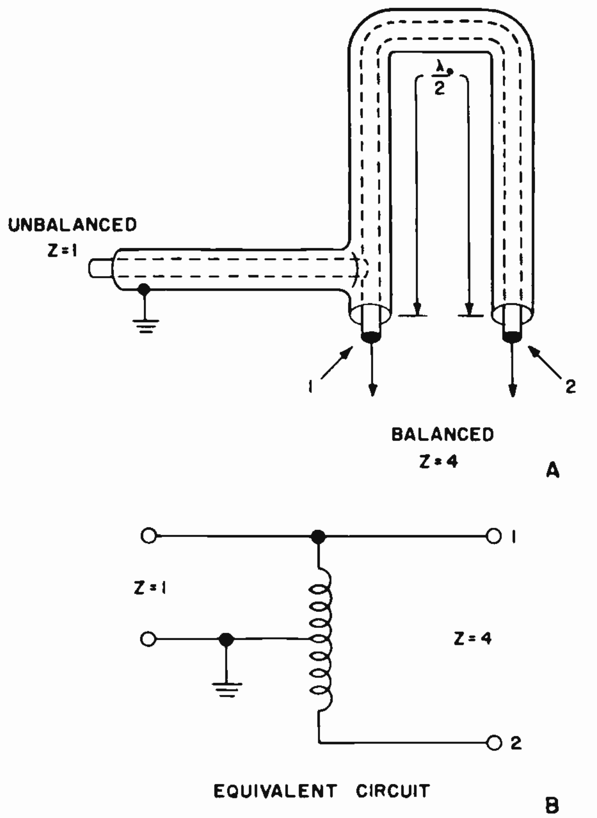


Figure 6-10. Phase Inverter as Line Balance Converter

very little loss provided the diameter or width of the guide is greater than the cut-off value. If these dimensions are below cut-off, then there is no longer any true wave propagation, but instead the fields are exponentially attenuated down the length of the guide. The attenuation per unit length, for any given mode, may be computed directly from theory. Furthermore, if the guide diameter is made small compared to the free-space wavelength, the attenuation is then substantially independent of frequency over a wide range. The mode which is most commonly used in attenuators is the $TE_{1,1}$. This mode is excited by a loop or waveguide and energy is taken out again by another loop or by direct connection to another waveguide (figure 6-11). The attenuation for this mode is as follows

$$TE_{1,1} \text{ db/cm} = \frac{32.0}{d}\sqrt{1-\frac{3.41d^2}{2\lambda}} \qquad (6\text{-}34)$$

where d is the diameter in cm of the cut-off attenuator. Since the attenuation is a linear function of the length, the distance between source and load can be calibrated directly in db. Attenuators of this type usually have a loss of 10 to 20 db at the position of maximum coupling or minimum length. In the immediate vicinity of the loops, modes other than the desired one exist, and as a result the attenuation is not a linear function of the length for short distances. The usable portion of the attenuation range begins at the point where the characteristic becomes linear. This will in general occur when the attenuator is long enough to obtain a loss of 25 to 35 db for the $TE_{1,1}$ mode.

Since a line terminated by a loop is substantially equivalent to a short-circuited line, an attenuator of this type presents a gross impedance mismatch looking in toward either the input or output ends. This is undesirable in many applications and may be circumvented in a number of ways. The most common scheme is to pad the input and output ends with about 10 db of lossy cable so that reflections from the loops are substantially damped out at the far ends of the cables.

Dissipative attenuators function by introducing resistive elements in the waveguide or coaxial line so as to dissipate energy in the form of heat. In coaxial lines the loss may be introduced either in the form of a series resistance or shunt conductance. Examples of the former are platinized glass-rod and carbonized rod attenuators. Shunt conductance types are ordinary lossy cable and sand loads.

The insertion of an attenuator may cause a considerable mismatch, which may be avoided in two ways. First, the dissipative element may be tapered so that the loss is introduced gradually. This scheme is more often used in waveguides than in coaxial lines. Second, a matching section of some sort may be used.

Lossy cable is a coaxial line made with a dielectric which has a high power factor. The attenuation of this cable is around 1.2 db/ft in the 10 cm band. The attenuation varies nearly linearly with both frequency and temperature, so that its use as a calibrated attenuator is to be discouraged.

Sand load attenuators are used as terminating sections to dissipate power. The space between inner and outer conductors is filled with a sand and carbon mixture which acts as the dissipative element.

A platinized glass attenuator is essentially a section of coaxial line in which the inner conductor is a length of glass tubing which has been coated with an extremely thin platinum film. The film is resistive and thereby produces attenuation. By controlling the thickness of the platinum film during manufacture various degrees of attenuation may be achieved.

The carbonized rod attenuator is a series resistance type. The center conductor is a quartz rod which has a thin deposit of carbon.

Another type of coaxial attenuator makes use of the r-f absorbing properties of "polyiron". This material is insulated powdered iron of the kind used extensively for magnetic cores at the lower radio frequencies. If the space between the inner and outer conductors of a coaxial line is filled with a molded slug of polyiron, attenuations of the order of 20 to 30 db/cm may be obtained at 10 cm wavelengths. Molded slugs of polyiron are extensively used in r-f line filters where it is desired to keep r-f from leaking out through power leads.

Waveguide dissipative attenuators are either of the resistor-strip type, or of the sort in which the entire guide is filled with a lossy material.

In the resistor strip type, a strip of bakelite coated with resistance material is placed longitudinally in the guide, with the plane of the strip parallel to the short side. When the guide is excited in the $TE_{1,0}$ mode the electric field is parallel to the plane of the strip so that current flows in the resistive material, and energy is dissipated in the form of heat. The attenuation is an inverse function of the surface resistivity of the strip, with the lower the resistivity, the greater the attenuation (within limits) and the greater the mismatch of the attenuator.

Since the electric vector varies from a minimum to a maximum across the guide, the amount of attenuation with a given strip may be varied by moving the strip laterally across the guide. Attenuations up to 70 db have been obtained with attenuators of this type, using strips 4 inches long.

Resistor strip attenuators are matched either by mechanically tapering the strip or by cutting notches in the end of the strip. The notched sections act as matching transformers. By making the taper long enough, it is possible to get very low standing wave ratios.

In the second type of waveguide dissipative attenuator, the guide is filled with a conductive material, generally a plaster-of-paris and graphite mixture, the proportions of which are adjusted to give the proper attenuation. Matching is accomplished either by means of a slow mechanical taper or by matching the ends. A bakelite plug has been used instead of the plaster-of-paris mixture with resulting attenuations of about 1 db/cm having been reported.

INSTALLATION NOTES

BENDS AND TWISTS. Occasionally it is necessary to employ considerable lengths of waveguide and, if the guide is unable to proceed in a straight line, some form of bended or twisted section must be used. In a rectangular guide, bends may be made in the H-plane by making the radius of curvature lie in the same plane as the magnetic lines of force, or in the E-plane where the radius is in the same plane as the electric lines of force. They may be smooth, gradual bends, or right angles where the outer corner of the guide is beveled at a 45° angle.

Smooth bends in rectangular waveguides are theoretically perfect since the reflections caused by them are very small. Actually, however, it is difficult to make a bend in such a way that the guide retains its proper width to avoid discontinuities. It is possible, when great care is taken, to have a reflection less than that corresponding to a standing wave ratio of 1.05 of power if the radius of the bend is equal to λ_g or more. Likewise, a waveguide may safely be twisted in order to rotate the electric field through 90° provided that the guide is not deformed in the process. The length of the twist should also be equal to λ_g at a minimum.

Bends in circular waveguides are more likely to cause trouble since the polarization is not defined by the shape of the guide and the wave is made elliptical in going around a bend. This effect may be compensated in any given system by squeezing the round guide in such a way as to remove the ellipticity, or by arranging an even number of bends that cancel each other. Flexible guide may also be used, although the attenuation is somewhat higher than in a solid guide, and some reflection will occur.

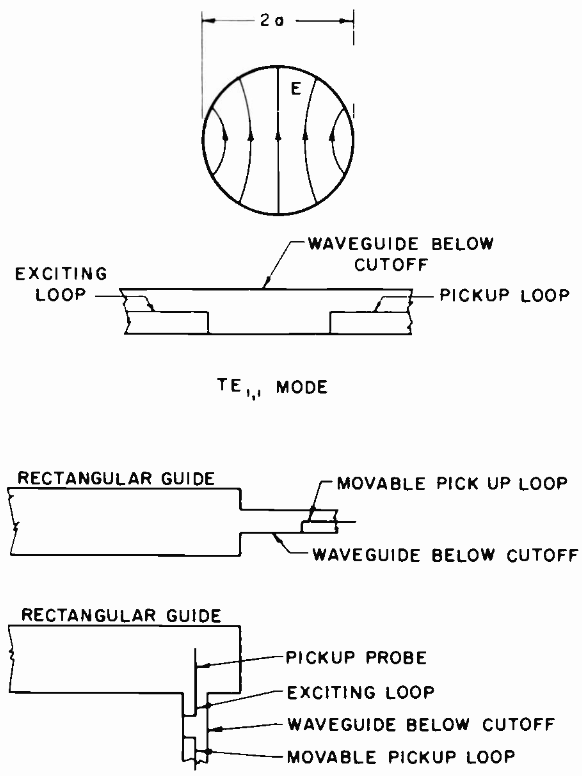


Figure 6-11. Cutoff Attenuators

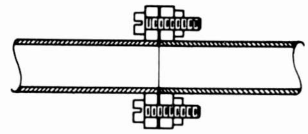
If it becomes necessary to bend a section of rigid coaxial line, the same care should be taken in order to prevent any discontinuities. Here the problem becomes more severe, since the inner conductor must at all times be evenly spaced from the outer conductor and the radius of the bend should, therefore, be large enough to allow the spacing to remain constant.

For most purposes, custom-made bends and twists are available for either waveguides or rigid coaxial lines, and should be used in preference to constructing them from straight sections.

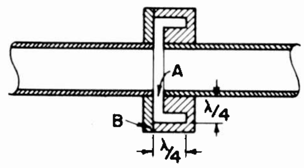
WAVEGUIDE JUNCTIONS. The losses and reflections from a junction between two sections of waveguide are negligible if care is taken to align the sections for good electrical contact. A simple joint may be constructed with the use of flat flanges soldered on the ends of the waveguide and bolted together as shown in figure 6-12A. Such a joint can be made more accurate than the average soldered joint if the flange surfaces are clean and if they contact uniformly. However, since clean, flat, parallel surfaces are hard to achieve and maintain, the more complicated choke-flange joint as illustrated in figure 6-12B may be used instead.

The L-shaped cavity may be considered a shorted half-wavelength transmission line which effectively passes the conduction current developed in the walls of the guide. Gap "A" in figure 6-12B is spaced $\frac{1}{2}\lambda$ from the shorted end of the L, and therefore the voltage across the gap is zero while the current is equal to the conduction current in the guide walls. Point "B" occurs at a $\frac{1}{4}\lambda$ point where the current is zero, and thus the contact at B need not be perfect. In fact, for some applications, there is no contact at all at point B.

PRESSURIZING. Transmission line losses caused by moisture or dirt entering a waveguide or coaxial line can be minimized by pressurizing the line. Pressurizing also helps prevent component breakdown at high-altitudes caused by corona discharge.



(A) FLANGE-TO-FLANGE JOINT



(B) CHOKE-FLANGE JOINT

Figure 6-12. Waveguide Joints

Usually the complete r-f line from the transmitter to the antenna is pressurized with dried air or nitrogen, the various joints being sealed with rubber gaskets. On some systems, even the driven antenna is sealed, in the case of a dipole, by means of a plastic cup which surrounds the dipole. If the line is a waveguide, the pressurized portion can be isolated by means of a thin sheet of mica mounted between a pressurized choke and flange joint. If the line is coaxial, a bead may be used as a pressure seal, although it is better to make the seal in a standard coupler. Such couplers are quite numerous and information relating to them is best found in the manufacturer's data.

In general, the installation of a transmission line system depends upon the r-f equipment being used with it, and the location of each unit. In such cases, the transmission line network is usually "tailor-made" to the rest of the equipment with complete instructions for installation and maintenance.

CHAPTER 7

ANTENNAS

INTRODUCTION

The basic principles of radio-relay antennas do not differ from those of any other antenna. As in an ordinary resonant circuit, changes in frequency cause changes in the current and voltage, which, in the case of an antenna, affect its gain, bandwidth, impedance, and radiation pattern. These four factors can be considered of practical importance, rather than of theoretical interest. A lack of knowledge of any of these factors may lead to unsatisfactory results when it is desired to choose an antenna for a specific application. In the following chapter these factors are considered with regard to radio-relay antennas, which must meet certain desirable characteristics if they are to be used successfully.

When developing the subject of antenna characteristics, another important thing to remember is the theory of reciprocity. In free space, the radiation patterns of a transmitting antenna and a receiving antenna with the same physical characteristics are identical. Therefore transmitting and receiving antennas may always be interchanged, this fact being called the *law of reciprocity*. For example, if two antennas are situated at a distance of several wavelengths apart, a current equal in both phase and amplitude will be received, no matter which antenna is used for transmitting. In the balance of this chapter, when the characteristics of an antenna are discussed, such characteristics will apply to both transmitting and receiving antennas, unless otherwise specified.

BASIC ANTENNA PRINCIPLES

ANTENNAS AS RESONANT CIRCUITS. A receiving antenna intercepts a portion of the directed wavefront radiated from a transmitting antenna. This wave front creates a field around the receiving antenna which induces a voltage in the conductors forming the antenna. As the changing electromagnetic waves go through a cycle, the induced voltage attempts to follow these changes, and, in so doing, reradiates a small portion of the intercepted energy.

Because of the current and voltage distributions which exist on an antenna, it will be found that at a certain frequency, the inductive and capacitive components will cancel, leaving only an effective resistance. The lowest frequency at which this occurs is the *fundamental resonant frequency* of the antenna, or when the antenna is a half-wavelength long. This will also occur at frequencies which are harmonically related.

Depending upon the point where a transmission line is connected to an antenna, an antenna may be considered as a series or parallel circuit. These two conditions are illustrated in figure 7-1. For a dipole antenna which has a transmission line connected to its center, the current is a maximum and the voltage is a minimum at fundamental resonance or the lowest resonant frequency. The ratio of voltage over current is then low, which means the impedance is also low, since

$$Z = \frac{V}{I} \text{ ohms} \qquad (7\text{-}1)$$

where

Z = Impedance in ohms

V = Induced E.M.F. in volts

I = Induced current in amperes.

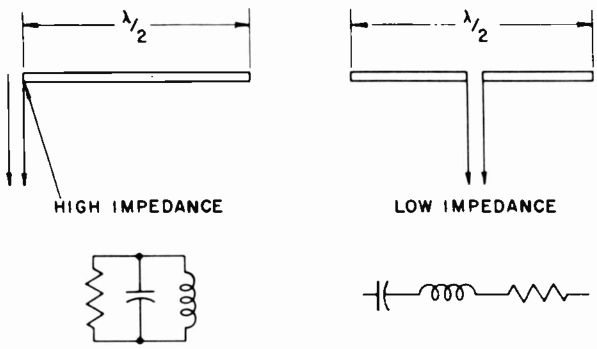


Figure 7-1. Antenna Feed Systems

This low impedance can be considered the result of resonating a series tuned circuit. For an antenna which has a transmission line connected at a half-wavelength point (as measured from one end), or is end fed, the voltage at this point is a maximum and the current is a minimum. The ratio of voltage over current is then high, resulting in a high impedance value. An end fed antenna may therefore be thought of as a parallel tuned circuit which has a high resistance at resonance.

When an antenna is required to receive or transmit signals over a wide frequency range, the conditions illustrated in figure 7-2 are present. The voltage and current distributions along the antenna depend mainly upon the length of the antenna with respect to frequency. At the fundamental and odd harmonic frequencies, the impedance is low, but at the even harmonic frequencies, the center impedance is high. Thus, it will be seen that an antenna required to receive a wide range of frequencies will present impedances which vary in a periodic fashion like a damped wave train, or alternately present low or high impedances at various frequencies.

At some resonant frequency, the antenna will present its lowest impedance. Under this condition, the inductive reactance and capacitive reactance of the circuit cancel, leaving the circuit series resonant. Conversely, when the antenna passes through a value of maximum impedance, the reactive components have cancelled, providing parallel resonance. This condition is sometimes called *anti-resonance*. Figure 7-3 shows the variations in the inductive and capacitive reactances of a practical dipole antenna. As the frequency increases progressively, there are cycles of series resonant and parallel resonant conditions constituting a series of impedance changes from low to high. The effect is a gradual variation of impedance made up of low to high impedance changes and gradually approaching 300 ohms. This effect is illustrated in figure 7-4.

Unless special means are employed to reduce these impedance excursions, an antenna can be matched to a 72-ohm transmission line only at its fundamental and approximately at the odd harmonics. The variation in impedance with frequency for a half-wave dipole is shown in figure 7-5. It will be noted that the minimum value is about 75 ohms and the maximum value about 750 ohms.

ANTENNA BANDWIDTH. One of the inherent difficulties in designing a satisfactory radio-relay antenna arises from the bandwidths generally required. Such bandwidths extend upward to 40 mc or more when it becomes necessary to transmit an extreme number of messages simultaneously. In general, an antenna can be broadbanded by changing its shape in some manner to obtain the desired characteristics. Probably the easiest method is to increase the diameter of the antenna conductors, thus preventing the impedance of the antenna from changing rapidly with frequency. Another method which can be used to broadband an antenna is to use a transmission line with a higher impedance than that of the antenna at fundamental resonance. By using such a line, the voltage developed at resonance is decreased, but the voltage developed at other frequencies is increased. Hence the response curve of the antenna is effectively flattened, with resulting broadband antenna characteristics.

Generally, the bandwidth of an antenna system can be arbitrarily specified by the frequency limits f_1 and f_2 at which the voltage standing wave ratio (VSWR) on the transmission line exceeds an acceptable value. (See Chapter VI.) At times the frequency bandwidth is also specified as the ratio of $f_2 - f_1$ to f_0, or in terms of per cent as given by the formula

$$B\% = \frac{f_2 - f_1}{f_0} \times 100 \qquad (7\text{-}2)$$

where

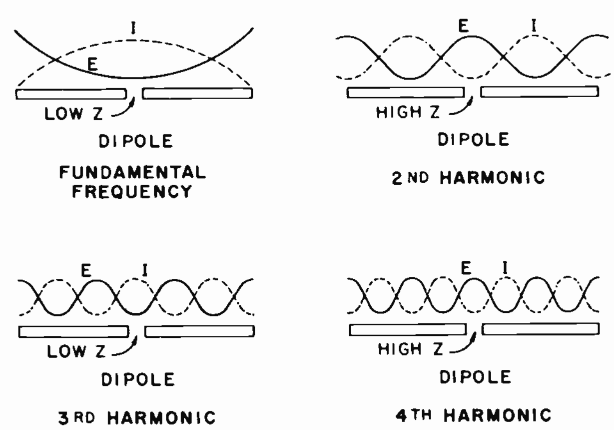


Figure 7-2. Dipole Current and Voltage Distribution

B% = Bandwidth as a percentage of f_0

f_0 = The center design frequency

RADIATING EFFICIENCY. For some applications, it is necessary to obtain a high gain-to-size ratio from an antenna of a given physical size, or to obtain a given gain with an antenna that is as small as possible. Such attempts invariably reduce the bandwidth of the antenna and also its radiating efficiency. Assuming no losses, the power radiated from an antenna would be equal to the power delivered to the antenna. Therefore, such power must be equal to the square of the rms current flowing on the antenna times a resistance, called the *radiation resistance*,

$$P = \left(\frac{I}{\sqrt{2}}\right)^2 \times R \qquad (7\text{-}3)$$

where

P = Radiated power in watts

$\frac{I}{\sqrt{2}}$ = Rms current

R = The radiation resistance

In actual practice, all the power that is delivered to a transmitting antenna is not radiated, but is dissipated in the form of heat because of a certain loss resistance. Thus, the ratio of the power radiated to the total power delivered to an antenna at a given frequency is defined as the *radiating efficiency*. The total terminal resistance (R_t) of an antenna may then be considered as

$$R_t = R + R_L \qquad (7\text{-}4)$$

where

R = The radiation resistance

R_L = The equivalent terminal loss resistance.

Therefore, the radiating efficiency of an antenna may be expressed by the formula

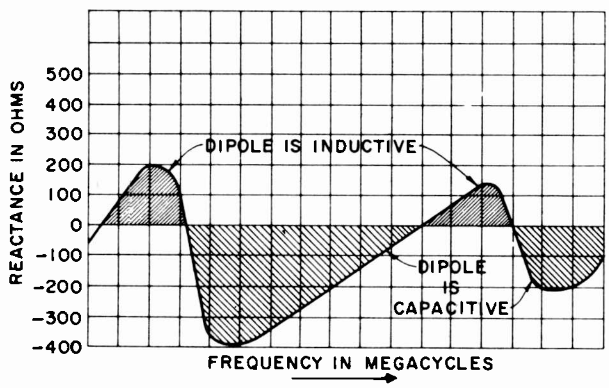


Figure 7-3. Dipole Reactance Variation

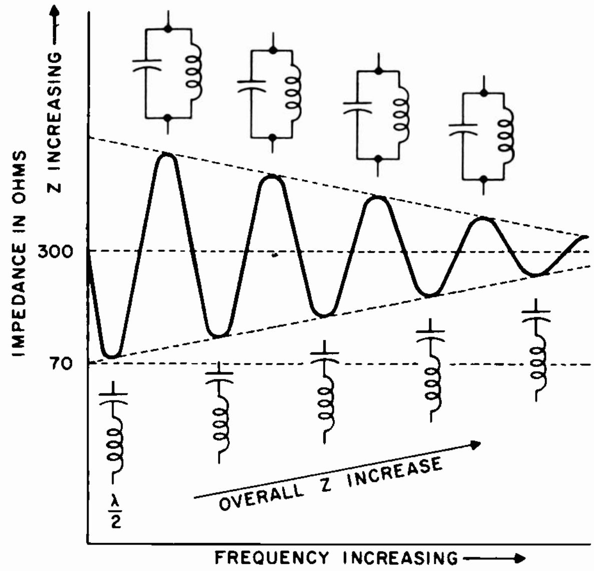


Figure 7-4. Dipole Impedance Excursions

$$\text{Radiating Efficiency } \% = \frac{R}{R + R_L} \times 100 \qquad (7\text{-}5)$$

Any antenna whose radiation resistance is large, as compared to any loss resistance, has a high radiating efficiency. Loss resistance may occur in the antenna conductors, supporting insulators, masts, or guy wires, with faulty antenna installation being a great contributor to such losses.

ANTENNA DIMENSIONS. As mentioned previously, an antenna is resonant when its length is a half-wavelength of the frequency used. This length, however, is not exactly a physical half-wavelength such as that which would be developed in free space; instead, it will be somewhat shorter.

In free space, electromagnetic waves travel at a constant velocity of 186,000 miles per second (3×10^8 meters/sec.). The r-f energy on an antenna, however, moves at a velocity considerably less than that of the radiated energy in free space because the metal comprising the antenna has a permeability greater than that of free space. Since the velocity of electromagnetic waves in any medium is determined by the formula

$$\nu = \frac{1}{\sqrt{\mu\kappa}} \text{ meters per second} \qquad (7\text{-}6)$$

where

κ = The dielectric constant of the material through which the wave is being propagated (Mks units).

μ = The permeability of the material (Mks units)

It will be seen that as the dielectric constant (κ) increases, the velocity of the wave decreases.

Because of the difference in velocity between the wave in free space and the wave on an antenna, the physical length of an antenna no longer corresponds to its electrical length. In other words, such an antenna is a half-wavelength electrically, but somewhat shorter than this physically. This fact is shown in the formula for the velocity of electromagnetic waves as expressed in terms of wavelength and frequency, which is

$$\nu = f\lambda \tag{7-7}$$

where, (ν) is the velocity, (f) is the frequency, and (λ) is the wavelength. Since the frequency of the wave on an antenna remains constant, a decrease in the velocity also results in a decrease in the wavelength.

The actual value of the permeability of an antenna, and thus its physical length, depends upon several factors. As the circumferences of an antenna's conductors increase, the permeability increases and the wave velocity is decreased. Such changes in wave velocity on an antenna are commonly called *end effect* because the ends of the antenna are made closer together physically than they are electrically. End effect is counteracted by making the actual length of the antenna shorter than the free-space wavelength, as expressed by the formula

$$L = K\,(492/f) \tag{7-8}$$

where (L) is the required length in feet for a half-wave dipole, and (f) is the frequency in megacycles. The factor (K) takes into consideration the end effect just mentioned, and may be found in figure 7-6 for various antenna conductor circumferences. For thin antenna conductors whose circumference is equal to 0.001λ or less, an average value of 0.95 may be used for K.

ANTENNA GAIN. The gain of an antenna is an indication of the antenna's concentration of radiated power in a given direction. It is the ratio of the power radiated in a given direction to the power radiated in the same direction by a standard antenna. As mentioned in Chapter II, an isotropic radiator is the basic, theoretical form of antenna, and is commonly used as a relative standard for all other antennas. However, a half-wave dipole is quite often used as a standard antenna instead of an isotropic radiator, and for this reason an antenna gain figure should always be followed with a note as to its reference antenna.

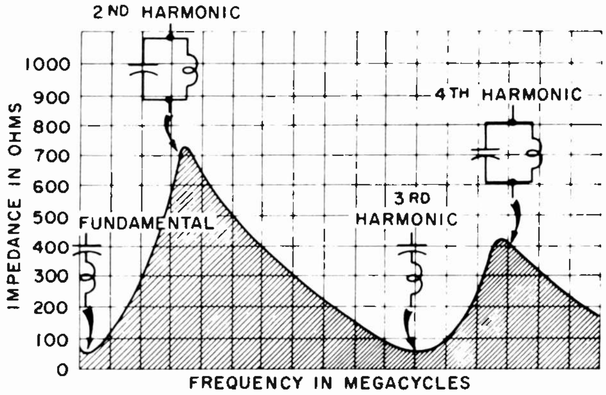


Figure 7-5. Dipole Impedance Variation

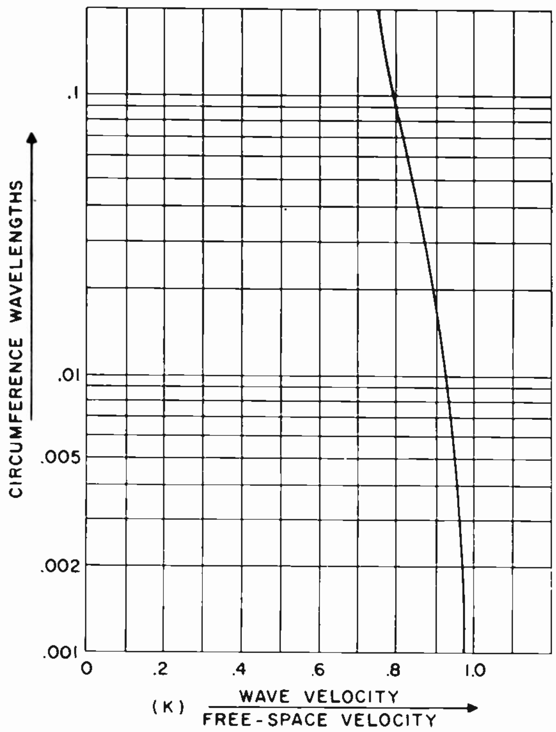


Figure 7-6. End Effect Factor

Since the effective power radiated from a transmitting antenna, or the available power at the output terminals of a receiving antenna are both dependent upon the antenna's effective area, the power gain of an antenna also depends upon the antenna's effective area. The effective area of an isotropic radiator is given by the formula

$$A_e = \frac{\lambda^2}{4\pi} \tag{7-9}$$

where

A_e = The antenna's effective area expressed in the same units as λ.

λ = Operating wavelength.

In this case, the gain of an isotropic antenna is taken as unity, and the relation between the effective area and the power gain of *any antenna* becomes

$$A_e = \frac{G\,\lambda^2}{4\,\pi} \qquad (7\text{-}10)$$

where

G = The power-gain of the antenna over an isotropic antenna.

If two identical antennas, one a transmitting and the other a receiving antenna, are placed a distance (D) apart (where $D > 2A^2/\lambda$), and A is the maximum aperture dimension, with their axes coincident, the gain of each antenna may be determined by the relation

$$G = \frac{4\,\pi\,D}{\lambda}\sqrt{\frac{P_r}{P_t}} \qquad (7\text{-}11)$$

where

G = The power gain of each antenna over an isotropic antenna.

P_r = Received power in watts at the antenna terminals.

P_t = Transmitter power in watts.

λ = Wavelength in meters.

To find the antenna gain in decibels, the following formula may be used:

$$G_{db} = 10\log_{10} G_{power-gain} \qquad (7\text{-}12)$$

If the gains of the two antennas are slightly different, then the value of G obtained with formula (7-11) satisfies the relations

$$G = \sqrt{G_1 \times G_2} \qquad (7\text{-}13)$$

where G_1 and G_2 are the gains of the two antennas. When G is known, it is possible to determine G_1 by placing each of the two antennas on a receiving mount, and measuring the power received from a third transmitting antenna. If the readings are P_1 and P_2, then

$$G_1 = G\sqrt{P_1/P_2}, \text{ and } G_2 = G\sqrt{P_2/P_1} \qquad (7\text{-}14)$$

ANTENNA DIRECTIVITY. An important factor to be considered when selecting an antenna is its directivity or lobe pattern, which actually determines antenna gain. Because an antenna may radiate in any direction, it would theoretically be necessary to show a three dimensional radiation pattern to obtain a true picture of its directivity. However, it is generally sufficient to show various sections of the lobe pattern, which will give the necessary information.

Figure 7-7 illustrates the doughnut-shaped pattern of an ordinary half-wave dipole, representing the radiated energy in all directions. For the horizontal dipole, this radiation is minimum in a direction extending from the ends of the dipoles; while in a direction broadside from the antenna, the radiation is maximum. Usually, an antenna directivity pattern is shown for vertical and horizontal planes, such as illustrated in figure 7-8. From the figure it can be seen that a half-wave dipole antenna has a non-directional pattern vertically, which means that it will respond equally well to any signal or noise arriving at any vertical angle. However, such a dipole does have some directivity in the horizontal plane or "looking down" upon the dipole. It is bidirectional in a direction at right angles to the antenna and, therefore, does not discriminate against signals arriving from the rear. For this reason, all antennas used in radio-relay applications are "backed-up" by a reflector which concentrates the antenna's directivity pattern to one direction only.

Since a half-wave dipole antenna is bidirectional in the horizontal plane, while an isotropic antenna radiates energy equally well in all directions, the dipole has a directive gain over an isotropic antenna. Directive gain is simply another way of expressing antenna power gain, and both terms may be used interchangeably. For example, because a dipole antenna is *more directive* than an isotropic radiator, the dipole will radiate *more power* in some directions and thus a *power gain* will be obtained. Numerically, the power gain of a half-wave dipole is equal to 1.64, or approximately 2.2 db over an isotropic radiator with a gain of unity.

DIRECTIVITY PATTERNS. Directivity consideration is very important in radio-relay work, and for most purposes horizontal directivity patterns are sufficient to give the desired information. In order to plot directivity characteristics, the antenna is placed on a rotatable stand and a low power source of r-f energy is placed some distance away. A receiver, which is designed especially for these measurements, is connected to the antenna and used to determine the voltage developed at the antenna terminals. In certain laboratory setups, the directivity or lobe

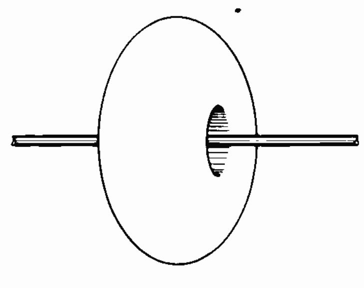


Figure 7-7. Dipole Field Pattern

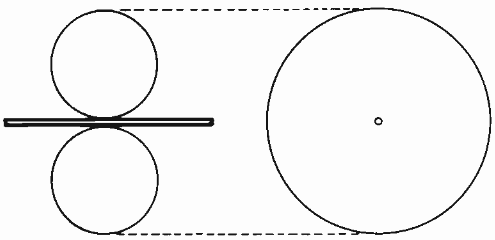


Figure 7-8. Horizontal and Vertical Directivity Pattern

patterns may be drawn by automatic curve tracing apparatus.

A special type of graph paper is used to plot antenna characteristics. As shown in figure 7-9, the center of the graph represents the location of the antenna, and around this point are drawn a series of concentric circles. These circles represent the magnitude of the voltage developed at the antenna terminals. Straight lines are also drawn, passing through the center of the graph, representing the orientation angle of the antenna. The point where the antenna lobe pattern intersects one of the lines and a concentric circle indicates the voltage developed at that particular orientation angle.

The example in figure 7-9 indicates that at 27 degrees the voltage developed by the antenna is 0.70 of the maximum voltage. This 0.70 voltage represents one half of the total power radiated in the forward direction, since

$$\text{E volts per meter} = \sqrt{P\,Z_0}\ \text{watts per meter}^2 \qquad (7\text{-}15)$$

The antenna beam width given by the example is, therefore, twice 27 degrees or 54 degrees. It is conventional to place the 0° point at the top of the graph paper, in line with the eye of the observer. Graph paper, as actually used, contains a greater number of straight lines and concentric circles, so that a greater accuracy in plotting can be obtained, and the example of figure 7-9 has been simplified to demonstrate the principle involved.

FRONT-TO-BACK RATIO. In any radio relay repeater, it is important that the signal transmitted is not picked up by its own receiving antenna. While this trouble can be minimized by several methods, its main cause is the energy radiated from the rear of the transmitting antenna. Figure 7-9 illustrates that, besides the main forward lobe, there is a second smaller lobe extending from the back of the antenna. The ratio of the maximum voltage or power effective towards the front to the maximum voltage or power effective towards the back is called the *front-to-back ratio*, and is usually expressed in db. Likewise, for a receiving antenna, the front-to-back ratio is a factor used to indicate its sensitivity to signals received at the rear with respect to signals received at the front. This is determined by measuring the voltage ratio at these two positions. For example, if an antenna develops 1 volt at 0°, and 0.1 volt at 180°, the front-to-back ratio would be 1/0.1 or 10, indicating that the antenna is 10 times more sensitive to signals arriving in the desired direction. Expressed as a formula, the front-to-back ratio becomes

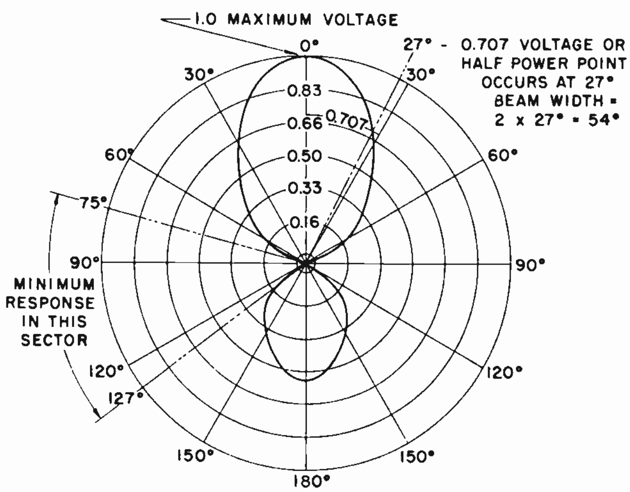
$$\text{Front-to-back Ratio (db)} = 20 \log_{10} \frac{\text{E volts per meter, front lobe}}{\text{E volts per meter, back lobe}} \qquad (7\text{-}16)$$

$$\text{Front-to-back Ratio (db)} = 10 \log_{10} \frac{\text{P watts per meter}^2\text{, front lobe}}{\text{P watts per meter}^2\text{, back lobe}} \qquad (7\text{-}17)$$

The less intense the back lobe, the larger the front-to-back ratio, and, consequently, the less interaction or feed back will be developed between transmitting and receiving antennas of the same repeater. Front-to-back ratios approaching 40 db are quite possible with parabolic reflectors, while for lens and horn antennas, the front-to-back ratios can be as high as 50 to 60 db.

EFFECTIVE RADIATED POWER. An important directivity characteristic for transmitting antennas is the effective radiated power (ERP), or the power radiated in a given direction due to antenna gain. Thus, ERP is simply an indication of a transmitting antenna's gain in terms of watts, and may be found by the formula

$$\text{ERP} = G_t \times P_t \text{ Watts} \qquad (7\text{-}18)$$



DIPOLE WITH REFLECTOR

Figure 7-9. Antenna Radiation Pattern

where

G_t = Power gain of the transmitting antenna over a reference source.

P_t = Power in watts at the terminals of the transmitting antenna.

If transmission line loss is neglected, the actual transmitter output power may be substituted. For example, a transmitting antenna with an input of 10 watts and a power gain of 5, will produce the same field strength in its direction of maximum radiation, as an isotropic antenna with an input of 50 watts.

PARASITIC ELEMENTS. It was mentioned previously that an antenna collects energy in its elements and reradiates a portion of this energy. This will occur independently of the type of transmission line used, and whether or not it is connected to the antenna. If two antennas are placed in close physical proximity, the reradiated energy will produce an interaction between the antenna elements. Therefore, a parasitic element may be regarded as a radiating element, not coupled directly to the transmission line of the antenna but which materially affects the antenna's directivity pattern. Likewise, antenna elements which have transmission line connections are known as driven elements.

The magnitude of the current in a parasitic element and its phase relation to the current in the driven element depends upon the tuning of the parasitic element. Such tuning is usually accomplished by adjusting the length (L) of the parasitic element. Depending upon the tuning of the parasitic element, it can either act as a reflector sending the maximum radiation in the $\phi = 180°$ direction as shown in figure 7-10, or as a director sending the maximum radiation in the $\phi = 0°$ direction. A reflector will generally be made at least 5% longer than the driven element, and a director approximately 5% less.

The following results can usually be expected when a reflector type of parasitic element is added to a half-wave dipole. The impedance of the combination is decreased below that of a simple dipole from 30 to 70% depending upon the length and spacing of the reflector. The directivity pattern is made essentially unidirectional at the resonant frequency, while the overall bandwidth is generally decreased. The gain of the array, however, is greater and as much as 4 db may be added by careful design.

Basically, the increased gain results as follows: the reflector element intercepts r-f energy the same as the driven element of the antenna. This energy is reradiated from the reflector and is intercepted by the driven element, thus adding to the received

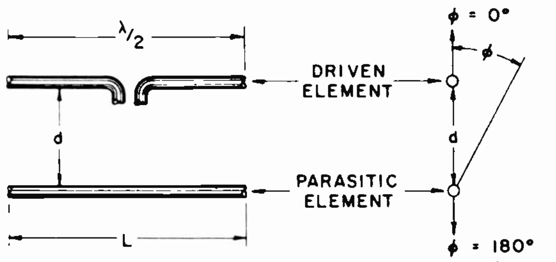


Figure 7-10. Antenna Array with One Parasitic Element

signal. Likewise, some of the energy reradiated from the driven element, reflected by the parasitic element, is intercepted again by the driven element. With fixed antenna dimensions, the action of the parasitic element is effective only over a portion of the bandwidth encountered, and as the frequency increases, the reflector becomes decoupled from the driven element.

YAGI ARRAYS. More than one parasitic element may be used with a driven element, although the net gain realized will be practical only at bandwidths up to approximately 10 mc in the UHF band, for example. The impedance of such a combination is also reduced considerably as more and more parasitic elements are added. An antenna which consists of two or more parasitic elements is commonly called a Yagi antenna or array. In this case, one of the parasitic elements is placed behind the driven element and is used as a reflector, while the other parasitic elements are placed in front of the driven member and act as directors.

With antenna arrays, such as the Yagi, it is important that the antenna present a reasonable impedance to the transmission line. A reasonable impedance may be regarded as an impedance value which approximates that of the transmission line intended for use with the antenna. For example, a half-wave dipole with one reflector and several directors may have an impedance as low, or lower than, 25 ohms, due to the influence of the parasitic elements.

In order to provide an increase in impedance, folded dipoles such as illustrated in figure 7-11 may be used. A standard folded dipole, with its upper and lower elements equal in diameter, has an impedance of nearly 300 ohms. When the upper element of the folded dipole is several times larger in diameter than the lower member, the impedance value becomes greater than 300 ohms, the exact value depending upon the ratio of the element diameters. If a very large initial impedance is required before the parasitic elements are added, a three element folded dipole may then be used, increasing the impedance value to about 675 ohms.

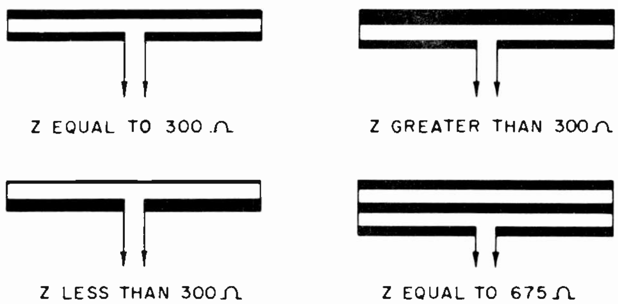


Figure 7-11. Folded Dipole Impedances

RADIO RELAY ANTENNA TYPES

LINEAR ARRAYS. An antenna array may consist of a single driven antenna with one or more parasitic elements, or it may consist of combinations of driven antennas, with or without parasitic elements. Of the latter, the simplest array consists of two driven half-wave dipoles which may be driven either in phase or in phase opposition. When the two elements are driven in phase, the direction of maximum radiation is normal to the common plane of the two elements, or *broadside*. Therefore an antenna array, consisting of two or more elements driven in phase, is termed a *broadside array*, and is the type most commonly used for radio-relay applications at frequencies below about 300 mc.

Illustrated in figure 7-12 is a typical vertical-plane pattern of a broadside array consisting of two vertical in-phase half-wavelength elements spaced one-half wavelength apart. For sake of comparison, the pattern of a single half-wave dipole with the same power input is shown by the dotted line. When the spacing between the elements of a two-element broadside array is $\lambda/2$, the gain realized is 3.9 db over a half-wave dipole, or 6.1 db over an isotropic antenna. In order to use such an antenna for radio-relay applications, a plane sheet reflector is commonly used to direct the radiated energy in a single direction. A typical broadside antenna, consisting of four half-wave dipoles operating in phase, is illustrated in

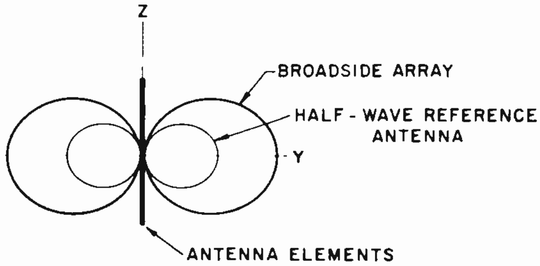


Figure 7-12. Vertical Directivity of Broadside Array

figure 7-13. Such antennas may have gains from 12 to 26 db over an isotropic antenna when operated in the frequency range from 44 to 300 mc.

When an array consists of two or more half-wave dipoles driven in phase opposition, the direction of maximum radiation is perpendicular to, and in the same plane as, the antenna elements. Such an array is commonly called an *end-fire array* and, when the spacing between elements is equal to $\lambda/4$ or less, the gain realized is at its highest value. For the case where only two driven elements are used, the maximum gain approaches 6.1 db over an isotropic antenna. End-fire arrays are sometimes called flat-top beam antennas since the array is commonly operated with both elements horizontal to the ground as illustrated in figure 7-14. Since end-fire arrays are essentially bidirectional, a reflector would have to be used to make them suitable for radio-relay applications. However, the use of a reflector with such an array would actually reduce the overall gain by changing the phase characteristics of the driven elements, since one would be nearer the reflector than the other. End-fire arrays, therefore, are seldom used as radio-relay antennas because it is impossible to obtain a high gain in one direction only.

REFLECTOR ANTENNAS. In order to modify the radiation pattern of an antenna for radio-relay applications, reflectors are generally used to increase forward gain and eliminate any backward radiation. Various types of reflectors are illustrated in figure 7-15, the most commonly used being the plate sheet reflector, the corner reflector, and the paraboloid reflector. Many other shapes are possible, depending upon the directivity pattern desired and the space-weight considerations.

Essentially, a reflector intercepts radio energy and reradiates it in another direction in the same manner as a mirror reflects a light beam. However, just as some light is scattered by a mirror, some of the power of the radio beam is scattered by the reflector. The amount of scattered power depends upon the smoothness and type of material used in a reflector. *Smoothness* here may be defined as the ratio between the size of the irregularities of the reflecting surface and the wavelength of the incident energy. If the irregularities of the reflector are small in comparison to the operating wavelength, then a good reflector is obtained. In actual practice, if the irregularities of the reflector are no greater than $\lambda/8$, satisfactory reflection will be obtained with only a small amount of scattering. Since most antenna reflectors are exposed to weather, a definite advantage is found in the use of a screen, or grill, which will offer little

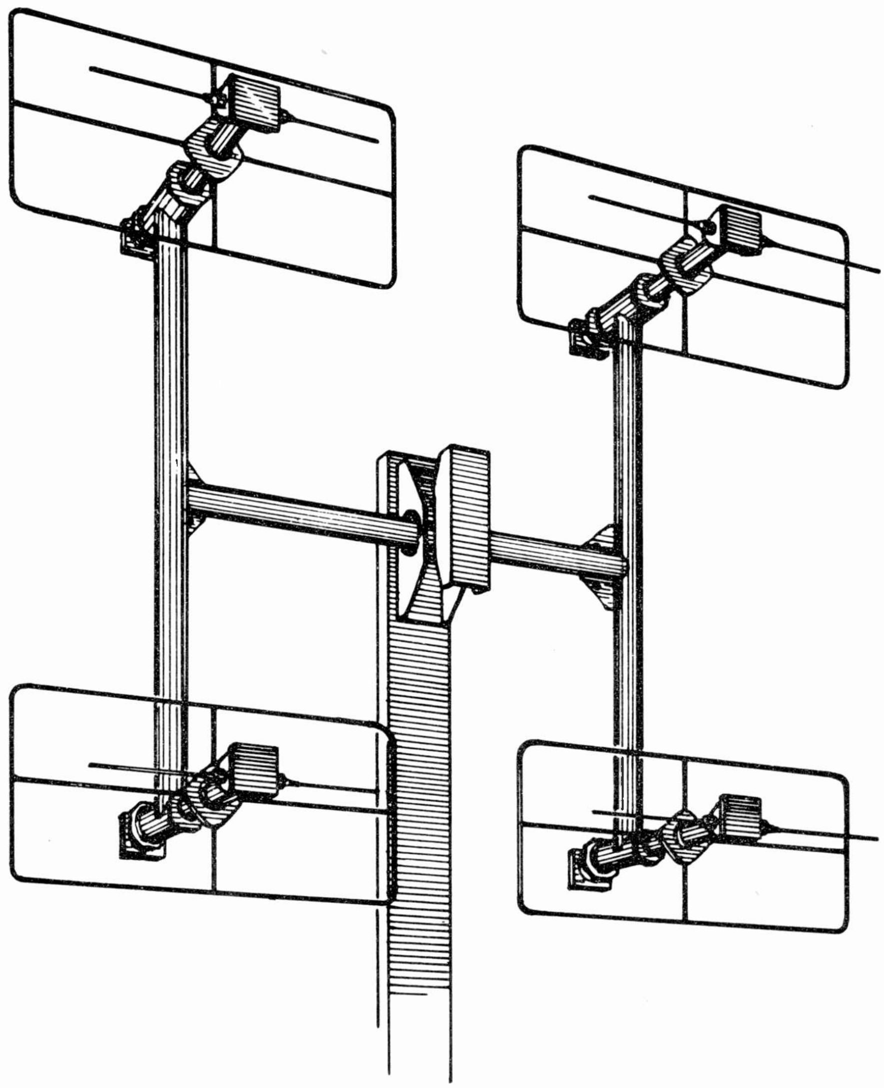


Figure 7-13. Four Element Broadside Array

resistance to wind but will still reflect energy satisfactorily.

Perhaps the simplest reflector, other than a thin parasitic element, consists of a flat sheet of metal, or wire mesh, such as illustrated in figure 7-15A. The gain realized with such a reflector is generally limited to about 9.2 db over an isotropic antenna when the spacing between the driven element and the reflector is equal to 0.1λ. A flat reflector does not appreciably change the radiation pattern, but only converts the normal bidirectional pattern into a larger unidirectional pattern by eliminating backward radiation. Therefore, a flat surface reflector only increases forward gain and does not "beam" radiation. The forward gain in field intensity as a function of the spacing (S) is presented in chart form under figure 7-15A. Provided that losses are negligible, very small spacings can be effectively used, although at such spacings the bandwidth is considerably narrowed.

By bending the sheet reflector into a cylindrical parabolic shape, higher directivity and gain will result. A cylindrical parabolic reflector is illustrated in figure 7-15B, showing a half-wave dipole located at the focus of the parabolic curve. If the distance across the aperture is several wavelengths, most of the directivity will be in the horizontal plane.

Even though high gains may be obtained with cylindrical parabolic reflector systems, they offer no advantages over a corner reflector, such as that illustrated in figure 7-15C. In practical form, a corner reflector consists of two flat reflecting sheets intersecting at right angles, thus forming a square corner. Corner angles greater or less than 90° can be used, although at angles much less than 60° no net advan-

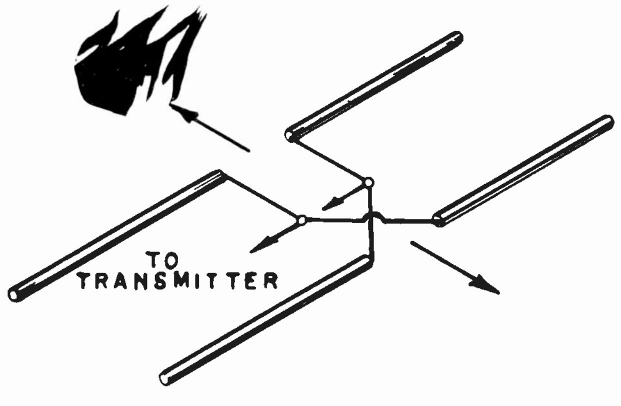


Figure 7-14. Endfire "Flat Top" Array

tage is realized. If a dipole radiator is located within the plane bisecting the corner angle and parallel to the intersection of the two sheets, gain figures from 10 to 15 db over an isotropic source can be obtained. Curves of gain versus dipole-to-corner spacing for corner reflectors of 90° and 60° are presented in figure 7-15C, assuming no losses. As with the plane sheet reflector, small spacings are objectionable because of bandwidth limitations, but on the other hand, too large a spacing results in low gain.

Since the corner reflector and the cylindrical parabolic reflector concentrates energy in one plane only, they are not often used as radio-relay antennas, although they are occasionally used, along with the plane sheet reflector, as passive reflectors. For reliable radio-relay operation, an antenna which has high directivity in both horizontal and vertical planes must be used in order to transfer a maximum amount of energy from one relay to another. The most common reflector type used to accomplish this is the *paraboloidal,* or *parabolic reflector* as illustrated in figure 7-15D.

The surface generated by the revolution of a para bola around its axis is called a paraboloid or parabola of revolution. If a half-wave dipole is placed either vertically or horizontally at the focus of such a reflector, considerable "beaming" action will take place. If the aperture of the parabolic reflector is very large in terms of λ, a spherical wave striking the reflector will be converted to a plane wave upon reflection. Unfortunately, a half-wave dipole does not provide an isotropic, spherical source of illumination, and it is impossible to provide uniform illumination

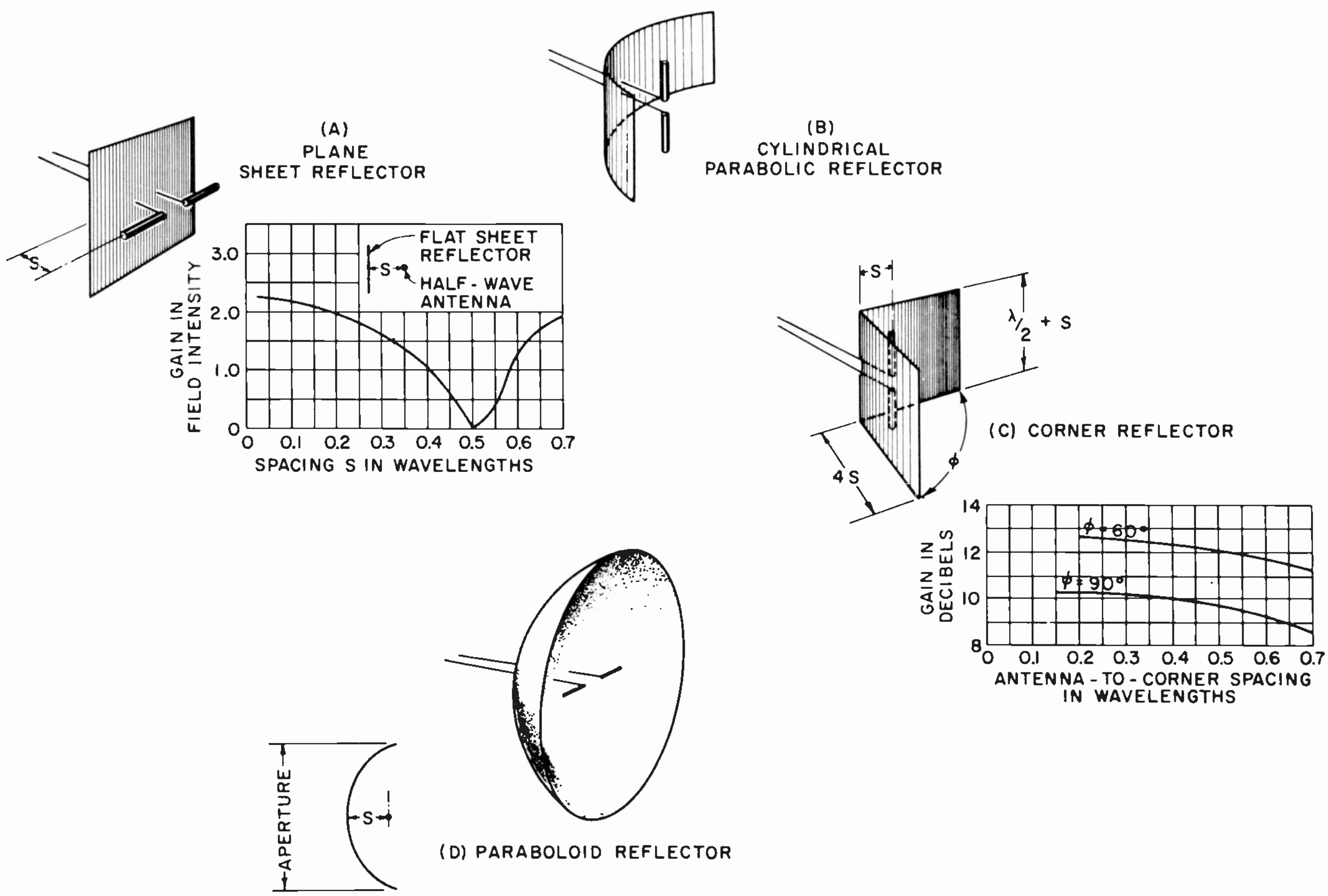


Figure 7-15. Reflector Types

of the "dish". The resultant beam, therefore, is elliptical in shape and somewhat broader than might be expected.

For a given reflector diameter, the characteristics of the reflected wave are best when a shallow dish is employed, that is, a parabola with a somewhat long focal length. Compared to a shorter focal length, however, the amount of incident illumination is then reduced. If the distance (S) between the vertex and the focus of the parabolic reflector (focal length) is an even number of quarter wavelengths, the direct and reflected radiation, in an axial direction from the source, will tend to be in phase opposition and therefore cancel the central region of the reflected wave. However, if the focal length is held within the limits given by the formula

$$S = \frac{\eta \lambda}{4} \qquad (7\text{-}19)$$

where η = 1,3,5,---, the direct radiation in an axial direction from the source will be in the same phase and tend to reinforce the central region of the reflected wave.

When the focal length of a parabolic reflector cannot be prescribed, direct radiation from the driven element can be minimized and reflector illumination increased by using a parasitic or small flat sheet reflector close to the dipole. This secondary reflector is located so as to concentrate most of the radiation from the dipole into the dish.

The major factor contributing to the gain, or beaming action, of a parabolic reflector is the total area of the reflecting surface. For a uniformly illuminated dish, the reflected beam angle is proportional to the wavelength and inversely proportional to the linear dimension of the aperture. The beam width between first null points for large circular apertures is given by the formula

$$\text{Beam width} = \frac{140}{D_\lambda} \text{ degrees} \qquad (7\text{-}20)$$

where D_λ = the diameter of the aperture in wavelengths. The beam width between half-power points for a large circular aperture may be found by the formula

$$\text{Beam width (P/2)} = \frac{58}{D_\lambda} \text{ degrees} \qquad (7\text{-}21)$$

The beam width given by formula (7-21) is actually an "equivalent beam angle" and is the cross section of the cone where the radiated energy is assumed to be concentrated. Actually, the beam of a directive parabolic reflector starts out in a cylindrical form, as illustrated in figure 7-16, to a certain point and

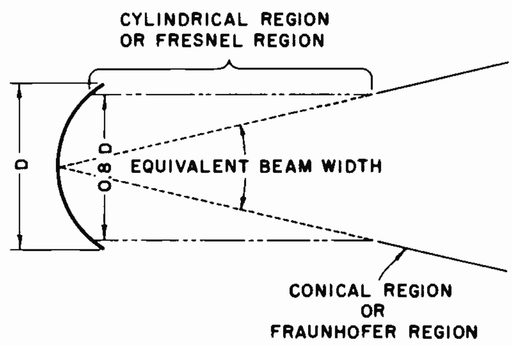


Figure 7-16. Parabolic Reflector Beaming Action

then diverges progressively into a conical beam. Such a beam has a cross section equal to the equivalent beam angle with its apex located at the center of the reflector.

The cylindrical portion of the beam is called the near field or Fresnel region and the conical part is called the far field or the Fraunhofer region. The cylindrical region has a substantially constant field and a parallel beam, so that any reflector with sufficient cross section that is placed in that region will reflect the radiated energy with very little loss. The cylindrical region's maximum distance is proportional to the square of the aperture diameter and inversely proportional to the wavelength as given by the equation

$$\text{Fresnel Region Length} = \frac{D_\lambda^2}{2\lambda} \qquad (7\text{-}22)$$

For example, at 2,000 mc, using a ten foot parabolic reflector, the Fresnel region is approximately 100 feet long, and at 8,000 mc with a five foot parabola, the length would be the same. However, a ten foot parabola at 5,000 mc would have a Fresnel region of 250 feet in length.

This definite Fresnel region is occasionally used to advantage by relay systems employing passive reflectors. In this type of equipment the transmitter and antenna can be located at ground level and the antenna signal beamed at a passive reflector mounted above the equipment. If the distance between the antenna and the passive reflector does not exceed the cylindrical region length, the passive reflector will then radiate the energy to a receiving source with a minimum of loss when reflected.

The gain of a large parabolic reflector, where the aperture is uniformly illuminated, may be found by the formula

$$\text{Gain (db)} = 10 \log_{10} 6D_\lambda^2 \qquad (7\text{-}23)$$

where D_λ = the diameter of the aperture in wavelengths. For example, a parabolic antenna whose

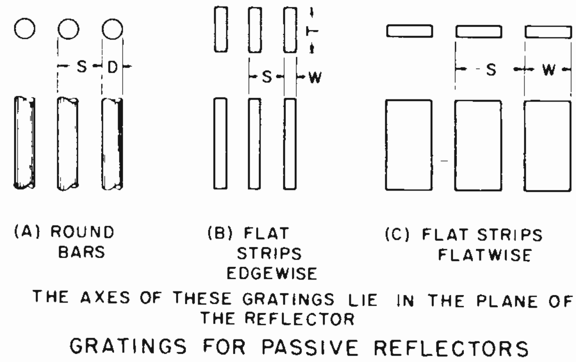


Figure 7-17. Gratings for Passive Reflectors

diameter is 15 wavelengths and is perfectly illuminated has a gain of nearly 32 db with respect to an isotropic antenna. Since, in actual practice, it is impossible to achieve uniform illumination, a more accurate answer will be obtained when calculating beam widths and gains by using an equivalent reflector diameter equal to 80% of the actual diameter.

Parabolic antennas of the type just described are suitable for use in the frequency range from 250 to above 13,000 mc with excellent results in all cases. At frequencies above about 3,000 mc, however, it becomes impractical to use half-wave dipole antennas as the driven elements, and some other means of driving the parabola must be used. At these higher frequencies, the horn type of antenna will give excellent results as the driven element, especially when the horn is fed by means of a waveguide transmission line.

REFLECTOR CONSTRUCTION. There is no rule that states that reflectors must be made of any particular pattern of openings. If wind resistance were no problem, the reflector could be made of sheet metal. Since winds will cause the reflector to move, unless rigidly mounted, reflectors are made of gratings

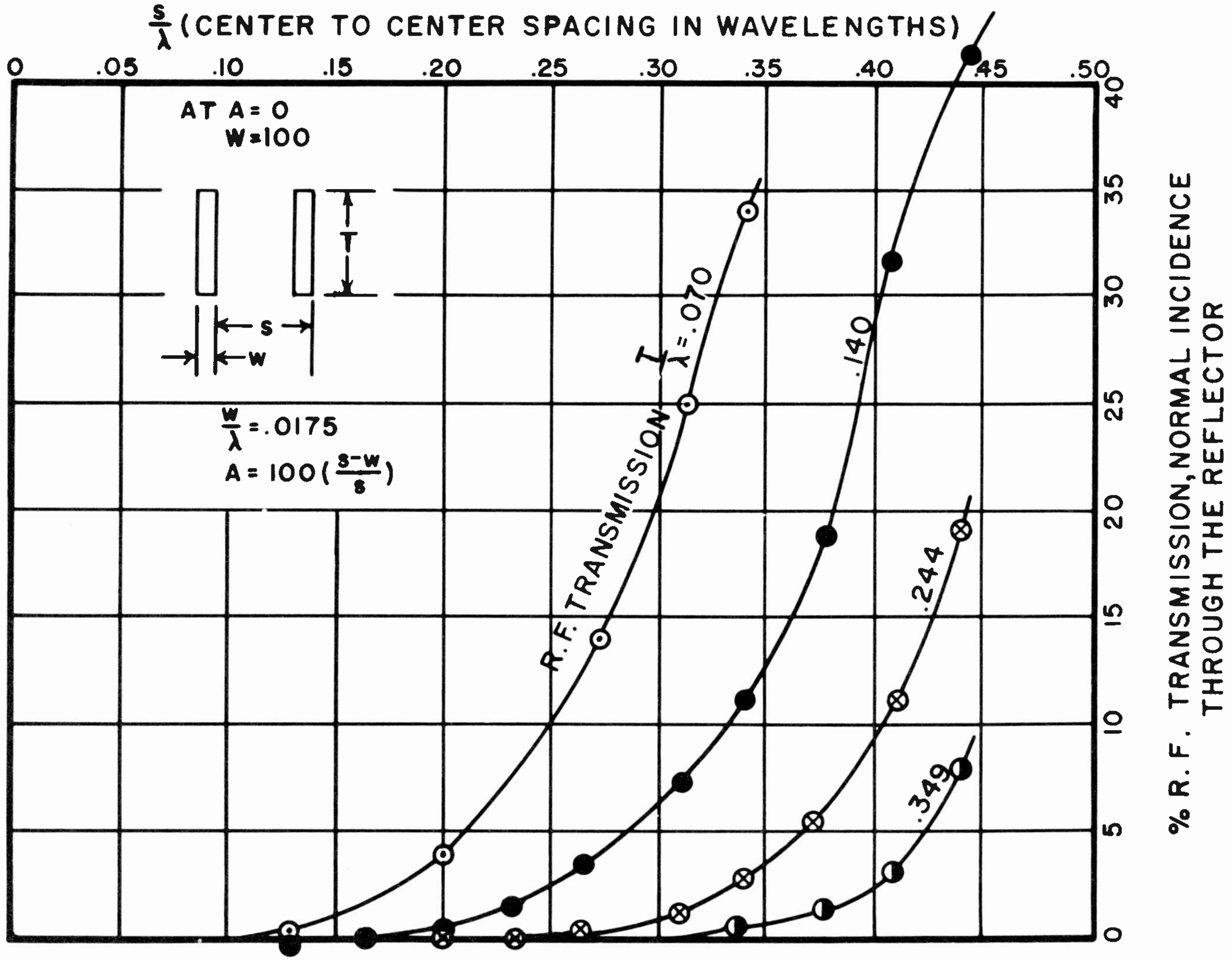


Figure 7-18. Reflector Power Loss

or screens that give little wind resistance, but still reflect almost all of the incident wave. The reflector may be a rigid or reinforced woven metal screen, or it may be a piece of reinforced perforated sheet metal.

The openings in either the punched sheet or screen are limited in size by the frequency of the incident wave. If the holes are square, as found in a screen, the edge length of the holes should be less than $\lambda/8$. Where a sheet is punched with round holes, the diameter of the holes should be less than $\lambda/8$. The power transmitted through the reflector can be kept within reasonable limits if the openings are kept small. For example, a copper screen woven of #17 B&S wire with openings ⅛″ square will reflect about 95.5% of the incident power at a wavelength of 9 cm.

Where greater rigidity for the reflector is needed, grates constructed of round, square, or flat bars, may be used instead of screen. The flat bars may be mounted edgewise, or flatwise (figure 7-17). When flat bars, or strips, are used, they are generally mounted edgewise, since better reflection with less wind resistance is obtained. The width of the strip is not of prime importance, except that, in general, the wider the strip, the better the reflection. Better reflection may also be obtained by using thicker strips. The spacing of the bars is of the greatest importance. Spacing of the gratings should be less than $\lambda(1 + \sin\theta)$ center to center, where θ is the angle of incident wave with the normal to the reflector. It is a good rule to limit the opening between the gratings to $\lambda/8$.

In constructing gratings of flat strips, the thickness may be any desired value. Figure 7-18 shows values for transmitted power (power lost through the reflector, instead of being reflected from it) for different openings between the strips. Lost power is also determined for different values of thickness. This diagram may be used in determining the thickness of the strips in designing a reflector.

Any grating, regardless of shape, is critical to polarization. In order that a minimum of power be lost through the reflector, the electric field vector (E vector) of the incident wave must be in the plane determined by the incident wave and one of the bars of the grating. In other words, if the wave is horizontally polarized, the bars must be installed horizontally. Figure 7-19 illustrates a reflector antenna designed for 300 mc and constructed with a wire grating.

HORN ANTENNAS. Several types of horn antennas are illustrated in figure 7-20; those on the left are rectangular horns energized from rectangular wave guides, while those on the right are circular types

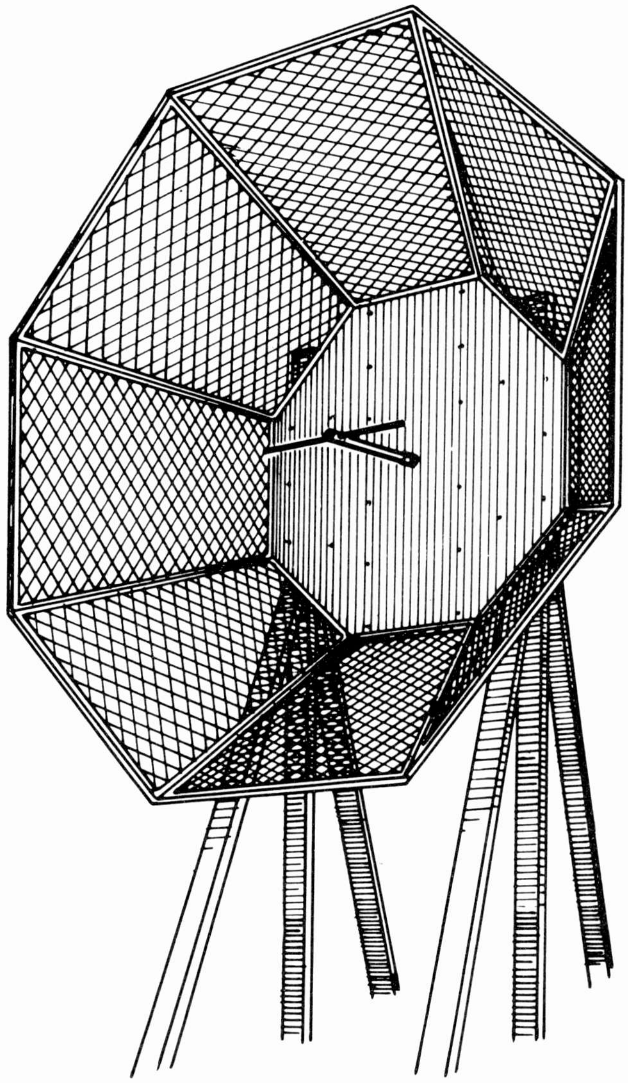


Figure 7-19. Low Frequency Reflector Antenna

energized from circular wave guides. The transition region between the wave guide and the aperture is sometimes given an exponential taper as illustrated in order to minimize reflections of the guided wave. However, horns with straight flares are more commonly used since they are easier to manufacture.

The radiation pattern of a horn antenna is largely determined by the horn aperture and the field distributions within the aperture. For a given aperture, the directivity is maximum with a uniform field distribution while variations in the field across the aperture decrease the directivity. To obtain an aperture distribution as uniform as possible, a long horn with a small flare angle is required, but for convenience the horn should be as short as possible.

The beam width and gain characteristics for circular horns are the same as for parabolic reflectors and formulas 7-20 through 7-23 may be used for such

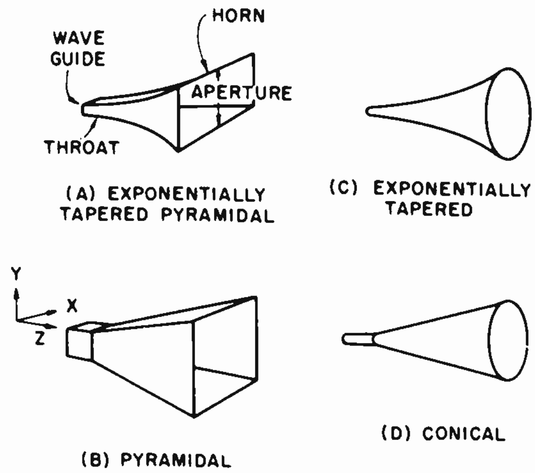


Figure 7-20. Horn Types

calculations. For rectangular horns of optimum length, the beam width between the first nulls is given by the following formula

$$\text{Beam width} = \frac{115 \text{ to } 172}{A_{E\lambda}} \text{ degrees} \qquad (7\text{-}24)$$

or

$$\text{Beam width (P/2)} = \frac{56 \text{ to } 67}{A_{E\lambda}} \text{ degrees}$$

(Between half-power points)

Likewise, the gain of a rectangular horn, uniformly illuminated is given by the formula

$$\text{Gain (db)} = 10 \log_{10} 4.5\, A_{E\lambda}\, A_{H\lambda} \qquad (7\text{-}25)$$

where

$A_{E\lambda}$ = E aperture in free-space wavelengths

$A_{H\lambda}$ = H aperture in free-space wavelengths

when a horn antenna is used to excite a parabolic reflector, it is usually in the form illustrated in figure 7-21, where the waveguide feeding the horn also acts as a support for the horn. When greater gains and higher front-to-back ratios are desired at frequencies above 3,000 mc. horn antennas are often used in conjunction with lens antennas which are situated in the aperture of the horn.

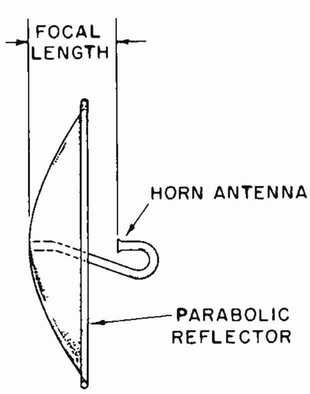


Figure 7-21. Horn Fed Parabola

LENS ANTENNAS. A lens antenna is a structure transparent to radio waves and with a relative dielectric constant different from unity, constructed in such a manner as to produce a desired radiation pattern. Such structures may employ dielectric lenses made of lucite or polystyrene, artificial dielectric lenses constructed of flat metallic strips, or metal plate lenses.

Dielectric lenses and artificial dielectric lenses *increase* the electrical path length of a spherical wave front, thus *retarding* the incident spherical wave as illustrated in figure 7-22A. Such lenses follow the laws of geometrical optics in that they may be designed by ray analysis. Metal plate lenses *decrease* the electrical path length of a spherical wave front and, therefore, *accelerate* the incident wave as shown in figure 7-22B. In this type of lens, metal plates are constructed parallel to the plane of the electric field and guide the waves in a fashion similar to that of a waveguide transmission line.

All lens antennas either convert a spherical wave front to a plane wave front or a plane wave front to a spherical wave front. Thus, when used as a receiving antenna, a lens collects a portion of the incident plane wave front and focuses it upon a primary antenna. When used as a transmitting antenna, a lens receives the energy radiated from a primary antenna and converts it to a concentrated beam.

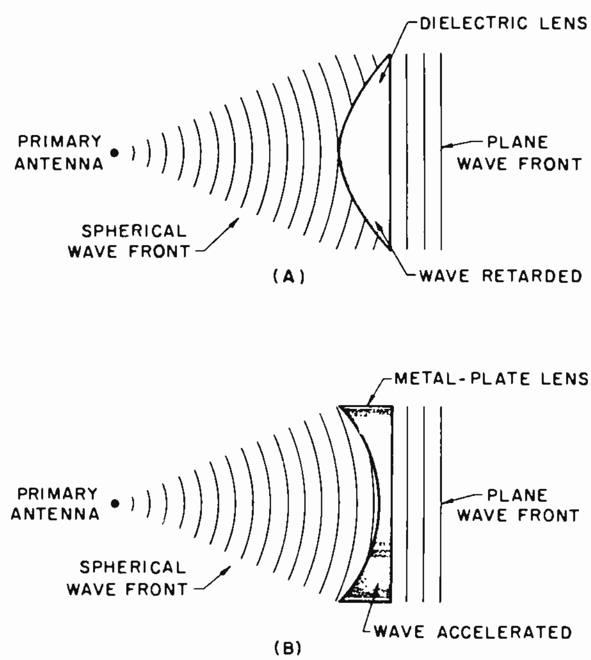


Figure 7-22. Lens Antenna Characteristics

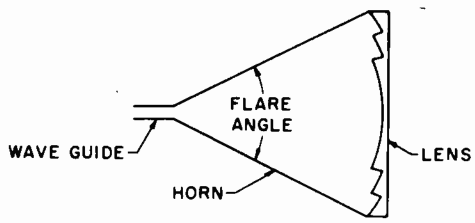


Figure 7-23. Lens-in-Horn Antenna

For maximum radiation efficiency, a horn antenna is generally used as a primary antenna for feeding lens antennas. The lens may be conveniently mounted in the aperture of the horn, as illustrated in figure 7-23, thus shielding the back surface of the lens. Lens antennas are in common use above 3,000 mc where their apertures can be made large in terms of wavelengths, and gains up to 40 db are possible in some instances.

CHAPTER 8

THE TERMINAL TRANSMITTER

INTRODUCTION

The terminal station of a radio relay system consists of a transmitter, a receiver, and associated equipment to send and receive information to and from the other terminal stations in the system. Previously, the means of radiating the desired signal and the effects of the atmosphere and earth on the signal have been discussed. In this chapter the means of generating this signal and combining it with the desired information in the form of a modulated r-f carrier will be considered.

The transmitter consists basically of an oscillator, modulator, and amplifier, the complexity of these units depending on frequency, power output required, and type of modulation used.

OSCILLATORS

Conventional oscillators using lumped capacitance and inductance operate satisfactorily, although with decreasing efficiency, up to a frequency of about 150 mc. Above this frequency special UHF and microwave oscillators must be used since conventional vacuum tube element spacing becomes an appreciable fraction of a wavelength in size, and even a straight piece of wire has enough inductance and capacitance to affect the circuit. At first the attempt was made to solve this problem by decreasing the tube spacing and by using connecting leads as short as possible, but this soon led to the limiting factor of arcing between elements. As frequencies increased above 500 mc, completely new types of circuits were developed in order to generate the required frequencies. In these oscillators, the tuned circuits become part of the tube in the form of resonant cavities, similar in effect to the tuned cavities used to produce audio tones in many musical instruments.

VHF OSCILLATORS. Components of VHF oscillators must be designed carefully; physical layout of parts and the use of short leads in circuit wiring are important. The use of lumped property components offers the advantage of small physical size with circuitry which is easily adapted to wide range tuning. Where space is not limited, distributed property components may be used increasing the efficiency and stability of operation.

Above 400 mc, distributed property components are used almost exclusively, since the stability and efficiency of lumped property components become too low, and the physical size of distributed property components becomes practical for use. At this point vacuum tube design is also considerably changed and special lighthouse, pencil, or disk-seal planar-type triodes are used. Figure 8-1 shows the cross section of a lighthouse tube. The effects of transit-time are reduced by making electrode spacing very small. One of the disadvantages of decreasing tube size is that less power may be handled because of the decrease in heat dissipating area. In one type of tube, the inverted lighthouse or "oil-can" tube, the power handling capacity has been greatly increased by making the plate of the tube the largest element. These tubes are designed to be inserted into a resonant cavity in such a way as to make the tube itself a portion of the cavity, thus eliminating lead inductance. Loss from radiation from the tube elements and leads is eliminated since the tube is completely shielded by the resonant cavity.

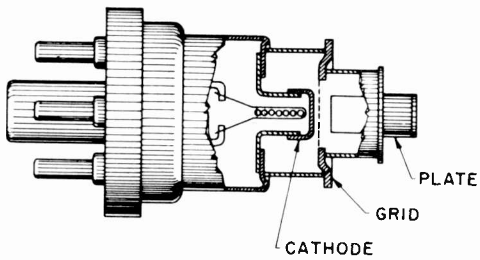


Figure 8-1. Lighthouse Tube (Cross-Section)

Figure 8-2 shows a lighthouse tube and cavity, the combination forming short-circuited coaxial lines in the cathode-to-grid and grid-to-plate circuits. The cavity forms the outside conductor, while the tube elements themselves form the inner conductor of the coaxial line. The lines are tuned by sliding pistons which change the length of the coaxial lines. Oscillation may be produced by providing regenerative feedback either by connecting input and output coupling loops together in proper phase relationship, or by building capacitance into the cavity so that the proper phase of feedback is established. Other types of UHF vacuum tubes may be used in similar circuits by properly designing the resonant cavity for the tube.

MICROWAVE OSCILLATORS. As microwave frequencies are approached the triode coaxial oscillators begin to fall off in efficiency and again new systems must be resorted to in order to produce oscillations at these new frequencies. Three general types of oscillators have been developed for use in these ranges: the klystron, the magnetron, and the traveling wave tube. While in the coaxial triode oscillators the problem of interelectrode transit time was encountered, these new types of oscillators take advantage of the finite speed of the electron.

The first type of microwave oscillator is the klystron. Two main categories of klystron oscillators exist: the reflex klystron and the 2-cavity klystron. The reflex klystron has low efficiency (about 1 per cent) and generates relatively low power. It is primarily used in receivers, local oscillators, and test equipment. Two-cavity klystrons are much more efficient (from 20 to 40 per cent) and can be designed to generate high power levels. Most of the limitations of conventional negative grid tubes do not exist in klystrons. The cathode and anode are outside the r-f field and therefore may be made as large as desired. The cathode to anode spacing is of the order of one inch so that extremely high voltages may be used without danger of internal arcing. The only limiting factor in the amount of power which may be produced by a klystron is the loss in the dielectrics making up the windows between the output cavity and the load.

Figure 8-3 illustrates the operation of a 2-cavity klystron. A stream of electrons from an electron gun passes through a resonant cavity called a buncher. This cavity is the input cavity and contains an r-f field corresponding to the signal input. In the case of an oscillator, the input is fed back with the proper phase relationship from the output or catcher cavity. The buncher either accelerates or retards the electrons in the stream depending on the portion of the r-f cycle. Following the buncher there is the drift space where the electron beam is unaffected except by the uniform accelerating force of the anode voltage. In this space the electrons form into bunches, the retarded electrons falling back and the accelerated electrons moving forward to the next bunch. When the electrons, now bunched, reach the catcher, they set up a varying electric field from which energy may be taken. The electrons themselves continue on to the anode, or collector, where their remaining kinetic energy is dissipated as heat. Figure 8-4 illustrates the bunching effect by means of a graph representing the distance-time relationship of each electron. The slope of each line represents the velocity of an electron emitted at the given time relationship

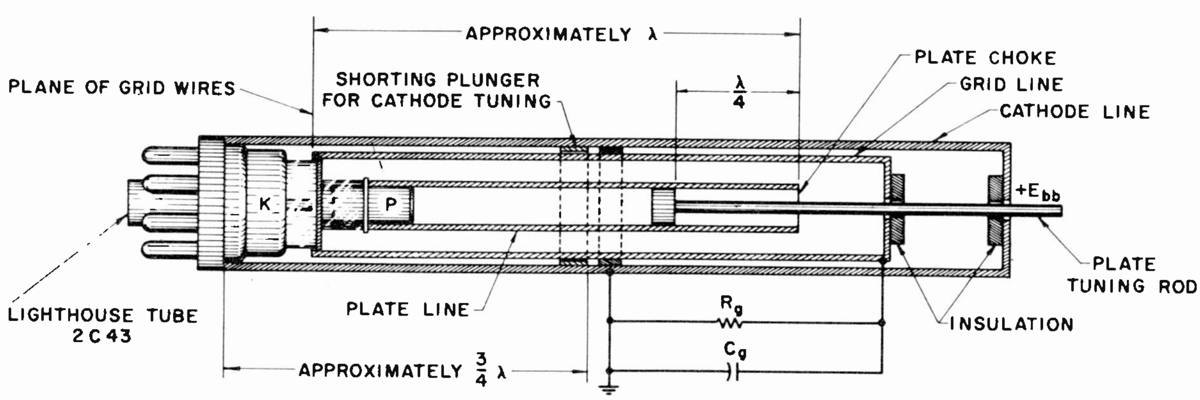


Figure 8-2. Lighthouse Tube Oscillator Circuit

with the input wave shown.

A modified version of the 2-cavity klystron is the reflex klystron oscillator. The operating principles remain nearly the same. In this klystron the coupling between the input and the output is accomplished by the electron beam itself. In fact, the two resonating cavities are replaced by one cavity which functions both as buncher and catcher. Figure 8-5 is a simplified diagram illustrating the operation of a reflex klystron oscillator.

The electrons are produced and accelerated by an electron gun as before. Then they pass through the cavity for the first time, being velocity modulated as in the buncher of a 2-cavity klystron. The electrons travel into the drift-space but instead of being further accelerated, they are in a uniform retarding field produced by the negative repeller plate. The electron beam slows to a stop and then reverses direction being accelerated back toward the cavity. As the bunched electrons pass through the cavity for the second time, they give up part of their energy to the cavity and are then stopped by the cavity, which also functions as the collector. The frequency of operation can be changed to a limited extent by changing the repeller voltage, thus changing the transit time in the drift space.

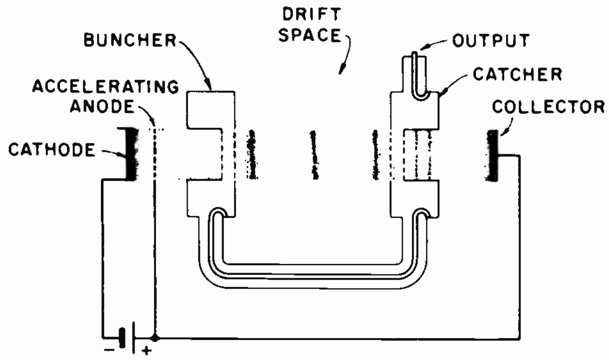


Figure 8-3. Two-Cavity Klystron Oscillator

The reflex klystron is less efficient than the 2-cavity klystron because a single resonator performs both functions of bunching and catching. On the other hand the single resonator tuning and the ease of electrical tuning by varying the repeller voltage make it better for use when only small amounts of power are required. Another advantage of the reflex klystron is its greater stability as compared to a 2-cavity klystron when used as a master oscillator. Figure 8-6 illustrates the efficiency of operation of klystrons at various frequencies. It can be seen that the efficiency of a klystron falls off slowly as the frequency is increased and that the maximum power is not determined by the wavelength.

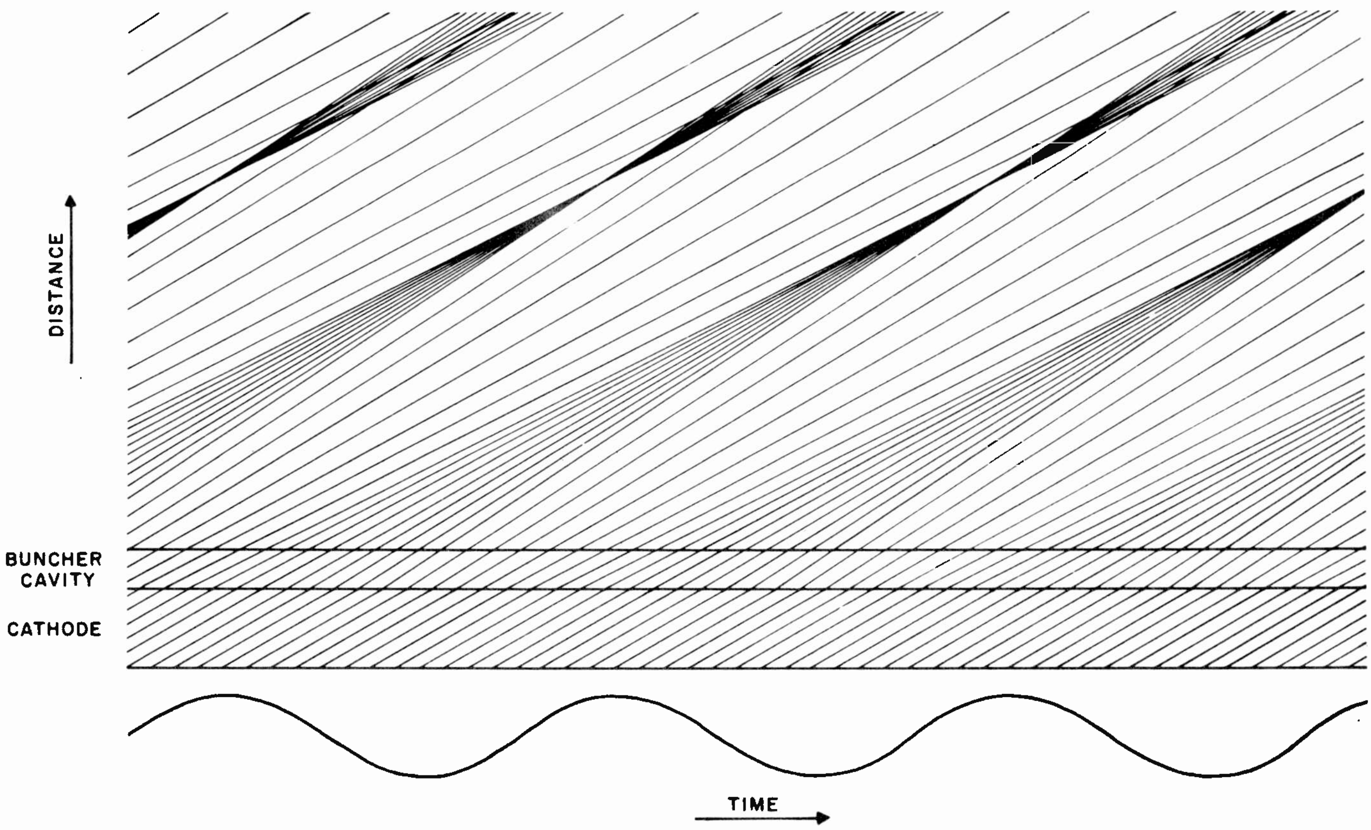


Figure 8-4. Bunching in Two-Cavity Klystron

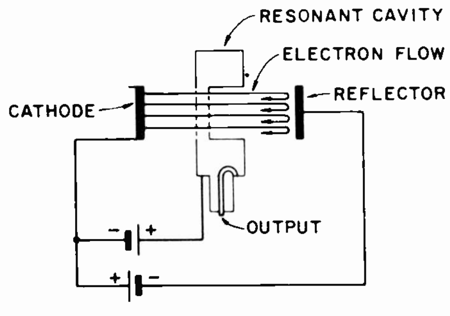
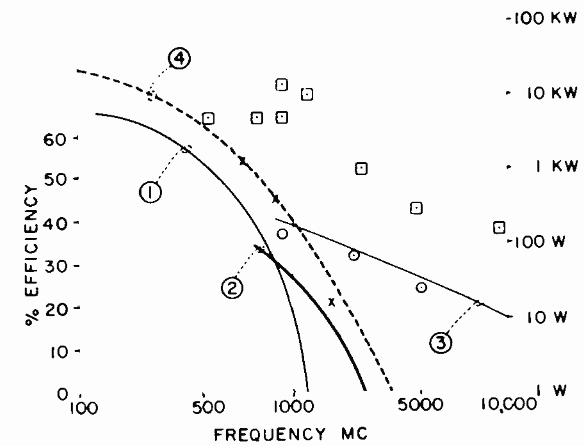


Figure 8-5. Reflex Klystron Oscillator

The second main type of microwave oscillator is the magnetron. To aid understanding of the principle of operation in a traveling-wave magnetron, the most commonly used, consider the movement of an electron in magnetic and electric fields. An electron moving at right angles to a magnetic field will be acted upon by a force perpendicular to both its direction of motion and the magnetic field. This force does not change the velocity of the electron but causes it to move in a circular path, the radius of the path being determined by the magnetic field strength and the velocity of the electron. An electron moving parallel to an electric field will be either accelerated, taking energy from the field, or retarded, giving energy to the electric field.

In the magnetron, figure 8-7, there are two electric



CURVE ①: EFFICIENCY VERSUS FREQUENCY FOR TYPICAL UHF TETRODE - 4X150G. (PLATE DISSIPATION 150 WATTS.)
CURVE ②: EFFICIENCY VERSUS FREQUENCY FOR TYPICAL UHF TRIODE - 2C39. (PLATE DISSIPATION 100 WATTS)
CURVE ③: TYPICAL EFFICIENCY OF KLYSTRONS VERSUS FREQUENCY (INDEPENDENT OF OUTPUT POWER.) THIS IS THE EFFICIENCY AT THE OPTIMUM FREQUENCY FOR EACH TUBE.
CURVE ④ (DOTTED): MAXIMUM POWER OUTPUT OF THE LARGEST COMMERCIALLY AVAILABLE NEGATIVE-GRID TUBE AT VARIOUS FREQUENCIES.
POINTS □ CW POWER OUTPUT OF VARIOUS KLYSTRONS (NOT THE LARGEST POSSIBLE.)

Figure 8-6. Efficiency-Frequency Diagram

fields operating. The first is a high voltage d-c field which accelerates the electron from cathode to anode. In addition, modifying the d-c field, is an r-f field at the operating frequency moving as shown in the diagram. The direction of the magnetic field, which is produced by a permanent magnet, is into the page. Under the action of these fields, the electrons emitted from the cathode will follow the path shown in figure 8-8 under certain conditions. The first of these conditions is that the magnetic field strength be high enough so that the electron, without the action of the r-f field, will not reach the anode, but will instead be curved back to the cathode. Secondly, the average electron motion to the right must equal the wave velocity to the right. The slowed wave velocity is produced by the action of the resonant cavities decreasing the wave velocity to a value which can be equaled by the electron flow.

Under these conditions, the electron emitted from point 1 will find itself in a retarding field and will not have enough energy to reach the cathode again before stopping. When it stops, short of the cathode, it again will try to reach the anode. Since it is moving to the right at the same rate as the r-f wave, it will again find itself in the retarding field. This process will continue, the electron giving up part of its energy to the wave each time until the electron finally reaches the anode. If an electron is emitted from the cathode at a point where it would take energy from the field, it is speeded up and driven back to the cathode, thus removing it from further interaction with the r-f field (figure 8-8, point 2). To produce regenerative feedback for oscillations, all that is necessary is to bend the magnetron around upon itself as shown in figure 8-9. The magnetron is basically a fixed frequency device, but certain of the newer types may be frequency modulated by changing the potentials on certain elements.

Anode power is normally applied to a magnetron in very short pulses of very high amplitude. Voltages of 40 kilovolts and currents of 100 amperes are not unusual in pulsed magnetron service. It is possible to produce a peak power of 2.5 megawatts at 3,000 mc with an efficiency as high as 50%. At frequencies of 25,000 mc, more than 50 kw may be obtained, but the efficiency will then fall to about 25%.

The traveling wave tube connected as an oscillator is the last main type of microwave oscillator. It is essentially an amplifier which uses an electron beam and an r-f wave traveling together in such a way that the wave accepts energy from the electron beam. In some ways it is similar to the linear magnetron discussed previously. The traveling wave tube consists of a helical coil inside a conductor. It may be considered as a coaxial cable with a helical inner conductor. The operation is as follows: the r-f wave to be amplified travels along the helical coil which greatly reduces the velocity of the wave. This slower velocity causes the r-f wave to travel at the same speed as an electron beam centered in the helical coil, which enables the r-f wave to accept energy from the beam. Figure 8-10 illustrates the r-f wave form and the electron beam path.

Assuming the electron velocity and the wave propagation velocity are the same, the electrons in the beam will be retarded or accelerated by the electric field. This will cause bunching to occur, with the electron bunches forming in alternate points of zero longitudinal electric field, as shown at points 1 in figure 8-10. In producing these bunches, as many electrons are retarded as are accelerated and no net transfer of energy is made in either direction. Since this would produce no amplification, some means must be found of obtaining a transfer of energy from the electron beam to the electric field. The bunches now form in the positions shown in the diagram. This can be done by a slight increase in the velocity of the electron beam. The electron bunches are now at a retarding point of the electric field and the electrons are retarded for a larger period of time than they are accelerated. This will produce a transfer of energy to the wave, and therefore the wave is amplified. A necessary addition to the traveling wave tube is some means of preventing the electron beam from spreading. This is done by using a longitudinal magnetic field. As long as the electron beam moves parallel to the magnetic field it has no effect on the electrons. When the electron strays from a parallel path, however, the magnetic field forces the electron back into the beam. By coupling the output to the input in the proper phase relationship, oscillation may be produced. The traveling wave tube can be used over a great range of frequencies with high gain, at a cost of low proper output and efficiency.

FREQUENCY STABILIZING SYSTEMS

An ideal oscillator would be one in which the frequency could be easily adjusted and once set, remain at that frequency regardless of temperature, output load, or voltage input. At low frequencies these conditions are relatively easy to approach, but as operating frequencies are increased, stability of operating frequency becomes more difficult to obtain. Even by attaining the same percentage of stability, which in itself is hard to do, serious frequency shifts may occur at microwave frequencies. A frequency shift of .01% at 1 mc is only 100 cycles which presents

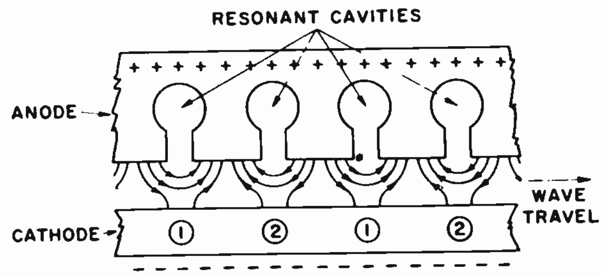


Figure 8-7. Electric Field in Linear Magnetron

no problems, but this same percentage shift at 10,000 mc is equal to 1 mc which is enough to interfere with satisfactory operation.

Three primary factors affect the operating frequency of an oscillator. There are, first, geometric factors in which the effective inductance and capacitance are changed directly through mechanical motion; second, pulling factors in which reactance is coupled into the oscillatory circuit from the load; and third, pushing factors in which reactance is introduced by changes in input conditions, such as voltage, current, or magnetic field.

There are two means of insuring stable operating frequencies. One is by the use of frequency stabilizers which tend to maintain a constant frequency of oscillations, and the other is by automatic frequency control systems which mechanically or electronically retune the oscillator when it shifts from a reference frequency.

FREQUENCY STABILIZERS. Magnetrons have high geometric stability except during warm-up or during periods of changing duty cycle as in certain types of modulation. This factor can usually be eliminated by maintaining the temperature of operation within $\pm 5°$ C during operation. Pushing factors are usually negligible up to 3,000 mc. Above this frequency highly stabilized power supplies may be used to counteract any frequency shift due to input voltage variations. Pulling effects are not normally found in microwave communication systems since none of the r-f system is required to move, thus changing the output load. However, to counteract any possible shift due to pulling, a stabilizer tuner

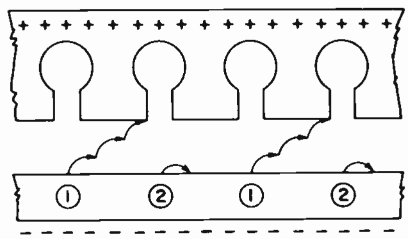


Figure 8-8. Electron Trajectories in Linear Magnetron

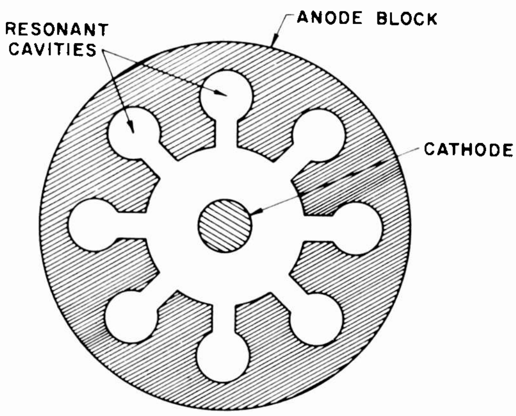


Figure 8-9. Cross-Section of Multicavity Magnetron

may be used on the output. This tuner consists of an external cavity resonator and a resistor in series, placed in parallel with the output line of the magnetron. The placement is critical and should be mounted at a standing wave node.

AUTOMATIC FREQUENCY CONTROL SYSTEMS. Sufficient frequency stability for communication may not always be obtained with the procedures given above. In such case, some means of automatic frequency control (AFC) must be used, either at the transmitter, the receiver, or both. Normally, in point-to-point operation, AFC is applied to the receiver so that it locks the receiver frequency to the transmitted signal.

AFC systems fall into two classifications: electronic and mechanical. These systems may also be classified according to whether they adjust the oscillator frequency to an absolute frequency or to a relative frequency. Terminal transmitters all fall under the classification of absolute frequency systems, while repeaters normally compare their operating frequency to the received signal.

An AFC system is a feed-back loop which samples the oscillator output, compares it with a reference, obtains an error signal, and uses this error signal to retune the oscillator. Electronic tuning may be used with the klystron and certain of the newer magnetrons, while mechanical tuning must be used on the traveling wave oscillator, most magnetrons, and coaxial triode oscillators, and may be used on klystrons.

Electronic tuning may be accomplished in the reflex klystron by varying the d-c voltage applied to either the accelerator grid or to the reflector. A tuning range of as much as ± 60 mc may be obtained in this way without seriously lowering output power. Special magnetrons, designed for frequency modulation, also may be tuned about $\pm 0.5\%$. In the spiral beam FM magnetron, electron guns with control grids are placed in several of the resonant cavities so that the electron beam is parallel to the magnetic field. By varying the intensity of these electron beams, the operating frequency may be changed slightly.

Mechanical tuning of any cavity oscillator may be accomplished by changing either the shape or size of the cavity, or by inserting a tuning probe in such a position as to appear inductive or capacitive. The mechanical changes may be operated either by an electric motor, or by a high temperature-coefficient-of-expansion rod heated electrically. An error voltage with which to retune the oscillator may be obtained in several ways. Most absolute methods depend on comparing the frequency of the oscillator to one or more resonant cavities. One method uses a pair of cavities and a pair of detector crystals in a circuit comparable to the discriminator used at lower frequencies. The output of this circuit will be a d-c voltage, the polarity depending on whether the frequency is high or low, and the magnitude depending on the magnitude of the error.

Another method uses a single cavity with a vibrating diaphragm moved by an electromagnet operating at the a-c input power frequency modulating the resonant frequency. The output of this cavity will be the oscillator frequency amplitude modulated by twice the power frequency if there is no error in the oscillator frequency. If there is an error in the oscillator frequency, the cavity output will be the oscillator frequency, amplitude modulated by the power frequency. The direction of the oscillator error frequency, high or low, will determine whether the modulation is in phase or 180° out of phase with the power or reference frequency. By demodulating the signal from the cavity and comparing it with the power source, the error voltage may be used to retune the oscillator.

The error voltage developed from the comparison circuits may be either a-c or d-c. If the error voltage is d-c, it may be used, directly or after amplification, to drive a d-c tuning motor or it may be applied to the voltage sensitive elements of the oscillator. If the error voltage is a-c it may be used, directly or after amplification, to drive an a-c 2-phase motor to tune the oscillator.

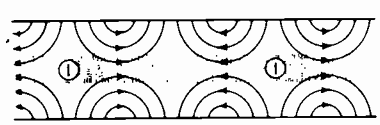


Figure 8-10. Electron Flow in Traveling Wave Tube

MODULATORS

Some means must be found of making the carrier frequency as generated by the oscillator convey information to the receiver. This is done by changing some characteristic of the carrier, such as the amplitude, frequency, or phase; or by turning the carrier on or off. The carrier may be modulated by a voice, tone, visual, or other desired signal.

MODULATION IN THE VHF RANGE. In the VHF ranges, where conventional vacuum tubes may be used, modulation is relatively simple.

1. Amplitude Modulation. Amplitude modulation is a change in the amplitude of the carrier which is directly proportional to the modulating signal. One hundred percent amplitude modulation exists when the amplitude of the radiated wave is (1) zero if the modulating signal is at its maximum negative value, and (2) twice the amplitude of the unmodulated carrier if the modulating signal is maximum positive. Figure 8-11 illustrates the waveforms of a carrier at different percentages of amplitude modulation.

In producing amplitude modulation, frequencies other than the carrier frequency are produced. These frequencies, called sidebands, are the sum and difference frequencies of the carrier and modulating frequencies. The sidebands for a 40 mc carrier amplitude modulated by a 5,000-cycle sine wave would be at 40.005 mc and at 39.995 mc thus requiring a 10-kc-wide pass band. In general the pass band required for AM signals is twice the highest modulating frequency. When complex wave forms such as voice signals are being transmitted, the harmonic frequencies making up the complex wave must also be considered if the received signal is to be undistorted. With 100% modulation by a sine wave, the sidebands each contain one-fourth the power of the carrier frequency.

There are two general methods of producing amplitude modulation. One is by changing the input to the plate of the tube and having the efficiency of operation remain constant, and the other is to allow the plate power input to remain constant, and change the efficiency of operation. Combinations of the two methods, in which plate power input and efficiency are both varied, are also used. Constant efficiency, variable input modulation, also known as plate modulation, is the most common and allows maximum output from a given transmitter tube. Such modulation requires that the modulator be able to furnish one-half as much power as is generated by the r-f system. The use of large and heavy components capable of handling the modulator power are therefore necessary. Constant input, variable efficiency systems may be accomplished by modulating the voltage applied to any of the other elements of the tube such as the cathode, control grid, screen grid, or suppressor grid. It is difficult to attain 100% modulation with good linearity in this manner. Compromise designs may be made in which both the plate input and one of the grids, usually the screen grid, are modulated. The overall efficiency, including the modulator, of the two AM systems is about the same.

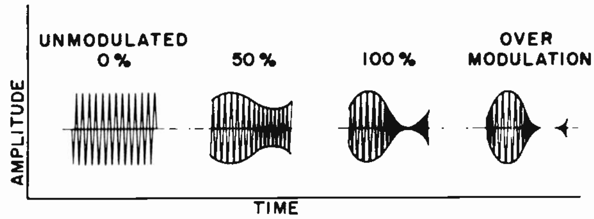


Figure 8-11. Amplitude Modulation Waveforms

Since the information is contained in the sidebands and not in the carrier and only one sideband is required, the wasted power in the other portions of the modulated carrier need not be transmitted. This system is known as single sideband (SS). The carrier is produced as usual and amplitude modulated at a low level in a balanced modulator which eliminates the carrier by phase canceling. One of the sidebands is removed by filtering and the remaining sideband is amplified and transmitted. Much higher effective power levels and decreased bandwidth requirements may be obtained in this way, but at the expense of more complicated transmitting and receiving equipment.

2. Frequency Modulation. Frequency modulation is a change in the output frequency of the carrier which is directly proportional to the amplitude of the modulating signal. In contrast to AM, one hundred percent modulation is not physically defined, but is fixed by regulation. Modulation percentage is the ratio of the frequency deviation to a fixed deviation which is set by FCC regulation, the amount of permissible deviation varying at different points in the frequency spectrum.

As in AM, frequency modulation of a carrier produces sidebands. In the case of FM however, a great many more are produced than with AM. The sidebands produced for a 40 mc carrier modulated by a 5,000-cycle sine wave would be not only at 40.005 and 39.995 mc, but theoretically, sidebands would also appear at 5,000-cycle intervals to infinitely high and low frequencies. Practically, the number of sidebands which will have enough power to be capable of causing interference will vary with the *modulation index*. The modulation index is the ratio of the frequency deviation of the carrier to the frequency causing that deviation. For example, a 75 kc deviation caused by a 15 kc signal will have a modulation index of 5. This will produce 8 effective sidebands on each side of the carrier frequency. The vectorial sum of the sideband and carrier power is the same as the power of the unmodulated carrier power regardless of the percentage of modulation. The transmitted power remains constant; only the distribution of the power among the sidebands and carrier varies with the modulation index.

Frequency modulation may be produced by two general methods. One method is direct, using a reactance tube to modulate the frequency of the oscillator. The reactance tube is a conventional vacuum tube connected so as to appear as a capacitive or inductive reactance that varies proportionally to the applied grid potential. With a reactance tube modulator, the ratio between the oscillator frequency and the amount of frequency deviation is small, and therefore a low-frequency oscillator followed by frequency multiplying stages must be used. These stages not only increase the carrier frequency but also increase the frequency deviation by the same amount.

One of the disadvantages of the direct method of producing FM is that crystal-controlled oscillators may not be used since reactance-tube modulators have negligible effect on their frequency. The second method of obtaining FM, indirect FM or phase modulation, produces FM following the oscillator which operates at a constant frequency. In contrast to direct FM, which changes the frequency of the oscillator itself, phase modulation shifts the phase of the constant frequency output of the oscillator. Since the phase of a voltage cannot be changed without also producing a frequency shift during the phase change, a form of frequency modulation is the final result.

The original method of producing phase modulation is still in use. This is the Armstrong system which uses a balanced modulator to produce two sidebands independent of the carrier as in single-sideband operation. By recombining these sidebands, phase shifted 90° with the original carrier, phase modulation of the carrier results. A second method of producing phase modulation consists of combining the carrier and the modulating voltage in a saturable inductor. This is a coil which reaches magnetic saturation at a very low current. The modulating voltage biases the coil so that the saturation point will vary in time (figure 8-12). The pulse output of the saturable inductor is amplified and shaped into a sine wave in following frequency amplifier stages.

In any of the methods of phase modulation, the frequency deviation produced will vary directly as the modulating frequency.

$$\Delta F = f\Delta\theta \tag{8-1}$$

where

ΔF = Frequency deviation

f = Modulating frequency

$\Delta\theta$ = Maximum angular shift of the carrier (radians)

A low frequency signal therefore will not produce as large a frequency deviation as a higher frequency of the same amplitude. This is compensated by the use of a predistorting filter which decreases the phase shift caused by the high frequency signals and allows equal amplitude high and low frequency signals to produce the same frequency deviation of the carrier.

Most of the received noise will be, in effect, high frequency signals, so pre-emphasis networks are used which effectively increase the transmission power of the high signal frequency, and de-emphasis networks are used in the receivers to restore the proper frequency response. Both predistorter and pre-emphasis networks are used in phase modulators. As in direct frequency modulation, only small amounts of frequency deviation can be obtained, and so low frequency oscillators are used followed by frequency multipliers to obtain a high modulation index. Relatively small amounts of modulating power are required for frequency modulation, and since a constant amount of power is handled by the tubes, they can be operated at maximum power levels without danger of overloading on modulation peaks.

MODULATION AT MICROWAVE FREQUENCIES. With microwave oscillators (see Table VIII-1), modulation becomes simpler in some ways and more difficult in others. The main problem in AM is to amplitude modulate the oscillators without producing frequency modulation at the same time. Frequency modulation is much simpler to obtain in microwave than in VHF transmitters.

1. Amplitude Modulation. Magnetrons and traveling wave oscillators may be successfully amplitude modulated by varying the power input by plate modulation. This method of modulation requires a fast acting electronic AFC system to counteract the FM also produced. Spiral beam frequency control magnetrons are well suited for AM. By applying the modulating voltage to both the anode and the spiral beam control grids, pure AM may be obtained. Plate

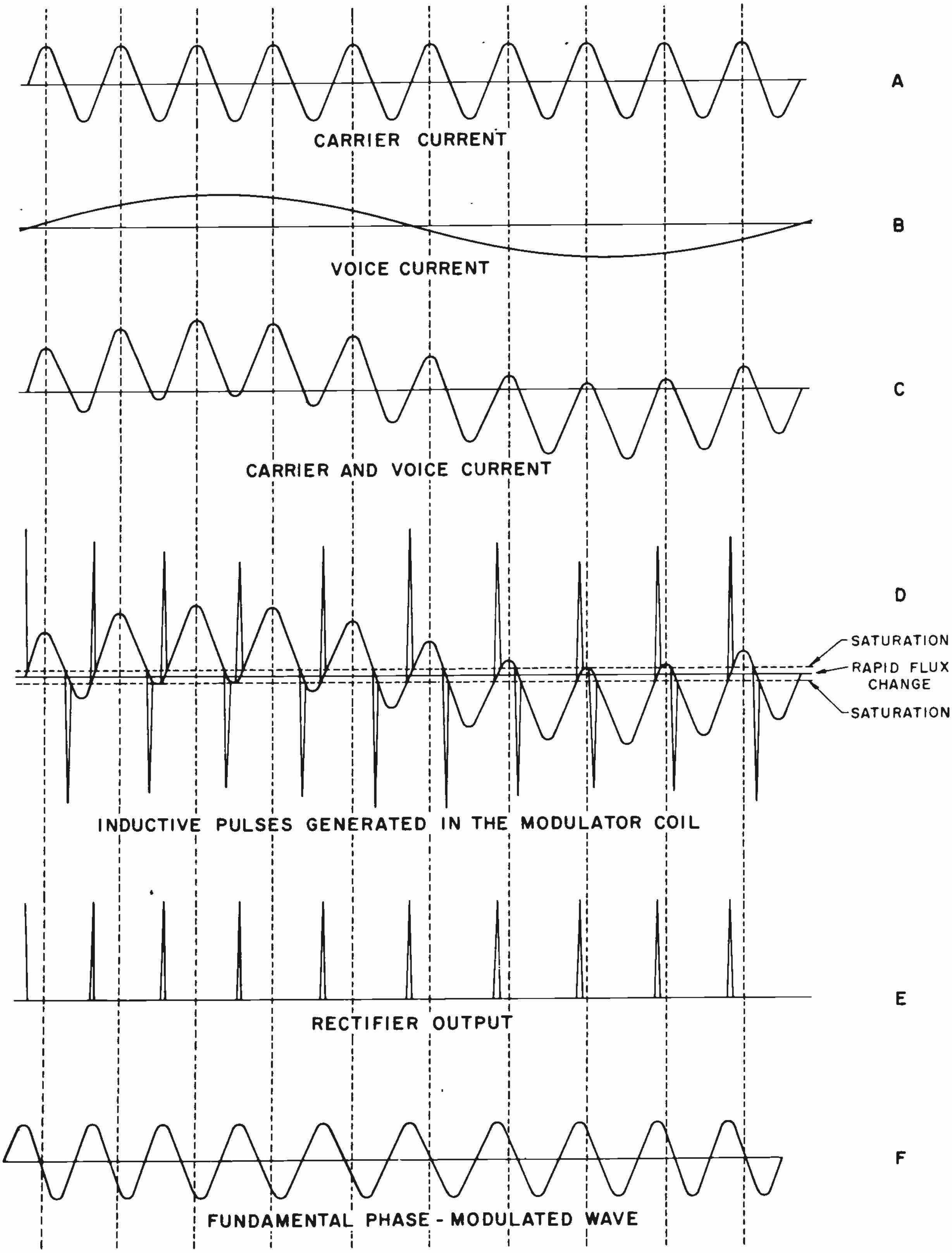


Figure 8-12. Waveforms Developed in Nonlinear Coil Modulator

TABLE 8-1: COMPARISON OF MICROWAVE TRANSMITTING TUBES

Type	*Freq. Range (mc)*	*Power Output*	*Cavity*	*Type Modulation*	*Efficiency*	*Type Output*	*Tuning*
Disk-seal triode	Up to 3300	500 W	External	Pulse, AM FM difficult	2%–20%	Coax	Mechanical
2-cavity klystron	900 to 21,000	5 W (avg.)	Part of tube	FM AM difficult	5%–30%	Coax or Waveguide	Mechanical
Reflex klystron	1,200 to 60,000	.005 to 1 W	Part of tube	FM AM difficult	Less than 10%	Coax or Waveguide	Mechanical and Electronic
Magnetron	700 to 60,000	60 KW (c.w.)	Part of tube	Pulse AM-FM difficult	50%	Waveguide	Mechanical
Traveling wave	500 to over 10,000	Low	External	Pulse AM FM difficult	10%–25%	Waveguide	Mechanical

modulation of klystrons, however, must be introduced at an amplifier stage following the oscillator, since the frequency variations produced cannot be controlled.

The method of varying the amount of r-f power allowed to reach the load may be applied to any of the microwave oscillators. In such a method an absorption cavity is used which can be controlled by an electron beam. The electron beam absorbs energy from the r-f wave passing through the cavity, thus controlling the amount of r-f power reaching the load. This method has a poor efficiency of operation, but modulating power requirements are low.

Another system for producing AM indirectly employs a subcarrier of 6 mc (or any other appropriate value) which is amplitude modulated in a balanced modulator circuit. The output consists of the upper and lower sidebands without the carrier. The lower sideband is filtered out and the remaining sideband is used to frequency modulate another carrier, known as the tree carrier, which is 6 mc below the operating frequency. A low modulation index is used so that the first order sidebands predominate. Since a subcarrier frequency of 6 mc is used, the first-order FM sidebands occur 6 mc above and below the tree carrier. The tree carrier and all sidebands except the first-order upper sidebands are rejected by filtering. A new carrier, at the operating frequency, is now injected in the proper phase to produce AM. The distortion produced by this system is very low.

2. Frequency Modulation. Frequency modulation of microwave oscillators is relatively easy to obtain. Klystron oscillators may be frequency modulated directly by placing the modulating voltage on the accelerator grid, or on the repeller plate in the case of the reflex klystron. Negligible power is required to modulate the oscillator and a large frequency deviation with very little AM can be obtained. Frequency modulation of magnetrons may be used in the special magnetrons using spiral beam frequency control. This method was discussed in the section on frequency stabilizing systems. As in FM klystrons, very little modulating power is required.

POWER AMPLIFIERS

The modulated signal may be passed to an amplifier to increase the amplitude of the outgoing signal. The same limitations of conventional circuits at microwave frequencies that applied to oscillators apply as well to amplifiers. In the microwave region no amplification is possible with conventional vacuum tubes and circuits; either the oscillator itself must supply enough power, or specially designed amplifiers must be used.

VHF AMPLIFIERS. R-F power amplifiers used at frequencies up to about 400 mc may be of the conventional type using lumped property components and standard vacuum tubes. As frequencies are increased, use is made of grounded grid circuits, in which the control grid is placed at r-f ground providing a shielding effect to reduce interelectrode capacitance. Grounded grid circuits have the disadvantage of requiring more grid driving power than grounded cathode circuits, although the additional power is not lost but appears in the output of the stage.

Since the efficiency of operation of oscillators and amplifiers decreases as the frequency increases, one method of obtaining greater overall system efficiency is through the use of frequency multipliers. The oscillator can now be operated at a much lower frequency than the output carrier frequency. Usually a buffer amplifier follows the oscillator so that no loading of the oscillator will take place. The amplifier stages, following the buffer amplifier, consist of amplifiers operated with a very high negative grid bias. This produces a large harmonic content in the output of the stage. By tuning the plate circuit of the tube to one of these harmonics, a frequency multiplying action takes place. Frequency multiplications of up to four times per amplifier stage are commonly used.

Another method of frequency multiplication which operates with much higher efficiency is the push-push amplifier. In this circuit two vacuum tubes are used with the grid circuits fed 180° out of phase and the plates of the tubes connected in parallel. Two pulses of plate current will appear in the plate load for each cycle of the input. By tuning the plate load for the second or fourth harmonic of the input frequency, high efficiency frequency multiplication takes place. The same action can be made to occur by connecting the grid circuits in parallel and the plate circuits in push-pull. When using FM, frequency multiplying also has the advantage of increasing the frequency deviation as well as the carrier frequency. However, there is the disadvantage of also multiplying the frequency instability of the oscillator.

In the transition range, from 250 mc to 1,500 mc, distributed property components come into wide use. Amplifiers using miniaturized vacuum tubes and tuned transmission lines are used in the lower portion of this frequency range, and disk seal triodes in coaxial circuits in the upper portion. The coaxial amplifiers are identical with the coaxial oscillators previously described except that the regenerative feedback has been eliminated.

MICROWAVE AMPLIFIERS. As frequencies are increased beyond the point where triode amplifiers will operate efficiently, klystron amplifiers and traveling-wave-tube amplifiers are used.

The klystron amplifier may be a 2-cavity klystron as used for an oscillator or a special amplifier klystron used for high power, which is known as a cascade or 3-cavity klystron. This tube is effectively two klystrons connected in cascade within the same envelope, with the catcher for the first section functioning as the buncher for the second section. The signal to be amplified is fed to the first cavity and the power output is taken from the third cavity. The second cavity is energized by the bunched electron beam and is not supplied with external r-f driving power. These tubes are capable of power gains up to 30 db, efficiencies of 30 to 40%, with a power output of 12.5 kw.

Traveling-wave-tube amplifiers are the second type of amplifier tube which may be used at microwave frequencies. The tubes have inherent regenerative feedback due to wave reflections in the tube. When designed for amplifier service, some means of attenuating the reflected wave must be provided. However, they are capable of large amplifications with a wide pass band and high efficiency.

INSTALLATION AND MAINTENANCE

Installation and maintenance procedures will vary with the equipment to be considered. Specific methods of installation are described in detail in the instruction manuals for each piece of equipment. The size and weight of the equipment, the permanency of the installation, the power input requirements, the type of operation, and the frequencies used are all factors which will affect the installation.

MAINTENANCE. Maintenance will, in general, follow the same lines regardless of the type of installation. There are two forms of maintenance which will be considered here—preventive maintenance and corrective maintenance. Preventive maintenance is any procedure which tends to reduce or prevent failure of the equipment. It should be a regularly scheduled check of operations. It can be broken down into six operations, abbreviated FITCAL, as follows:

F—feel: May be used to discover loose connections, plugs, and sockets; and overheating of components.

I—inspect: Visual inspection to discover leakage of sealing compound or liquid dielectric from components and corrosion of metal parts. Discoloration or blistered paint on metal surfaces is also an indication of overheating.

T—tighten: Tighten any loose connections, plugs, sockets, or fastenings to the proper tension.

C—clean: Cleanliness around high voltage or points at high r-f potential is very important. Dust and moisture decreases the voltage breakdown point of dielectrics and can easily cause failure of the equipment.

A—adjust: Adjustment of necessary mechanical parts and controls to insure proper operation.

L—lubricate: Lubrication of moving parts according to the proper lubrication order.

By performing preventive maintenance at regular intervals many failures of equipment can be prevented before they have a chance to make the equipment inoperative.

Corrective maintenance involves the test and repair of equipment after a breakdown has occurred. Spare units are usually available to replace the defective equipment if continuous operation is required. Test procedure remains the same for any equipment. The first step is to localize the defect to a unit, then to a stage, and finally to the defective component itself. Voltage and resistance charts and waveform diagrams, which will be found in the maintenance section of the instruction book for the equipment, are very helpful in this localizing procedure. After replacement of the defective component, performance tests should be run to insure proper operation of the equipment.

A military radio-relay system calls for compromises in design in order to achieve simplicity, minimum weight, and maximum performance. Operating frequencies vary from 45 to 13,000 mc. Power consumption is a few watts in the most compact equipment increasing to hundreds of watts in fixed equipment, while power output varies from a fraction of a watt to about 250 watts. A minimum of 30 watts is considered necessary for radio-relay links of 50 miles. Reception is possible up to 10 miles for portable equipment, with 50 mile ranges accepted as standard for transportable terminal equipment. Under ideal siting and propagation conditions, ranges of 100 miles have been reported.

CHAPTER 9

THE TERMINAL RECEIVER

INTRODUCTION

At frequencies ranging from 44–13,000 mc a variety of differences in receiving conditions are found as compared to lower frequencies. The fluctuation noise of tubes and circuits in the receiver becomes greater than external noises, such as atmospheric disturbances and man-made interference. Therefore, receiver noise is one of the chief limitations of receiver sensitivity at such frequencies. The necessity of handling greater receiver bandwidths is also of great importance, both for determining the maximum rate at which information can be received and as a controlling factor of the total noise encountered. These greater bandwidths will increase the effective noise power originating in the resistors and tubes of the receiver as well as the received external noise (see Chapter 13).

SUPERHETERODYNE RECEIVER PRINCIPLES

The main type of receiver used at these frequencies, as at lower frequencies, is the superheterodyne. Variations of the individual stages occur as the frequency is increased, but the general principles of operation remain the same.

A superheterodyne receiver (see figure 9-1) basically operates by heterodyning, or mixing, the received r-f signal with a locally generated r-f voltage obtained from the local oscillator. These two voltages are combined in a non-linear device such as a vacuum tube or crystal rectifier known as the mixer, producing, in addition to the original frequencies, the sum and difference frequencies. This process is identical with amplitude modulation as used in transmitters.

The difference frequency, or the lower sideband, is selected and amplified by a fixed tuned intermediate frequency (i-f) amplifier system. The i-f amplifier frequency remains the same in a given receiver regardless of the incoming signal frequency. This is accomplished by changing the local oscillator frequency so that the difference frequency between the desired signal and the local oscillator signal remains constant.

After amplification, the i-f signal is demodulated. The information and the carrier are separated and the signal containing the information is passed on to the video amplifier. Here the signal is amplified enough to be used as desired.

RECEIVER CHARACTERISTICS

SENSITIVITY. Receiver sensitivity may be defined as the minimum signal strength capable of producing a given value of signal output. So defined, the sensitivity is a measure of the overall gain of the receiver. However, random fluctuations, or noise, originating in the first stages of the receiver establish a limit to the useful sensitivity in VHF and higher frequency receivers. As a result, receiver sensitivity must also be defined in terms of the minimum signal which will produce a useable signal-to-noise ratio.

SIGNAL-TO-NOISE RATIO. The relationship between a received signal and the prevailing noise is expressed as a ratio, referred to as the signal-to-noise ratio (S/N). Generally, in the frequency range from 44 to 13,000 mc, internal noise originating in the initial stages of the receiver is the determining factor (see Chapter 13). However, at lower frequencies atmospheric and man-made noise can become troublesome.

BANDWIDTH. Closely associated with the sensitivity of a receiver is the bandwidth of its circuits. The bandwidth of a communication system is defined as the range of frequencies within which the signal level does not decrease more than 3 db below the maximum value of the signal. In services such as television and multi-channel radio-relay systems, receivers must possess a much greater bandwidth than that required for simple radiotelegraph or radiotelephone operation. For example, bandwidths of 6 mc are required in television receivers and up to 15 mc or more in certain radio-relay receivers. To achieve this wide-band response with high gain is a very difficult problem, since the product of the gain and the bandwidth of an amplifier is a constant. It is essentially a problem of initial design, and circuits as well as tubes and other components must be carefully selected.

SELECTIVITY. The selectivity of a receiver is its ability to differentiate between the desired signal and signals at other frequencies. Selectivity is usually defined as the ratio of the sensitivity for desired signals to the sensitivity for undesired signals, expressed in decibels. The sensitivity of the receiver for a range of frequencies about the desired frequency is often plotted as a selectivity curve whose shape is primarily determined by the response of the i-f amplifiers.

IMAGE RATIO. The image ratio at the receiver is a measure of the rejection of that frequency which, when mixed with the local oscillator signal, produces the same i-f as the desired signal. If the local oscillator is operated below the signal frequency, the image frequency will be the same amount below the local oscillator frequency that the signal frequency is above. The image rejection must take place prior to the mixer in the r-f amplifier or preselector stages.

R-F AMPLIFIERS

Receivers used at frequencies below 400 mc conventionally use amplifiers as the first stage or stages following the antenna. Advantages that may be gained from the use of an r-f amplifier include suppression of image response, improvement in signal-to-noise ratio, and isolation between the antenna and the local oscillator to prevent local oscillator radiation.

LUMPED-CONSTANT CIRCUITS. When conventional coils and capacitors are used at frequencies ranging from 44 to 400 mc, a point is reached where the capacitance has been reduced to that inherently present in the tube, socket, wiring and coil, and the inductance is as small as possible for the required L/C ratio. Upon replacement of the tube, therefore, there will generally be a different input capacitance, the difference being a large enough fraction of the total circuit capacitance to render inaccurate a calibrated tuning dial or other control arrangement. Reduction of the tuning capacitance to the point that it consists of only inherent minimums also has the objection that if automatic gain control is applied to the stage, the resulting variation of tube input capacitance in triodes may also cause excessive detuning.

Lumped-constant tuning, with a lumped inductance but only distributed capacitance, has been widely used in military equipment operating up to 300 mc, and the limitations can be largely eliminated by the use of pentode vacuum tubes. The tuning is done inductively by inserting into the coil a sleeve of silver-plated metal to reduce the inductance, or by employing a slug of compressed powdered iron to increase the inductance. For covering a wide fre-

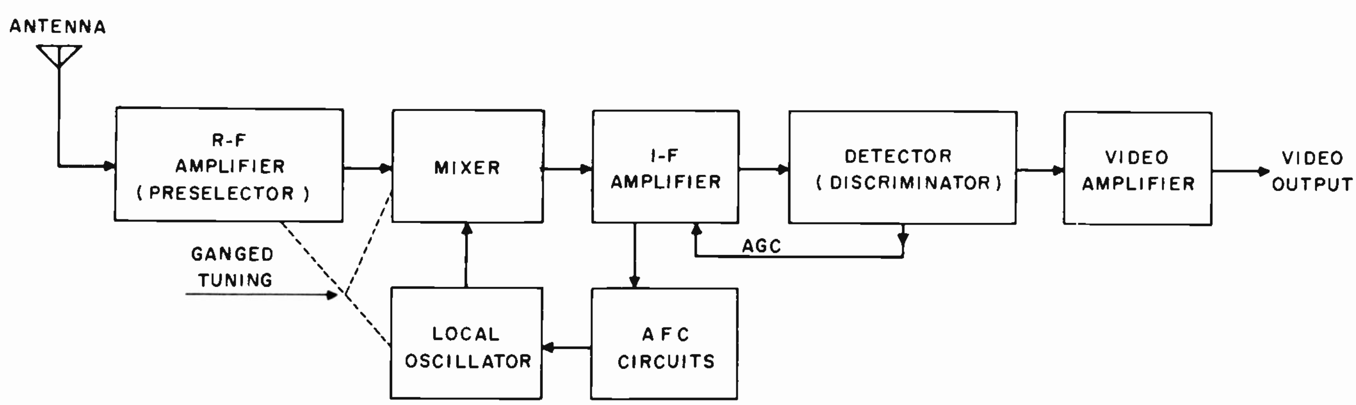


Figure 9-1. Block Diagram of Superheterodyne Receiver

quency range, a composite core with iron at one end and silver plating at the other may be used to advantage.

GROUNDED-GRID TRIODE. For signals in the frequency range of 100 mc to approximately 1,000 mc, advantages may be found in using a grounded-grid triode as an r-f amplifier, especially if the receiver is intended for wide-band operation. In a grounded-grid amplifier the high side of the input is connected to the cathode and the low side to the grid, which is grounded. The output is taken between plate and ground in the normal manner.

The gain of a grounded-grid amplifier tube is given by the formula

$$\text{Voltage Gain} = \frac{R_L(\mu + 1)}{R_p + R_L} \qquad (9\text{-}1)$$

where

R_L = Load resistance, ohms

μ = Amplification factor of the tube

R_p = Internal plate resistance, ohms.

If the load resistance is much less than the internal tube resistance, and if μ is much greater than unity, the gain simplifies to $R_L g_m$, where g_m is the tube transconductance. The value of the grounded-grid stage is essentially that the input and output circuits are fairly well shielded from each other, as in a screen-grid tube, while at the same time the low noise level of a triode is retained. Slightly less than critical coupling into a grounded-grid stage gives the best signal-to-noise ratio.

CASCODE AMPLIFIERS. The *cascode* (not to be confused with *cascade*) amplifier has been developed as a method of obtaining adequate gain with a low noise figure. The simplified circuit shown in figure 9-2 uses two triodes effectively in series. V1 functions as a grounded cathode amplifier which is directly coupled to V2, a grounded-grid amplifier. This arrangement combines the desirable features of both triodes and pentodes, while eliminating some of the undesirable features of each. It has the high gain, high impedance input, and stability of a pentode, with the low noise figure of a triode. Neutralization is not required but may be used to improve the noise figure further.

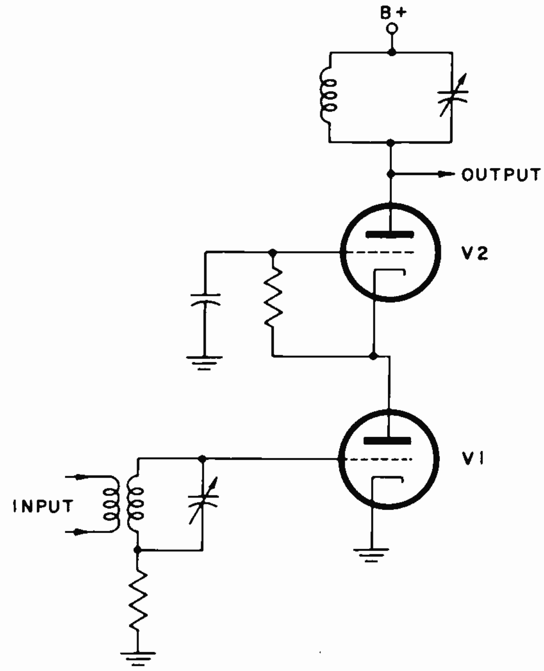


Figure 9-2. Cascode Amplifier

TUNING

UHF-VHF TUNERS. For tuning in the general region of 300 to 1,300 mc, especially over limited ranges, transmission line sections can be used. A $\lambda/4$ line shorted at the far end acts as a parallel tuned circuit at the near end, while other lengths of shorted lines act as high-Q inductors or capacitors. Balanced parallel-rod lines can be used up to 700 mc, and coaxial lines above this.

When signals above 500 mc are to be received, it is a distinct advantage to change from the conventional type of tube to the lighthouse version (see figure 8-1) and usually at the same time to employ a coaxial type of transmission line for tuning rather than the parallel-rod type. The lighthouse tube is constructed with flanges or belts around the edge so that large, low-inductance connections can be made to the coaxial type of transmission line. The short lengths of these coaxial transmission lines and the large circular contacts required for connection to the lighthouse tubes give equipment of this type an appearance which has led to the term "plumbing" as the usual informal name. In wide-band, r-f amplifier service, the limit of usefulness for lighthouse tubes is approximately 1,000 mc, although for local oscillators of the coaxial type the range can extend to over 3,000 mc. Two other types of UHF tubes are also available, the acorn tube and the door-knob tube, which are so named because of their shapes and sizes. Both types are available in diodes, triodes, and pentodes.

BUTTERFLY TUNERS. Tuners of the butterfly type, such as illustrated in figure 9-3, are continuously adjustable over a wide frequency range. Butterfly tuners of moderate size achieve tuning ratios as

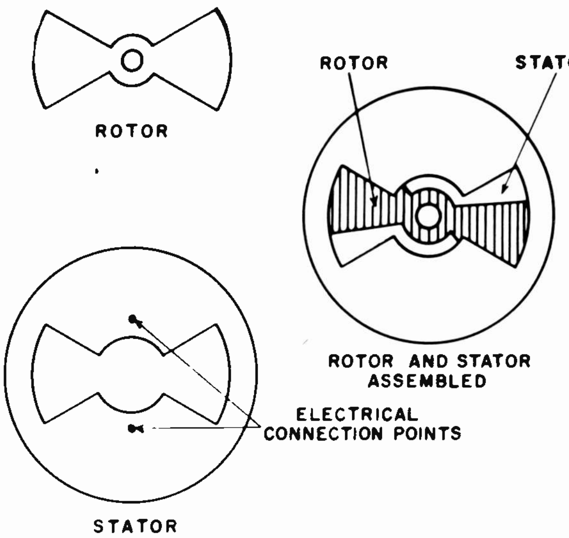


Figure 9-3. Butterfly Tuner

great as 5:1 in frequency, while larger sizes may be designed to achieve even greater ratios. These tuners have been used at frequencies as low as 40 mc and as high as 1,000 mc or more. The construction is somewhat similar in appearance to a variable capacitor, with the exceptions that (1) the circuit inductance is built in and consists chiefly of the circular-strap portions of the stator plate, (2) both connections are made to the stator, and, therefore, (3) there are no sliding contacts. When the rotor is turned towards increased meshing, the frequency is lowered by the increased inductance and capacitance. The rotor turns through an angle of 90 degrees for its full frequency range.

Butterfly tuners are suitable for use where wide frequency ranges must be covered. In design for such service, the size and number of plates are made as large as possible while still obtaining the necessary top frequency and freedom from spurious modes of operation. The necessary low frequency is then obtained by reducing the air-gap clearance as far as required.

MICROWAVE TUNERS. At some point in the frequency band about 1,500 mc, it is found that the use of an r-f stage in a receiver results in a lower signal-to-noise ratio than without the stage. Klystrons will produce amplification at these frequencies and are useful as transmitting amplifiers, but they are too noisy for use in receivers as r-f amplifiers. Tube noise is important only when small signal levels are being handled. If the signal strength is of the same order as the noise produced in a stage, a large amount of distortion will result. However, if the signal level is of much larger amplitude than the noise, little if any effect will be produced. For this reason, noise produced in amplifiers is of importance only in low power level stages such as the initial stages of a receiver. Traveling wave tubes are expected to become useful as r-f amplifiers in receivers as they are developed. Experimental tubes have been made which have as low a noise figure as the crystal mixers now used for the first stages in receivers above 500 mc. The receivers now used for frequencies above this point have as their first stages (1) the antenna, (2) a tuned circuit or preselector, and (3) a crystal mixer for converting the signal to a lower intermediate frequency. The preselector is composed of one or more tuned circuits which reject all frequencies except the desired frequency band. Butterfly tuners may be used for frequencies below 1,000 mc and coaxial cavities from about 500 mc up.

In general, a preselector is fed from a low-impedance source, the antenna, and is coupled to the higher impedance of a crystal mixer. The tuner must therefore act also as an impedance transformer to obtain maximum energy transfer. It must have a pass band that is at least as wide as the pass band of the i-f amplifier and an off-frequency attenuation which produces the largest amount of image and harmonic rejection that can be obtained. If the desired attenuation characteristic cannot be obtained by a single-tuned circuit and still maintain the required pass band, double- or triple-tuned circuits must be used.

LOCAL OSCILLATORS

The local oscillator used in the receiver must operate at a frequency similar to the received frequency. The exact frequency of the oscillator is the desired signal frequency plus or minus the intermediate frequency. The local oscillator may operate either at a higher or a lower frequency than the received signal, since the difference frequency will remain the same. The oscillators used up to 500 mc may be conventional vacuum tube oscillators, while coaxial oscillators are used up to about 3,300 mc. Klystron oscillators, usually of the reflex variety, are used exclusively above this frequency.

MIXERS

The device used to produce the i-f from the received signal and the local oscillator signal is known as the mixer or first detector. Any nonlinear device may be used for this purpose, although at frequencies above 100 mc triode and diode vacuum tubes and crystal rectifiers are usually used. At frequencies up to about 500 mc, triodes such as the type-2C40 lighthouse tube may sometimes be used in spite of the large amount of noise generated, since the relatively high conversion gain makes it possible to obtain a higher S/N ratio than if a diode or crystal rectifier were used. Conversion gain is defined as the ratio of the mixer i-f output to the r-f signal input and is usually expressed in db. Above 500 mc crystal mixers are appreciably superior to any triode mixers now available because of the small amount of noise generated. The conversion loss of type-1N26 crystals is low, even at 23,000 mc, with a maximum conversion loss varying from 5.5 to 7.5 db in the frequency range from 3,000 to 16,000 mc. The maximum noise ratio, or the amount of noise generated by the mixer compared to the noise produced in an equivalent impedance resistor, varies from 1.5 to 2.5 in the same frequency range. Silicon crystal mixers have the disadvantage that they damage easily if overloaded.

I-F AMPLIFIERS

The difference frequency produced in the mixer may actually be any desired value. Selection of the frequency to be used in a given receiver will depend on several factors. A high i-f is desirable to eliminate image response, but the selectivity and noise figure of an amplifier becomes worse with an increase in frequency, and ganged tuning of the local oscillator and preselector becomes increasingly more difficult. The choice of the intermediate frequency is therefore a compromise between the desired image-rejection ratio, on one hand, and the desired sensitivity, selectivity, and circuit simplicity on the other.

A method of obtaining both good image-rejection by using a high i-f, and good selectivity and simplicity of circuits through the use of a low i-f, is the double-conversion superheterodyne receiver. A high i-f is first produced to obtain a good image-rejection characteristic. This frequency is then further reduced by a second frequency conversion, bringing the i-f down to a frequency that can be easily amplified and that provides good selectivity. The oscillator used to produce the second i-f may be crystal controlled since it heterodynes with a fixed frequency. The disadvantage of this system is the additional local oscillator and mixer stages required.

I-f amplifiers are required to have a wide band width and high gain. These factors are mutually antagonistic as shown in the equation below for a single-tuned amplifier stage.

$$G \times B = \frac{g_m}{2 \pi C} \tag{9-2}$$

where

G = The voltage gain per stage at the center frequency

B = Band width of the amplifier stage

g_m = Transconductance of the tube

C = Total stage capacitance.

$G \times B$ is known as the gain-bandwidth product and is a constant for a given tube and associated components. Using a type-6AK5 pentode, the gain-bandwidth product is about 65.2 mc. With a bandwidth of 10 mc, the available gain is only 6.52. When a number of identical single-tuned amplifier stages are connected in cascade, the overall bandwidth becomes

$$\text{Overall B} = (\text{one-stage bandwidth}) \times \sqrt{2^{1/n}-1} \tag{9-3}$$

This limits the overall bandwidth available. For instance, with one of the best tubes available for use in i-f amplifiers, the 6AK5, and obtaining a gain of 100 db, a commonly needed gain, the maximum bandwidth theoretically available is 6.9 mc which in many cases is inadequate. In addition, 23 stages are required to accomplish even that result.

It is easily seen that some means must be found to obtain increased bandwidths with high gain. This is accomplished by stagger tuning. In this technique, the stages are tuned to several different frequencies, which makes wide-band operation with high gain practical.

DETECTORS

Some means of separating the carrier from the desired information must follow the i-f amplifier. The second detector or demodulator is used for this purpose.

If AM is being received, a simple rectifier and filter circuit is used. Another way of looking at amplitude demodulation is that the carrier and its sidebands are heterodyned together and the difference frequency recovered as the information. When receiving single-sideband signals, the carrier must be injected at this point to heterodyne with the received sideband. This is usually done with a crystal controlled oscillator operating at the intermediate frequency.

The detectors used for FM are often called discriminators and utilize completely different principles. A discriminator is defined as a device which produces a d-c voltage proportional to the frequency of an input signal. This may be accomplished in a number of ways. One method uses two tuned circuits, one tuned above the center frequency and one tuned below, to obtain two i-f voltages whose amplitudes depend directly on frequency. These voltages are then rectified and combined so that zero voltage output is obtained at the center frequency. A difference voltage of the two i-f voltages which is proportional to the frequency is obtained when the frequency of the i-f signal is above or below the center frequency. A discriminator of this type is sensitive to amplitude variations in the input and so a clipper-limiter stage must be used preceding the discriminator to remove any AM. Another method of obtaining frequency discrimination is the use of the phase detector. This operates by comparing the phase relationships of two signals, one the i-f signal and the other the i-f signal phase shifted in a resonant circuit tuned to the center frequency. The amount of phase shift will be proportional to the frequency of one signal. A limiter stage is also required for the phase detector.

To eliminate the need for a limiter stage, several types of discriminators have been developed which will not respond to amplitude changes. The first of these is the ratio detector. Instead of the two rectified voltages of the ordinary discriminator being combined with opposite polarity, they are combined so as to add. The sum of the voltages is kept constant by a large value capacitor. This eliminates any amplitude variations of the i-f signal and makes the use of a limiter stage unnecessary. The ratio of the two voltages changes as the frequency input to the ratio detector changes and the output is taken from one of the rectifier loads.

Another type of discriminator, which combines a highly efficient limiter and special type of discriminator in one envelope is the gated-beam detector using a 6BN6 vacuum tube. This tube has an extremely sharp cut-off and saturation characteristic and will operate with as little as 1 volt input to the limiter grid. The discriminator action takes place due to the gating characteristics of the tube. Two control grids are present and each must have a positive signal applied for plate current to flow. The i-f signal is applied to each of these grids, to the first, or limiter grid, directly, and to the second, or quadrature grid, through an L-C phase shifting network. The phase of the quadrature grid voltage will depend on the frequency of the i-f voltage. By comparing these voltages in the gating action of the tube the average output current will be inversely proportional to the frequency of the input. Several other discriminator circuits have been developed, but those described above are in widest use.

AFC SYSTEMS

Receiver stability is even more important than transmitter stability. In contrast to the AFC systems used in transmitters, which tend to keep the operating frequency at an absolute value, receiver AFC systems usually compare the receiver's operating frequency to the received signal. This is done by using a discriminator in the i-f channel to develop the error signal used to retune the oscillator. An FM receiver may use the same discriminator for signal detection and AFC provided the signal variation is averaged over a period of time so that the oscillator does not attempt to follow the modulation, but only slow frequency drifts. With the exception of the development of the error signal, AFC is obtained by the same methods as in transmitters.

AGC SYSTEMS

Since received signal strength will vary over a wide range some means of automatic gain control (AGC) is required in the receivers. This is accomplished by rectifying the received carrier and averaging the voltage over several cycles of modulation. This voltage is then used as grid bias for the i-f amplifiers so that as the average signal strength increases, the gain of the i-f amplifier is decreased. The bias is usually not applied to the first amplifier stage since the stage tends to become noisier with AGC applied. With some types of modulation, such as for television, the rectified i-f signal is not proportional to the carrier level, or the ratio of peak-to-average signal may be too large to make practical the use of the average value to control the receiver gain. Special techniques must be used to provide AGC in these cases.

VIDEO AMPLIFIERS

The detected modulation must be amplified for further use. If the modulating signal takes the form of pulses, an increase in S/N ratio can also be obtained at this point by pulse shaping circuits. Video frequencies are those frequencies from near zero to frequencies of several megacycles. The video amplifier must amplify this complete frequency range, which in some equipment is up to 20 mc, with uniform response. For pulse transmission, the bandwidth in mc must be at least twice the reciprocal of the pulse duration in microseconds if the pulse is to approximate its original shape. Phase shift characteristics are as important as band widths, since phase distortion will change the pulse shape by phase shifting the harmonics which make up the pulse difference angles.

R-C coupled pentode amplifiers are most satisfactory for these video amplifiers. The high frequency response is limited by the output capacitance, the input capacitance of the next stage, and the distributed capacitance of the wiring. These three capacitances appear in parallel and shunt the load resistor, decreasing its impedance and therefore the stage gain, at high frequencies. The low frequency response is limited primarily by the time constant of the grid circuit which must be long compared with the lowest frequency to be amplified.

High frequency response is increased by

1. decreasing the plate load resistors
2. using tubes with low grid-to-plate capacitance
3. carefully laying out components and wiring
4. decreasing the grid resistors
5. using multiple low gain stages
6. using negative feedback
7. using compensating networks.

Shunt compensation uses an inductor in series with the plate load resistor of the stage. As the frequency increases, the impedance of the inductor increases, compensating for the decrease in impedance of the capacitance in shunt with the load. Shunt compensation will increase the flat response of a stage by a factor of about 10. Series compensation uses an inductor in series with the coupling capacitor to the next stage. The inductor is selected so that it is series resonant with the input capacitance of the next stage at a frequency higher than the desired high frequency cut-off of the amplifier. As the frequency is increased the voltage drop across the input capacitance increases, thus compensating for the high frequency loss in the preceding stage. Series compensation can increase the flat response of the amplifier by a factor of about 15. Combination compensation using both shunt and series compensation may be used to increase the flat response of the amplifier by a factor of about 18. Low frequency compensating networks are composed of a resistor and capacitor in parallel, connected in series with the load resistor. The high and low frequency compensating networks operate independently and do not interfere with each other. Low frequency compensation is important primarily to minimize phase distortion.

PULSE RESTORATION. The first stages of the video amplifier may be used when receiving pulse transmissions to reshape the pulses and remove the noise characteristics. Figure 9-4 illustrates the wave shapes of the pulse as received, and after clipping and limiting. Most of the noise distortion is removed in

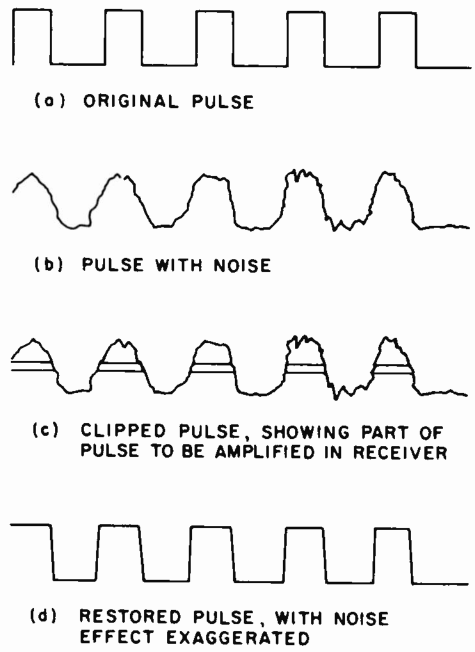


Figure 9-4. Pulse Restoration

this process. By increasing the band width of the system and using shorter pulses, greater reduction of noise can be made.

DIRECT DETECTION RECEIVERS

The simplest type of receiver is one in which a crystal detector is connected to the antenna and the video output is amplified to a suitable level for use. No local oscillator or mixer is used. Selectivity is provided by a tunable or fixed-frequency r-f filter inserted between the antenna and the detector.

An application of this type of receiver is in small portable receiver units where sensitivity and selectivity may be sacrificed to obtain simplicity of design. Receivers of this type are capable of detecting signals with powers as low as 4×10^{-9} watts using a bandwidth of 10 mc at a carrier frequency of 3,000 mc. At 10,000 mc and a bandwidth of 10 mc, a signal of 8×10^{-9} watts may be detected. These sensitivities are lower than for comparable superheterodyne receivers which may have a sensitivity of the order of 10^{-12} watts at the noise threshold.

CHAPTER 10

REPEATERS

INTRODUCTION

Since radio transmission in the VHF and UHF bands are essentially short range, some form of relay must be provided if operation over a great distance is required. The relays used in radio-relay communication are called repeaters. They are essentially just what the name implies—they receive a signal from the terminal transmitter or another repeater, amplify it, and then transmit it onward to another repeater or terminal receiver.

TYPES OF REPEATERS

Two general groups of repeaters are used—active and passive. Active repeaters actually amplify the signal and re-radiate it, while passive repeaters are nothing more than reflectors whose only gain is that due to the change in direction of the radiation. Contained within the group of active repeaters are a number of types that vary according to their use and method of operation.

Obviously two terminal radio equipments back-to-back, the receiver output from one feeding the transmitter input to the other, could be used as a repeater station. This is done at what is called a terminal repeater, but then only for reasons of rearranging the traffic channels within the baseband. The baseband is composed of the multiplexed signals used to modulate the carrier. On a radio-relay system of several hops, it is preferable for reasons of performance and cost to employ a heterodyne type repeater. In this repeater the incoming carrier is heterodyned down to a frequency which may be easily amplified, and after amplification heterodyned back up to the transmitting frequency.

It is often necessary to drop or insert traffic at a repeater station to feed a branch, or a local user, without resorting to a terminal repeater. This is done in a drop repeater which has special provisions for extracting or "dropping" certain traffic from either or both directions and inserting complementary traffic for duplex (two-way simultaneous) transmission. Figure 10-1 is a simplified block diagram of the various types of active repeaters.

THROUGH REPEATERS. Straight-through repeaters, in which the received carrier is amplified directly making them the simplest type of repeater that could be designed, are not used for several reasons. One reason is that noise-free amplifiers are not available for frequencies above about 1,000 mc at present. In addition, it is desirable to change the frequency of transmission to prevent regenerative feedback due to coupling between the antennas.

A heterodyne through repeater is the simplest type of active repeater in use. Figure 10-2 is a block diagram of one type of heterodyne repeater. The signal is received and heterodyned to the intermediate frequency, which can be amplified efficiently, and a second mixer is then used to increase the frequency to the desired output carrier frequency. This is then amplified and transmitted with no demodulation and subsequent remodulation required.

By the addition of another mixer stage the frequency of operation of the repeater is dependent only on the received frequency. The same oscillator is used to heterodyne the frequency first down and then up, thus compensating for any frequency instability. A second crystal-controlled oscillator, equal in frequency to the frequency difference desired between

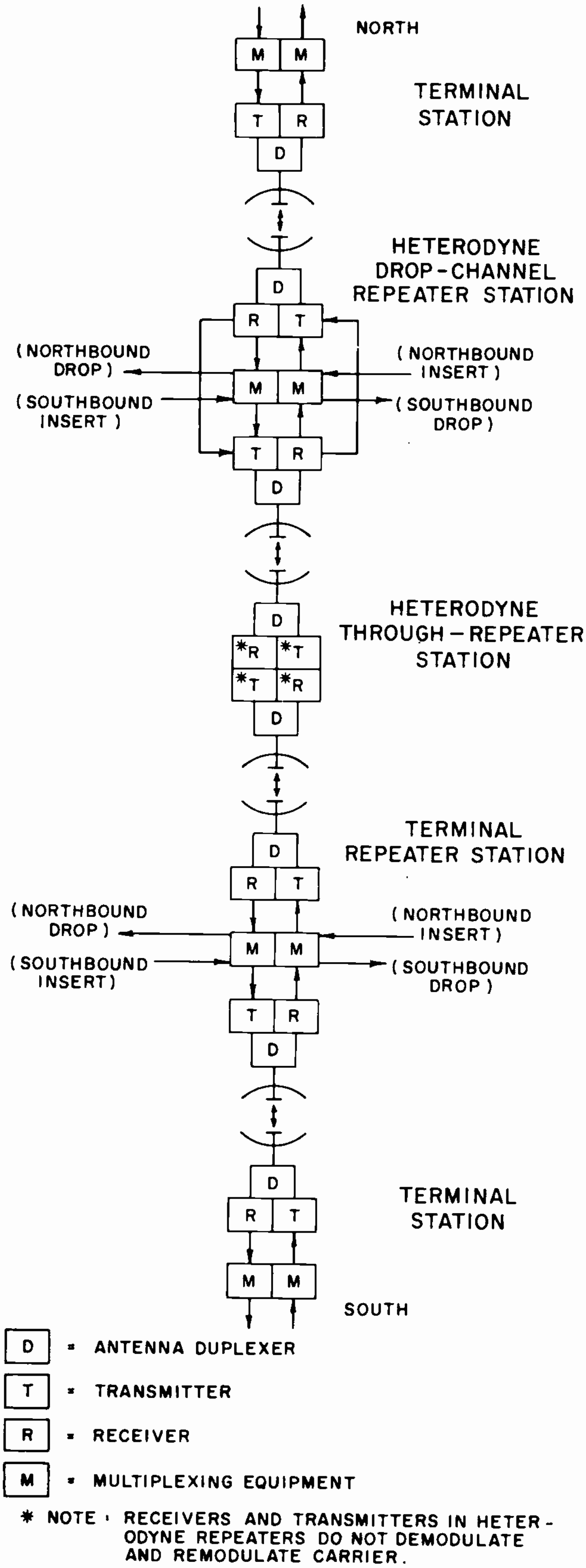


Figure 10-1. Types of Repeaters

the receiver and the transmitter, is used to change the operating frequency, thus requring the additional mixer stage.

A third type of through repeater may be used in which the same oscillator is used as the local oscillator and the transmitting oscillator. The signal is received, demodulated, and the video signal then used to modulate the oscillator. The transmitted and received signals therefore have a difference frequency equal to the intermediate frequency. If this system is used for frequency modulated signals, the i-f bandwidth must be twice the bandwidth of the received signal. This is necessary because the oscillator must be arranged in such a way as to shift frequency when modulated in the direction opposite to the frequency shift of the received signal. If this is not done the oscillator will try to follow the received signal and the modulation index will be reduced by a factor of two through the repeater.

DROP REPEATERS. To drop channels, all that is required is to place a parallel branch on the receiver i-f amplifier and pass this signal to a demodulator and baseband amplifier to recover the entire signal spectrum originally passing through the main system. Those channels intended for local use are then selected by appropriate multiplexing equipment. All of the channels dropped at such a repeater also continue through the system to the next terminal station. No means is available for stopping the channels at a drop repeater, but it has the advantage that simpler multiplexing equipment may be used than is required for a terminal repeater. Insertion of channels is accomplished by modulating one of the heterodyning oscillators as shown in figure 10-2.

One-way channel insertion or dropping, that is, inserting or dropping a channel toward or from one terminal only, is most common although in some cases the channel may be inserted in both directions.

TERMINAL REPEATERS. A terminal repeater station is one in which all signals are completely demodulated and repeated by remodulation. In its simplest form it consists of two complete terminal equipments, the receiver output from one feeding the transmitter input of the other through multiplexing equipment.

It is clear that in a system having several drop channel repeaters, where traffic is inserted and dropped, a substantial portion of the baseband would become loaded with traffic traversing only a part of the main truck. It may become necessary to reuse some of the baseband on portions of the system, so a terminal repeater is used to recover these channels or to redistribute the channelling layout. At such repeaters the baseband is demodulated, those channels intended for local use or onward transmission are selected, and the others eliminated. Beyond this point, the channel space used up to this point by the eliminated channels can be used again. The original channelling structure may be maintained or complete rechannelling may take place.

SERVICE UNITS

Service units are used at both repeaters and terminals to provide a common party line for the use of operating and maintenance personnel. Usually the first channel of the system is kept open for this purpose in frequency division multiplexing systems, since multiplexing equipment is not required for access to this channel (see Chapter 11). Ringing circuits are also provided to permit calling any desired point in the system.

In many cases repeaters are unattended, requiring that some means of signaling equipment failures be provided. The same equipment used to signal failure of the equipment can also be used to automatically switch in standby equipment keeping service interruption due to equipment failure to a minimum. Delay circuits are used to prevent fault indication from more than one repeater at a time. A transmitter failure could show up as receiver failure at the next terminal if these delay circuits were not used. Fault indicators are placed at attended stations at distances which depend on the repeater spacings and the accessibility of the stations.

PASSIVE REPEATERS

The passive repeater is used in microwave radio relay work to direct the beam round some obstacle that cannot be removed. It occasionally happens that no choice exists for the location of a repeater station, except one that contains an obstruction in the beam path. When this happens, one solution is a passive reflector. Since even the best of the passive reflectors attenuate the signal considerably, they should not be used except as a last resort.

Two types of passive reflectors are used, depending upon the terrain. These are plane reflectors and passive repeaters. If the repeater stations are so located that the signal can be bounced off a plane reflector to the receiving station, then the plane type will do. However, where a high ridge intercepts the beam, or a situation exists where a plane reflector cannot be made to reflect the signal toward the receiver, the passive repeater is used. This type consists of two parabolic antennas, one facing each repeater station, and connected together by means of a waveguide or a coaxial cable. The operation here is somewhat different from that of the plane type of

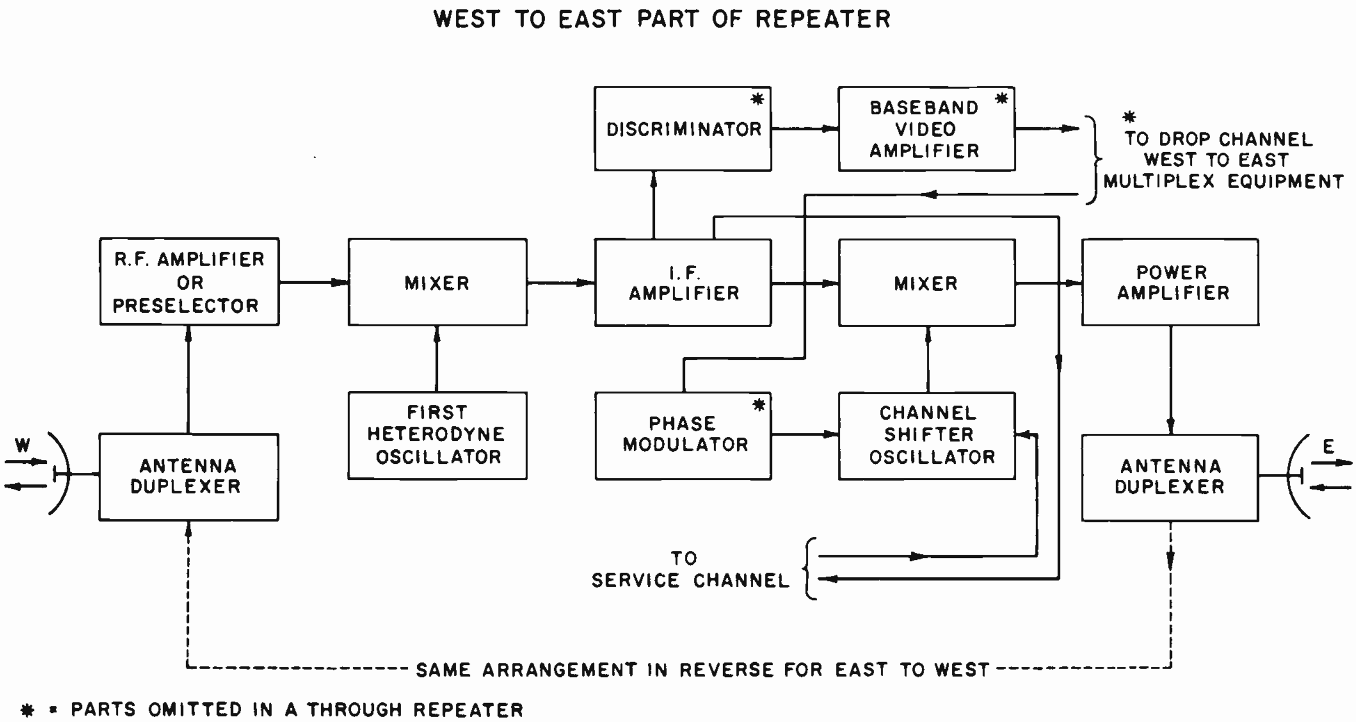


Figure 10-2. Heterodyne Repeater

reflector. The signal picked up by the receiving antenna is transmitted through the waveguide or coaxial cable to the transmitting antenna of the passive repeater, and thus sent on to the receiving repeater. The passive repeater is merely a repeater station with no amplification power. It has the drawback of causing considerable attenuation, since only the amount of energy falling on the passive receiver parabola is used, and there is additional attenuation due to the interconnecting waveguide and passive transmitting antenna. Another disadvantage of this type is loss due to refraction. Since the passive repeater transmits less than the power received, attenuation caused by the beam bending away from the receiving antenna can cause considerable fading. For these reasons, the passive repeater should be used only on extremely short hops, where attenuation in the repeater and beam bending cannot have any serious effect. It is most often installed in locations where terrain makes maintenance and inspection of a repeater station difficult.

CHAPTER 11

MULTIPLEXING

INTRODUCTION

Each frequency range of the radio spectrum has certain advantages that make it desirable for some particular type of communication. Higher frequencies are desirable for relaying information that may require a wide bandwidth. At higher frequencies, highly directive antennas may be employed, allowing low power transmitters to direct a signal to the receiver, and thus obtain a good signal-to-noise ratio. In using these frequencies for relay work, the ability to focus the radio beam gives an efficient method of transmitting information over long distances. The relay system is often more economical than a wire transmission line, and more reliable, since the repeater stations are less affected by storms than wire lines.

The wide bandwidth available at higher frequencies may be used where the message has a wide range of frequencies, such as television, or where it is desirable to send a number of low frequency messages simultaneously, as in telephone communication. Where several messages such as voice, facsimile, or teletype are to be combined, two general methods are available, and the combining of the messages is called multiplexing. The methods used for multiplexing are frequency division, and time division, and may modulate the carrier or subcarrier with either amplitude modulation or frequency modulation.

FREQUENCY DIVISION

Frequency division consists in separating the various messages to be combined so that each message is limited to a certain frequency range, with no overlapping between these ranges. This is accomplished by first limiting the audio bandwidth to definite values. This bandwidth may vary, according to the desired quality of speech reproduction. By filtering out all frequencies below 200 cycles per second, and above 3,000 cycles per second, the audio bandwidth becomes 200 to 3,000 cycles per second. This frequency range is entirely adequate for intelligibility, although it is definitely substandard reproduction. An audio range of 100 to 3,500 cycles per second sounds far more natural to the listener, and is the standard set for non-military telephone service.

Groups of telephone circuits must be combined in such a way that they may be separated at the receiving terminal and routed to their respective destinations with a minimum of crosstalk. This is done by modulating a separate subcarrier for each voice channel, with the subcarrier spacing a little more than twice the highest audio frequency to be transmitted. Such a method gives entirely satisfactory communication, but requires a rather wide band. This band may be halved by using single-sideband modulation, which requires a carrier spacing of a little more than the highest audio frequency, instead of twice the audio frequency. The most common method of frequency-division multiplexing is single-sideband, suppressed carrier, and it will be the type considered here.

Frequency separation of voice for transmission is accomplished by modulation with an audio frequency in such a way that the various messages are separated uniformly over a range of frequencies. This range of frequencies, which includes the spacing between messages, becomes the modulation bandwidth. Obviously, the audio signals can not modulate the carrier directly, for the messages would be hopelessly scrambled. They can be separated quite easily, however, by heterodyning each message with a different audio frequency, and filtering out the lower sideband. For example, four messages are to be multiplexed by single-sideband multiplexing, each message having an audio range of 100 to 3,500 cycles per second. The first message would not need to be heterodyned with an audio signal, but could remain unchanged. The second message to be multiplexed has a frequency of 4,000 cycles heterodyned with it, so that its range of frequency now becomes 4,100 to 7,500 cycles per second. Combining these two messages gives a bandwidth of 100 to 7,500 cycles per second. The messages are easily separated at the receiver by using a bandpass filter that will pass the frequency range of one message, and block all others. In the same way, a third message would be heterodyned with 8,000 cycles per second, to produce a range of 8,100 to 11,500 cycles, and a fourth message would have a frequency range of 12,100 to 15,500 cycles per second. The modulation bandwidth now becomes 100 to 15,500 cycles per second.

There is no reason for limiting the group of messages to four. Theoretically, as many messages as the allotted bandwidth at the carrier frequency will carry may be used, with the bandwidth in the United States being regulated by the FCC. Where many messages are to be transmitted by the same carrier, the messages may be multiplexed as described for four messages, or they may be broken into groups of any convenient number with each group modulating a separate subcarrier. The subcarriers, in turn, modulate the carrier for transmission. Needless to say, the subcarriers must be spaced so that no overlapping of frequency will occur at any point in the ratio spectrum. The combined messages may amplitude modulate or frequency modulate the carrier (or subcarrier), as desired.

PULSE TIME MODULATION (PTM)

INTRODUCTION. As stated earlier in this chapter, messages may be combined for transmission by the same carrier by separating the messages either in frequency, or time. In order to separate the messages in time, each message is transmitted for a brief period of time, in regular sequence. This process is called sampling. The receiver separates the samples, and reconstructs the original messages from the samples.

PULSE AMPLITUDE MODULATION (PAM). One system of time-division multiplexing commonly used, and the basis of nearly all time-division systems, is pulse-amplitude modulation, (PAM). In this system, the signal is periodically sampled, and a pulse obtained whose height, or amplitude, is proportional to the amplitude of the signal at the instant of sampling. As an illustration of PAM, consider a curve plotted on a sheet of graph paper (figure 11-b). By determining the ordinates (amplitude) of the curve at uniformly spaced points along the abscissa (time), the curve may be readily plotted. The pulses are uniformly spaced in time, with their amplitudes determined by the amplitude of the signal. There is no necessary limit to the maximum number of points in plotting a curve, but there is a minimum number if the curve is to be properly reconstructed. A similar limitation exists in regard to the number of samples to be taken in properly reconstructing the audio signal. Experiments on pulse modulation have shown that a sine wave must be sampled at least 2½ times each cycle if it is to be reproduced by the receiver. The sampling rate is therefore determined by the highest audio frequency tone to be transmitted, which is dependent upon the quality of reproduction desired. If the highest frequency to be transmitted is 3,200 cycles per second, sampling must be taken at a rate of 8,000 per second.

In the above paragraph, only one of the multiplexed messages has been considered. Where four

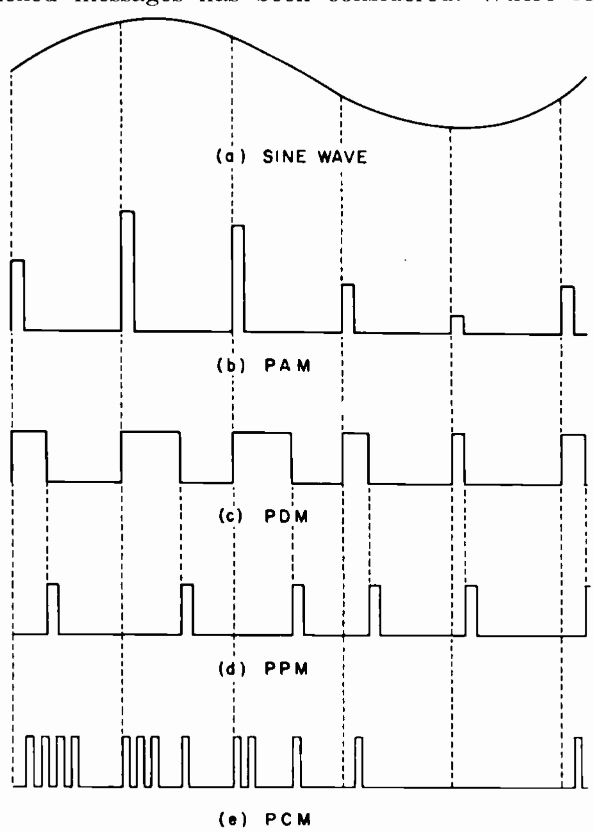


Figure 11-b. Types of PTM

messages are to be transmitted, each of the four must be sampled 8,000 times, with a total pulse repetition rate of 32,000 per second. The sampling is done in rotation, so that every fourth pulse will be taken from the same message. One other consideration must be taken into account, namely, identifying the message to be sent to any particular circuit at the receiver. Not only must the message be properly reconstructed by the receiver, but it must also be routed correctly. In order to identify the messages, that is, identify which is channel one, two, three and four, some form of coding must be included in order to send each channel to its proper destination at all times. This may be accomplished by using a special pulse for channel #1. If such a pulse were twice the width of the other three pulses in the group, the width could be used to establish the start of each group of pulses. Another method used is to add an extra pulse to each group, with the extra pulse as a marking pulse, to identify the start of each pulse group.

PAM, like frequency-division modulation, depends upon the amplitude of the signal as the major characteristic to be transmitted. While these systems are the simplest forms of multiplexing, they are most likely to be affected by noise. Atmospheric noise picked up by the repeaters, as well as internal noise in the repeaters themselves, will be transmitted along with the signal, with each repeater adding its own increment of noise. This may be partially overcome by using a S/N ratio for each repeater that is considerably higher than the S/N ratio for the system as a whole, a remedy that requires more powerful transmitters or closer spacing of the repeaters, or both. There are other forms of pulse time modulation, such as pulse duration modulation, that are less affected by noise than PAM (figure 11-1a).

PULSE DURATION MODULATION (PDM). The ideal pulse for use in time division multiplexing is a square-wave pulse, with zero rise and decay time. This ideal wave shape is closely approximated in the transmitter, but becomes distorted by the time it reaches the terminal receiver. This distortion is caused partly by the time constants of the transmitters, and receivers in the system, and partly by the noise added to the signal. In PAM-AM systems, this distortion can produce complete unintelligibility, unless a very high S/N ratio is maintained for each repeater. Other forms of time division modulation are available, which are less affected by distortion, particularly the distortion produced by noise. One of these systems is pulse duration modulation, (sometimes called pulse length modulation, or pulse width modulation). PDM, as the name implies, transmits the amplitude information not by the height of the pulse, but by its duration. Any variation in the amplitude, such as produced by noise, will have less effect on the signal than when PAM is used (figure 11-1c).

The signal to be multiplexed may be sampled at the same rate and in the same manner as for PAM. These pulses may then be used to determine the duration, or width, of the pulses to be multiplexed, so that the signal amplitude will be expressed by the duration of the pulse rather than its amplitude. In PDM the period of time allotted for each pulse is uniform, usually being measured from the leading edge of the pulse or the center of the time interval. Where the leading edge of the pulse marks the start of each period the position of the trailing edge will be determined by the amplitude of the signal at the instant of sampling.

One disadvantage of PDM is the irregular power demand on the transmitter. Since the carrier is transmitting only during the pulse periods, the power requirements of PDM, with its varying duration, will result in the transmitter output being irregular. In systems using uniform pulses uniformly spaced, the power required of the transmitter will remain constant. Another disadvantage of PDM, which is shared by PAM, is the use of the pulse shape to determine the amplitude of the signal. Any distortion of the pulse shape will therefore affect the intelligibility of the message to some extent, although PDM is much less affected by distortion than PAM.

PULSE POSITION MODULATION (PPM). The most desirable method of pulse modulation is one whereby the message is transmitted by some means other than by the shape of the pulse itself. This can be accomplished in several ways. One method employs pulses of uniform height and width, but displaced in time from some base position according to the amplitude of the signal at the instant of sampling. Pulse position modulation (PPM), or pulse phase modulation, employs a uniform pulse which transmits the signal information by means of the pulse position within an allotted period of time (figure 11-1d). The pulse period is the same as for PDM, which is determined by the initial rate of sampling. The trailing edge of the PDM pulse may be used to determine the location of the pulse in PPM, as shown in figure 11-1d. Where PDM is used to trigger the pulses for PPM, or where PDM is used to modulate the carrier, the duration of the pulse must never be greater than the period of time allotted for each pulse. Actually it will be somewhat less, to allow decay time for one pulse, and rise time for the following pulse, so that no crosstalk may occur.

Another system of PPM shifts the pulse toward either edge of the time period, with the pulse in the center of its allotted time interval in the event that no message is being transmitted. A time interval between pulses must also be employed, to reduce crosstalk to a minimum.

PULSE CODE MODULATION. Each type of time division multiplexing mentioned in this discussion is less affected by noise than the types preceding, and more affected by noise than the types that follow. On such a scale, pulse code modulation (PCM), sometimes called pulse numbers modulation, belongs last, since it is least affected by noise. As with PPM, the information is not determined by the shape of the pulse, and therefore the S/N ratio will be much greater than for PAM.

PAM employs pulses determined by the amplitude of the signal, and thus an infinite number of steps, or degrees of amplitude, may be expressed by the pulses. In PCM, an infinite number of pulses would be required to express an infinite number of amplitudes. Fortunately, it is quite possible to express various degrees of amplitude without using an infinite number of amplitudes. Because of the inability of the human ear to distinguish slight variations in amplitude, the original signal may be reproduced satisfactorily with surprisingly few degrees of amplitude. These degrees of amplitude, or steps, may be increased in several different ways. The increase from one step to the next may be uniform, or may vary with a geometric or logarithmic ratio, as desired. Since the ear is insensitive to variations of amplitude much less than double the initial amplitude, the steps are usually logarithmic, or arranged so that each step has double the value of the preceding step.

The next consideration is the required number of steps to indicate the amplitude of the signal satisfactorily. A single step, for example, would indicate the presence or absence of a tone, and nothing more. Obviously, one step would not be sufficient to convey intelligence. If two steps were used, the tone would then be either loud, or soft, or non-existent. In practice, it has been found that if the amplitude of a sine wave, from its minimum (in a negative direction) to its maximum (in a positive direction) is divided into 31 parts, satisfactory speech reproduction will be obtained. This means, then, that half of the 31 steps indicate the negative part of the wave and half indicate the positive half. In actual practice zero amplitude would be represented by the 16th step, and each half-wave divided into 15 parts with each step outward from zero amplitude increasing logarithmically, or having twice the value of the preceding step.

There is no magic in the number 31, only compromise. If the amplitude is expressed in fewer steps, accurate reproduction of the original waveshape will be impossible. Where it is desirable to recreate the original signal very accurately more steps may be used, requiring more pulses, and therefore a wider bandwidth. By limiting the number of steps to 31 satisfactory reproduction may be obtained and the number of pulses kept within a reasonable amount.

If sampling as indicated above is used, with 15 steps above the 15 below the zero amplitude value, obviously 31 values of the amplitude must be transmitted. If bandwidth were no problem, this could be done merely by using 31 pulses, each expressing a definite degree of amplitude, with the first pulse indicating maximum positive amplitude and the 31st pulse indicating the maximum negative amplitude. Since 31 pulses would require an unduly wide bandwidth, the amplitude must be coded in such a way that the 31 steps of amplitude may be represented by fewer than 31 pulses for each sample of the original signal. This may be done by use of only five pulses operating five binary counters, with the counters operated in regular sequence by the groups of five pulses. Let the values of the counters be 16, 8, 4, 2, and 1, respectively. An amplitude of 31 would be represented by all five pulses $(1 + 2 + 4 + 8 + 16 = 31)$. An amplitude of 13 would be presented by 8, 4, and 1. The code would therefore require all 5 pulses for 31, and only 3 pulses for 13, hence the reason for each pulse triggering a particular binary counter. Where a value is not needed, as in the case where the amplitude of 13 is to be transmitted, no pulse will appear. Thus, only the second, third and fifth pulses will be transmitted to indicate an amplitude value of 13.

There are two other methods of pulse time modulation: delta modulation, and random pulses. Neither is in general use at present, but deserves mention in this section. Delta modulation transmits information concerning the message by employing a pulse to indicate an increase in amplitude, and omitting the pulse to indicate a decrease in amplitude. It has the advantage of giving exceptionally low distortion where the increments are small, which advantage is more than offset by the wide bandwidth (100 kc) required. Random sampling is done by use of a random pulse oscillator, similar to a noise generator, so that the sampling rate is approximately the same as for other types of pulse time modulation. Decoding at the receiver is accomplished by employing a delay line in the transmitter, with paired pulses. The spacing between the paired pulses, one positive and the other negative, is different for each message and identifies each message. This is a comparatively new method and is not in general use.

CLIPPING. The effects of noise on the signal are greatly reduced by using pulses whose shape is independent of the information being transmitted. While the pulse will be distorted by noise, the message itself will be unaffected unless the pulse becomes badly distorted. The effect of noise is shown in figure 11-2b, a condition that would produce almost complete unintelligibility in a PAM system. The pulses may be improved by the process known as clipping, whereby the top and bottom of each pulse is removed leaving only the center section (figure 11-2c). By amplifying the remaining segment of each pulse, the original pulse may be approximately restored. The amplitude of the pulse actually used amounts to no more than 5 or 10% of the original pulse. Such a process not only removes the noise in the pulse itself, but also most of the noise that might affect the rise or decay time of the pulse. While the addition of the clipper circuits and amplifier circuit increases the equipment necessary, the S/N ratio may be correspondingly raised with satisfactory results. Further, the clipper circuits need not be installed at each repeater but only at the terminal receivers.

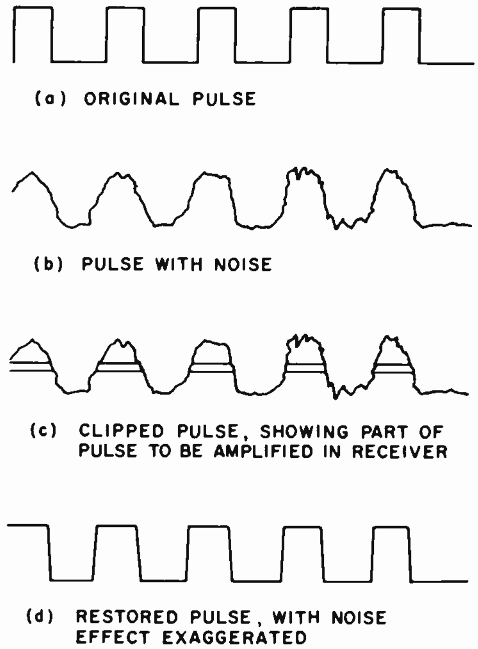


Figure 11-2. Pulse Restoration

SIGNAL-TO-NOISE RATIO

The process of combining audio messages in order that more than one message may be transmitted over a single modulation bandwidth is divided into two parts. First, the messages are combined with each other in the process called multiplexing, and then the multiplexed messages are used to modulate the carrier. Such a system is referred to as double modulation. Where the bandwidth of the carrier permits, several such carriers and their bandwidths may modulate still another carrier of higher frequency. This is called triple modulation, and the intermediate carriers are called subcarriers. The process could be repeated still further, if desirable, with the limit being determined by the permissible bandwidth of the final carrier. Triple modulation is most frequently used, with quadruple modulation rarely, if ever, used at present.

Many combinations of multiplexing and modulating the subcarriers and carriers are available, and abbreviations are used in order to designate the type of double or triple modulation in concise form. For example, when PAM is used to amplitude modulate a subcarrier, which in turn frequency modulates the carrier, the triple modulation that results is expressed as PAM-AM-FM. In the same fashion, PCM-FM-FM indicates that pulse code multiplexing frequency modulates the subcarriers, which in turn frequency modulate the carrier. Double modulation is used where there is no subcarrier, and, of course, one term will be omitted. Thus, PDM-AM indicates that pulse duration modulation amplitude modulates the carrier.

The number of combinations of multiplexing and modulating the subcarriers and carriers is almost infinite, and listing all of the possibilities, or even most of them, would be impractical. Many of the combinations would be comparatively worthless, and while others might be used, any advantage in S/N ratio, or increase in the number of messages, would not offset the added complexity of the equipment. In discussing S/N ratios only the most commonly used types of multiplexing and modulation will be presented.

Frequency-division multiplexing may amplitude modulate or frequency modulate the carrier, and single-sideband modulation may also be used. The S/N ratio is different for each, and is also affected by triple modulation, when used. In presenting the S/N ratios, three types of frequency multiplexing (AM, FM, and SS, or single-sideband), and four types of time division (PAM, PDM, PPM, and PCM) are considered. These are presented in the table on S/N ratios, figure 11-3, together with the common methods of modulating the subcarriers and carriers. In order to reduce the various S/N ratios to a common denominator, they will be compared to the S/N ratio for single channel AM, with an S/N ratio of unity.

SPECIALIZED CHANNELS

DROPS AND ORDER WIRE. In any system of communication where more than one message is to be transmitted, it is seldom that all of the messages will travel from one end of the system to the other. In many instances, some of the messages will be sent from an intermediate point to a terminal, or from one intermediate point to another. Messages may be dropped at the repeater stations for transmission to some nearby point, or they may be picked up for transmission to some other point on the relay system. Where a message is dropped at a repeater station, an empty channel will exist from that station onward unless it is used to transmit a message from that station to the terminal.

In most relay systems, one voice channel will be withheld from voice transmission in order to use it as a service wire, or order wire. This channel will inform the nearest terminal station of a fault existing at a relay station, together with some indication of the type of fault existing. A complicated system of blocking is employed, so that two repeaters cannot "report" trouble at the same time, which would happen in the event of the failure of one station disrupting the operation of an adjoining station.

AUDIO CHANNEL WIDTH. While this discussion has stressed the use of voice channels ranging in frequency from 100 to 3,500 cycles per second, there is no reason why several adjacent voice channels cannot be combined, if desired, in order to transmit a wider range of audio signals than are ordinarily used in telephone communication. Many repeater stations are designed to transmit both voice and

S/N RATIO FOR DOUBLE MODULATION, COMPARED TO A S/N RATIO OF UNITY FOR SINGLE CHANNEL AM

SYSTEM	S/N RATIO R.M.S. SUMMATION	SYSTEM	S/N RATIO R.M.S. SUMMATION
AM - AM	$\frac{1}{4\sqrt{n}}$	PAM - AM	$\frac{1}{\sqrt{n}}$
AM - FM	$\frac{\sqrt{3}\,B}{16 f_c \sqrt{n^3}}$	PAM(±)-FM	$\frac{\pi B}{12 f_m \sqrt{n^3}}$
FM - AM	$\frac{B\sqrt{3}}{16 f_m \sqrt{n^3}}$	PPM - AM	$\frac{2B}{9 f_m \sqrt{n^3}}$
FM - FM	$\frac{3B}{16 f_m \sqrt{n^3}}$	PPM - FM	$\frac{2\sqrt{B}}{3n\sqrt{f_m}}$
SS - AM	$\frac{1}{\sqrt{n}}$	PDM - AM	$\frac{\sqrt{B}}{n\sqrt{6 f_m}}$
SS - FM	$\frac{\sqrt{3}\,B}{2 f_c \sqrt{n^3}}$	PDM - FM	$\frac{\pi\sqrt{B}}{n\sqrt{3 f_m}}$
SS-FM-FM	$\frac{B\sqrt{3}}{10 d_s \sqrt{2n^3}}$	PCM - XX	∞

AFTER V.D. LANDON

WHERE:

- B = EFFECTIVE NOISE BANDWIDTH OF THE RECEIVER THROUGH THE R.F. AND I.F. CIRCUITS
- d_s = THE ROOT-MEAN-SQUARE DEVIATION OF THE MAIN CARRIER FOR ALL CHANNELS
- n = TOTAL NUMBER OF CHANNELS
- f_c = SPACE BETWEEN SUBCARRIERS FOR SS-AM; ½ SPACE BETWEEN SUBCARRIERS FOR AM-FM
- f_m HIGHEST MODULATION FREQUENCY FOR EACH CHANNEL

Figure 11-3. Single Channel S/N Ratios

television, with both modulating the same carrier. Similarly, a voice channel may be broken down into subchannels for telemetering, or telegraphy. The basic method of multiplexing will in any case remain the same.

CHAPTER 12

INTERFERENCE

INTRODUCTION

The intelligibility of a received signal is limited by noise, which, in the broadest sense, is any type of interference. *Interference* can be regarded, therefore, as noise resulting from the broad group of disturbances classed generally as static, either natural or man-made. The term noise may include continuous or discontinuous disturbances, the frequency with which such disturbances occur determining the character of the noise.

Discontinuous disturbances may be classified in two groups. Relatively infrequent, clearly separated pulses, such as ignition noises, are called *impulsive noises*. However, if the pulses occur so rapidly that they overlap and are not clearly distinguishable, then the noise is *random*. Tube and thermal agitation noises are examples of random noises, as are atmospheric noises.

Continuous disturbances are primarily the result of signal harmonics radiated from transmitters on the same or adjacent channels and harmonics of other transmitters or oscillators. This type of disturbance is usually classified as *selective interference* or interference whose energy is concentrated in a narrow band of frequencies. Radiation from unshielded diathermy apparatus is a good example of selective interference.

Figure 12-1 illustrates the main sources of radio noise as plotted against operating frequency and average field strength. The curves were drawn with respect to a carrier bandwidth of 10 kc, and are shown for comparative purposes only. The ranges and intensities of noises given are for the United States or regions of similar latitude.

NATURAL NOISES

ATMOSPHERIC. At frequencies below about 100 mc, atmospheric and precipitation noises are the most important types to be considered. Nearly all atmospheric noise is considered to originate in the lightning flashes associated with electrical storms. As received from a distant point, this noise may be expected to display the same propagation characteristics as do ordinary skywave transmissions propagated from a distant source by way of the ionosphere. In addition to these expected propagation characteristics, there are the variations to be expected as a result of different concentrations of electrical storm activity in various parts of the world, and the variation with frequency in the noise intensity radiated from individual lightning flashes. The variations in noise intensity are approximately inversely proportional to the square of the operating frequency, and the absolute noise intensity is proportional to the square root of the bandwidth. Thus, at a given location, the atmospheric noise level is composed of noise from nearby centers of noise generation such as local electrical storms, plus noise which has been propagated by way of ionospheric reflections from one or more of the principal centers of noise generation, such as the electrical storm centers in equatorial Africa, Central America, and the East Indies. The location and activity of the various noise centers vary with the time of day and season, and occasionally with the 11-year sunspot cycle. At frequencies above 100 mc, atmospheric noise originates from local sources only, since energy reflections from the ionosphere rarely occur at these frequencies. Generally, the higher the frequency, the less the effect of atmospheric noises on a received signal.

Precipitation noise arises from charged particles striking a receiving antenna. Snow, ice, rain, and dust particles are generally electrically charged, and, when such particles are wind-blown against a receiving antenna, the antenna will acquire a non-uniformly distributed electrical charge. During turbulent air conditions, the magnitude and polarity of the charge varies, causing the antenna to acquire a charge of one sign, then rapidly lose it, and so on. When the antenna loses electrical charges to the atmosphere, corona and electrical sparks often occur and cause electromagnetic radiations known as *precipitation static*. Precipitation static depends upon the operating frequency, and, at times, may be a major source of radio noise. Generally, the higher the operating frequency, the less noticeable will be the precipitation static.

COSMIC AND SOLAR. Cosmic radio noise of extraterrestrial origin is known to be the principal source of interference to reception under many circumstances at frequencies under 100 mc. This noise has much the same characteristics as the random noises of the atmosphere. The sources of cosmic noises are not uniformly distributed over the sky but tend to be concentrated in several regions on the celestial sphere, the primary source being in the region Scorpio-Sagittarius near the center of our galaxy. Consequently, when received on a directional antenna, the noise varies from hour to hour and from day to day. The cause of this noise is not yet understood. Some investigators believe it is r-f radiation from eruptions similar to spot eruptions on our sun, occurring on all the stars in the galaxy, while others have considered it to originate in free-field electron transitions occurring throughout all of space between the stars.

Cosmic noise intensity varies in proportion to the operating frequency of the receiver in use. Therefore, the signal intercepted by a receiving antenna must have a field intensity greater than that of the noise in order to be intelligible. The required field intensity, incident upon a half-wave dipole, for intelligible reception in the presence of cosmic noise is expressed approximately by the formula

$$E = \frac{2}{\sqrt{f}} \text{ Microvolts per meter } (\mu v/m) \qquad (12\text{-}1)$$

where

f = The operating frequency in megacycles (18 mc $<$ f $<$ 160 mc).

For frequencies above 160 mc, the field intensity of cosmic noise is extremely small and can be disregarded.

At the time cosmic noise was first discovered, an attempt was made to determine if similar phenomena were observable in connection with our sun. Negative results were then obtained. Relatively recently, however, measurements of radio noise at 200 mc, made with a sensitive receiving system and a high gain directive antenna, disclosed the fact that the sun also acts as a source of noise.

The theory of *thermal radiation* predicts that any black body will radiate a certain amount of energy at all wavelengths. If the theory is applied to the sun, assuming that the sun is a black body with a temperature of 6,000 degrees absolute, the intensity of radiation at 200 mc would be below the threshold level of present-day receiving systems. The intensities which have actually been measured at 200 mc are of magnitudes which correspond to a solar temperature of nearly a million degrees Kelvin. Furthermore, at times of ionospheric storms, which are associated with large sunspots, intensities of radiation have been observed corresponding to temperatures of a thousand million degrees Kelvin.

By the use of directional antennas, it has been observed that the bursts of noise emanating from the sun are circularly polarized, either right-handed or left-handed, depending on whether the sunspots are in the sun's northern or southern hemisphere. This polarization occurs only for noise bursts of high intensity; normal radio noise coming from a quiescent sun appears to have random polarization.

Measurements have also been made of solar radiation at frequencies of 3,000 to 24,000 mc, and these also give distinct evidence of radio noise coming from the sun. The intensity of this noise, however, corresponds rather closely to that which would be expected from thermal radiation, with effective black-body temperatures of the order of 20,000 degrees absolute. Some plausible and complete theories have been formed regarding the nature of solar noise and the frequency distribution of its intensity, but only a few experimental measurements have been performed toward confirming it. Except at the time of large sunspot eruptions, solar noise is important only on very high frequencies when high gain antennas are actually pointed at the sun; consequently it usually need not be considered in relation to practical problems of propagation.

MAN-MADE NOISE

ELECTRICAL NOISES. Electrical interference is caused by the operation of electrical apparatus other

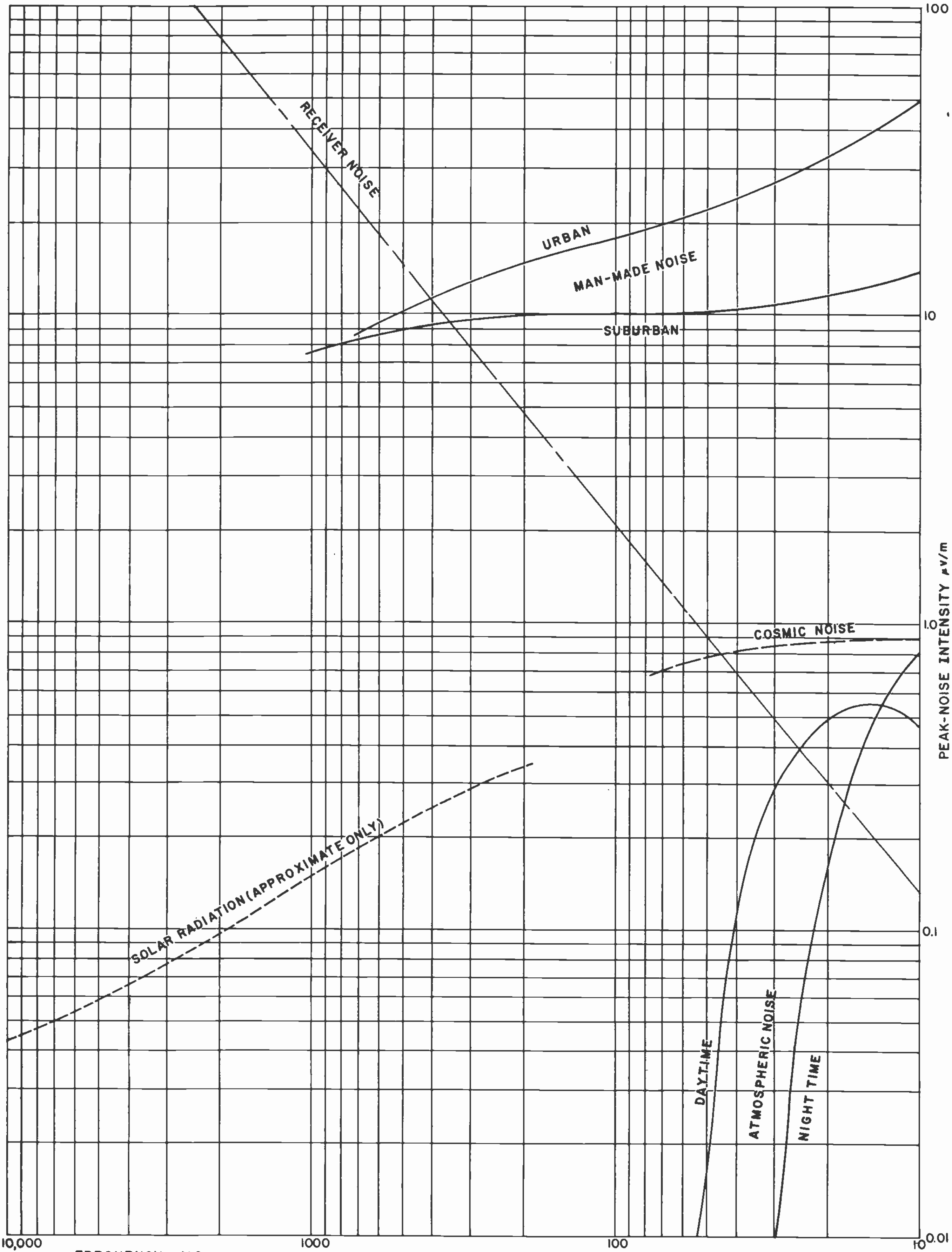


Figure 12-1. Sources of Radio Noise

than transmitters and oscillators. Man-made electrical noise can be generated by a great variety of electrical devices such as power-line discharges, motor-brush sparking, and ignition systems. The noise so generated may be propagated to the receiving antenna direct from the noise source or be radiated from power lines located near the receiver site. Electrical noise may also reach the receiver by direct transmission *through* the power lines.

NOTE

In general, electrical equipment of good quality is constructed so that any radiating noise is of a minimum value. However, unless such equipment is installed and maintained properly, satisfactory design is of no avail. Home appliances, in particular, are often poorly maintained and may cause noise because of poor contacts or loose wires. Methods of shielding electrical equipment are considered in the following chapter.

High-voltage transmission or power lines can cause excessive radio interference unless they are constructed and maintained properly. The main source of transmission line noise is *corona,* which is a discharge of electricity that originates from sharp points and corners of the line. Corona may also form on insulator tie wires, pins, and hardware because of faulty insulator design. Since corona is a high-potential electrical discharge, small sparks or arcs of electricity may occur anywhere along the line where there is a sharp point or corner and so introduce electrical spark noise. Power-line systems cause noise in a variety of ways, including arcs at fuses and switches, overloaded transformers, or simply as a carrier of appliance noises.

At frequencies above 30 mc, ignition noise is an extremely objectionable disturbance. The peak ignition noise produced by 90% of the vehicles passing 100 feet from an antenna 35 feet high has been found to be $9/\mu$ v/m for 450 mc and nearly $20/\mu$ v/m for 30 mc. As ignition interference is impulsive noise, each pulse shock-excites the usual r-f tuning circuits of the receiver. The resulting oscillations have the frequency of the receiver's tuned circuits and die out exponentially, with each disturbing pulse producing a damped train of oscillations. The damped pulse intensity dies away considerably before the occurrence of the next pulse. The field intensity of electrical noises is greatest in densely populated and industrial areas, and, according to the best available information, the peak field strength of such noise increases substantially as the receiver bandwidth is increased. The man-made noise curves as given in figure 12-1 show typical median values for the United States and are interpreted to mean that 50% of all receiving sites will have lower noise levels than the values given. Conversely, 90% of all receiving sites located in the United States will have noise levels less than seven times these values.

SELECTIVE INTERFERENCE. In addition to electrical noise, the efficiency of radio communication systems can be limited by the interference of undesired signals radiating from transmitters on the same or adjacent channels to the one desired, unshielded oscillators, harmonics of other transmitters, diathermy machines, or r-f heating generators. Interference from diathermy apparatus is perhaps the worst disturbance of this kind, and signal strengths in excess of 100μv/m are commonly found in large cities. Diathermy and induction heating systems radiate a relatively narrow-band noise with the field strength of such noise being the same, for instance, in a 100-kc bandwidth receiver as in a 10-kc bandwidth receiver. The FCC is attempting to reduce diathermy interference by assigning several frequency bands for the use of such equipment; however, many older machines not conforming to the new regulations are still in service.

If signals from two or more transmitters are picked up by a receiver tuned to one frequency, the interference so introduced is usually referred to as *cross modulation.* Cross modulation may arise from poorly designed, improperly shielded receiver circuits, or from external sources. External cross modulation generally arises from poor contact between various extended conductors in the vicinity of the receiver. Such contacts act as frequency converters or linear detectors which produce new r-f signals that may reach the receiver either by conduction or by reradiation. Conductors which may form contacts acting as frequency converters include house electric wiring, BX-cable sheath, antenna and ground leads, water and gas pipes, metal lath, and rain spouts.

RECEIVER NOISE

THERMAL AGITATION. Random noises in the form of resistance and tube noise are always present to some degree in even the most carefully constructed receiver. These noises produce a characteristic hiss-like sound in loud speakers and "grass", or an extremely irregular sweep, on oscilloscopes.

The irregular motion of electrons in any resistor and in the space current of any tube produces a fluctuation type of noise. These irregular motions result in small fluctuations of voltage and current, the instantaneous value of such voltage or current obeying the probability law. There is a certain small probability that the value of instantaneous voltage or current will fall at any one point in particular, and the probability that a very large value will occur is extremely small. However, the *probability of a large noise amplitude increases as the bandwidth of the equipment is increased.* Although the fluctuations have no definite wave form, there is a definite *rms value* and a definite *average value* for all positive and all negative noise amplitudes. The average positive or negative values are equal and mount to 80% of the rms value, while the instantaneous amplitudes exceed the rms value approximately 32% of the time. The noise so produced by fluctuations of voltage and current establishes a limit to the amount of amplification that can be successfully employed in any receiving system.

If the random motion of electrons in a resistor is produced by molecular agitation due to temperature, the resulting disturbance is called *thermal-agitation noise*. The same effect occurs in a parallel resonant circuit or other passive network, including the receiving antenna, to the extent corresponding to the resistive component of the impedance. The magnitude of the open circuit rms voltage for a given temperature may be found by the formula

$$E = 0.0074 \sqrt{R\ B\ T} \text{ microvolts } (\mu v) \qquad (12\text{-}2)$$

where

R = Circuit resistance in ohms
B = Equipment bandwidth in megacycles
T = Temperature in degrees Kelvin

Values of H for various ranges of resistance and bandwidth based on normal room temperature are plotted in figure 12-2.

An important quantity associated with thermal-agitation noise is the available power output from such a source. The available power or the maximum power which can be delivered is obtained by matching a resistive load to the internal circuit resistance in which the noise is generated. The available power from thermal-agitation noise is given by the formula

$$P = 1.38 \times 10^{-11}\ B\ T \text{ microwatts} \qquad (12\text{-}3)$$

A method for determining available noise power at room temperature is given by the formula 4.0×10^{-21} watts per cycle of bandwidth.

Random noise originates in tubes chiefly because of the irregular nature of electron emission at the cathode, and, if the tube has a screen grid, the inconsistently changing distribution of current between plate and screen. Irregular electron emission was first

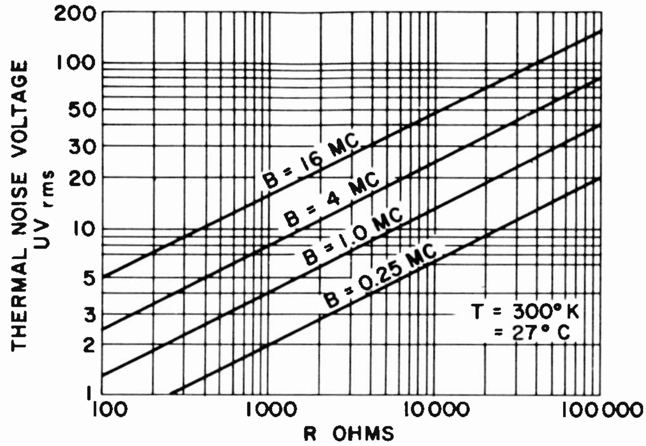


Figure 12-2. Thermal Agitation Noise

discussed by Schottky along with resistance noise, all under the name "shot effect" from the similarity of the noise to the sound of many small pellets falling upon a hard surface. *Shot effect* is now defined as tube noise resulting from random cathode emission so that resistance noise is not included. The term *partition noise* is generally applied to tube noise resulting from the irregularities in current distribution between plate and screen. Maximum tube noise is obtained when the temperature of the cathode approaches saturation, or the point at which no space charge occurs. When temperature saturation is present in a tube, all electrons emitted from the cathode are immediately drawn to the plate. In fact, a diode operated in this condition is a convenient standard noise source for use in measuring the noise figure of a receiver. The presence of space charge in electron tubes is very fortunate, for it makes amplification possible and greatly reduces the tube noise.

Tube noise is commonly represented by using an equivalent resistance which, if connected at the control grid of a noiseless tube, would produce the same thermal noise output. Values of equivalent grid resistance or *grid noise resistance* of various tubes have magnitudes within the following ranges:

Triode amplifiers..........	200- 2,300 ohms
Triode mixers.............	900- 6,000 ohms
Sharp-cutoff pentode amplifiers...............	700- 7,000 ohms
Gradual-cutoff pentode amplifiers..............	2,400- 14,000 ohms
Pentode mixers............	2,800- 35,000 ohms
Hexode and heptode mixers.	190,000-300,000 ohms

Mixer noise increases beyond these values if the oscillator injection is insufficient. If the amplifier gain is reduced by the application of more negative grid bias, the effective input noise increases, but the output noise decreases. Equivalent noise resistances may be substituted in formula 12-2 in order to determine the rms noise voltage so generated.

Equivalent grid noise resistance does not apply if the tube is used in a feedback circuit or if it has any input loading. To avoid these limitations, the internal cathode resistance of the tube can be considered as the noise source. In this case the temperature of the internal cathode resistance is approximately 0.6 times the absolute temperature of the heated cathode. After application of the 0.6 factor the temperature is approximately twice the absolute room temperature for oxide-coated cathodes, three times for thoriated filaments, and five times for pure tungsten filaments. The term generally applied to the internal cathode resistance is called *cathode noise resistance* and is measured in ohms.

For triodes, the cathode noise resistance may be found by the formula

$$R = \frac{1}{g_m + g_p} \text{ ohms} \qquad (12\text{-}4)$$

where

g_m = Tube transconductance
g_p = Plate conductance, or the reciprocal of the a-c plate resistance R.

For multielement tubes, the cathode noise resistance is the a-c resistance as measured from the cathode to all other electrodes joined together by capacitors while operating at their normal d-c potentials. Noise in multielement tubes is increased by partition noise, the sources of such noise being considered as due to the internal cathode resistance being at a higher temperature than the initial 0.6 value.

NOISE FIGURE. The sensitivity of a receiver cannot be described in terms of thermal-agitation noise alone, because the set noise at the output is several times the pure thermal noise. This increased set noise arises from amplifications and additions of thermal noise as it passes through the receiver stages. Therefore, another quantity, called the noise figure, is used to describe receiver noise. The *noise figure* (N_f) of a receiving system is defined by the formula

$$N_f = \frac{P_{N-o}/P_N}{P_{S-o}/P_S} \qquad (12\text{-}5)$$

where

P_N = The noise power as given by formula 12-3 from the antenna that is delivered to the receiver.
P_{N-o} = The noise power at the receiver output.
P_S = The signal power from the antenna as delivered to the receiver.
P_{S-o} = The signal power at the receiver output.

The ratio P_{S-o}/P_S is called receiver gain (g) and is not to be confused with antenna gain (G). By using formula 12-3, formula 12-5 may be restated as

$$N_f = \frac{7.25 \times 10^{10}\, P_{N-o}}{B\, T_g} \qquad (12\text{-}6)$$

If the receiver consists of several stages in cascade, including converters, amplifiers, and detectors, the overall noise figure can be compounded from the noise figures and gains of the individual stages or components by the means of the formula

$$N_f = N_{f1} + \frac{N_{f2-1}}{g_1} + \frac{N_{f3-1}}{g_1 g_2} \cdots \frac{N_{fk-1}}{g_1 g_2 \cdots g_k} \qquad (12\text{-}7)$$

where

N_f = The overall noise figure
N_{fk} = The noise figure of the k^{th} stage
g_k = The gain of the k^{th} stage.

Equation 12-7 simply shows that most of the receiver noise originates in the early stages of reception. It also shows, of course, that noise picked up at later stages is much less amplified by the system than the noise from the first stages.

In equipment specifications, the noise figure is usually expressed in the decibel scale as *decibels above thermal noise*. Actual noise figures vary from a few decibels above thermal noise for receivers operating at frequencies up to 300 mc, to larger values at the higher frequencies.

SIGNAL-TO-NOISE RATIO. Since the intensities of both signal and noise can be measured in terms of rms voltage or current, or in terms of power, such a comparison is used to express reception conditions. The ratio of signal intensity to noise intensity at the output of a receiver is called the *signal-to-noise ratio* (S/N), and is expressed in decibels. For signal and noise intensities that are measured in rms voltage *directly at the output terminals of a receiver*, the S/N ratio in decibels may be found by the formula

$$S/N\ (\text{db}) = 20 \log S/N \qquad (12\text{-}8)$$

where

S = Signal intensity in rms voltage (Peak voltage for impulsive noises)

N = Noise intensity in rms voltage (Peak voltage for impulsive noises)

For signal and noise intensities that are measured *at the output terminals of a receiver* in terms of power, the S/N ratio in decibels may be found by the formula

$$S/N\ (db) = 10 \log P_S/P_N \qquad (12\text{-}9)$$

where

P_S = Signal Power in watts
P_N = Noise Power in watts

The signal-to-noise ratio of a receiving system can best be analyzed by finding all the various noise sources and then referring them to a single point. When an equivalent noise resistance is used at the receiver input to express the output noise of the first tube of a series, noise sources beyond this point are also referred back to this same input. Thus, a single equivalent noise resistance can be used to express all the noise sources at or beyond the first tube.

When there is external noise present at the input of the first tube, the optimum signal-to-noise ratio depends on the product of external noise and equivalent noise resistance. Even when the input resistance is infinite or negative, the optimum signal-to-noise ratio is limited either by this external noise or by the bandwidth of the input circuit. The signal-to-noise ratio is also inherently limited by the receiving antenna and associated transmission lines which develop noise of their own. The antenna noise may be considered equal to the thermal-agitation noise developed in an equivalent dummy-antenna resistance.

The signal-to-noise ratio of a radio receiver at ultra-high frequencies, or higher, is primarily dependent on the above-mentioned noise relations for tubes and circuits. However, in addition to these noise relations, consideration must be given to the many types of signal modulation used. For example, there is a great difference between the signal-to-noise ratio in one of the telephone channels in a multiplex system and the carrier-to-noise ratio at the final detector. The received carrier-to-noise ratio depends upon a number of factors. First, the propagation path should be clear of any obstructions such as hills and trees. Other factors which determine the carrier-to-noise ratio include transmitter power, antenna gain, wavelength, distance between stations, receiver bandwidth, and receiver noise figure.

With either amplitude modulation or frequency modulation, the signal-to-noise ratio at the output of the receiver is not the same value as the carrier-to-noise ratio at the final detector. It is common practice to compare the signal-to-noise ratios of the various multiplexing systems to that obtained with amplitude modulation. The ratio of signal to noise ratio, comparing a single channel of a multiplex system to a single channel AM system having the same average radiated power is called the *wide-band improvement* of a multiplex system.

In the case of amplitude modulation the equation for the signal-to-noise ratio is given by the formula

$$\frac{S}{N}\ (AM) = M \sqrt{\frac{B_{i-f}}{2B_{a-f}}}\ \frac{E_c}{N_{ld}} \qquad (12\text{-}10)$$

where

B_{i-f} = The intermediate frequency bandwidth
B_{a-f} = The audio frequency bandwidth
M = The modulation index
E_c = The carrier voltage
N_{ld} = The noise ahead of the last detector

In frequency division multiplexing, where each signal modulates a separate sub-carrier and the sub-carriers are so spaced in frequency as to prevent overlapping, the ratio of signal-to-noise voltage ratios as compared to a single channel AM system is given by the formula

$$\frac{S/N\ (xx\text{-}xx)}{S/N\ (AM)} = K_1 M_1 M_2 K_2 \qquad (12\text{-}11)$$

where

(xx-xx) = Type of modulation
K_1 = 1 for AM, $\sqrt{3}$ for FM
M_1 = Modulation index of the signal on each channel for AM, or the deviation ratio if FM.
M_2 = Modulation index of the sub-carrier or individual signal applied to the main carrier
K_2 = The second modulation constant.

The above formula may be simplified by using an AM signal-to-noise ratio of unity in the denominator. Figure 3 in Chapter 11 lists fourteen formulas of S/N ratios for double modulation as compared to a S/N ratio of unity for single channel AM.

CHAPTER 13

INTERFERENCE ELIMINATION

INTRODUCTION

Unwanted interference picked up by a receiving antenna along with the desired signal is, in effect, part of the modulated r-f wave. Such interference is then amplified and demodulated together with the desired signal. Therefore, noises which lie within the acceptance band of the receiver will be present at the receiver output in spite of all possible precautions taken in the initial system design.

Every type of receiving system in the frequency range from 44 to 100 mc has a definite signal-to-external-noise ratio below which the intelligibility of the signal is insufficient for reliable operation. If the signal-to-external-noise ratio of the received wave is below this minimum value, one of the means available to obtain an intelligible signal is to increase the effective radiated power of the transmitter. Occasionally, advantage may be taken of the receiving antenna's directivity, making it blind to the sources of the noise and responsive only to the desired signal. This may be accomplished only when the noise and signal directions are at wide angles to each other, and when suitable antenna directivity is possible.

In the frequency range above 100 mc, where the external noise intensity is very low, there are different transmission characteristics to be considered. In this range, very weak signals may be utilized by the use of high receiver gain, since the received field intensity is not limited by external noise. Receiver noise, however, can be a limiting factor in this case, with the limitations increasing directly with the operating frequency. Under these conditions, the signal-to-noise ratio of the receiver is always improved by increasing the signal power delivered to the receiver, although excessive power is too often employed in order to override engineering deficiencies, thus causing unnecessary interference with other services.

NATURAL NOISE REDUCTION

ATMOSPHERIC. Atmospheric noise varies throughout the day and season and with operating frequency, the effects of such noise decreasing with increasing frequency. For the frequencies ranging from 44 to 100 mc, local electrical storms are the main source of atmospheric noise, since noise from the principal noise centers of the world is rarely ever propagated by way of the ionosphere at these frequencies. It is good practice, however, to make atmospheric noise measurements at locations where a receiving antenna is to be installed. The measurements should be made in such a way as to reveal the diurnal and yearly range of noise intensities at the frequency ranges intended for use. Noise must always be taken into account when choosing a site for a receiving station, and both atmospheric and man-made noise may be involved. In the frequency range above 100 mc, it is seldom necessary to make atmospheric noise measurements at receiver locations since even local electrical storms do not ordinarily cause any interference.

At times, precipitation noise may become serious in some locations. Sandstorms, dry snow, and wind-blown fog are well-known causes of precipitation static interference. Such noise can be reduced substantially by embedding all metallic antenna and feeder system components in a low-loss dielectric in order to prevent the charged particles which strike the antenna from discharging directly to the metal. By preventing exposure of the metal antenna components to the charged particles, a reduction of several decibels in noise level can be realized during conditions of precipitation noise. The addition of a d-c ground to the antenna system, if one may be added without upsetting the r-f response, may also tend to eliminate precipitation noise.

COSMIC AND SOLAR. Radio noise of extraterrestrial origin has much the same characteristics as atmospheric noise, and its intensity varies in proportion to the operating frequency. Formula 12-1 in Chapter 12 may be used to determine the necessary field intensity for intelligible reception in the presence of cosmic noise. The formula is limited to an upper operating frequency of approximately 160 mc, since the field intensity of cosmic noise is extremely small at higher frequencies and can, therefore, be disregarded.

Radio noise emanating from the sun interferes with radio-relay operation only under certain conditions. For example, a receiving antenna that is directed horizontally towards the east may intercept solar noise at the particular time when the sun is just appearing over the horizon. Similarly, a receiving solar noise emanating from the setting sun. However, antenna directed toward the west may intercept as this phenomenon occurs only a few minutes a day and, at the most, several days a year, the noise so intercepted need not be considered. At the time of large sunspot eruptions, solar noise may then be more noticeable, although the use of highly directive antennas offsets the noise considerably.

MAN-MADE NOISE REDUCTION

ELECTRICAL NOISES. A receiving site, for equipments operating in the frequency range from 44 to 300 mc, must be as free as possible from electrical noises. The site should be an adequate distance from large cities or other places that might be sources of electrical noises. Automotive ignition systems are also a major source of electrical noise and substantial distance between receiving sites and well-traveled highways should be allowed. The distance may be determined by direct measurement of such noise at various locations. The amount of man-made noise that can be tolerated at a particular receiving site depends upon the prevailing natural noise levels, and any man-made noise present at the site should always be substantially less than the natural noise received during low-noise periods. At a well-selected site, intelligible reception should be limited only by natural causes.

The suitability of a receiving site will often depend upon the direction from which the noise is propagated. If the dominant noise interference is always from a direction substantially different from that of the desired signals, antenna directivity can be employed to favor the signal and discriminate against the noise. The man-made noise may also arrive at the site at a much lower angle than the incoming signal, in which case the antenna may be tilted slightly off horizontal in order to give the best response to the incident signal and a relatively low response to the noise.

There are times when a receiving site must be located in a city or other region of severe man-made noise. It may be necessary to attempt to suppress noise from troublesome dominant sources in the neighborhood of the receiving station in order to obtain tolerable system performance. Especially troublesome are the clicking noises which arise from elevator power controls where both the receiver and the elevator relays are located in the same building. The preventive measures that can be successfully applied depend upon the nature of the device and the kind of noise it emits, *assuming that the source can be located.* Loose wires, poor contacts, worn motor brushes, and transmission line corona are perhaps the most common electrical noise sources. Poor grounding and shielding of electrical devices may also be a major reason for excessive electrical noise. Bypassing sparking contacts with suitable capacitors and the use of power line filters may be remedies for certain types of noise. Commercial filters are available for this purpose. In many instances, 0.05 μf capacitors, connected across the power line at the point where it enters the electrical equipment causing the problem will prove effective. A brief summary of a number of forms of electrical interference, together with a description of the general noise they produce is given below. The summary is not meant to be complete, but rather to serve as an indication of types of noise sources.

A. *Whirring or Whining Noises* (Often accompanied by crackling)

Adding Machines	Electric Sewing Machines
Barbers' Clippers	Elevators
Beauty Parlor Equipment	Farm Lighting Plants
Cash Registers	Field Telephone Magnetos
Dental Equipment	Floor Polishers
Dish Washers	Generators
Dough Mixers	Hair Dryers
Electric Addressographs	Humidifiers
Electric Blowers	Motor Generators
Electric Fans	Portable Electric Drills
Electric Refrigerators	

Printing Presses	Valve Grinders
Toy Electric Trains	Washing Machines
Vacuum Cleaners	

B. *Rattles, Buzzes, and Rapid Clicking Noises*

Automotive Ignition Systems	Dial Telephones
Buzzers and Trembler Bells	Elevator Controls
Dental Equipment	Sewing Machines
	Vibrating Rectifiers

C. *Heavy Buzzing or Rushing Sounds*

Air Purifiers	Neon Signs
Battery Chargers	Oil Burner Ignition Systems
Diathermy Machines	R-F Induction Heaters
Flour Bleachers	X-Ray Equipment

D. *Crackling or Spluttering Noises*

Bad Connections	Poor Contacts in Power Wiring
Defective Light or Power Sockets	(Occasionally intermittent breaks in connecting cables)
Elevator Controls	Trolley Cars
High-Tension Lines	
Partially-Grounded Power Lines	

E. *Clicking*

Electric Typewriters	Mercury Arc Rectifiers
Elevator Controls	Ovens
Flashing Signs	Telegraph Relays
Incubators	Traffic Signals

A portable receiver with a loop antenna may be used as an interference locator. With the receiver operating, gradually rotate the loop antenna. If a point is found at which the interference diminishes or vanishes, the direction of radiation will then be at right angles to the loop. Once the direction of interference is found, it is not too difficult to move toward the direction in which the interference increases.

While this method of interference detection is the most satisfactory, it is only applicable in a few instances where the major interference comes from one particular source which can be satisfactorily suppressed. In urban areas, the source of the noise is usually not easy to determine, and when it is, the cost of filters is prohibitive. It is therefore best to follow the procedure outlined below for improving the antenna system.

1. Use a more directional, high-gain antenna.
2. Attempt to find an antenna position which gives a minimum noise pickup.
3. Check antenna height for maximum signal pickup.
4. Make certain that the antenna, transmission line, and receiver are correctly matched and also well-balanced to ground.
5. Make use of antenna tilting, if possible, making the antenna blind to the noise sources and responsive only to the desired signal.

SELECTIVE INTERFERENCE. The signal generated by a transmitter, diathermy machine, or other r-f oscillators operating on the same or adjacent channels to the one desired, can have a component which will interfere with the desired signal. The component can be a harmonic of the operating frequency, a parasitic oscillation, or it can be produced by cross-modulation. For example, an improperly operating c-w transmitter may give rise to spurious emissions with sidebands that may extend many hundreds of kilocycles either side of the steady carrier frequency. These large sidebands are generally caused by a large, rapid frequency shift as the key is closed or opened.

In the various stages of a transmitter, the effectiveness of direct harmonic radiation may be minimized by short, direct leads and the isolation between stages afforded by link coupling. Good grounds and complete shielding of the lower-power stages should be obtained if possible. Direct harmonic radiation from a high-powered final stage is minimized in the same manner, although greater care must be taken because of the magnitude of the currents and larger-size components. Steps necessary to prevent harmonic currents from finding their way to the transmission line and the antenna include low-pass filters and tuned traps (transmission-line stubs) interposed between the final tank and the transmission line. It cannot be emphasized too strongly that if harmonic currents are permitted to flow to ground by way of long paths, such as the a-c line, the transmission line, or the antenna, the resulting radiation will be a more dangerous source of interference than direct harmonic radiation from the transmitter stages. Likewise, shielding should not be used as a convenient ground for general use.

Flexible coaxial line is ideal for use as a tuned trap or rejection stub, since the field produced by out-of-phase currents flowing from the inner to outer conductor is contained within the cable. Coaxial stubs therefore strictly limit any direct radiation and may be coiled up for convenience. A quarter-wavelength stub, figure 13-1 shorted at the far end presents a very high resistive impedance across its unshorted terminals at the operating frequency. Such a stub bridged across the output terminals of the offending transmitter effectively shorts the even harmonic frequencies. The most common stub lengths and hook-ups are given in table 13-1.

If additional harmonic attenuation is needed in a particularly troublesome situation, a second rejection stub may be placed along the main transmission line. The second stub is spaced one-quarter wavelength of the undesired harmonic from the first stub. Transmission line stubs may also be used at the receiver input terminals and, in some cases, are a practical means of selective interference elimination when such interference lies within a narrow frequency band. The stub is cut to a length which is resonant at the frequency of the unwanted carrier and bypasses it from the receiver input. Several commercial wavetraps are also available and are sometimes very effective.

Interference from diathermy apparatus may be particularly troublesome at certain site locations. Although most modern diathermy machines are equipped with r-f filters so that radiation into the power line is reduced, many machines will still interfere with wide-band receiving systems such as microwave relays and television receivers. When the frequency of the diathermy apparatus is nearly the same as the i-f frequency of a nearby receiver, a buzzing sound may be heard over the entire receiver bandspread tuning range. Little can be done at the receiver site to remove this type of interference, although additional shielding of the i-f amplifier and better front end selectivity may be tried. If the diathermy machine can be located, the owner should be asked to cooperate by shielding the machine, by keeping the machine on its correct operating frequency, and by grounding it properly, all as prescribed by the FCC.

SHIELDING. Undesired electromagnetic energy in the form of noise or r-f waves may be confined by blocking all undesired methods of propagation. Such a process is called *shielding*. The placement of all the

TABLE 13-1: SHORTING STUB LENGTHS

Harmonic to be Eliminated	*Stub Length in Terms of Operating Wavelength*	*Conditions*
Even Harmonics	$\lambda/4$	Shorted at far end.
3rd Harmonic (two stubs needed)	$\lambda/6$ $\lambda/12$	Shorted at far end. Open at far end.
5th Harmonic (two stubs needed)	$\lambda/5$ $\lambda/20$	Shorted at far end. Open at far end.

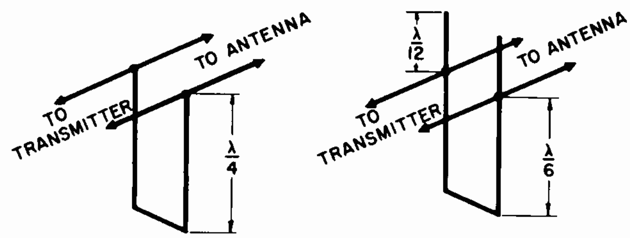


Figure 13-1. Harmonic Rejection Stubs

components of an undesired radiation source within a grounded shielding box confines the *electric field* within the box. Even if the box contains holes, slots, or is constructed of wire screening, the shielding of the electric field is still almost complete, provided the resistance of the screen is very low. In most cases, copper and aluminum are entirely satisfactory shielding materials for the electric field. Copper screening is satisfactory if it is bonded to copper strips at the edges, in order to insure a good low resistance contact to each wire of the netting, despite a faulty contact between individual wires.

Generally, *magnetic fields* are produced by the flow of current along a conductor. Since high-frequency r-f currents are propagated along the surface or skin of a conductor, no magnetic field can be produced external to a solid metal box, regardless of whether the box is grounded or not. Strong currents propagated along the inner surfaces of the box generate a second magnetic field which neutralizes the first internal field, making the total field external to the box zero. However, small holes in the box will permit penetration of the magnetic field to the outside. The magnitude of the flux penetrating a hole is a function of the diameter of the hole, and for small holes (less than $\lambda/8$) external flux exists only in the immediate vicinity of the hole. Slots, even if narrow, can be extremely dangerous if the magnetic lines of force are parallel to the slot. On the other hand, if the magnetic lines of force are at right angles to the slot, the slot will be no more harmful than a row of holes. Slotted shielding is usually avoided because of the difficulty of determining the direction of the magnetic lines of force, and any small holes necessary for ventilation should be carefully located with respect to the components. Individual secondary shields placed around components that produce large magnetic fields may also be used to advantage.

RECEIVER NOISE

NOISE FIGURE. While at frequencies below 300 mc, man-made noises are the dominant types of interference in cities, receiver noise is prevalent in rural areas, and is dominant down to 44 mc, and possibly still lower. Only recently has a good receiver been built which will amplify the received signal at radio frequencies above 500 mc. Hence, it is customary to convert the signal to a lower frequency before it is amplified. There are two methods of doing this. The most direct approach is to rectify the signal and change it to an audio wave which may be amplified. However, in the majority of cases it is more convenient to change the frequency of the signal to some intermediate frequency, and first amplify the intermediate frequency, demodulating it after a large amplification. There are three main reasons for this:

1. It is possible to detect weaker signals.
2. Greater selectivity is obtained.
3. It is more desirable to have the receiver gain distributed between two frequencies, thereby avoiding feedback troubles.

It is well known that if an amplifier has enough gain, a random voltage appears at the output terminals. As discussed in Chapter 12, this random voltage produces noise in the output of the receiver. For any receiver, there will be a definite amount of input signal power which is needed to make the output signal power equal to the output noise power. The *noise figure* of a receiver is the difference, in db, between the actual noise present in a receiver and the thermal noise of a resistance equal to the input load connected to the receiver. Normally the noise figure of a receiver is expressed in decibels above thermal noise.

CRYSTAL MIXERS. Ordinary heterodyne radio receivers use tubes for first detectors. Many tubes have been tried for the same purpose in microwave reception but all have proved quite noisy. In the 3,000-mc band, tubes in service seem to develop about 25 db more noise than the theoretical noise power, while the best of these tubes will develop 16 db. The receiver noise situation is even worse for reception at 10,000 mc. As a result attention has been turned to the crystal detector, once widely used at ordinary radio frequencies but for many years completely superseded in this field by tubes.

A crystal used in this fashion is generally called a "mixer". Several housings have been designed for crystals mounted in coaxial cables operating at 3,000 mc. Such devices must have inputs for signal, local oscillators, and an i-f output. There also must be provisions for bypassing the r-f, and a filter must be used to prevent the r-f from escaping through the i-f line.

The excellence of performance of a crystal used as a frequency converter will be determined primarily by the conversion loss of the crystal and by the noise which it introduces into the receiver. It has been found experimentally that the power from the local oscillator causes the crystal to generate more noise than an equivalent resistor at room temperature. The output noise ratio (N_o) of a crystal is defined as the ratio of the available output noise power of the crystal to the theoretical available noise power as developed by an equivalent resistor.

A weak signal from the receiving antenna becomes still weaker when it has become converted by the crystal to the intermediate frequency. The *conversion loss* (L) of a crystal is defined as the ratio of the available r-f power as applied to the crystal over the available i-f power output from the crystal, and is always greater than unity. The conversion loss and output noise ratio of the crystal are related to the receiver noise figure by the expression:

$$N_{f-r} = L_c(N_o + N_{f(i-f)} - 1) \qquad (13\text{-}1)$$

where

N_f = The receiver noise figure
N_o = The crystal output noise ratio
$N_{f(i-f)}$ = The noise figure of the i-f amplifier
L_c = The crystal conversion loss

It has been found that L_c and N_o vary so widely from crystal to crystal that it is necessary to measure each unit and accept it for use only if it falls within specified limits.

The effect of the application of d-c bias on crystal performance has been noted experimentally. It is found that negative bias is undesirable in that the noise output is increased and the signal reduced. However, an improvement is obtained with a small positive bias. This improvement also varies widely from crystal to crystal and may be as much as 1 db in the noise factor of a receiver having an i-f noise figure of 5 db. The improvement occurs mainly in the reduction of output noise and is most evident in extremely noisy units, where there is usually an improvement in the signal output also. The most favorable bias appears to be in the neighborhood of +0.10 volt. Crystal detectors should be tested at odd intervals of time, since the slightest overload may seriously impair the operation of such a detector.

CHAPTER 14

COMPUTATIONS OF WAVE PROPAGATION FOR RADIO RELAY APPLICATION

INTRODUCTION

When engineering radio relay systems, it is important to have a relatively accurate means of predicting circuit performance and reliability. Two frequency ranges will be considered in this chapter, the first ranging from 30 to 500 mc, and the second ranging from 500 to 13,000 mc. In the higher range of frequencies, generally speaking, present equipment limitations restrict the length of any individual hop in a system to line-of-sight distances. In the range 30 to 500 mc, however, it is entirely possible to work over distances considerably in excess of optical. Therefore, the lower frequency range appears to represent the optimum frequency range for radio relay systems designed for a moderate number of channels. For a greater number of channels, or for television applications, a higher frequency range, where greater bandwidths are readily available, should be used.

THE BULLINGTON NOMOGRAMS

One of the most useful tools available at the present time for the computation of path losses is the set of nomograms prepared by Bullington.* These nomograms (figures 14-11 through 14-19) are based on a smooth spherical earth with a standard atmosphere, and have been designed to take into consideration reflection, refraction, and diffraction which are the fundamental properties of wave mechanics.

*P.I.R.E. October 1947—Bullington: Radio Propagation at Frequencies Above 30 MC.

Provided these nomograms are used with a reasonable degree of care, relatively accurate predictions of system performance can be made for all but those cases where the system extends well beyond the line of sight. For these latter cases, however, there is now a considerable amount of experimental evidence indicating that field strengths may be expected which are consistently higher than the predicted values given by the nomograms.

When predicting system performance, it is essential to know the received field to be expected over a long term period for some percentage of the total time. Actually, it is not considered practical to predict field strength values closer than 5 db, because of the variations in the received field experienced in practice. On the other hand, if the predicted values are not within 5 db, the relay systems involved would have to be engineered more conservatively, with a consequent increase in capital cost.

In most applications, it is thought that the field strength value experienced 90% of the total time is a good one to aim at. It might well be asked why this value is taken, rather than 99.99%. *The 90% value* *has been chosen arbitrarily as a good engineering compromise between economy and efficiency, and is considered to be satisfactory for all but very special applications.* To achieve a 99.99% value, it may be necessary to provide a field intensity at the receiver which is as much as 15 to 20 db higher than the field intensity for a 90% value, thus involving the installation of quite a number of additional repeaters in a multi-hop system.

It must be remembered that the 90% value is an expected field strength value and not the reliability of the system. For example, if calculations show a certain field strength may be expected for 90% of the total time, and the fade margin of the receiver is 40 db below this value, a fade of 20 to 30 db would be well within the compensating abilities of the receiver. In this case, the total reliability of the system would be much higher than 90%, since a fade of over 40 db is required for a complete outage. Only when the field strength value expected for 90% of the time is equal to the threshold level of the receiver, will the reliability also be 90%. (See "reliability", Chapter 5.)

The problem now is to know whether the Bullington nomograms are adequate to give an assessment of expected field strength within 5 db for 90% of the total time for any given radio relay circuit. To check this it is necessary to apply the Bullington data to as many cases as possible where results of field intensity measurements over a relatively long term period are available. This has been done, and the results of the analysis are indicated below.

PATHS WITHIN THE OPTICAL RANGE. For all such paths, the profiles will be taken to be drawn using an earth's radius which is 4/3 times the true radius. This, of course, provides an allowance for refraction effects in a standard atmosphere. Paths within this category include those whose range would be optical but for an intervening obstruction, other than that due to the earth's curvature.

Nine such paths, for which reasonably long term field intensity measurements are available, have been analyzed and excellent correlation exists between the predicted and measured field intensities received 90% of the time. A summary of these values is given in table 14-1, and shows that in no case is the predicted value 2 db greater than the measured 90% value. As would be expected also, the difference between the 50% and 99% measured values for such paths is relatively small. In only one case does it exceed 6 db.

It should be emphasized, however, that the method of applying the Bullington nomograms can in some cases materially influence the value of the predicted field. An example is the Empire State Building to Hauppauge circuit, the profile of which is shown in figure 14-1.

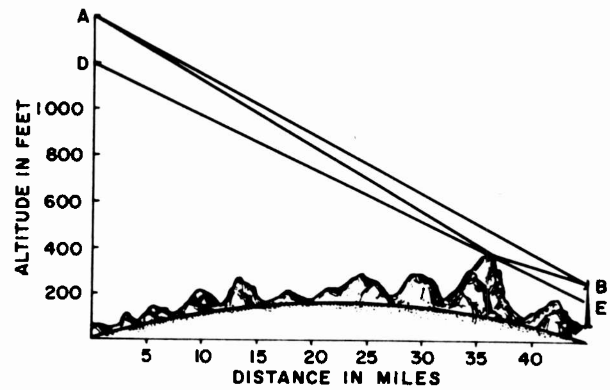


Figure 14-1. Empire State Building to Hauppauge Circuit

At first sight it would appear that the considerable elevation at each end will give a field at B due to a transmitter at A which should approach the free-space value. However, a closer check indicates that peak C falls within the first Fresnel zone (see nomogram H in Appendix I for determining first Fresnel zone clearance). Peak C is, therefore, a possible point of serious reflection to be taken into account. By considering an imaginary flat earth DCE running through point C in such a manner as to make the angle of incidence equal to the angle of reflection, the antenna heights AD and BE will be found. By using these antenna heights in conjunction with the nomogram in figure 14-14, we find that such a procedure gives a predicted field which corresponds to the 99% value and is within 1.8 db of the measured 90% value. Any other method of applying the nomograms to this particular profile, such as taking an increased effective height at each end on a plane earth basis, and including the shadow loss from figure 14-19, gives a predicted field considerably in excess of the measured field.

This method of drawing an imaginary plane earth through any point on the profile which is within the first Fresnel zone has been applied to all the profiles examined under this heading and, as can be seen from table 14-1, has yielded remarkably good correlation with the measured fields.

OBSTRUCTED LINE OF SIGHT. The procedure followed for the obstructed line of sight paths will be more readily understood by examining the two profiles shown in figure 14-2.

In the Alliford Bay to Marble Island circuit, the antenna A is on the edge of a cliff and has complete first Fresnel zone clearance in the immediate foreground. On the other hand, the antenna at B does not have first Fresnel zone clearance in the immediate foreground. The plane earth line has therefore been taken as DC, giving antenna heights of AC and BD. The shadow loss from figure 14-19, using distance d and height h, has then been added to the plane earth loss given by figure 14-14.

In the Bridgeport to Riverhead circuit, the antenna at B does not have first Fresnel zone clearance, whereas the antenna at A does. The plane earth line, therefore, has been taken as running tangential to the earth's surface at F from the base D of the antenna at B. This gives antenna heights AC and BD to be used for assessing the plane earth loss from figure 14-14. Distance d and height h are used in assessing the shadow loss from figure 14-19.

Summarizing, therefore, it would seem that contours for clear line of sight and obstructed line of sight paths can be analyzed to give a predicted value of field intensity which can confidently be expected to exist for 90% of the time, provided that the following basic principles are observed.

1. Draw all profiles with an earth's radius of 4/3 times the true radius.
2. For paths where complete first Fresnel zone clearance exists over the entire path, take the free-space path loss given by figure 14-10.
3. For clear line-of-sight paths where first Fresnel zone clearance does not exist over the entire path, draw a plane earth through possible points of reflection (whether they are in the immediate vicinity of the antennas or well away), and use the effective antenna heights so given to find the plane earth loss from figure 14-14.

TABLE 14-1

COMPARISON OF THEORETICAL COMPUTATIONS AND LONG TERM FIELD INTENSITY MEASUREMENTS FOR LINE OF SIGHT CIRCUITS

Reference	*Hop*	*Freq. in MC/S*	*Distance in Miles*	*Predicted Value of Field Intensity in Db above 1 μv/m*	*Percentage of Time Measured Fields Indicated (Db above 1 μv/m) Are Exceeded* 1%	50%	90%	99%	99.9%	*Duration of Measurements*
CLEAR LINE OF SIGHT PATHS										
RCA Review June 1946	Prince Rupert–Alliford Bay	45	98	26	47.3	42.9	40.3	36.6	33.3	1 week
RCA International Field Report	Tel Aviv–Judian Hills	250	26.5	55	59.5	56	53	42	36	3 months
RCA International Field Report	Tel Aviv–Mt. Carmel	250	50	48	57	53.5	51	47.5	40	3 months
RCA Report EM 61-17	Empire State–Hauppauge	288	42.5	51	63	55.7	52.8	50.9	49	3 months
RCA Review June 1946	Alliford Bay–Marble Island	45	30	27.5	30	27.8	24.7	23.8	23.6	1 week
RCA Review June 1946	Bella Coola–White Point	45	55	13	24.8	20.1	17.6	14.88	14.1	1 week
RCA Report F 43-118	Alpine–Hauppauge	93.1	40.5	61.5	69	64	62	60.5	53	15 months
RCA Report F 43-121	New York–Hauppauge	203.75	41.9	51.5	60	55	53.5	52.3	51	15 months
RCA Report F 43-119	Bridgeport–Riverhead	534.75	33	52.8	62	55	53	49	42	21 months

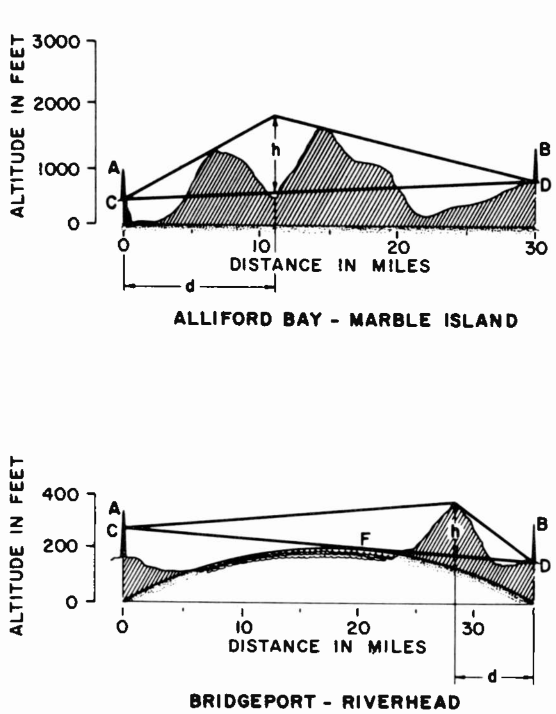


Figure 14-2. Alliford Bay to Marble Island Circuit

4. For obstructed line-of-sight paths, use the same concept of a plane earth as indicated in 3, and increase the path loss due to the obstruction as given by figure 14-19.

5. The possible exception to 4 is the case of an obstructed line-of-sight path which would come under the free-space category but for the obstruction. Such a path would be analyzed by using figures 14-10 and 14-18.

PATHS BEYOND THE RADIO HORIZON. Bullington has indicated that the received power at points far beyond the horizon is relatively independent of frequency, antenna heights, and weather effects. From experimental data he has produced two empirical curves, which are illustrated in figure 14-20, giving a value for the expected medium field in terms of db below the free-space value for varying distances. It is of interest to compare these curves with some experimental curves published by Gerks.* These curves are reproduced in figure 14-3, and on them have been superimposed the free space and the Bullington empirical curves for the particular conditions that apply.

This procedure shows that there is a good degree of correlation between the work of Gerks and Bullington for paths extending well beyond the optical range. Bullington's statement that the field intensity well beyond the horizon is independent of the antenna height is also well substantiated by the bunching beyond the optical range of curves A, B, C, and D, which are curves for receiving antenna heights ranging from 10 feet to 15,000 feet.

Let us now turn to the paths beyond the optical range which have been considered. Nine such paths for which reasonably long term field intensity measurements are available (periods from one week to 41 months) have been analyzed, and the results are indicated in table 14-2. In each case, two predicted values have been shown: one calculated on the basis of the Bullington nomograms, and the other calculated from his empirical curves as shown in figure 10. These latter are median value curves, and would therefore be expected to yield results reasonably close to the measured median (50%) values.

Examination of table 14-2 indicates that for 6 of the 9 paths considered, the predicted values of field strength (using figure 14-20) are within 3.8 db of the measured median value. Of the next 3 paths, the predicted value for one is within 6 db of the median value experienced; and for the last two paths, the predicted values correspond to the fields experienced 99.9% of the time.

Consequently, with the exception of the latter two paths, there is a very good degree of correlation between the median values predicted using figure 14-10, and the measured median values. Insofar as these latter two paths are concerned, it will be observed that the duration of measurement was only one week. Also, as both paths are over water, it is possible that a high degree of atmospheric stratification may have existed during the tests. Such atmospheric conditions would result in higher median levels of field intensity for a short term than would be experienced over a long term.

A comparison of the measured values in table 14-2, and the values calculated using the Bullington nomograms, indicates that no consistent correlation is possible for these paths. The problem is, therefore, to decide which method should be applied in computing path losses for any particular circuit extending beyond the optical horizon.

Bullington has indicated* that his nomograms provide a result which agrees reasonably well with experimental results if the correction for earth's curvature is less than 20 or 30 db. If we assume that 20 db is the maximum allowable curvature correction when using the nomograms, it is possible from figure 14-15 to plot a curve which will indicate the distances

*P.I.R.E.—November 1951—Gerks: Propagation at 412 mc from a high power transmitter.

*P.I.R.E.—January 1950—Bullington: Radio Propagation Variations at UHF and VHF.

TABLE 14-2

COMPARISON OF THEORETICAL COMPUTATIONS AND LONG TERM FIELD INTENSITY MEASUREMENTS FOR CIRCUITS EXTENDING BEYOND THE RADIO OPTICAL HORIZON

Reference	*Hop*	*Freq. Mc/s*	*Distance in Miles*	*Calculated Value of Field Intensity—Db above 1 μv/m* Bullington Nomograms	Fig. 10	*Percentage of Time Measured Fields Indicated (Db above 1 μv/m) Are Exceeded* 1%	50%	90%	99%	99.9%	*Duration of Measurements*
RCA Review June 1946	Prince Rupert–Massett	45	76	36	24.5	42.9	35.7	34.1	32.3	31.2	1 week
RCA Review June 1946	Prince Rupert–Langara Island	45	108	36	18.2	35.9	28.4	26.3	22.4	19.5	1 week
Frequenz Jan. 1952	Berlin–Harz	60	133	46.5	28	—	30.8	21.6	7.3	—	12 months
RCA Report F 43-118	Alpine–Riverhead	93.1	67	46	44	59	43	38	29	17	16 months
RCA Report F 43-118	Boston–Riverhead	92.9	127	−6.7	17.8	38	18.5	7	−2	—	22 months
RCA Report F 43-118	Boston–Hauppauge	92.9	150	−10.5	13.3	38	16	7	−2	—	15 months
RCA Reports F 43-106 F 42-121	New York–Riverhead	203.75	69.5	29.5	36	55	30	22	17	12	41 months
RCA Report EM 61-17	Empire State–Riverhead	288	70	28.2	25.4	48	24.5	22	18	—	3 months
PIRE Nov. 1951 (Gerks)	100 Mile Smooth Earth Path −1 KW Radiated	400	100	−57.5	8.8	44	5	−9	−20	—	11 months

for varying frequencies beyond which the nomograms should not be used. Figure 14-4 illustrates a curve which has been produced from figure 14-15, assuming that the antenna heights are restricted to the values given in Scale A of the figure. It is now of interest to see which of the paths listed in table 14-2 could be analyzed by using the Bullington nomograms when the limitations of figure 14-4 are taken into consideration. The actual distance beyond the radio optical path length for each case, and the limitations imposed by figure 14-4, are listed in table 14-3.

The radio optical distance in each case has been calculated using the equation

$$D_0 = \sqrt{2h_r} + \sqrt{2h_t} \qquad (14\text{-}1)$$

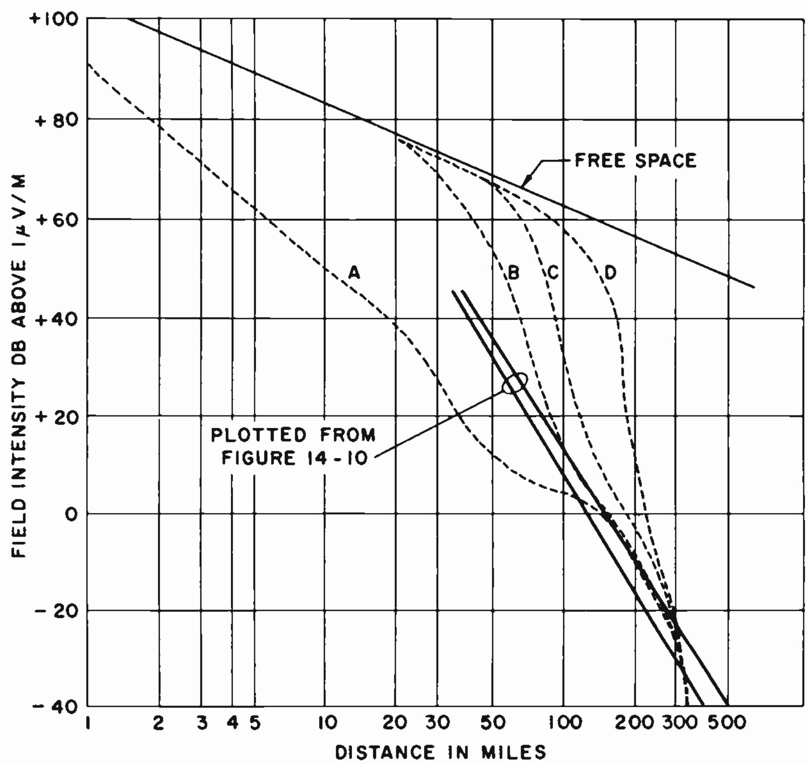
where

D_0 = Radio optical distance in miles (profile drawn on an earth's radius 4/3 times true radius).
h_r = Receiving antenna height in feet.
h_t = Transmitting antenna height in feet.

Formula 14-1 may be found as nomogram J in Appendix I.

An alternative method is to use figure 14-17 with the antenna heights at each end taken above an average smooth earth. The addition of the two distances to the horizon gives the radio optical horizon.

Examination of table 14-3 indicates that circuits 1, 4, and 8 are the only ones for which figure 14-14 clearly indicates permissible use of the Bullington nomograms. In these three cases it will be seen by referring to table 14-2 that the nomograms give predicted values within 3.7 db of the measured median values. For circuits 2 and 3, the actual distance beyond the radio optical horizon coincides with the



CURVES A, B, C AND D, SHOW MEDIAN VALUES OF FIELD STRENGTH FOR VARIOUS RECEIVING ANTENNA HEIGHTS. FREQUENCY 410 MC/S. EFFECTIVE RADIATED POWER 1 KW. TRANSMITTING ANTENNA HEIGHT 40 FEET.

A - EXPERIMENTAL CURVE - RECEIVING ANTENNA 10 FEET
B - " " " " 1,800 "
C - " " " " 4,500 "
D - " " " " 15,000 "

Figure 14-3. Curves by Gerks

limiting values given by figure 14-4, and the predicted values, using the nomograms, are seen from table 14-2 to be 7.6 db in excess of the measured median values. The predicted values could, therefore, be classed as unreliable. Of the other 4 circuits, the nomogram predictions for numbers 5, 6, and 9 are seen to be well at variance with the measured median values, while for circuit 7, the nomogram prediction is within 1.2 db of the median value.

It is felt that the above analysis justifies the use of figure 14-4 as an indication of the limitations of the Bullington nomograms. However, where the actual distance beyond the radio optical horizon is within about 10 miles of the limits set by the figure, it is considered advisable to compute the path losses on the basis of the nomograms and figure 14-20, and select the lower value given as the expected median value.

All the above information on paths extending beyond the horizon indicates that predicted values should be considered as median values. The question now arises as to what fading margin should be allowed to bring such values up to the level to be expected 90% of the time.

An examination of table 14-2 shows that the difference between the measured fields received 50% of the time and 90% of the time vary in a rather random manner. If these differences are plotted against frequency, as shown in figure 14-5, an arbitrary curve could be drawn through the points as shown. It is recognized, however, that there is insufficient data to justify the shape of this curve, but it is felt that it provides a reasonably adequate indication of the 90% fading range to be expected.

Summarizing the situation for paths extending beyond the radio optical range, the above evidence indicates that field intensities for 90% of the time can be predicted to within 5 db by adopting the following procedure:

1. Check the smooth earth radio range for the particular antenna heights involved (figure 14-17).

TABLE 14-3

COMPARISON OF DISTANCE BEYOND RADIO OPTICAL PATH LENGTH AND LIMITING VALUES FROM FIGURE 14-4

Circuit No.	*Hop*	*Freq. MC/S*	*Total Path Length*	*Radio Optical Path Length*	*Actual Distance Beyond Radio Optical*	*Limiting Value from Figure 14*
1	Prince Rupert–Massett	45	76 miles	63 miles	13 miles	31 miles
2	Prince Rupert–Langara Island	45	108	77	31	31
3	Berlin–Harz	60	133	105	28	28
4	Alpine–Riverhead	93.1	67	52	15	24
5	Boston–Riverhead	92.9	127	39	88	24
6	Boston–Hauppauge	92.9	150	49	101	24
7	New York–Riverhead	203.75	69.4	47.7	21.7	18
8	Empire State–Riverhead	288	70	61.8	8.2	16
9	100 Mile Smooth Earth Path-(Gerks)	400	100	13.4	86.6	14

2. If the distance of the actual circuit beyond the smooth earth radio range is greater than the limits set by figure 14-14, use figures 14-10 and 14-20 to obtain the median path loss. Increase this path loss by the fading margin taken from figure 14-15 to obtain the value expected 90% of the time.

3. If the distance of the actual circuit beyond the smooth earth radio range is within 10 miles of the limits set by figure 14-4, calculate the expected median path loss from the Bullington nomograms and also from figures 14-10 and 14-20. Select the most conservative value of path loss so given and increase by the fading margin taken from figure 14-5 in order to obtain the 90% value.

4. If the actual circuit exceeds the smooth earth radio range by a lesser amount than indicated in step 3 above, use the Bullington nomograms to assess the path loss. Consider the value given as a median value and increase by the fading margin taken from figure 14-5 in order to obtain the 90% value.

CONCLUSIONS. It is believed that the above analysis of long-term propagation measurements has shown that VHF and UHF path losses can be predicted with sufficient reliability to engineer radio relay systems for satisfactory operation. A summary of the recommended procedures and their application to a series of typical profiles is included in the following portion of this chapter.

For optical paths, and for paths that would be optical but for an intervening obstruction, it is con-

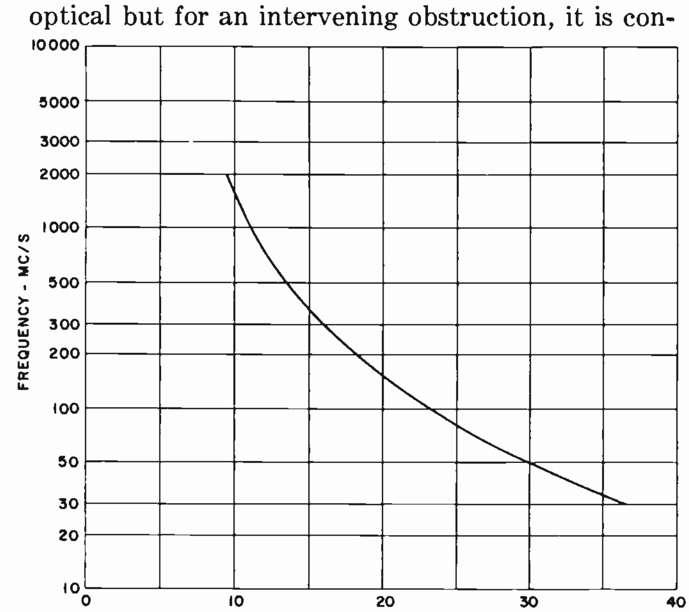


Figure 14-4. Limitations of Bullington Nomograms Beyond the Radio Horizon

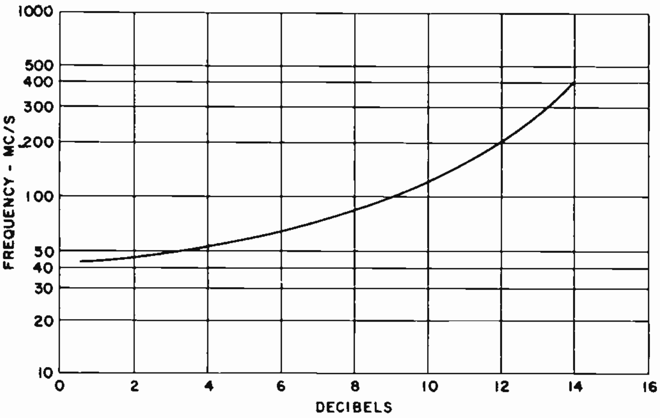


Figure 14-5. Fading Margin

sidered that procedures outlined will give a field intensity which is within 5 db of the value exceeded for 90% of the time over a long term period.

For paths beyond the optical range, at frequencies up to 500 mc, it is considered that an equivalent degree of reliability can be expected in the prediction of actual median field intensities. The assessment of the fading range to convert this predicted median value to a value expected over a long term period for 90% of the time cannot, however, be made with as much certainty. For this reason, the arbitrary curve shown in figure 14-5 has been drawn in a conservative manner through the practical points plotted. Further long term studies will be necessary before a high degree of reliability can be assured in the assessment of this fading range. In the meantime, however, the allowances which figure 14-5 indicate are considered to be adequate for use in systems engineering.

One point on which very little information has been found is the performance of circuits where both terminals and the intervening terrain are at altitudes above 10,000 to 20,000 feet. Under these circumstances, it is questionable whether the concept of a standard atmosphere (which is the basis for an earth's radius of 4/3 the true radius) is still applicable. In addition, there is the question of a possible "Obstacle Gain" on non-optical paths. Until more information is available on these points therefore, it is suggested that profiles for such circuits be drawn on a true earth's radius and predicted accordingly.

Paths such as those mentioned in the last paragraph do not often occur, while paths containing a clear line-of-sight, or partially obstructed line-of-sight, are most commonly encountered. In order to determine the suitability of a radio path under such conditions, the following steps are arranged to determine, first, whether the path is clear optical, obstructed optical, or non-optical. Having determined this, the next step is to determine whether or not the path attenuation under the condition found is sufficiently low to permit use of the sites. It must be borne in mind that equipments will usually be in kit form, that is, the relay systems will not be supplied with a variety of antenna units, power supplies, and so on, in order to utilize particular sites. If, with the available equipment, a site will give satisfactory path attenuation and reliability, then, other things being equal, the site is practical. Otherwise, a site must be chosen that will conform with these requirements. In order to explain the use of these steps and nomograms, the contour map is used to show two possible sites for relay stations, with a third site which is already established by the location of other relay sites (not shown) along the system.

SITING EXAMPLE

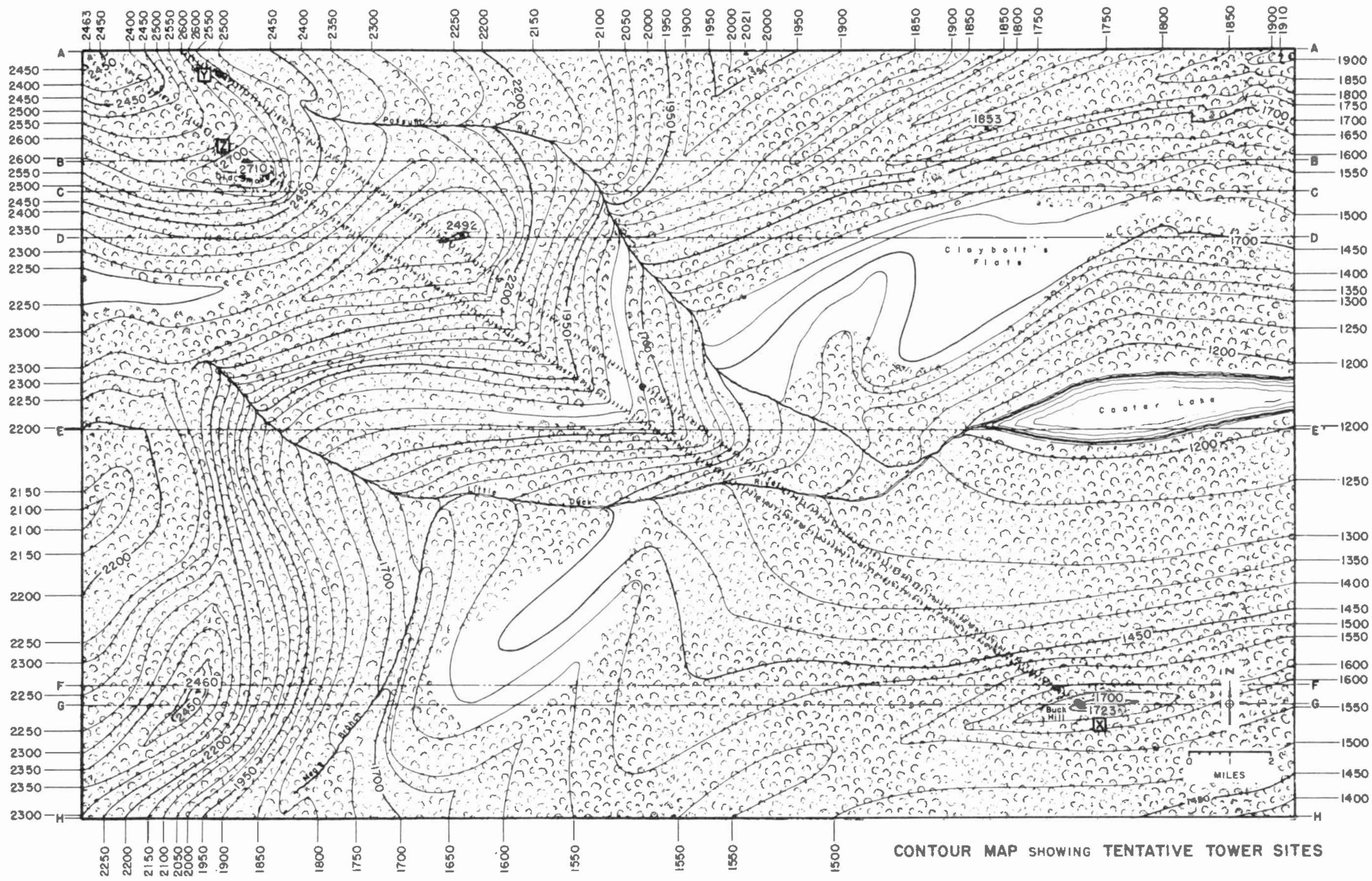
In order to show the relationship between the area in which radio relay stations are to be installed, and the actual reception conditions to be expected, the example starts with the contour map illustrated in figure 14-6. This map shows rolling country in which site X has already been established, and either site Y or site Z yet to be chosen. In regard to a readily obtainable source of power, site Y is considered as the better of the two. However, the path Y-X has the disadvantage of an intervening rise of ground about seven miles southeast of site Y, which may cause excessive shadow loss. On the other hand, site Z, which has an obviously clear line of site on the path Z-X, requires an additional two miles of power line construction.

CONSTRUCTION AND USE OF PROFILES. Clearance above hills, trees, or buildings should be determined by plotting the profile of the land below the transmission path, using standard 4/3 earth profile paper.

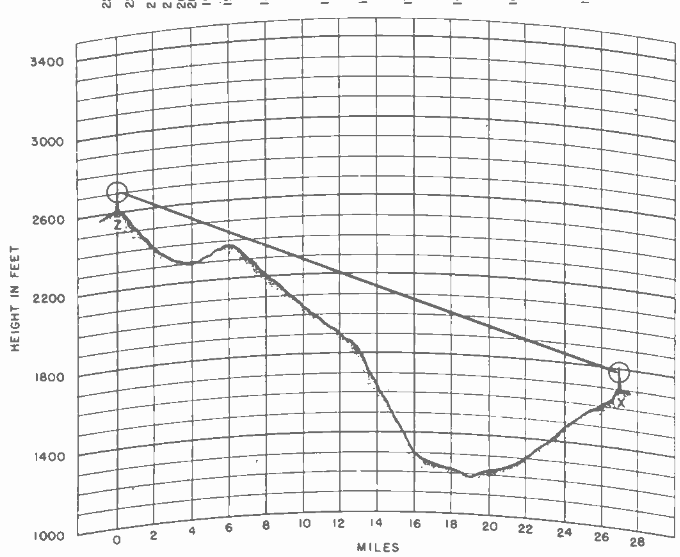
NOTE

> Unless the path has been purposely cleared, tree heights must be taken into consideration, either by plotting the profile along the tree tops, or by adding the average tree height to the bare earth profile.

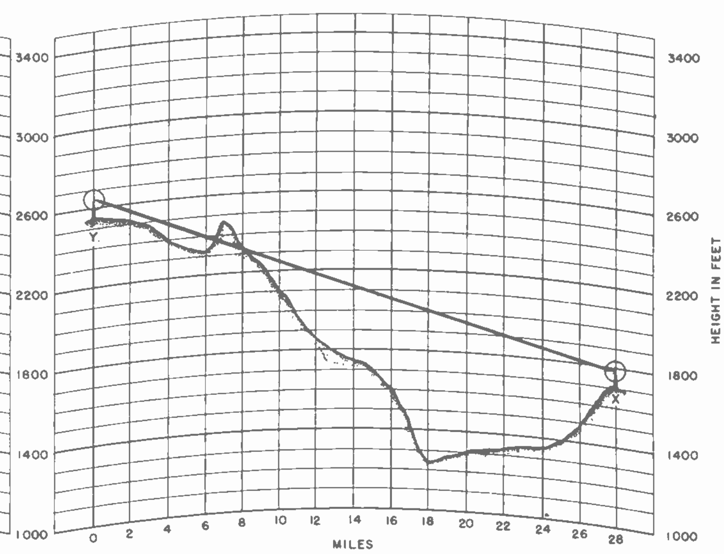
Since the profile is drawn on 4/3 earth profile paper, the radio beam must be a straight line between the antennas at the two sites. The distance between this line and the path profile will be the path clearance, or extent of obstruction, and may be read in feet on the elevation scale. General path classification may then be determined by referring to figure 14-7, which shows three types of paths, (1) clear line of sight,



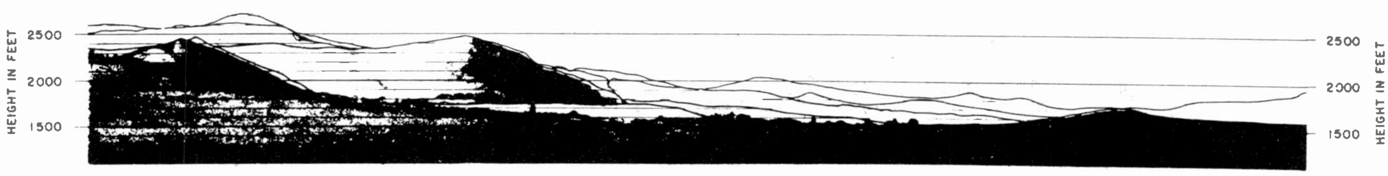
CONTOUR MAP SHOWING TENTATIVE TOWER SITES



LEG Z-X, CLEAR LINE OF SIGHT
SHOWN ON 4/3 EARTH PROFILE PAPER



LEG Y-X, OBSTRUCTED LINE OF SIGHT
SHOWN ON 4/3 EARTH PROFILE PAPER



VIEW OF AREA FROM THE SOUTH
DEVELOPED FROM PROFILES A THROUGH H

Figure 14-6. Contour Map and Profiles

(2) obstructed line of sight, and (3) non-optical. The first two classifications are self-explanatory, and the third simply refers to a path which extends beyond the radio horizon. These three general path classifications are further classified into individual types characterized by varying amounts of first Fresnel zone clearance in the first two cases, and by distance in the third (non-optical) case. Each type is illustrated in figures 14-8, 14-9, and 14-10, respectively.

Path Y-X, therefore, is classified as an obstructed line-of-sight condition, since the profile shows an obstruction extending one hundred feet above the line of sight. No allowance for trees is necessary in this case, as the contour map shows the obstruction to be a bare hill. Likewise, path Z-X is classified as a clear line of sight path.

PATH Y-X COMPUTATIONS. Since path Y-X has been classified as an obstructed line of sight, we may refer to figure 14-9 to determine the proper method of estimating path attenuation. In this case, the path would have first Fresnel zone clearance except for the

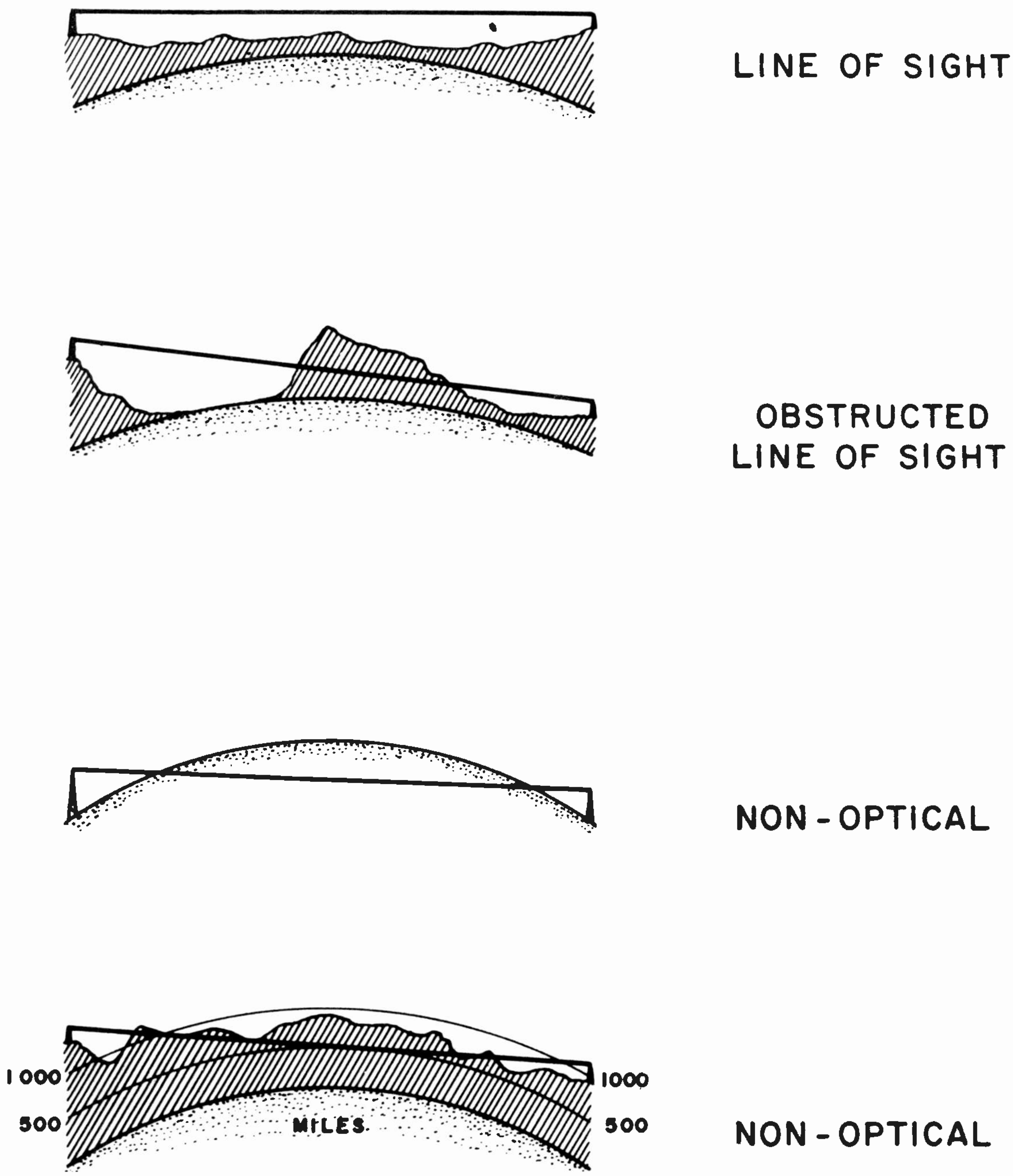


Figure 14-7. Types of Paths

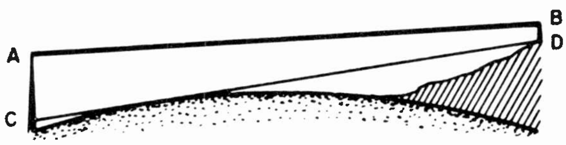
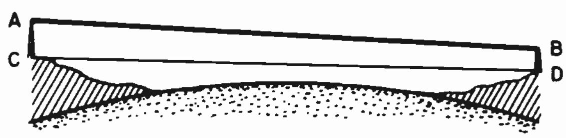
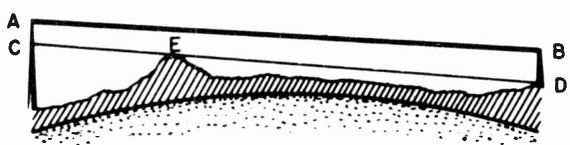
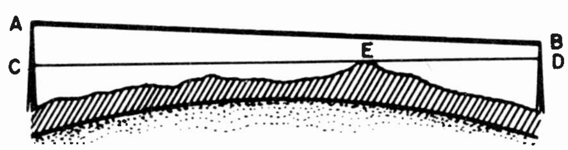
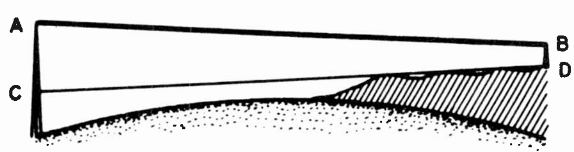
PROFILE	TREATMENT
 SMOOTH EARTH WITHOUT FIRST FRESNEL ZONE CLEARANCE	**A** DRAW TANGENT TO SMOOTH EARTH, SUCH THAT AC:BD = CE:ED. USE EFFECTIVE ANTENNA HEIGHTS AC AND BD WITH NOMOGRAM FIG. 14-14.
 FIRST FRESNEL ZONE CLEARANCE IN VICINITY ANTENNA A, BUT NOT B	**B** DRAW PLANE EARTH LINE CD. USE EFFECTIVE ANTENNA HEIGHTS AC AND BD WITH NOMOGRAM FIG.14-14.
 NO FIRST FRESNEL ZONE CLEARANCE IN VICINITY ANTENNAS A & B	**C** DRAW PLANE EARTH LINE CD. USE EFFECTIVE ANTENNA HEIGHTS AC AND BD WITH NOMOGRAM FIG.14-14.
 NO FIRST FRESNEL ZONE CLEARANCE AT E, NOR IN VICINITY ANTENNA B	**D** DRAW PLANE EARTH LINE CED. USE EFFECTIVE ANTENNA HEIGHTS AC AND BD WITH NOMOGRAM FIG.14-14.
 FIRST FRESNEL ZONE CLEARANCE IN VICINITY A & B, BUT NOT AT E	**E** DRAW PLANE EARTH LINE CED, SUCH THAT AC:BD = CE:ED. USE EFFECTIVE ANTENNA HEIGHTS AC AND BD WITH NOMOGRAM FIG.14-14.
 FIRST FRESNEL ZONE CLEARANCE IN FRONT ANTENNA A, BUT NOT B	**F** DRAW PLANE EARTH LINE CD. USE EFFECTIVE ANTENNA HEIGHTS AC AND BD WITH NOMOGRAM FIG 14-14.

Figure 14-8. Clear Radio Paths

obstruction, and therefore should be calculated as noted under path A of figure 14-9. Here there are two sources of attenuation: free-space attenuation, which occurs in all clear line of sight paths with first Fresnel zone clearance, and the shadow loss caused by the obstruction. As noted in the figure, the total path loss is found by adding the free-space loss obtained from figure 14-11, to the shadow loss found by use of figure 14-18. For path Y-X, figure 14-11 shows a free-space loss of 126 db (at a frequency of 1,800 mc, and path length of 28 miles). Figure 14-18 shows a shadow loss of 16 db (100 foot obstruction, 7 miles from an antenna, 28-mile path length), giving a total path attenuation of 142 db.

PATH Z-X COMPUTATIONS. Path Z-X has been classified as a clear line of sight, according to figure 14-7. The next step is to determine whether or not first Fresnel zone clearance exists over the entire path, by use of nomogram H or figure 14-18. If first Fresnel zone clearance does not exist, further classification is necessary, and figure 14-8 is used. In this case, however, adequate clearance is available, and the total path attenuation may be found from figure 14-11, a path loss of 124 db (at a frequency of 1,800 mc, and path length of 27 miles).

SYSTEM CALCULATIONS. The two possible transmission paths must now be compared in order to determine whether or not they will be acceptable, and which is the better as regards attenuation. Any peculiarities of the site that will effect signal attenuation should also be taken into account. Equipment specifications, taken from manufacturers' data, are as follows:

Equipment Data

Make and type of system	RCA CW-20A
Carrier frequency	1,800 mc
Transmitter power output	3 watts
Transmission line	RG-17/U (attenuation, 6.7 db per 100′)
Antenna type	6′ parabolic reflector
Receiver input	116 db below transmitter output (min.)

Site X Data

Antenna height	100 feet
Transmission line length	150 feet

Site Y Data

Antenna height	100 feet
Transmission line length	200 feet

Site Z Data

Antenna height	100 feet
Transmission line length	125 feet

COMPUTATIONS, LEG Y-X

The attenuation and gain found in various components must be determined in db values in order to properly compare the usefulness of the two sites in respect to signal attenuation. The power output of the transmitter, 3 watts, may be expressed in terms of db above one watt. This is done by the formula

$$\text{db} = 10 \text{ Log } P_1/P_2 \qquad (14\text{-}2)$$

where

P_1 = Transmitter output in watts
P_2 = Reference level of one watt.

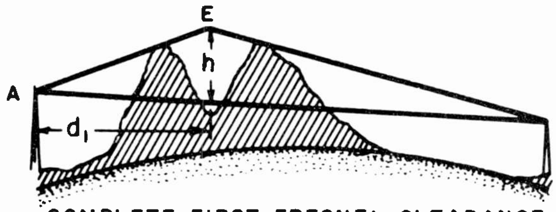
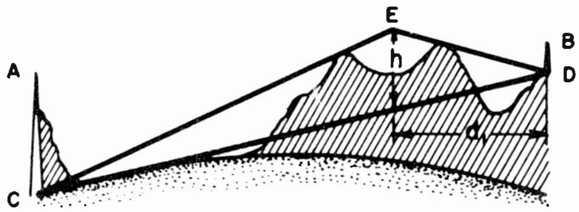
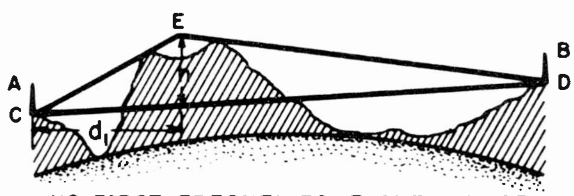
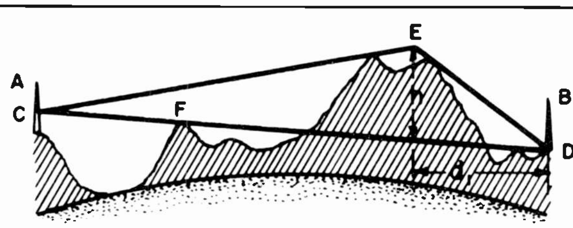
Thus the db output of the transmitter becomes 4.8 dbw, or 4.8 db above one watt.

Transmission line loss at site Y is determined by the attenuation of the transmission line in db per 100 feet. For RG-17/U, this attenuation is 6.7 db per 100 feet, giving a line loss of 13.4 db, which may also be expressed as a negative gain, −13.4 db. Power supplied to the antenna will then be the dbw of the transmitter, less the attenuation of the transmission line, or 4.8 dbw −13.4 db, or −8.6 dbw. This may also be stated as 8.6 db below one watt.

Antenna gain must be figured, since with any directional antenna, the effective radiated power in the beamed direction will be greater than the power in the same direction from a non-directional antenna. Nomogram K in Appendix I will give the antenna gain for a 6 foot parabolic antenna at 1,800 mc, as 26 db over a half-wave dipole. The power output from the antenna will then become −8.6 dbw + 26 dbw, or +17.4 dbw. Path attenuation, already determined, is 142 db. Power received at the site X antenna will then be the antenna output, 17.4 dbw, less the path attenuation, 142 db, or −124.6 dbw. The antenna gain, 26 db, added to this figure, gives the power supplied to the transmission line, or −98.6 dbw. Transmission line loss between the site X antenna and receiver is 10 db (for 150 feet of RG-17/U line), and the power delivered to the receiver at site X becomes −108.6 dbw. Although this value is calculated with site Y as the transmitting end, the same attenuation will occur in signals transmitted from site X to site Y.

COMPUTATIONS FOR Z-X LEGS

The calculations for this leg of the system are figured exactly as for the Y-X leg, and will be shown in tabular form. Attenuation is indicated by a minus sign, gain by a plus sign. For the sake of comparison, the Y-X computations are also shown.

PROFILE	TREATMENT
 COMPLETE FIRST FRESNEL CLEARANCE EXCEPT FOR OBSTRUCTION	**A** USE NOMOGRAM FIG. 14-11 TO COMPUTE FREE SPACE LOSS. INCREASE BY SHADOW LOSS FROM NOMOGRAM FIG. 14-18, USING DISTANCE d_1 AND HEIGHT h.
 FIRST FRESNEL ZONE CLEARANCE IN VICINITY ANTENNA A, BUT NOT B	**B** USE EFFECTIVE ANTENNA HEIGHTS AC AND BD WITH NOMOGRAM FIG. 14-14, TO COMPUTE PLANE EARTH LOSS. INCREASE BY SHADOW LOSS FROM NOMOGRAM FIG. 14-19, USING DISTANCE d_1 AND HEIGHT h.
 NO FIRST FRESNEL ZONE CLEARANCE IN VICINITY ANTENNAS A AND B	**C** USE EFFECTIVE ANTENNA HEIGHTS AC AND BD WITH NOMOGRAM FIG. 14-14, TO COMPUTE PLANE EARTH LOSS. INCREASE BY SHADOW LOSS FROM NOMOGRAM FIG. 14-19, USING DISTANCE d_1 AND HEIGHT h.
 NO FIRST FRESNEL ZONE CLEARANCE AT F, NOR IN VICINITY ANTENNA B	**D** USE EFFECTIVE ANTENNA HEIGHTS AC AND BD WITH NOMOGRAM FIG. 14-14, TO COMPUTE PLANE EARTH LOSS. INCREASE BY SHADOW LOSS FROM NOMOGRAM FIG. 14-19, USING DISTANCE d_1 AND HEIGHT h.
 FIRST FRESNEL ZONE CLEARANCE IN VICINITY ANTENNA A, BUT NOT B	**E** USE EFFECTIVE ANTENNA HEIGHTS AC AND BD WITH NOMOGRAM FIG. 14-14, TO COMPUTE PLANE EARTH LOSS. INCREASE BY SHADOW LOSS FROM NOMOGRAM FIG. 14-19, USING DISTANCE d_1 AND HEIGHT h.

NOTE: RADIO PATHS NOT SHOWN.

Figure 14-9. Obstructed Radio Paths

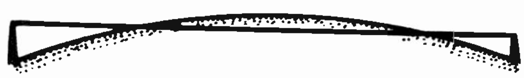
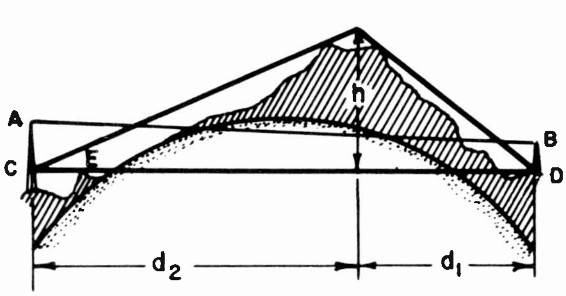
PROFILE	TREATMENT
 SMOOTH EARTH	**A** USE BULLINGTON NOMOGRAMS FIGURES 14 - 11, 16 AND 17, TO OBTAIN THE SMOOTH EARTH PATH LOSS.
 IRREGULAR TERRAIN	**B** OBTAIN THE SMOOTH EARTH PATH LOSS FROM NOMOGRAMS FIGURES 14 - 11, 16 AND 17, AND INCREASE BY THE SHADOW LOSS GIVEN BY FIGURE 14 - 19. TO OBTAIN THE DISTANCE d_1 AND HEIGHT h FOR USE WITH FIGURE 14 - 19, PLOT THE PROFILE ON RECTANGULAR CO-ORDINATE GRAPH PAPER.
	**C** DRAW PLANE EARTH LINE CED. USE EFFECTIVE ANTENNA HEIGHTS AC AND BD WITH NOMOGRAM FIGURE 14 - 14, TO COMPUTE PLANE EARTH PATH LOSS. INCREASE BY SHADOW LOSS GIVEN BY FIGURE 14 - 19, USING DISTANCE d_1 AND HEIGHT h.

Figure 14-10. Non-Optical Paths

Item	*Z-X Leg*	*Y-X Leg*
Transmitter output.....	+4.8 dbw	+4.8 dbw
Transmitter line loss....	−8.4 db	−13.4 db
Antenna gain (xmtr)....	+26.0 db	+26.0 db
Path attenuation.......	−124.0 db	−142.0 db
Antenna gain (at X)....	+26.0 db	+26.0 db
Transmission line loss (at X)...............	−10.0 db	−10.0 db
Power Level at the Receiver...............	−85.6 dbw	−108.6 dbw

These are the values to be expected at least 90% of the time; that is, the signal strength will have these values, or better, 90% of the time. Since the receiver is capable of amplifying the signal, this figure is conservative, and a greater attenuation may be received and satisfactory reception will still result.

The CW-20A system requires a path attenuation of not more than 116 db between transmitter and receiver, as given in the specifications for the system. Both paths, then, are entirely acceptable, since both have considerably less attenuation than the system will accept. The choice between these two sites would not, therefore, be determined by the received signal strength, but by accessibility. However, if the accessibility of both sites is approximately equal, the link with the least path attenuation should be used.

If the strength at the receiver were required in microwatts, instead of in dbw, the dbw figure may be converted to power in microwatts as follows:

$$\text{Power level (in dbw)} = 10 \text{ Log } P_1/P_2 \quad (14\text{-}3)$$

where

$P_2 = 1$ watt.

Since the power for the Y-X leg is 108.6 dbw, the power level referred to watts is

$$-108.6 = 10 \text{ Log } P_1$$

and $P_1 = 7.3 \times 10^{-10}$ watts.

Referenced to microwatts, this value becomes $7.3 \times 10^{-10} \times 10^{6}$, or 7.3×10^{-4} microwatts input to the receiver.

GENERAL SUMMARY

1. Draw the profile for the path concerned on an earth's radius 4/3 times the true radius. If suitable profile paper is not available a 4/3 earth's radius curve can quickly be drawn on rectangular coordinate graph paper using the following formula:

$$h = 0.5\, d_1 d_2 \quad (14\text{-}4)$$

where

h = height in feet under the curve at distance d_1 miles from the initial end, and d_2 miles from the far end.

If for any reason it is desired to draw a profile on a true earth's radius the formula becomes $h = 0.667 d_1 d_2$.

2. Check whether the path can be classed as line of sight, obstructed line of sight, or non-optical, as illustrated in figure 14-7. For the first three profiles, the classification is obvious, for the fourth profile, the line from A to B does not intersect the sea level earth's curvature. However, all the terrain is above the 500-foot level, and this would therefore be considered as the smooth earth line. Since AB intersects this 500-foot line, the path would be classed as non-optical.

Line of Sight Paths

1. For line of sight paths check whether first Fresnel zone clearance exists. This may be found by using nomogram H in Appendix I.

2. If first Fresnel zone clearance exists, obtain the path loss from Bullington nomogram figure 14-11 for free-space propagation when 1 watt is radiated between half wave dipoles. Correct for actual transmitter power and antenna gains proposed, and consider this the value to be expected 90% of the time.

3. If first Fresnel zone clearance does not exist draw a plane earth line CD, as indicated in the typical cases shown in figure 14-7 and obtain the plane earth loss for one watt radiated between half wave dipoles from Bullington nomogram figure 14-14 using antenna heights AC and BD. Correct for actual transmitter power and net antenna gains proposed, and consider this path loss as the value to be expected 90% of the time.

Obstructed Line of Sight Paths

1. For obstructed line of sight paths ignore the obstruction initially, and draw a plane earth line CD in the same manner as in paragraph 2.3 above. Use the base CD to form a shadow triangle CDE as indicated in the typical cases shown in figure 14-9.

   Use Bullington nomograms figures 14-14 and 14-19, or 14-11 and 14-18, as indicated in the typical cases in figure 14-8 to obtain a value of path loss for one watt radiated between half wave dipoles. Correct for actual transmitter power and net antenna gains proposed, and consider this path loss as the value to be expected 90% of the time.

Non-Optical Paths

1. For non-optical paths check the value of the radio optical range given when using the effective antenna heights above a smooth earth. As indicated previously, smooth earth is not necessarily taken at sea level. The optical range for 4/3 earth's radius is given by the formula (see nomogram J in Appendix I):

$$d_0 = \sqrt{2h_1} + \sqrt{2h_2} \qquad (14\text{-}5)$$

where

d_0 = optical range in miles

h_1 and h_2 = effective antenna heights in feet.

Alternatively, the radio optical range can be obtained by adding the distances from the top of each antenna to the smooth earth horizon as given by figure 14-17.

2. If the length of the path beyond the optical range is greater than the limits set by figure 14-4, compute the path loss for one watt radiated between half wave dipoles using figures 14-11 and 14-20. Correct for actual transmitter power and net antenna gains proposed, and consider this path loss as the median value to be expected.

   If it is desired to assess the value to be expected 90% of the time, increase the path loss by an amount indicated by figure 14-5.

3. If the distance of the actual circuit beyond the optical range is within 10 miles of the limits set by figure 14-4, calculate the path loss for one watt radiated between half wave dipoles, both from the Bullington nomograms as indicated in figure 14-10 and from figures 14-11 and 14-20. Correct the most conservative value so given for actual transmitter power and net antenna gains proposed, and consider this path loss as the median value to be expected.

   If it is desired to assess the value to be expected 90% of the time, increase the median value by an amount indicated by figure 14-5.

4. If the actual circuit exceeds the optical range by a lesser amount than indicated in step #3 above, calculate the path loss for one watt radiated between half wave dipoles from the Bullington nomograms as indicated in figure 14-10. Correct for actual transmitter power and net antenna gains proposed, and consider this path loss as the median value to be expected.

   If it is desired to assess the value to be expected 90% of the time, increase the median path loss by an amount indicated by figure 14-5.

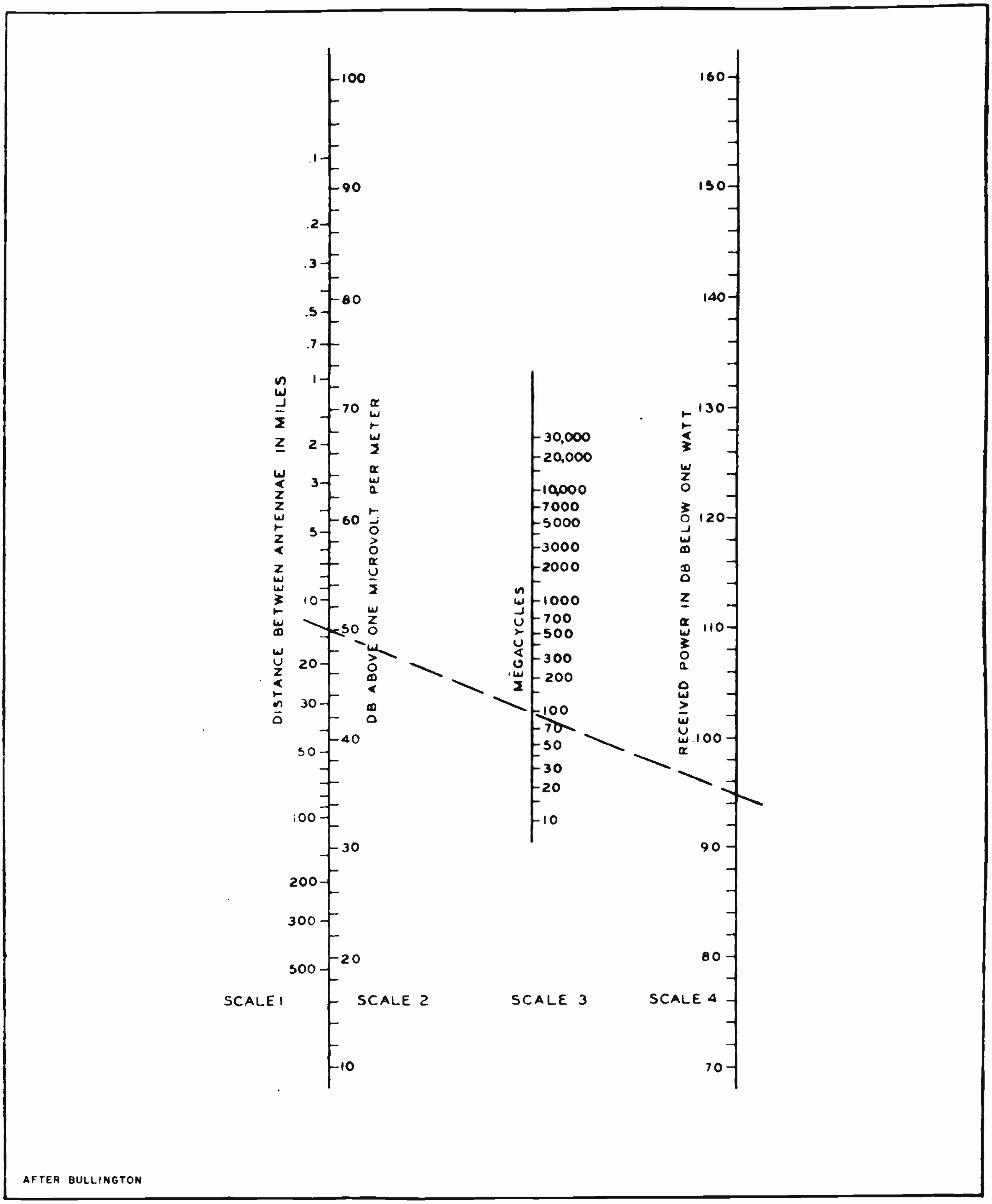


Figure 14-11. Free Space Field Intensity and Received Power between Half Wave Dipoles, 1 Watt Radiated

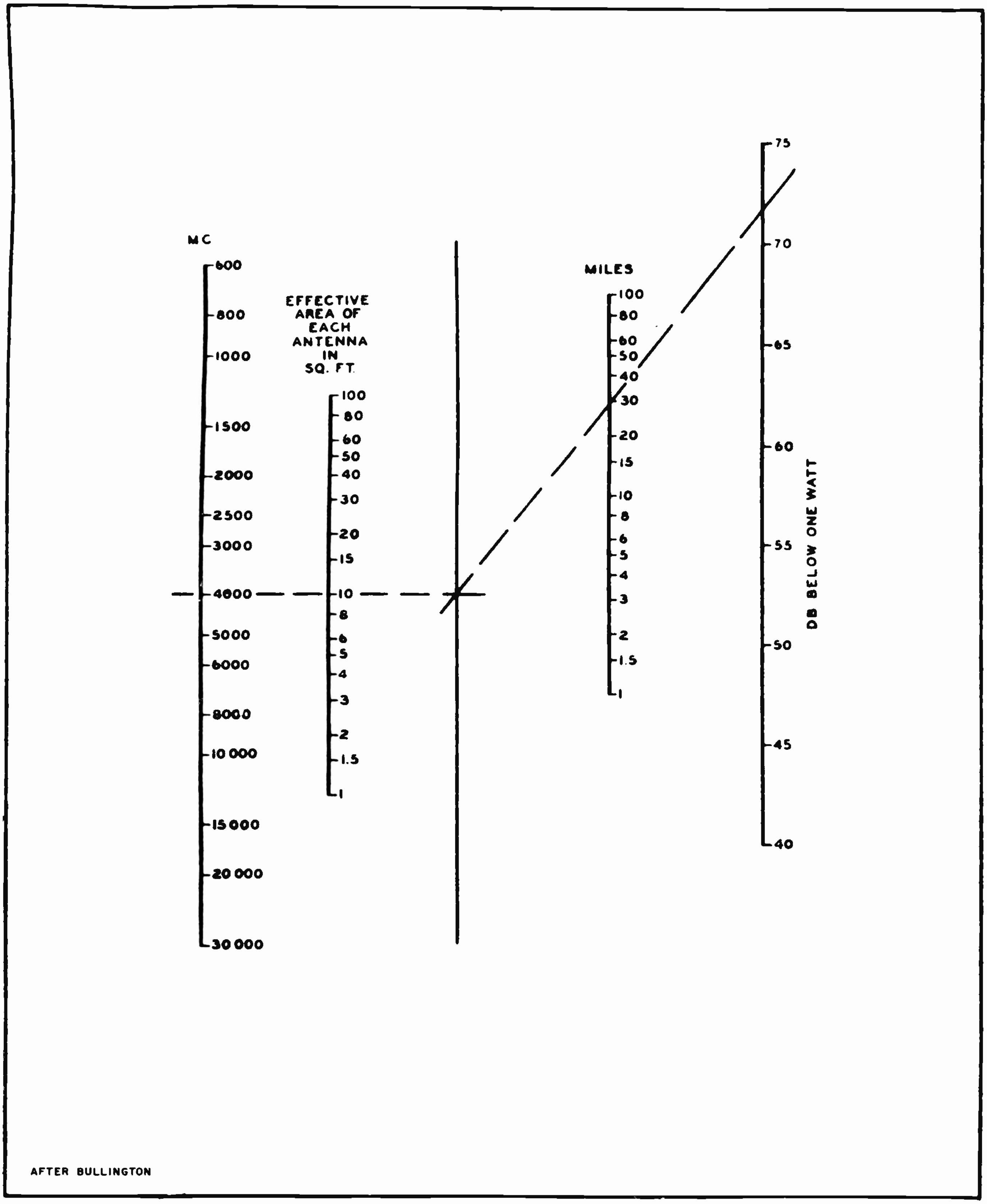


Figure 14-12. Received Power in Free Space between Two Antennas of Equal Effective Areas, 1 Watt Radiated

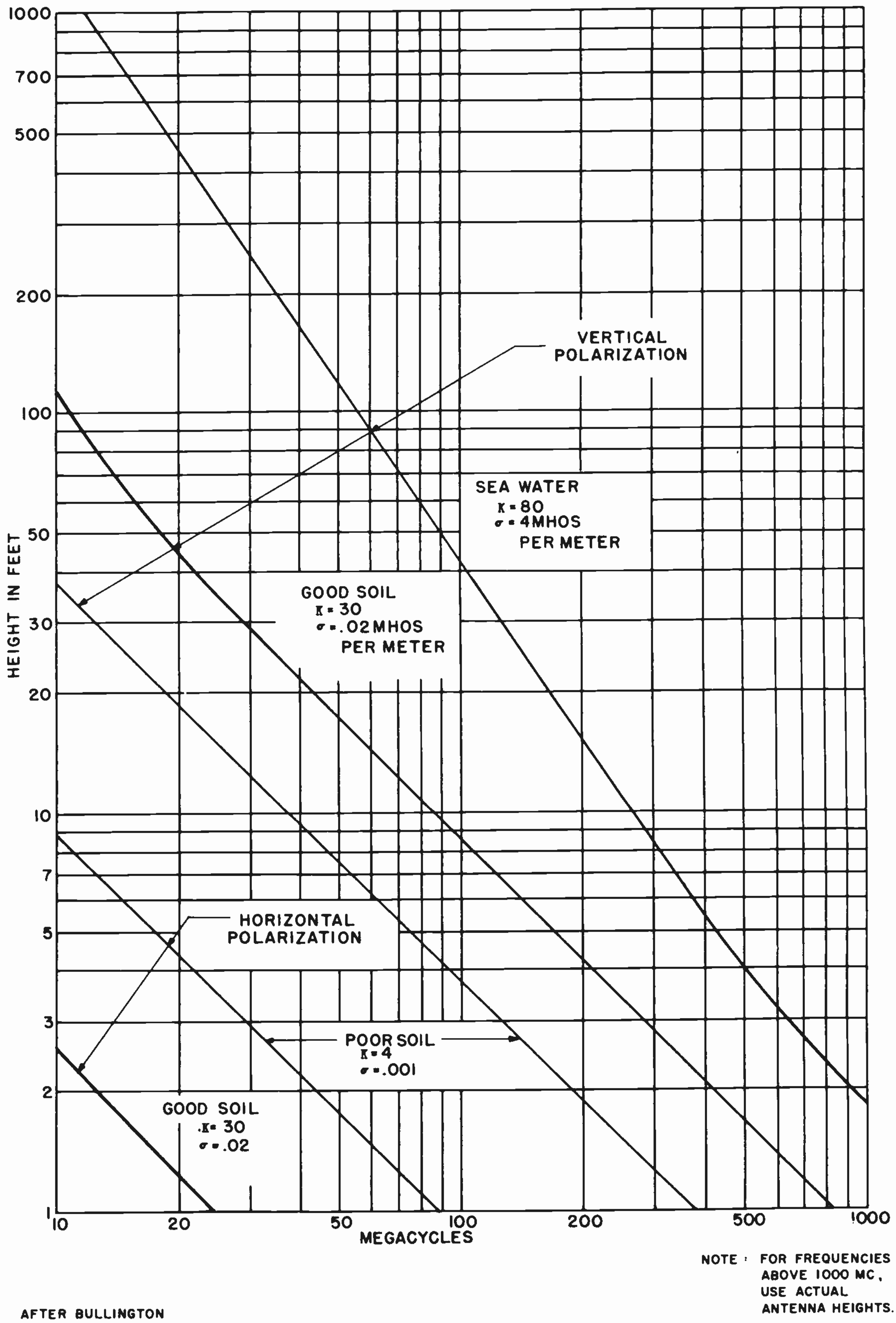


Figure 14-13. Minimum Effective Antenna Height

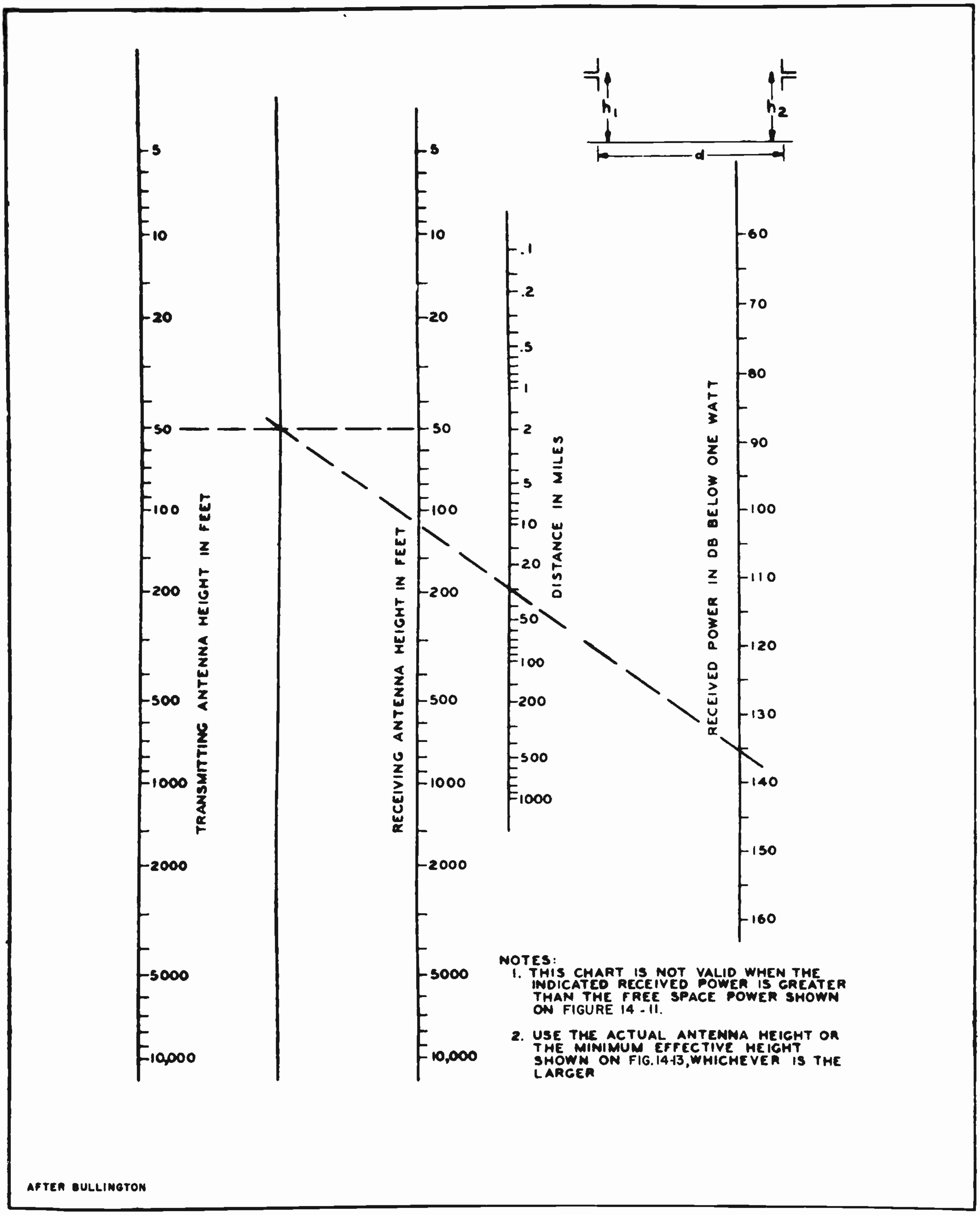


Figure 14-14. Received Power over Plane Earth Between Half-Wave Dipoles, 1 Watt Radiated

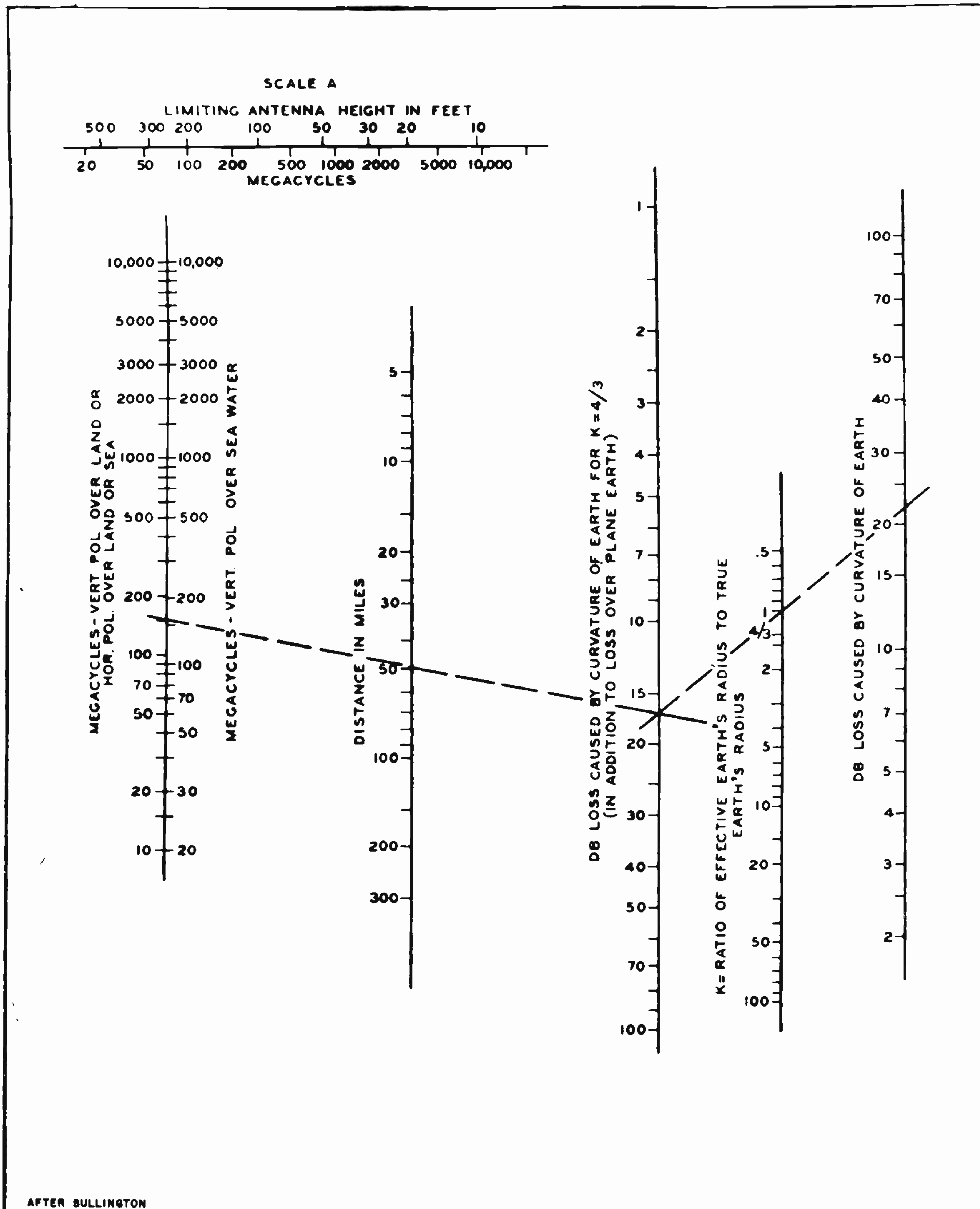


Figure 14-15. Diffraction Loss Caused by Curvature of the Earth Assuming Neither Antenna Height is Higher than Shown on Scale A

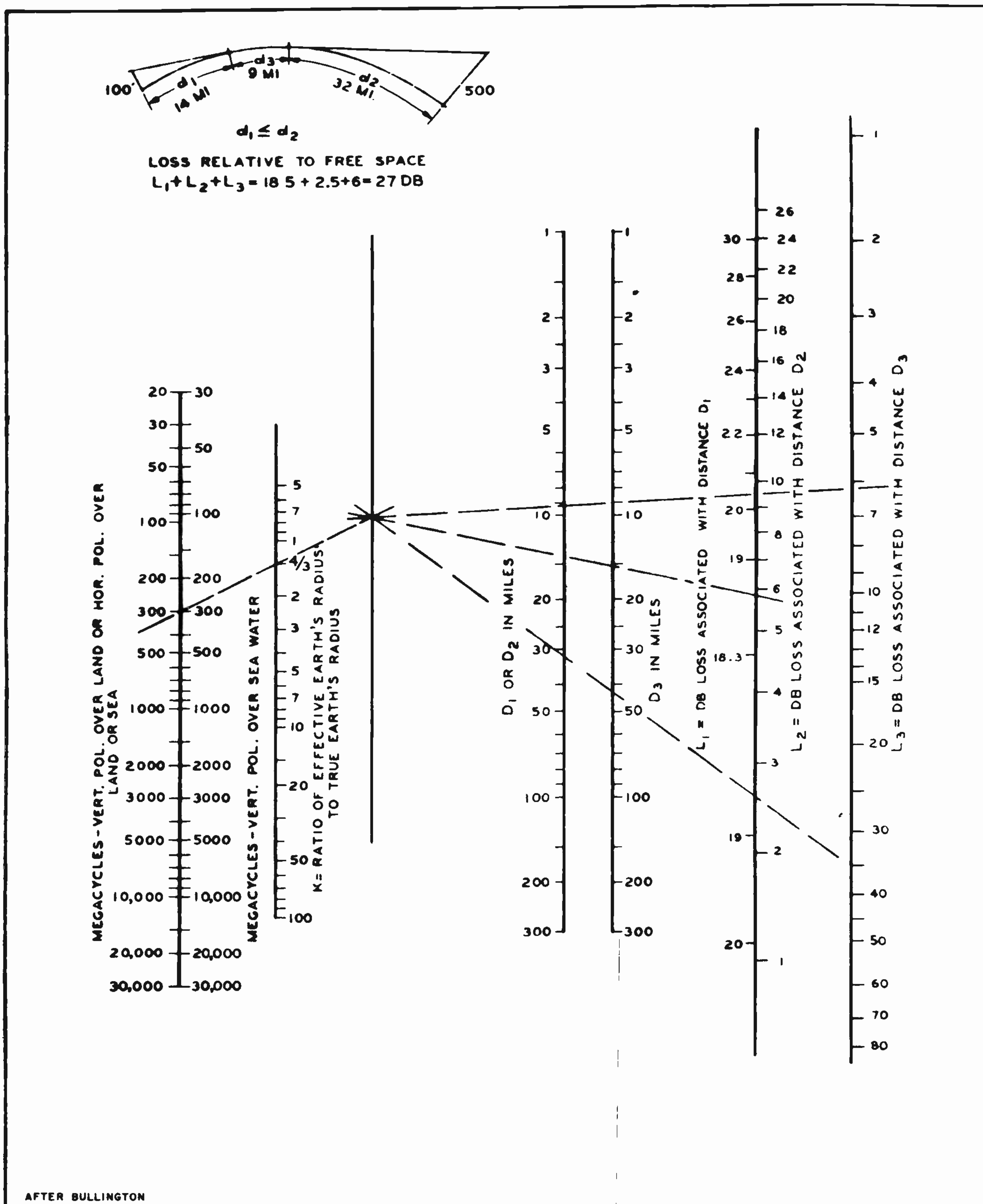


Figure 14-16. Db Loss Relative to Free Space Transmission at Points Beyond Line of Light over a Smooth Earth

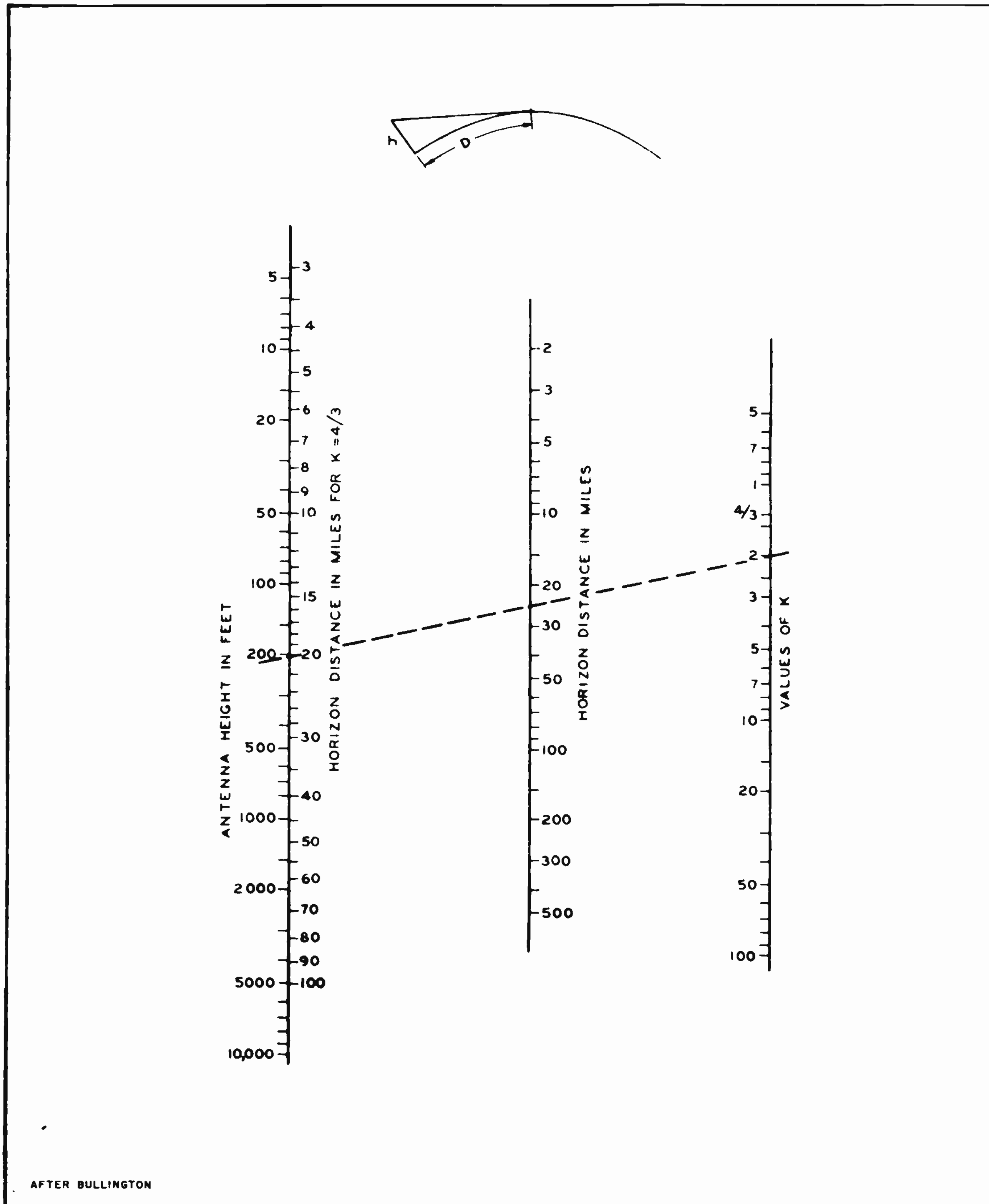


Figure 14-17. Distance to Horizon

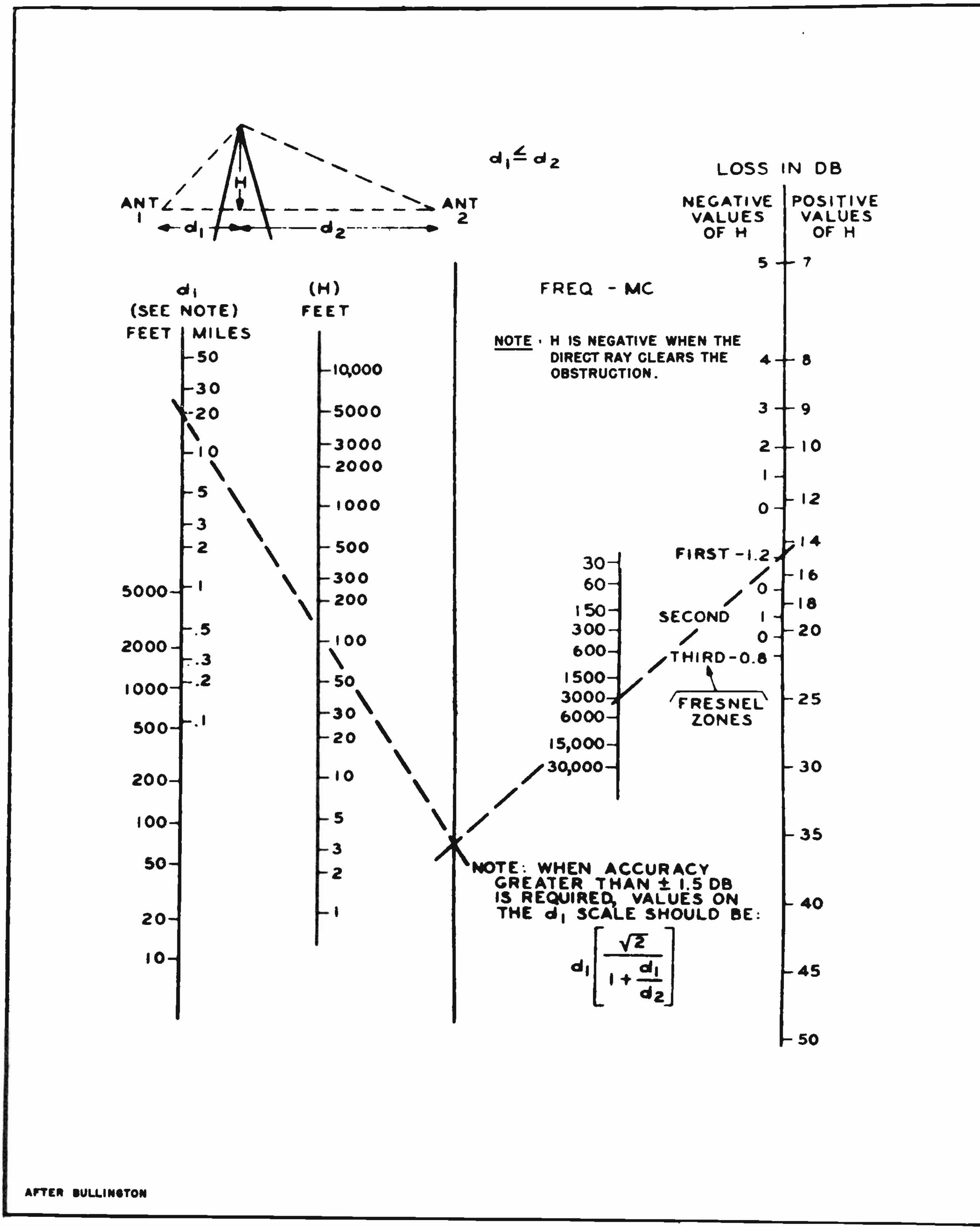


Figure 14-18. Shadow Loss Relative to Free Space

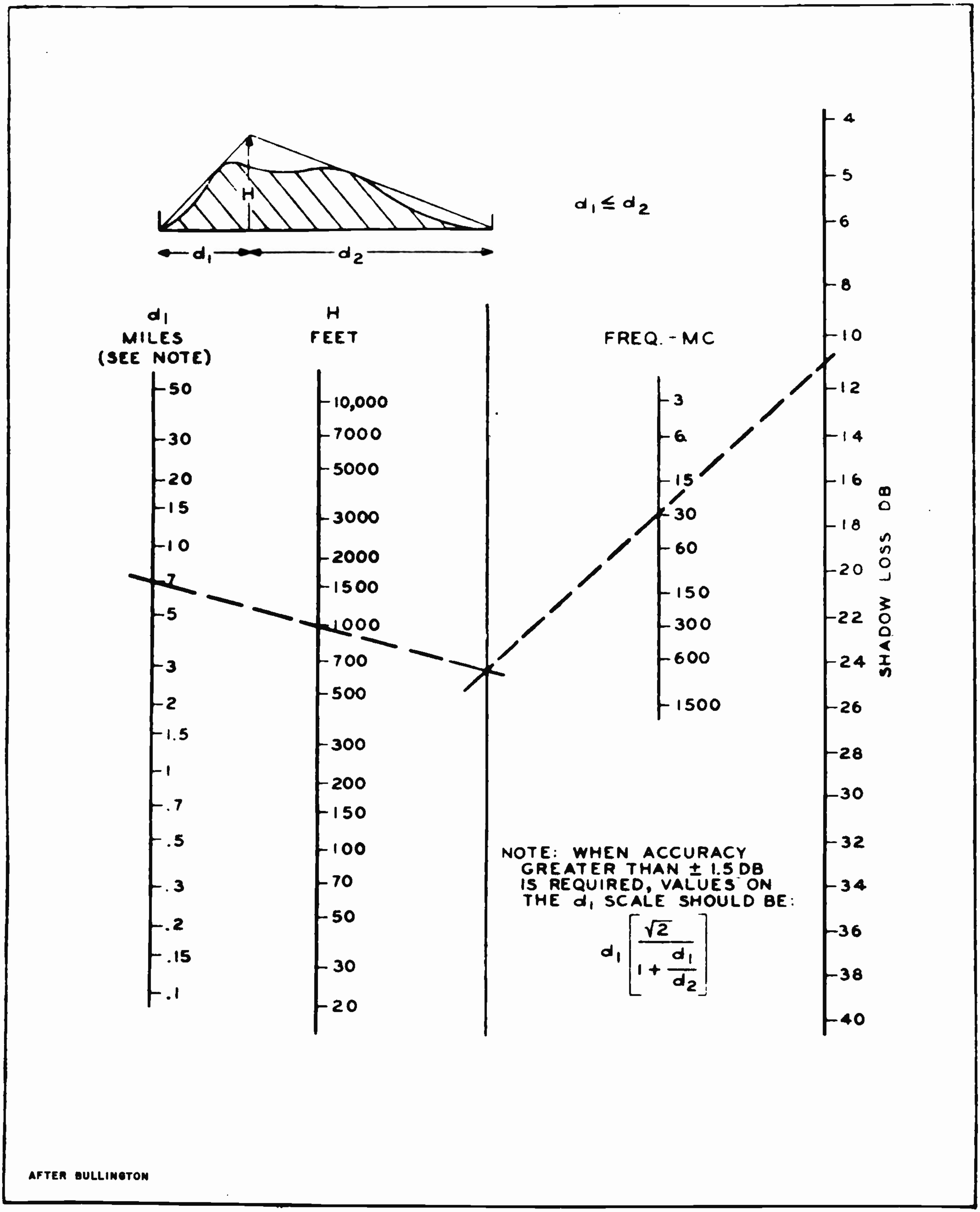


Figure 14-19. Shadow Loss Relative to Smooth Earth

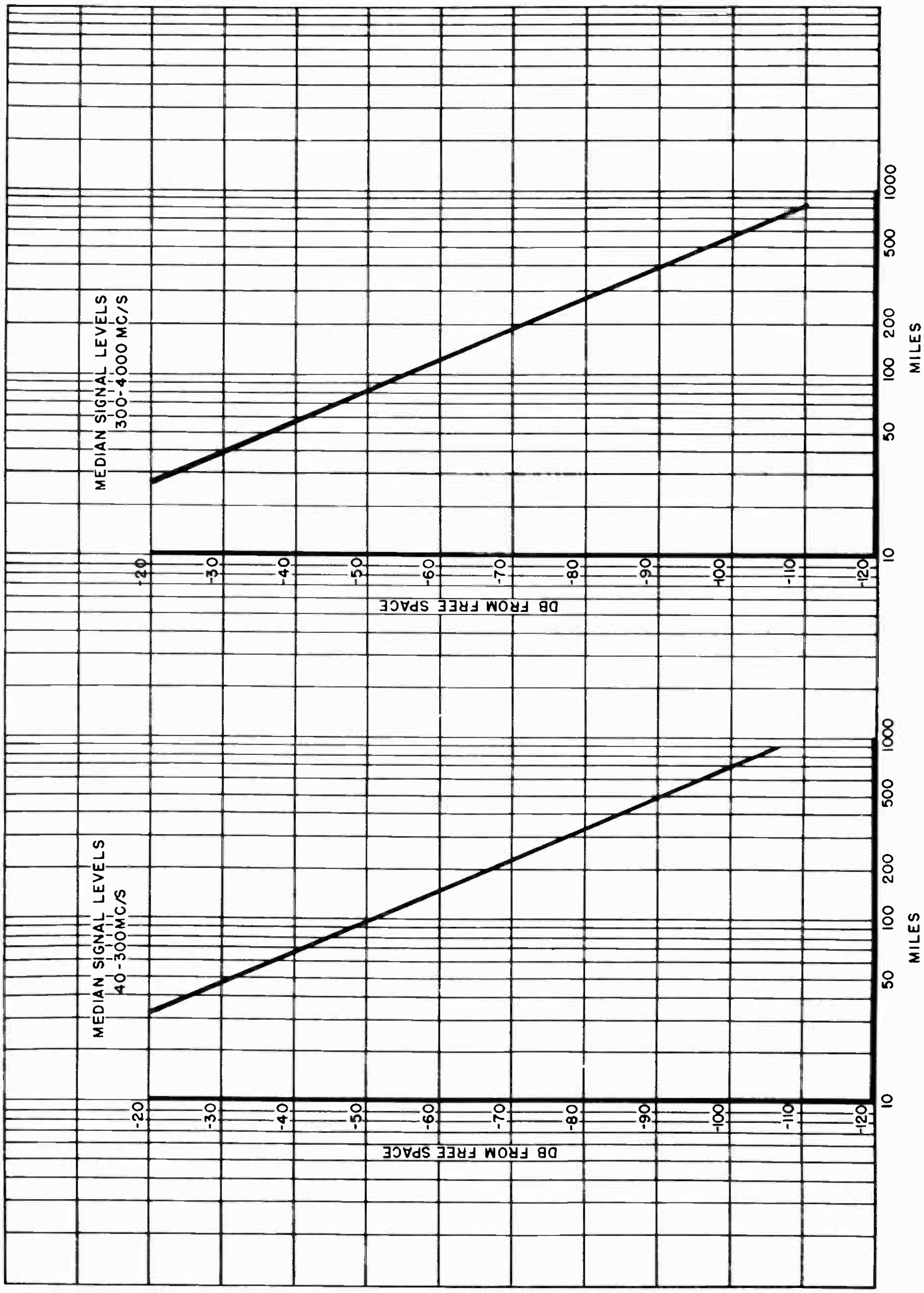


Figure 14-20. Median Signal Levels Relative to Free Space

Figure 14-20. Median Signal Levels Relative to Free Space

CHAPTER 15

IONOSPHERIC TRANSMISSION (30 TO 100 MC)

INTRODUCTION

As mentioned in Chapter 1, an electromagnetic wave may be propagated in the form of three different waves, the *ground wave*, the *ionospheric wave* or sky wave, and the *tropospheric wave*. Waves traveling through the troposphere (ground waves and tropospheric waves) are affected by atmospheric conditions, while the ionospheric waves are affected almost entirely by ionospheric conditions. It is the purpose of this chapter to discuss the problems connected with ionospheric waves, and the effects of ionospheric conditions on radio propagation in the frequency range from 30 to 100 mc.

IONOSPHERIC PROPAGATION

THE EARTH'S ATMOSPHERE. Although the earth's atmosphere outward to over 250 miles has been described many times, a description will be included in this section for convenient reference. The lower layer of the atmosphere, extending upward approximately 6 miles, is called the *troposphere*. This layer contains almost all of the gases comprising the atmosphere, and all of the phenomenon known as weather. Storms, warm or cold fronts, and air masses do not rise above this level.

Above the troposphere, extending outward to about 30 miles and containing a high ozone content, is the *stratosphere*. This region is characterized by a complete lack of weather, and a nearly constant temperature of -57 degrees C. As far as is now known, the stratosphere does not affect radio transmission directly; however, since it acts as a reservoir of heat, it does affect the weather in the lower troposphere, and thus indirectly affects radio propagation. The stratosphere also absorbs almost all of the ultraviolet rays from the sun that have passed through the ionosphere, and thereby protects life on the earth from too intense radiation.

THE IONOSPHERE. In the regions above the stratosphere, the air particles become widely scattered, and consequently when electromagnetic radiations from the sun bombard the molecules composing these air particles, they become ionized. Such radiations from the sun include ultraviolet light, X-rays, gamma rays, and cosmic rays, the most predominant ionizing agent being ultraviolet light.

At high altitudes, the air particles are widely scattered, and consequently less recombination occurs between the free electrons and the positive ions. Therefore, these regions remain ionized for considerable periods of time. The region of ionization which affects radio propagation lies between 30 and 250 miles (50 to 400 km) above the earth's surface and is called the *ionosphere*.

The ionization in the ionosphere is not uniformly distributed with altitude but is stratified, and there are certain definite layers in which the ionization density is sufficient to reflect radio waves. Such layers have no sharp boundaries and their altitude and density varies with the time of day, season, and sunspot activity.

The first layer, known as the *D layer*, exists only at times during the day at a height of about 30 to 50 miles, while the second or *E layer* exists at a height of approximately 50 to 90 miles. The third and largest layer, termed the *F layer*, exists as a single layer at night while during the day it separates into two layers, *F_1 and F_2*. At night, the F layer is present at an altitude of approximately 100 to 250 miles. During the day, the F_1 layer is at a height of about 120 to 150 miles, and the F_2 layer is at a height of 150 to 220 or more miles in the summer and from about 90 to 190 miles in the winter.

According to Dr. W. A. Miller,* the stratification of the ionosphere into definite layers is produced as follows:

D layer
(1) The first ionization potential of molecular oxygen.
(2) The ionization of the metal atoms in the upper atmosphere, sodium in particular.
(3) Possibly the ionization of nitrous oxide.
All of the above are sufficiently abundant, and have a low enough ionization potential to permit ionization.

E layer
(1) First ionization potential of molecular oxygen.
(2) Second ionization potential of molecular oxygen.
(3) Collision between two atoms of oxygen resulting in a positively ionized oxygen molecule, and a free electron.
(4) Pre-ionization of molecular oxygen due to the strong absorption bands in the energy range 12.2 to 13.55 ev.
(5) Absorption of high energy photons.

F_1
(1) The second ionization potential of molecular nitrogen.
(2) The first ionization potential of atomic oxygen.

F_2
(1) The first ionization of molecular nitrogen.
(2) The first ionization potential of atomic oxygen.
(3) The second and third ionization potentials of atomic oxygen.
(4) The first ionization potential of atomic oxygen forming the F. layer, together with a height dependent recombination coefficient and/or tidal effect.

E_S (Sporadic E)
Possibly produced by the ionizing impact of meteoric particles, or, in high latitudes, by bombardment by such fast solar corpuscles as produce auroras.

Radio signals reflected from the ionosphere are not all reflected from the same part of the ionosphere. Each of the ionized layers reflects different radio frequencies, allowing some frequencies to pass

*EM 62-47, The Solar Study Program of the RCA Laboratories, Rocky Point, Long Island.

MAJOR CHARACTERISTICS OF IONIZED LAYERS

	HEIGHT OF MAX. IONIZATION	MAX. AV. ELECTRON DENSITY (N_e) ELECTRONS / CM^3	TIME OF OCCURRENCE	N_e SUNSPOT MAX. / N_e SUNSPOT MIN.	DENSITY OF NEUTRAL PARTICLES PER CM^3	FREQUENCY (f_c) AT 1945 EQUINOX
D LAYER	60 KM (37 MI.)	1.5×10^4	DAYTIME ONLY	2.0	8×10^{15}	—
E LAYER	100 KM (62 MI.)	1.5×10^5	24 HRS.	1.50	6×10^{12}	3.7 MC
F_1 LAYER	200 KM (125 MI.)	2.5×10^5	DAYTIME ONLY	1.56	1×10^{11}	4.9 MC
F_2 LAYER	300 KM (186 MI.)	1.5×10^6	24 HRS.	4.00	2×10^{10}	10.5 MC
E_s LAYER	100 KM - (62 MI.)	SPORADIC PATCHES OF IONIZATION DENSITY FAR HIGHER THAN NORMALLY FOUND AT THIS ELEVATION. PREDOMINATES IN SUMMER, BUT NOT PREDICTABLE FOR f_c, OR TIME OF OCCURRENCE.				

Figure 15-1. Major Characteristics of Ionospheric Layers

through, and reflecting others. Figure 15-1 shows the main characteristics of each layer of the ionosphere. It must be understood that the values given in the table are of necessity only approximate, since these values are subject to diurnal, seasonal, and solar changes. Since the degree of ionization is dependent upon energy from the sun, the ionization would be expected to be greatest in the upper ionized layers, and so the table shows it to be. As the ultraviolet rays penetrate the ionosphere, fewer and fewer reach the lowest ionized layers, and a correspondingly lower degree of ionization in these layers results. At night, when the solar rays are blocked from the ionosphere, changes take place in the layers. The D layer disappears entirely, and the F_1 layer merges with the F_2 layer. The E layer and the F layers show a lower degree of ionization and their height above earth increases.

Changes in the ionosphere are geographical as well. At higher latitudes the degree of ionization is not as great as at the equator. Another change is the variation of ion density in the layers due to change in seasons. Just as the amount of effective sunlight received during winter is far less than the amount received during summer, so is the amount of ultraviolet radiation, and consequently, we find that the layers have a greater ion density during summer than during winter.

At times of excessive solar activity, the degree of ionization is again changed. There is a relationship, not clearly understood at present. between sunspots and density of the ions. Since sunspots represent greater increased activity in the sun, the ionization of the various layers increases, just as might be expected from such extra activity.

Another unusual source of radiation from the sun is solar flares. These flares, most easily seen during an eclipse, are vast quantities of ionized gases, flung outward perhaps hundreds of miles from the sun's surface. Many astronomers believe that some flares reach out as far as the earth. These flares produce *sudden ionospheric disturbances*, abbreviated *S.I.D.*, which cause rapid, short-lived fades in the ionospheric wave. If such a flare should occur near the center of the solar disk facing the earth, the resultant fade is usually followed some 8 to 36 hours later by an ionospheric storm. The relationship between S.I.D. and the ionospheric storms which follow is not clearly established, but it is believed that the fade is due to an unusually great emission of ultraviolet rays (which travel with the speed of light). The following storm is believed due to a simultaneous emission of solar corpuscles, probably calcium ions, which travel far slower than the speed of light.

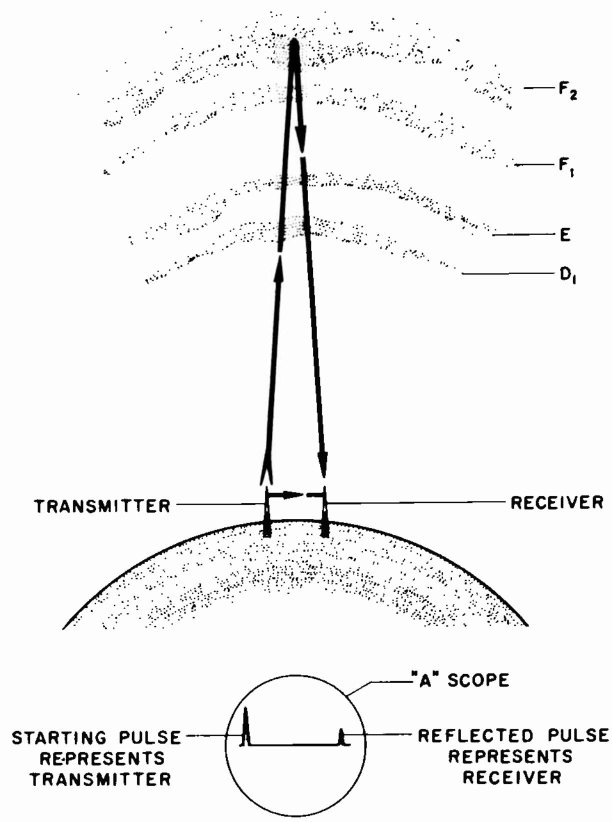


Figure 15-2. RF Sounding of Ionosphere

RADIO TRANSMISSION. The effect of the sun's activity on the ionosphere determines the *critical frequency* (f_c), which is defined as the highest frequency that may be reflected from a particular ionized layer. Any higher frequency will pass through or be captured by the layer, rather than be reflected, as illustrated in figure 15-1. This critical frequency is also affected by the ion density of a layer, and the greater the ion density, the higher the critical frequency will be. It naturally follows, then, that a radio signal might have a frequency higher than the critical frequency of the E layer, and pass through, and yet be reflected from the F_2 layer. Figure 15-1 shows that this is exactly the way different frequencies are reflected. Any activity of the sun resulting in an increase of ultraviolet rays will produce higher ionization of the layers, and therefore a higher critical frequency for each layer.

The critical frequency is important in radio transmission since it determines the best frequency to be used in reflecting a radio signal from the ionosphere.

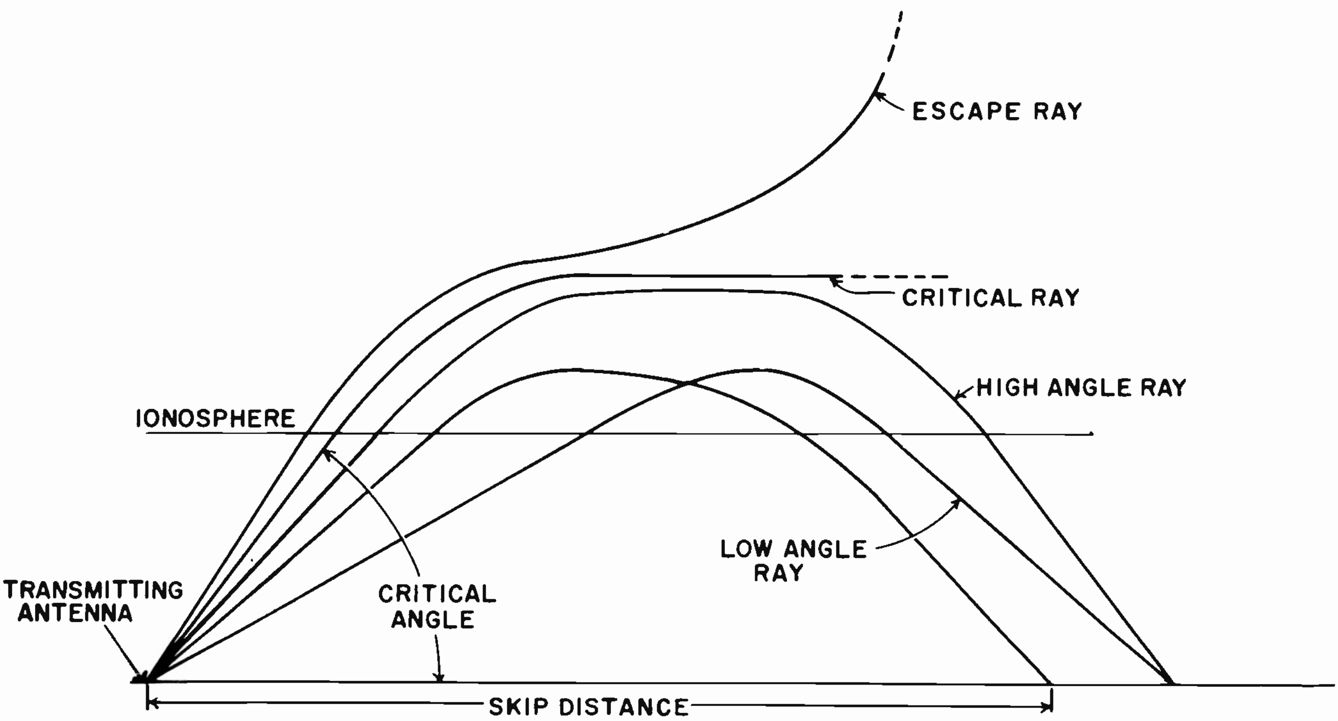


Figure 15-3. Ionospheric Dispersion

The best frequency is the highest frequency that may be reflected from the ionosphere, and is called the *maximum usable frequency* (MUF). This frequency is given by a formula known as the "*secant law*", or

$$\text{MUF} = f_c \sec \Theta \qquad (15\text{-}1)$$

where

f_c = Critical frequency
Θ = Angle of incidence

Since the ionosphere is subject to a continually changing bombardment of ultraviolet rays, the critical frequency, and consequently the MUF, is also subject to change. When a radio message of any length is to be transmitted, it is entirely possible that the critical frequency may decrease during the time of transmission due to a change in the ionized layer, causing the received signal to fade. In order to prevent this from happening, a lower frequency is used, called the *optimum traffic frequency* (FOT). This is taken as 90% of the MUF, and allows for a reasonable amount of fluctuation in the critical frequency. However, it does not allow for diurnal changes, which are usually taken into consideration by periodically changing to a different working frequency according to the time of day.

There is another term used in connection with ionospheric reflection that might well be mentioned here, called the *lowest useful frequency* (LUF). Attenuation of the reflected radio signal by the ionospheric layers is determined by the frequency of the signal, with the greatest attenuation occurring at the lowest frequency. Thus, the lowest useful frequency represents a compromise between a frequency that may be transmitted without interruption due to ionospheric changes, and frequency without excessive loss caused by ionospheric attenuation.

The critical frequency is determined by reflecting a radio pulse at various frequencies from the ionosphere. The returned pulse is picked up by a receiver located close to the transmitter. As the frequency of the transmitted signal is increased, a point will be reached where no pulse is returned, thus giving the critical frequency of that layer. By timing the interval between transmission and reception of the pulse, the distance the pulse has traveled may be determined, and will be twice the height of the reflecting layer. Since the signal does not travel as fast through the ionized layer as through the atmosphere, there will be an error in the measured height. The height found by timing the pulse will be the *virtual height;* that is, the height of the layer if it were a flat reflecting plane. This value is more useful than the true height, because it is the height that is used in determining the angle to beam the signal in order for the signal to reach the receiver.

The values of the critical frequency for any layer are different at different parts of the earth in accordance with changes in the ionosphere. For long-distance transmission it is necessary to know the critical frequency at other points than the transmitter location. This is done by use of recording equipment that continually beams pulses of varying frequency vertically into the ionosphere. The returned pulses are recorded, and show both the critical frequency for the layers of the ionosphere, and the virtual height of these layers (figure 15-2).

There are a great many such recording instruments throughout the world, and data from them may be used to construct iso-frequency charts, showing the critical frequency for each layer above different parts of the world. These maps may be used for predicting the MUF so that the most satisfactory transmitting frequency may be chosen. It should be noted here that, as might be expected, the receiver and transmitter may be on far distant parts of the earth, with different critical frequencies for each location. In such a case, a value for the critical frequency as found at one-third of the distance from the receiver to the transmitter is used as being the best compromise value for the radio path.

Most of the ionospheric reflection takes place in either the E layer, or the F_2 layer. However, there is another layer that may be used for ionospheric transmission up to approximately 100 mc, namely, the E_S or *sporadic E layer*. This layer, at the same level as the E layer, consists of highly ionized areas and may be used to obtain exceptionally good transmission over medium distances. The sporadic E layer has one serious drawback, however, in that it is extremely unstable and therefore cannot be accurately predicted, as compared with the E and F_2 layers.

Transmission to an extremely distant point often takes place by means of multiple reflections, either between an ionized layer and the ground, or between two ionized layers. Such an excursion of a radio wave from the earth to the ionosphere and back to the earth is called a *hop*, and the distance from the transmitter to the point at which the wave returns is called the skip distance (see figure 15-3). The skip distance is defined as the minimum distance at which radio waves of a specified frequency can be transmitted at a particular time by means of reflection from the ionosphere. However, ionospheric waves are frequently received at less distances than the skip distance by way of sporadic and scattered reflections.

CHAPTER 16

EQUIPMENTS TEST & CALIBRATION

INTRODUCTION

The testing and calibration of high-frequency communication equipments are tasks requiring considerable skill, and should not be attempted without careful preparation. However, a good knowledge of equipment circuit design, together with the use of systematic methods of service and proper test equipment, will greatly reduce the possibility of error. Many types of test equipment have been developed in order to simplify the tasks of testing and calibration. A description of some of the present standard instruments is listed in a later section of this chapter.

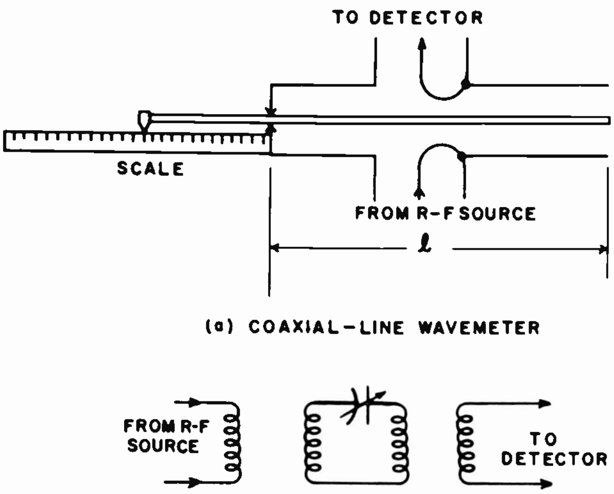
Perhaps the most exacting high-frequency measurements are those concerning frequency, standing waves, and r-f power levels. These measurements are important in all communication equipments and, in particular, in radio-relay systems, where they contribute a great deal toward maximum reliability and overall system performance.

HIGH FREQUENCY MEASUREMENTS

WAVEMETERS. To interpret most high-frequency measurements, it is essential to have an accurate knowledge of the wavelength or of the operating frequency of the equipments in use. Frequency is the most fundamental quantity in the sense that in any fixed linear system the frequency of oscillation has the same value for all parts of the system. However, wavelength is generally easier to measure, and by adapting the resonant-line technique to high-frequency circuits, satisfactory measurements can be made.

A coaxial-line wavemeter of the kind shown in figure 16-1a may be used at high frequencies, provided that the mean value of the circumferences of the inner and outer conductors is less than one wavelength. The frequency range would be limited to an upper frequency of approximately 3,000 mc. This method of wavelength measurement eliminates all wave modes other than the principal, or transverse electromagnetic (TEM), wave for which the wavelength in a coaxial line is nearly equal to the corresponding free-space value. The theory of operation of such a wavemeter is illustrated by an equivalent circuit, as shown in figure 16-1b. Any coupling between the source and the detector is normally negligible, but is increased to a large amount when the intermediate tuned circuit is brought into resonance.

In the coaxial-line wavemeter, the tuned circuit is replaced by a variable length of coaxial transmission line, short-circuited at one end and open-circuited at the other. A small amount of r-f energy is fed into the line by means of an excitation loop (marked "source" in figure 16-1a. Resonance is then established by adjusting the length of the inner conductor, using a detector coupled to another loop opposite to the excitation loop. A resonant condition will occur when the length of the inner conductor is approximately one-quarter wavelength. Resonance can be obtained at even wavelength points ($\lambda/2$, λ) if the line is short-circuited at both ends. If more than one resonance point can be obtained on the wavemeter, it is self-calibrating, since the distance between successive resonances is one-half wavelength. If only one resonance point can be obtained, the wavemeter must



(a) COAXIAL-LINE WAVEMETER

(b) EQUIVALENT CIRCUIT

Figure 16-1. Coaxial Line Wavemeter

be calibrated, and it is important that the coupling loop be fixed in position if accurate results are to be obtained. Increased accuracy may be obtained by taking equal readings on either side of the maximum resonant position, and the use of a vernier scale is highly recommended. Figure 16-2 illustrates four types of coaxial and transition wavemeters. The transition wavemeter is a cross between coaxial and cavity wavemeters.

Since a cavity resonator may be employed as a tuned circuit in high-frequency circuits, it follows that such a resonator may also be used as a wavemeter. Although a cavity-resonator wavemeter utilizes various wave modes of oscillation, its operating theory is much the same as that of a coaxial line wavemeter shorted at both ends. A cylindrical cavity is perhaps the most practical for use as a wavemeter because it is easier to machine accurately.

There are two main classes of wave modes that can be usefully employed in a cavity-wavemeter, the $TM_{0,1,0}$ mode and the $TE_{0,1,1}$ mode. A typical wide-band wavemeter which operates in the $TM_{0,1,0}$ mode is illustrated in figure 16-3A. A set of four wavemeters of this kind can be made to cover a frequency range from 200 mc to 10,000 mc, with an accuracy better than one part in ten thousand. The design of the wavemeter illustrated in figure 16-3A is based on the very large change in characteristic impedance which occurs between the two sections of coaxial line.

The use of the $TE_{0,1,1}$ mode as applied to wavemeters has several important advantages over the $TM_{0,1,0}$ mode, and are summarized as follows:

1. The amount of energy dissipated on the cavity walls is extremely low, and therefore a high Q-value results.
2. There is no radial current present on the end cavity walls.
3. The TE_0 form of wave comprises only modes which have no angular variation of field strength, and therefore does not introduce spurious points of resonance.

A high Q-value is an essential feature of the resonance type of wavemeter, and a cavity excited in the $TE_{0,1,1}$ mode is superior in this respect to any other. With the absence of radial current flow in the TE_0 wave, there is no need to establish good electrical contact between the end walls and the cylindrical surface. In fact, the absence of such a contact will tend to suppress any other modes of oscillation in the cavity which might otherwise arise. If one of the end plates is made slightly smaller in diameter than the cylinder, and arranged to move axially over a limited range, the resonant frequency of the cavity may be varied and the cavity may be used as a wavemeter.

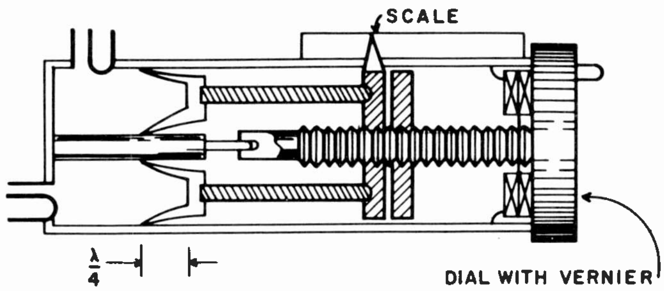
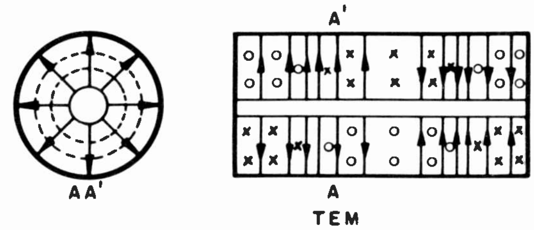
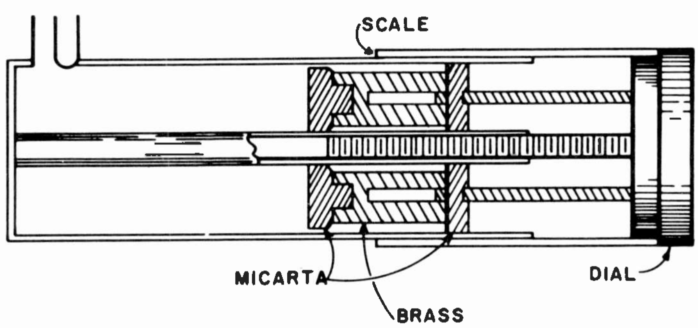
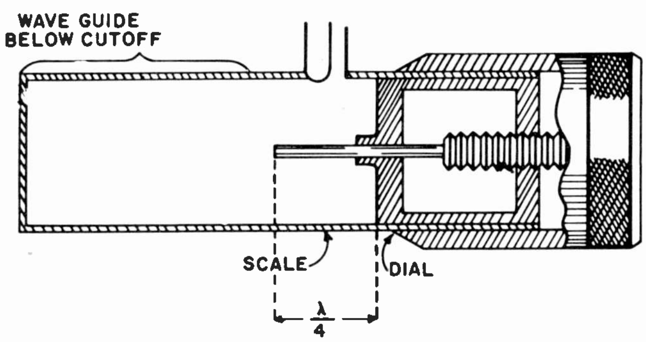
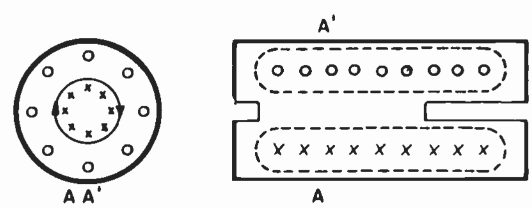
Wavemeters excited in the $TE_{0,1,1}$ mode are illustrated in figures 16-3B and 16-3C. Coupling to the wavemeter in figure 16-3B is provided by means of small loops similar to those in the coaxial type of wavemeter. An alternative coupling method, which has the advantage of avoiding spurious wave-mode excitations is illustrated in figure 16-3C. Cavity-resonator wavemeters are usually silver-plated on the inside to improve their Q-value and for accurate work the enclosure should be hermetically sealed or equipped with a dehydrator such as silica-gel.

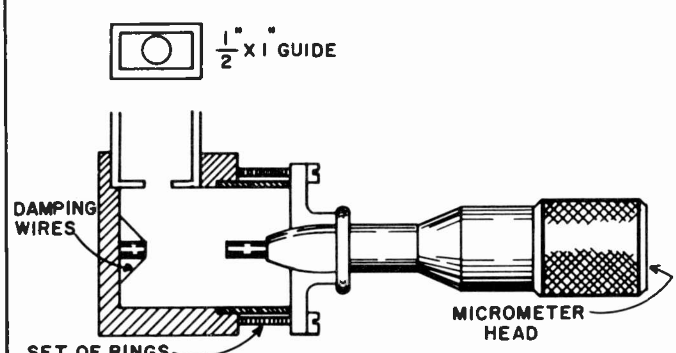
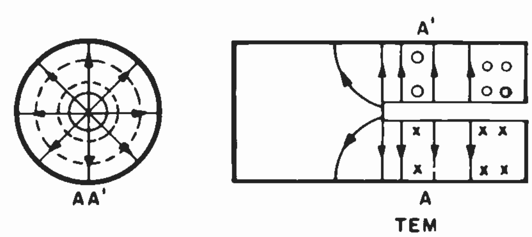
The calibration of a cavity wavemeter is theoretically calculable from the dimensions of the cavity. However, such dimensions are usually not known to a sufficient degree of accuracy because effective cavity size changes with temperature. A copper cavity may change in resonant wavelength about three parts in ten thousand between summer and winter, and such a change is by no means negligible at frequencies above 3,000 mc. A cavity-resonator wavemeter should, therefore, have the calibration temperature recorded on the instrument, and if the working temperature differs greatly from that value during operation, appropriate allowance must be made.

Cavity-resonator wavemeters may be calibrated by either of the two methods listed below:

1. By the use of a heterodyne frequency meter.
2. With a coaxial-line wavemeter.

The first method is perhaps the most accurate system of calibration. Power from an r-f oscillator is fed into a matched waveguide, which is equipped

<table>
<tr><th></th><th>TYPE</th><th>MODE OF OSCILLATION</th></tr>
<tr><td rowspan="3">COAXIAL</td><td>
</td><td rowspan="2">
</td></tr>
<tr><td>
</td></tr>
<tr><td>
</td><td>

$TE_{0,1,1}$</td></tr>
<tr><td>TRANSITION COAXIAL TO CYLINDRICAL</td><td>
</td><td>
</td></tr>
</table>

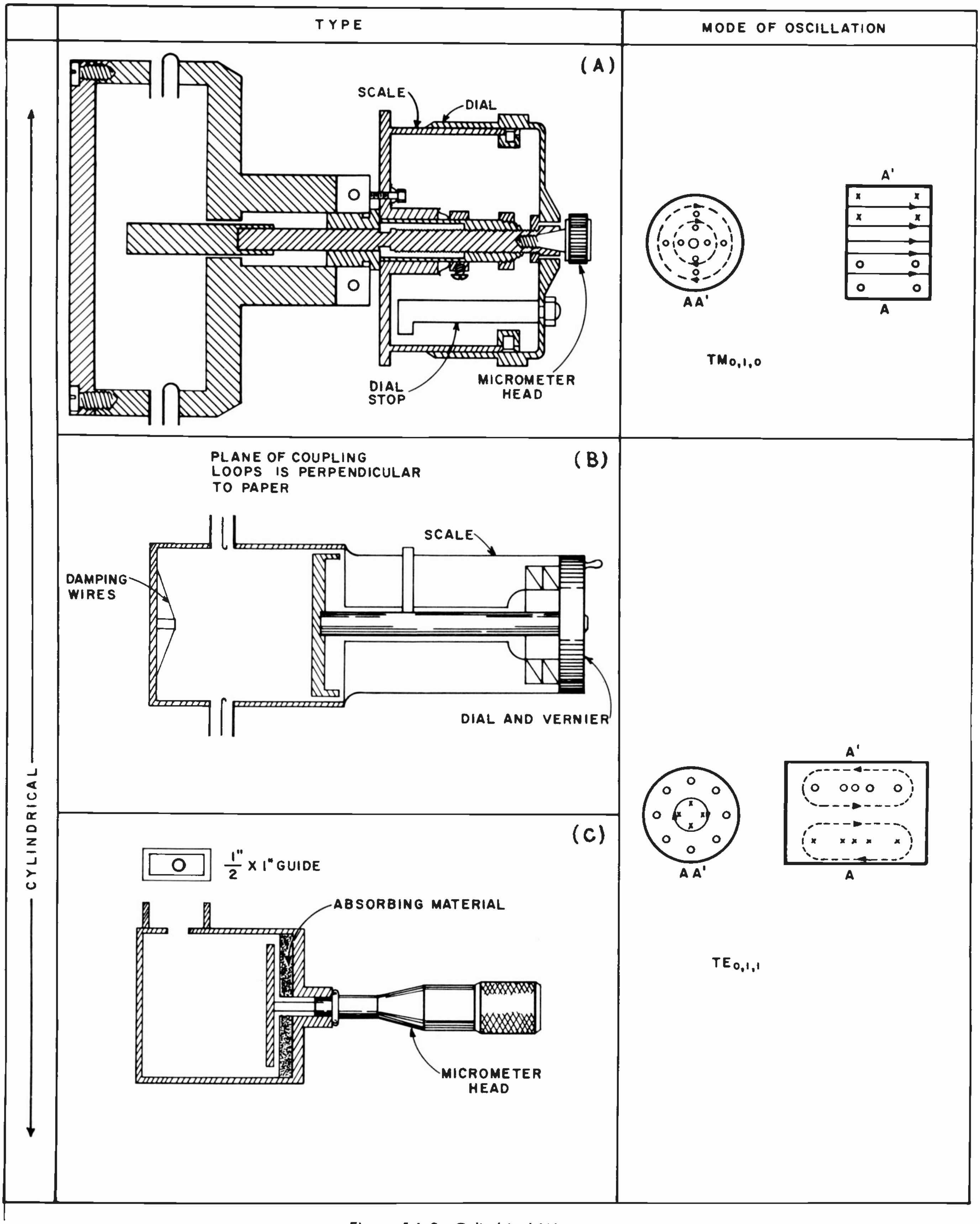


Figure 16-3. Cylindrical Wavemeters

with short probes to couple the cavity and the heterodyne meter to the guide. The oscillator frequency is varied and measured by the heterodyne frequency meter while dial readings of the cavity wavemeter are taken. A calibration chart is then prepared and should include the date, temperature, and humidity at the time of calibration. In the heterodyne method of calibration, the component to be measured has its frequency decreased to a predetermined value by heterodyne action, and is then amplified and measured at this frequency. The frequency of the unknown component is determined from the heterodyne frequency that must be used in order to change the unknown frequency to the known fixed frequency.

Cavity calibration may be made as described above by substituting a coaxial-line wavemeter or another cavity of known calibration in the place of the heterodyne frequency meter. Although the accuracy obtained with the use of these alternative standards is considerably lower than that obtained with the heterodyne meter, the calibration procedure can be carried out much faster.

The accuracy of a wavemeter depends upon the accuracy with which its scale is divided, and also upon the loaded Q-value. For example, if an accuracy of ± 0.001 cm is desired, the scale must be fine enough to read at least three decimal places, and the sensitivity of the wavemeter must be great enough to determine when the tuning is 0.001 cm or less from resonance. In other words, the frequency response curve of the wavemeter must drop to its "half power" points within 0.001 cm of the resonant wavelength. The necessary loaded Q-value of a wavemeter may be found by the formula

$$Q_L = \frac{\lambda}{\Delta\lambda} \qquad (16\text{-}1)$$

where

Q_L = The necessary loaded Q-value
λ = Operating wavelength in centimeters
$\Delta\lambda$ = Desired accuracy in centimeters

Therefore, at 10 cm, for an accuracy of ± 0.001 cm, the loaded Q-value must be

$$10/2\ (0.001) = 5{,}000.$$

STANDING - WAVE MEASUREMENTS. An important part of high-frequency measurement technique is the precise determination of the standing-wave pattern in transmission-line systems. If the distribution of the electric or the magnetic field of such a system is known, the reflection coefficient, and therefore the impedance at any point along the system, may be found. The field distribution arising from a standing wave on a length of transmission line may be recorded by an instrument known as a *standing-wave indicator*. In a waveguide, a narrow longitudinal slot cut into one face of the guide enables a suitable detector to be loosely coupled to the field within the guide and moved axially. A similar method is used to measure standing waves along a coaxial transmission line system. The detector is coupled to a length of solid coaxial line by means of a longitudinal slot or evenly spaced holes cut along the line. In any case, the slot length should be longer than one wavelength. Coupling is generally accomplished by means of a short electric probe which is inserted into the longitudinal slot and moved along its length by means of a close-fitting spline. Such an arrangement is illustrated in figure 16-4. The position of the sliding carriage along the guide must be accurately known, and a vernier scale is usually provided for this purpose.

Several types of detectors are available for standing-wave measurements, but those commonly used may be grouped into two classes:

1. *Square law detectors* which are comprised of crystal rectifiers, bolometer probes, or thermocouples, and
2. *Linear detectors*, or crystal mixers with one or more stages of i-f amplification.

Figure 16-5 illustrates a typical square law crystal detector which is frequently used as a standing-wave indicator. The probe provides loose coupling with the electric field in the transmission line, and the voltage intercepted is then rectified by the crystal and measured with a microammeter. This type of probe is frequency sensitive and needs adjustment when the operating frequency is changed.

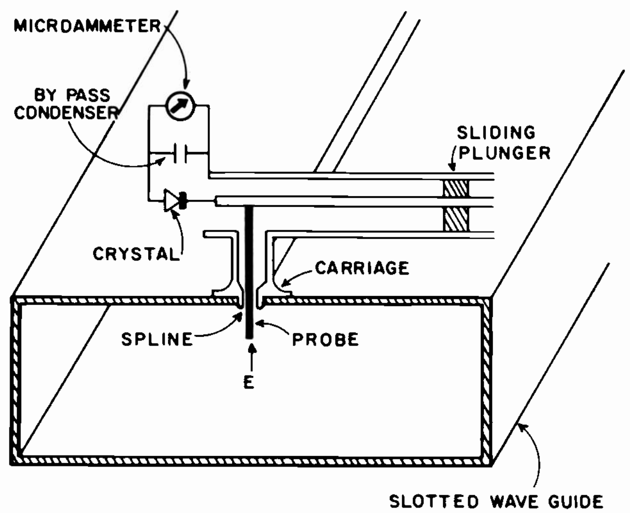


Figure 16-4. Standing Wave Indicator

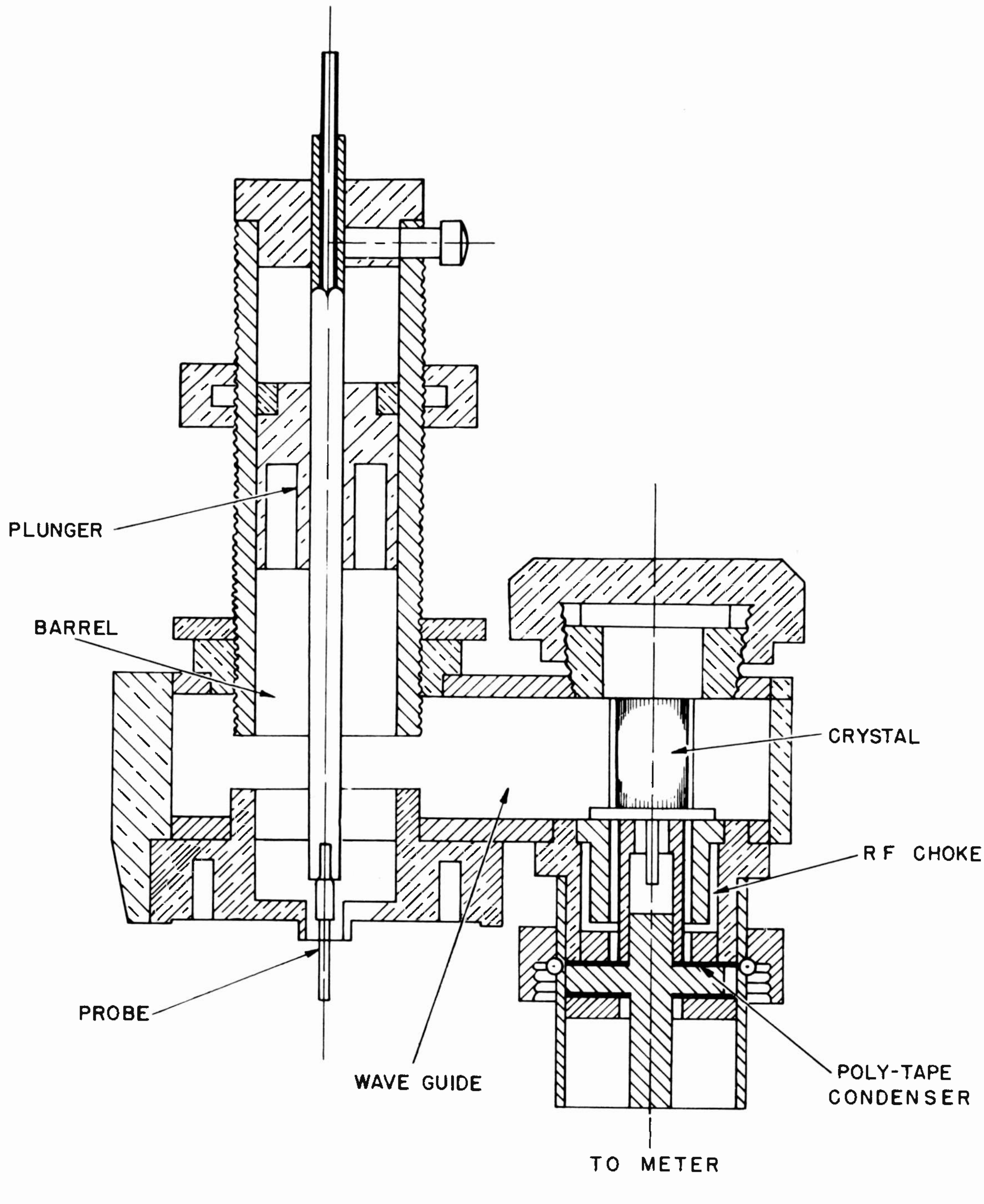


Figure 16-5. Crystal Probe

By substituting a "bolometer" fuse for the crystal in the indicator described above, a tunable bolometer probe is then obtained. "Bolometer" is a name given to a large class of components used in r-f measurements whose resistance changes with changes in temperature. A platinum bolometer fuse is illustrated in figure 16-6, with the protective glass cover and the upper cap removed. Temperature changes, and consequently, resistance changes in the platinum wire are brought about by the absorption of r-f power, such changes being measured with a suitable amplifier. A schematic of the circuit usually employed to accomplish this is shown in figure 16-7. Bolometer probes are normally used only for high-level, pulsed-power measurements, although they are adaptable to low-level work if sufficient r-f energy or extremely sensitive amplifiers are available.

The thermocouple is important only because it provided the first means of measuring standing waves. Such a device usually consisted of a sensitive vacuum thermocouple mounted in a small cavity and fed by a probe in much the same way as previously described. However, the thermocouple junctions were extremely hard to match to the probe, very delicate, and sometimes quite sluggish and insensitive because of their high thermal capacity. Thermocouple probes have almost entirely given way to the crystal and bolometer types as described above.

For nearly all standing-wave measurements, a point of minimum field strength must be employed as an origin of reference. The location of such minima may be found by sliding the standing-wave indicator along the longitudinal slot until a minimum reading is indicated on the meter. More accurate readings may be taken by recording the indicator positions for equal meter readings on each side of the minimum and then finding the average of these positions.

The voltage standing-wave ratio of any transmission line is the ratio of the maximum to the minimum electric field strength and is expressed by the formula

$$\text{VSWR} = E_{max}/E_{min} \qquad (16\text{-}2)$$

With a square law detector, the formula for the voltage standing-wave ratio becomes

$$\text{VSWR} = \sqrt{(D_{max}/D_{min})} \qquad (16\text{-}3)$$

where D_{min} and D_{max} are the minimum and maximum deflections of the meter used. With a linear detector, formula 16-2 may be used to determine the standing-wave ratio. If the law of the detector is unknown, a calibration curve must be plotted, giving the relative values of the electric field strength as a function of meter deflection.

Since the determination of the standing-wave ratio by the instruments described above involves taking two readings and calculating their quotient, it has been found necessary to devise various direct-reading instruments for use in the field. Such apparatus is generally less accurate than the other, but is considerably easier to use for fast measurements.

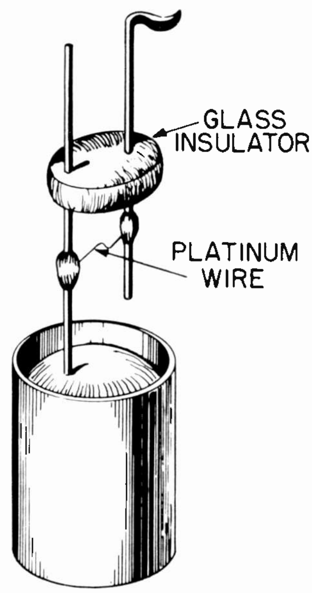


Figure 16-6. Bolometer Littelfuse

For application to high-power systems, a neon-tube indicator may be used. It consists of a long, thin, glass tube containing neon at a reduced pressure and arranged to penetrate a short distance into the transmission line in the same fashion as a probe. Excitation within the transmission line then has the effect of ionizing the neon gas in the tube so that a visible glow extends outside the guide to a height proportional to the product of the electric field concerned. As the tube is moved along the transmission line, the presence of a standing wave will then be shown by the variation in height of the ionized column. The effect is as though the standing-wave pattern were plotted as the tube is moved along the line. If the neon tube is calibrated initially, maxima and minima may be read directly, but with little accuracy. Other types of direct-reading indicators include reflectometers, phase-shifters, and the "magic-tee" waveguide bridge.

R-F POWER MEASUREMENTS. High frequency r-f power measurements are generally classified into two groups, high-level measurements and low-level measurements. Several devices are now available for measuring high level r-f power, some of which are outlined below.

1. Continuous flow calorimeter.

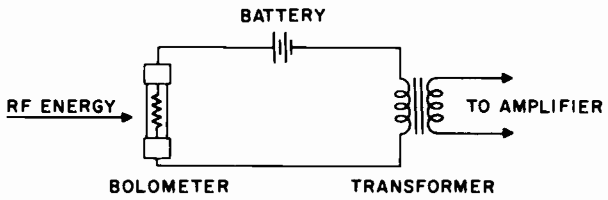


Figure 16-7. Bolometer Bias Circuit

2. Johnson power measurer.
3. Low-level power measurers equipped with an attenuator.

The continuous flow calorimeter is probably the simplest to use and the easiest to build. Figure 16-8 illustrates the construction of such a device as intended for use with a coaxial transmission line.

A section of coaxial transmission line is short-circuited at the far end and closed at the other by means of a thick dielectric bead which is usually made of glass or micalex. The length of the device should be such that the distance from the bead to the short circuit is several wavelengths. This is in order to insure complete absorption of the r-f energy. The length of the dielectric bead should be a quarter wavelength of the operating frequency, and the dielectric constant of the bead is chosen so that

$$Z_d = \sqrt{Z_{air} \times Z_{H_2O}} \qquad (16\text{-}4)$$

where

Z_d = The impedance of the dielectric-filled line.
Z_{air} = The impedance of the air-filled line.
Z_{H_2O} = The impedance of the water-filled line.

Water flows continuously through the chamber and is heated by the r-f energy. Thermocouple probes are placed in the water inlet and in the outlet to measure the rise in temperature of the water flowing through the cavity. To find the r-f power, it is necessary to measure the rate of flow and the change in temperature of the water. The r-f power may then be found by the formula

$$W = 4.19\ VT \qquad (16\text{-}5)$$

where

W = R-F power in watts
V = The volume of water flowing through the cavity per unit time
T = Temperature changes of the water.

In practice, T is not measured directly but rather, the current in microamperes flowing through the thermocouple. The difference in temperature of the two thermocouple junctions is a linear function of the current. Therefore, it is necessary to know the calibration of the thermocouples and meter. The constant may be obtained by placing the thermocouple junctions in water baths which differ in temperature by T_0 and noting the deflection, d, of the meter used. Formula 16-5 may then be rewritten as follows:

$$W = 4.19 \frac{T_0}{d} VD \qquad (16\text{-}6)$$

where

$\frac{T_0}{d}$ = The calibration constant for the components used
D = The deflection of the meter during the power measurement.

Since the expression $4.19 \frac{T_0}{d}$ in formula 16-6 is a constant for a given set of components, the r-f power may then be found by the formula

$$W = AVD \qquad (16\text{-}7)$$

where

$A = 4.19\ T_0/d$

Thus, once the constant A is known, the r-f power absorbed in the termination can be obtained by multiplying the constant by the rate of water flow, V, and the meter deflection, D.

There are several advantages of measuring r-f power by this method:

1. The measurement is absolute.
2. The rise in temperature can be made small, thereby limiting the energy lost by radiation and conduction from the transmission line.

However, in order to maintain a constant rate of water flow, it is necessary to use some device to supply water at a constant pressure. The only serious disadvantage of the method is that such a device must completely terminate a line and, in so doing, absorb 100% of the r-f power. A continuous flow calorimeter therefore cannot be used as a power

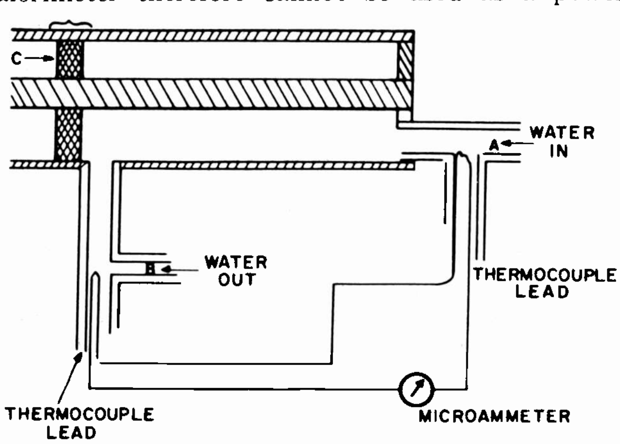


Figure 16-8. Coaxial Water Load

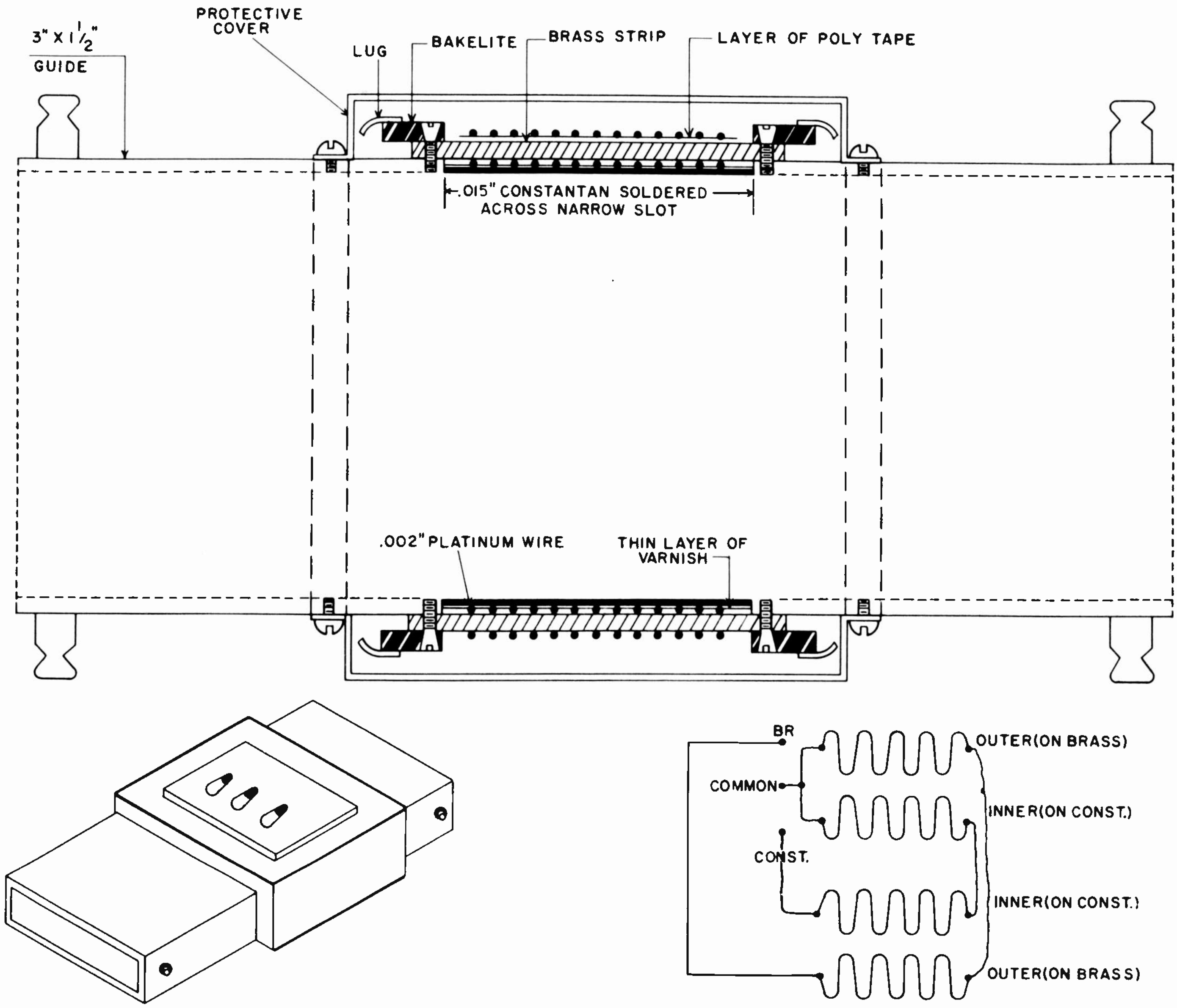


Figure 16-9. Johnson Wattmeter

monitor except by the use of a mechanical switching arrangement.

A waveguide calorimeter is electrically the same as that for a coaxial line, and is matched to the guide by a quarter-wavelength transformer. The transformer consists of a section of guide whose electrical length is one quarter-wavelength of the operating frequency, and is filled with a material whose dielectric constant is the geometric mean of those of air and water.

A type of r-f power measure which may be used as a constant monitor is the *Johnson Wattmeter*. This type of meter requires only one percent of the r-f energy in the transmission line as compared to the 100 percent needed for the constant flow calorimeter.

Figure 16-9 illustrates a type of Johnson wattmeter which is designed for use with a waveguide transmission line. A small longitudinal section of the waveguide wall is removed and replaced with a piece of high-resistance material. A sheet of constantan 0.015 inch thick can be used. At high r-f power levels, enough heat is dissipated in the constantan to produce an appreciable temperature change. Since the conductivity of the waveguide is greater than that of the constantan, there will be a measurable difference in the temperature of the two materials. This difference in temperature is used to give an indication of the r-f power level. The difference in temperature is measured by a platinum resistance thermometer which consists of 200 ohms of 0.002 inch platinum wire wound on the constantan section and another 200 ohms wound on the guide. The two windings serve as two arms of a Wheatstone bridge as shown in figure 16-9.

If the bridge is initially balanced, any change in temperature of one of the coils with respect to the other will unbalance the bridge. The unbalanced condition is caused by the change in resistance of the platinum wire around the section of constantan with temperature. The unbalanced current in the bridge is directly proportional to the power in the line, and a straight line calibration curve can be drawn by measuring only one point with an accuracy within 5 percent. Since there is a time delay between the turn-on of the power and the thermal equilibrium of the measuring device, there is a time interval of between one and three minutes after the turn-on before readings can be taken.

Low-level power measurements generally cannot be carried out by the methods described above because the temperature rise becomes too small for accurate determination. However, three common detectors may be used to advantage for such measurements; crystals, bolometers, and thermistors. The crystal detects by virtue of its rectifying property, while the bolometer and thermistor detect through a resistance change induced by the heating effect of absorbed r-f power. The thermistor consists of a resistance element made from certain metal oxides which have a high negative temperature coefficient, that is, the resistance of the thermistor decreases with increasing temperature. Its sensitivity is approximately ten times that of a bolometer, and it has both mechanical ruggedness and good overload characteristics. Therefore, for accurate work, the thermistor is the most desirable of the three detectors mentioned.

Precise measurements of small r-f power levels are generally made with the use of balanced bridges. The bridge is first balanced with a known amount of direct current, after which an unknown amount of r-f power is applied to the detector, unbalancing the bridge again. Balance is restored by removing some of the direct current from the detector, and the change in direct current power is used to calculate the magnitude of the r-f power added. Most bridge circuits are so arranged that the difference in direct current power between the two balanced conditions is exactly equal to the magnitude of the r-f power.

An unbalanced bridge may also be used to measure low-level r-f power, and is easier and faster to use than the balanced bridge. Such a bridge is initially balanced on direct current, and the deflection of the bridge meter is noted when the r-f power is added. The meter scale is previously calibrated in r-f power to give a direct reading r-f microwattmeter. For instance, with a thermistor detector and a low resistance microammeter, it is possible to have a full scale deflection represent 1 milliwatt with a calibration that is appreciably linear.

CURRENT STANDARD TEST EQUIPMENT

The following list is, *for the categories covered*, the approved list of recommended general purpose test equipment for use with prime equipment in the field of electronics. The models and types listed are the present standard equipment already developed and able to be procured as of 1 January 1953. The categories are taken from a comprehensive functional classification of test equipment as promulgated to the Services by Research and Development Board Memorandum RDB-EL 213/1 of 19 March 1952. The main headings of these categories are as follows:

Category	*Title*
1	Voltage and current measuring equipment
2	Frequency measuring equipment
3	Waveform measuring equipment
4	Signal generating measuring equipment
5	Field intensity measuring equipment
6	Impedance and standing wave measuring equipment
7	Amplifying equipment for measuring purposes
8	Time based measuring and counting equipment
9	Nuclear energy test and measuring equipment
10	Combination and group test set
11	Associated devices for electronics test equipment
12	Miscellaneous test equipment
13	Calibrating equipment for electronics test equipment
14	Power measuring equipment

1. VOLTAGE AND CURRENT MEASURING EQUIPMENT

Equipment	*Range*
1.1 Voltage and/or Current Meter	
1.1.1 Voltmeter	
1.1.1.1 Electronic Voltmeter	
1. ME-6/U, (BuS) (Electronic Voltmeter)	0.5 to 5; 5 to 50; 50 to 500 mv; 0.5 to 50; 50 to 500 VAC, 10 cps to 250 kc.
4. TS-580/U, (SigC) (D-C Amplifier)	0.1 to 1.0 VDC
1.1.1.2 A-C Voltmeters	
5. Type IS-185, (SigC) (Voltmeter)	0 to 150V; 25 to 2400 cps; 0.75%
6. ME-64/U, (SigC) (Voltmeter)	0 to 150/600V; 25 to 125 cps; 2%
1.1.1.3 D-C Voltmeters	
7. TS-443/U, (SigC) (Voltmeter)	0 to 3/150V; 0.25%
8. AN/URM-45, (BuA) (Voltmeter)	0 to 50 mv; 0 to 15/30V; 2% on 15V scale; 2½% on 30V scale
1.1.1.5 A-C D-C Voltmeters	
9. TS-340/U, (SigC) (Voltmeter)	0 to 150/300/750V; 25 to 125/1000 cps; 25 to 2500 cps; 0.5%
1.1.2.3 D-C Ammeter	
10. TS-60/U, (SigC) (Test Meter) (Also known as SigC Type I-139 Test Set)	0 to 1 ma, 75 ohm impedance; 3%
1.1.2.4 Recording Ammeters	
12. TS-584/U, (SigC) (Milliammeter Recorder)	0 to 5 ma; 0 to 10 ma with ext shunt, accuracy 1%
13. RD-49/U, (BuS) (Recorder-Milliammeter)	0 to 1 ma DC; 1%
1.1.2.5 A-C D-C Ammeter	
14. ME-65/U, (SigC) (Ammeter)	0 to 2/5/10/20/100/200 amp; Accuracy 1% DC, 1 to 2% AC at 15-135 cps
1.1.3 Multimeter	
1.1.3.1 Electronic Type Multimeter	
1.1.3.1.1 Electronic Volt-Ohm-Meter	
15. ME-25/U, (BuS) (Multimeter)	0-1/25/10/25/100/250 VDC-AC; 100/200/500/1000 ma DC; 0 to 1K/10K/100K/1M/10M/1000M ohms; DCV ±4% up to 250V, ±5% 250 to 1000V, ±6% 5000V; VAC ±5%; DC ±5%; Resist ±3° arc.
1.1.3.1.2 Electronic Volt-Ohmmeter	
16. AN/PRM-15. (SigC) (Multimeter) (small, battery operated) (Input impedance 11 megohms DC)	0-2.5/10/25/100/250/1000 VDC at 3%; 0-1K/10K/100K/1M/10M/100M ohms at 3%
17. AN/PRM-16, (SigC) (Multimeter) (small, line operated) (Input impedance 11 megohms DC)	0-2.5/10/25/100/250/1000 VDC at 3%; 0-1K/10K/100K/1M/10M/100M ohms at 3%

Equipment	*Range*
18. TS-505/U, (SigC) (Electronic Multimeter) (Input impedance 20 megohms up to 400 VDC, 50 megohms at 1000 VDC, 6 megohms shunted by 2 MMF for ACV; 4% DC, 6% AC, 4% Resistance; 0 to 500 MC on ACV)	0-2/4/10/20/40/100/200 ACV; 0-2/5/10/20/40/100/200/400/1000 DCV; 0-1000/10,000/100K/1M/10M/100M/1000M

1.1.3.2 Non-Electric Type Multimeter

1.1.3.3.1 Non-Electric Volt-Ohm-Ammeter

19. TS-297/U, (SigC) (Multimeter)	1000 ohms/volt 0-4/10/40/100/400/1000VAC & DC; 0-1K/10K/100K ohms; 0-4/40/100/400 ma DC; DC ± 3%; AC ± 5%
20. TS-352/U, (WADC) (Multimeter)	20,000 ohms/volt DC and 1000 ohms/volt AC & DC up to 1000 VDC; 0-2.5/10/50/250/500/1000 VAC and DC; 0-5000VDC; 0-0.25/2.5/10/50/100/500/2/5K/10K ma DC; 0-1K/10K/1M/10M ohms

1.1.3.2.3 Non-Electronic Volt-Ammeter

21. ME-67/U, (WADC) (Multimeter)	0-10/25/50/100/250/500 AC; 0-175/350/700 V; 50-70 cps; 3%
22. ME-1/U, (SigC) (Multimeter) (Clamp Type)	0-15/60/150/600 AC; 0-150/600 VAC; 50-70 cps; 3%
23. 1-50, (SigC) (Volt-Ammeter)	0-3/15/150 VDC; 0-3/15/30 DCa; 1%

1.2 Tester

1.2.1 Electron Tube Testers

24. TV-7/U, (SigC) (Electron Tube Test Set) (Receiving tubes)
25. AN/USM-23, (BuS) (Electron Tube Test Set) (Geiger-Mueller and Voltage Reg. Tubes)
27. TV-2/U, (SigC) (Tube Tester) (Receiving and small transmitting tubes)
28. AN/USM-31, (BuS) (Electron Tube Test Set) (Receiving and small transmitting tubes)
29. AN/UAM-( ) (ERDL) (Image Converter Tube Test Set)

1.3 Combined Tester and Meter

1.3 Tube Tester-Meter

30. TV-6/U, (SigC) (Electron Tube Test Set) (Electrometer tubes, including high resistance ohmmeter)

2. FREQUENCY METERS

GROUP I. Accuracy + .00001% or + 1 cycle whichever greater

32. AN/URM-18, (SigC) (Frequency Calibrator)	10 kc to 100 mc .0005%
33. FR-38/U, (BuA) (Frequency Meter)	0-50 mc Accuracy: same as Group II, ± one count
34. FR-67/U, (SigC) (Frequency Meter)	20 cycles to 100,000 cycles Accuracy, same as Group II ± one count
35. AN/USM, (BuS) (Frequency Meter)	15 kc to 30 mc 0.005%
36. FR-4/U, (SigC) (Frequency Meter)	100 kc to 20 mc .001%
37. FR-5/U, (SigC) (Frequency Meter)	10 mc to 100 mc .001%
38. FR-6/U, (SigC) (Frequency Meter)	100 mc to 500 mc .001%

Equipment	*Range*

GROUP III. Accuracy + .01% to + .001%

41. TS-186/UP, (BuS) (Frequency Meter)	100 mc to 10,000 mc

GROUP IV. Accuracy + .05% to + .01%

42. TS-174/UP, (SC) (Frequency Meter)	20 mc to 280 mc, .04%
43. TS-175/U, (SC) (Frequency Meter)	80 mc to 1,000 mc, .03%
44. BC-221, (SigC) (Frequency Meter)	125 kc to 20 mc, .02% (SCR-211)

GROUP V. Accuracy Less Than 0.05%

45. FR-63/U, (SigC) (Frequency Meter)	25 cycles to 60 kc, 2%
46. TS-494/U, (WADC) (Frequency Meter)	49 cycles to 51 cycles, 0.5%
47. I-209, (SigC) (Frequency Meter)	58 cycles to 62 cycles, 3%
48. TS-328/U (SigC) (Frequency Meter)	380 to 420 cycles, 3%
49. TS-117/GP, (SigC) (Wavemeter Test Set)	2,400 mc to 3,400 mc, 0.5%
50. AN/PRM-10, (RADC) (Test Oscillator Set) (Grid-Dip Meter)	2-400 mc, 1-12%
51. FR-39/U, (SigC) (Wavemeter)	55 to 400 mc ± 2%, Absorption Type
52. FR-40/GSM-1, (SigC) (Frequency Meter)	50 to 70 cps ± 0.5%
53. I-129, (SigC) (Frequency Meter Set)	1.5 to 41 mc ± 3%, Absorption Type
54. TS-480/U, (SigC) (Frequency Meter)	0.5 to 16 mc ± 2%, 16-150 mc ± 3%, Absorption Type

3. WAVE FORM MEASURING EQUIPMENT

3.1 Cathode Ray Oscilloscope

74. OS-8/U, (BuS) (Oscilloscope)
(Technician type Communications)

3.2 Synchroscope

75. TS-34/AP, (BuS) (Oscilloscope)
76. AN/USM-24, (BuS) (Oscilloscope)
(Bench type; Supersedes TS-239/UP)

3.4 Analyzer

3.4.1 Spectrum Analyzer

80. AN/UPM-33, (BuA) (Spectrum Analyzer)	8,970-9,630 mc, Formerly TS-148/UP
81. TS-333/AP, (BuA) (Spectrum Analyzer)	23.5-24.5 kmc
83. AN/UPM-17, (WADC) (Spectrum Analyzer Set)	10-16,500 mc, fixed

4. SIGNAL GENERATING EQUIPMENT

4.1 Signal Generators

4.1.1 Signal Generators—Unmodulated

GROUP A. Up to 30 KC (VLF)

84. TS-382A/U, (WADC) (Audio Oscillator)	20-200,000 cps, 2% accuracy
86. TS-535/U, (BuS) (Signal Generator)	7-160 kc, 0.3% accuracy

GROUP B. 30–300 MC (VHF)

87. TS-47/APR, (SigC) (Test Oscillator)	40-500 mc, 1%

Equipment	*Range*

4.1.2 Signal Generators, Modulated

4.1.2.1 Amplitude or Pulse Modulated

GROUP A. 30 KC–30 MC (LF, MF, HF)

Equipment	Range
88. AN/URM-25, (BuS) (RF Signal Generator Set)	10 kc - 50 mc
89. TS-606/U, (WADC) (Signal Generator)	85 kc - 40 mc, 10 watts output

GROUP B. 30–300 MC (VHF)

Equipment	Range
90. AN/URM-26, (BuS) (RF Signal Generator Set)	3-405 mc
91. TS-497/URR, (SigC) (Signal Generator)	2-400 mc
92. TS-510/UR, (BuA) (Signal Generator)	25-440 mc
93. TS-608/U, (WADC) (Signal Generator)	40-400 mc, 5 watts output

GROUP C. 300–3,000 MC (VHF)

Equipment	Range
94. TS-406/U, (BuA) (Test Oscillator)	1,000-3,500 mc; Buzzer type shock excited, battery operated, 5 lb.
97. AN/URM-49, (BuA) (Signal Generator)	400-1,000 mc, formerly TS-418/U
98. AN/URM-64, (BuA) (Signal Generator)	900-2,000 mc, formerly TS-419/U
99. AN/URM-61, (BuA) (Signal Generator)	1,800-4,000 mc, formerly TS-403/U

GROUP D. 3,000–30,000 MC (SHF)

Equipment	Range
100. TS-508/U, (BuA) (Test Oscillator)	3,000-11,000 mc
103. AN/URM-52, (BuA) (Signal Generator)	3,800-7,500 mc, formerly TS-621/U
104. AN/URM-44, (BuA) (Signal Generator)	7,000-10,500 mc, formerly TS-622/U

4.1.2.2 Frequency or Sweep Modulated

Equipment	Range
108. TS-452/U, (RADC) (Signal Generator)	5-100 mc, maximum sweep 5-55 mc in 6 steps depending on center frequency
109. AN/URM-48, (SigC) (Signal Generator)	20-100 mc, dev. 0-15 kc internal; 0-50 kc external. Includes CW Crystal-controlled IF Frequency, formerly SG-12/U

4.1.2.3 Combined AM or PM and FM or Sweep Modulated

Equipment	Range
113. ————, (BuA)	Boonton Type 207A 0.1-55 mc; 0.3-55 mc with 200 kc carrier deviation, used with Boonton Type 202-B
115. ————, (BuA) (FM Signal Generator)	Boonton Type 202-B; 54-216 mc; FM deviation 0-80/240 kc; AM 30% and 50% calibrated, variable 0-50% includes AF Oscillator with 8 fixed frequencies for either AM or FM modulation
116. AN/USM-16, (WADC) (Signal Generator)	10-440 mc, formerly TS-437(XA)/U, deviation ± -.- 125 to ± 7.5 MC

4.3 Pulse and Square Wave Generators

4.3.1 Pulse Generators

Equipment	Range
117. AN/USM-27, (BuS) (Signal Generator)	0.5 to 11 usec; 1 to +40V at 70 ohms, 1 to +50V at 100 ohms and 1 to −70V at 100K ohms, 100 to 400 pps

Equipment	*Range*
118. SG-30/UP, (BuA) (Pulse Generator)	0.2-20 usec, 5-50V into 5,000 ohms when shunted by 2,000 mmf, 20-20,000 pps, includes scope
119. AN/UPM-15, (WADC) (Pulse Generator)	0.15 to 110 usec; 0 to ±200V; Impedance − 0.25 ohms below 20 mv, 2.5 ohms from 20 to 200 mv, 50 ohms from 0.2 to 2V, 250 ohms from 2 to 20V, 2,500 ohms from 20 to 200V; single and double pulses; 30 to 10,000 pps; formerly TS-592/U

4.3.2 Square Wave Generators

120. TS-583/U, (SigC) (Square Wave Generator)	20-10,000 cps impedance 500 ohms each side of ground; Attenuator 0-75 db in 5 db steps; max output 75V

5. FIELD INTENSITY MEASURING EQUIPMENT

5.6 Noise Field Intensity Meter

121. AN/PRM*L°, (BuS) (Radio Test Set)	0.15-25 mc
124. TS-587A/U, (BuS) (Noise-Field Intensity)	25-400 mc
125. AN/URM-17, (BuS) (Radio Test Set)	375-1,000 mc
126. AN/URM-6, (BuS) (Radio Test Set)	14-250 kc
127. AN/URM-3, (SigC) (Interference Measuring Set)	0.15-40 mc

6. IMPEDANCE AND STANDING WAVE RATIO MEASURING EQUIPMENT

6.1 Impedance Measuring Equipment

6.1.1 Resistance Meters

6.1.1.1 Low Voltage Ohmmeters

130. ZM-17/U, (WADC) (Ohmmeter)	0.005-5 ohms, 2%, battery-operated
131. ZM-4/U, (SigC) (Resistance Bridge)	0.001-100,000 ohms, up to 1 megohm with ext. batt., 0/2%
132. Navy Type 60089, (BuS) (Vacuum Tube Megohmmeter)	0-10,000 megohms, battery-operated
133. ZM-9/U, (SigC) (Resistance Bridge)	0.1-10,000,000 megohms in five ranges; AC operated

6.1.1.2 Working Voltage Ohmmeters

134. AN/PSM-1, (BuS) (Insulation Test Set)	0-100 megohms, 500V, hand crank
135. I-48, (SigC) (Test Set)	2-1,000 megohms, 500V, hand crank
136. AN/PSM-2, (BuS) (Insulation Test Set)	0-1,000 megohms, 500V, hand crank

6.1.3 Capacitance Meter

137. ZM-3/U, (SigC) (Analyzer)	5uuf-10,000 uf: 1.1-100/10,000 megohms; leakage current 0-50 ma; 0-600V

6.1.4 Combination Type Impedance Measuring Equipment

6.1.4.4 Q Meters

138. TS-617/U, (SigC) (Q Meter)	50 kc-75 mc, similar to Boonton 160A
139. ————, (BuS) (Q Meter)	30-210 mc, similar to Boonton 170A

Equipment	*Range*
6.1.4.5 Combination R.C.L. Meter	
140. TS-460/U, (SigC) (Impedance Bridge)	1,000 cycles-0.001-11 megohms, 1 uuf-1,100 uf, 1 mh-1,100 h
141. ZM-11/U, (BuS) (Capacitance, Inductance, Resistance Bridge)	10 mmf-1,000 uf, 1 ohm-10 megohms, 200-1000/10,000 megohms insulation resistance, 100 mh-10 h
142. Navy Type 60094, (BuS) (RF Bridge)	400 kc-50 mc; 0-1,000 ohms Resistance, 0-5,000 ohms Reactance at 1 mc
144. TS-269 ( )/UR, (SigC) (Test Set)	20-1,000 kc, 0-111 ohms, 40-4,100 uuf, ± 5%

10. COMBINATION AND GROUP TEST SET

10.1 Combination Test Set (One Main Case)

(Includes signal generator, frequency, and power meter unless otherwise indicated)

145. TS-155/UP, (SigC) (Signal Generator)	2,700-3,400 mc, does not include frequency meter
148. TS-147/UP, (BuS) (Test Set)	8,520-9,600 mc, frequency modulated
150. TS-223/AP, (BuA) (Test Set)	23,500-24,500 mc
151. AN/UPM-14, (BuA) (Radar Test Set)	34,000-35,600 mc

11. ASSOCIATED DEVICES

11.12 Voltage Dividers	
194. TS-89/AP, (WADC) (Voltage Divider)	20 KV, 100:1 & 10:1, 15% accuracy
195. TS-265/UP, (SigC) (Voltage Divider)	50 KV, 100:1 & 10:1
196. TS-453/UP, (RADC) (Voltage Divider)	100 KV; 500:1, 100:1, and 10:1, 5% accuracy
11.13 Electronic Switch	
197. TS-433 ( )/U, (SigC) (Electronic Switch)	10-2,000 cps switching range, 0-25 kc response, ± 1 db, 10X Gain

12. MISCELLANEOUS TEST EQUIPMENT

12.6 Telephone, Telegraph, and Teletypewriter Test Equipment

12.6.1 Teletypewriter Test Equipment

198. TS-658/UG or TS-659/UG, (BuS) (Teletypewriter Test Set)
199. TS-383/GG, (SigC) (Distortion Test Set)
200. TS-2/TG, (SigC) (Test Set)
201. Sig C Type I-193, (SigC) (Test Set)
202. TS-660/UG, (SigC) (Teletypewriter Test Set)
203. TS-657/FG, (BuS) (Teletypewriter Test Set)
204. ——————, (SigC) (Monitoring Printer) (Teletype Corp. Model 14)
205. TS-611/FG, (SigC) (Teletypewriter Test Set)
206. TS-577/FG, (SigC) (Telegraph Monitor)
207. Sig C Type I-181, (SigC) (Test Set)

Equipment	*Range*
208. ——————, (BuS) (Relay Test Panel) (W.E. Type 111A)	

12.12 Test Sets

12.12.8 Crystal Test Sets

211. TS-710/TSM, (SigC) (Crystal Impedance Meter)	10 kc-500 kc
212. TS-537/TSM, (SigC) (Crystal Impedance Meter)	75-1,100 kc, p/o AN/TSM-3
213. TS-330/TSM, (SigC) (Crystal Impedance Meter)	1-15 mc, p/o AN/TSM-3
214. TS-683/TSM, (SigC) (Crystal Impedance Meter)	10-100 mc

12.12.10 Thermopile, Bolometer, and Photoconductive Cell Test Sets

215. AN-UAM( ), (BuS) (Thermopiles and Bolometer Test Set)

12.19 Infrared Test Equipment

217. TS-724/SAC, (BuS) (Infrared Test Set) (Infrared Communications)
218. AN/UAM-1, (BuA) (Infrared Test Set) (Homing Equipment)
219. TS-532/ASM, (BuA) (Test Set) (Homing Equipment)

13. CALIBRATING EQUIPMENT FOR ELECTRONICS TEST EQUIPMENT

13.1.3 Multimeter Calibrator

222. TS-656( )/U, (SigC) (Meter Tester)	0-2.5/5/10/25/50/100/250/500 VDC, VAC; 0-100/250/500 DC ua; 0-1/2.5/5/10/25/50/100/250/500 DC ma; 0-12; ± 2% DC, ± 5% AC

14. POWER MEASURING EQUIPMENT

14.1 DC and Line Power Frequency Power Measuring Equipment (0-3,000 cps)

223. TS-445/U, (SigC) (Wattmeter)	DC and 25-125 cps; 5w to 15kw; 75 to 300 V, 5%
224. TS-430/U, (SigC) (Wattmeter)	0-30 kw; 25-125 cps

14.2 Audio and Radio Frequency Power Measuring Equipment

GROUP A. 3–30 KC (VLF)

225. ME-22/PCM, (SigC) (Decibel Meter)	− 45 to + 26 dbm for 600 ohm circuit; 0/2-35 kc
226. TS-585/U, (SigC) (Output Meter)	0.1-5,000 mw; 0.03-10 kc
227. TS-629/U, (BuS) (Audio Level Test Panel)	− 40 to + 20 dbm 600 ohm circuit, − 20 to + 20 dbm 12,500 ohm circuit, 0 db = 1 mw, 600 ohm circuit, 0.03-15 kc

GROUP B. 30–3,000 MC

229. TS-118/AP, (SigC) (Radio Frequency Wattmeter)	2-500 w, 20-1,400 mc; Pulse
230. TS-107/TPM-1, (BuO) (Wave and Power Meter Set)	0.5-120 mw; 900-1,350 mc; includes frequency meter; Pulse

	Equipment	*Range*
231.	TS-125/AP, (SigC) (Power Meter)	0-200 mw; 2,400-3,335 mc; Pulse
232.	TS-206/AR, (WADC) (Radio Frequency Wattmeter)	50-1,000 w; 20-60 mc
233.	AN-URM-43, (BuS) (RF Wattmeter)	0-15/60 w; 100-500 mc, formerly ME-11/U
234.	AN/URM-19, (RADC) (Bridge Summation)	0-100 mw avg, 2-20 w peak; 20-1,000 mc; Pulse
235.	AN/URM-22, (RADC) (Bridge Summation)	0-5 w avg, 1-10 kw peak; 20-1,000 mc; Pulse
236.	TS-305/UP, (WADC) (Power Meter)	215-230 mc, 0-1 kw Peak Power Pulse Peak Power Meter
237.	TS-226/AP, (WADC) (Power Meter)	400-425 mc, 0-1 kw; 500-530 mc, 0-10 kw; Peak Power Peak Power Meter
238.	AN/URM-20, (RADC) (Bridge Summation)	0-100 mw, avg. 2-20 w peak; 1,000-4,000 mc; Pulse
239.	AN-URM-23, (RADC) (Bridge Summation)	0-5 w avg; 0.1-10 kw peak; 1,000-4,000 mc; Pulse
242.	ME-27/AP, (WADC) (RF Wattmeter)	0.1-1,000 mw; 8,500-9,600 mc; Pulse
243.	TS-230/AP, (BuS) (Frequency-Power Meter)	0.1-1,000 mw; 8,500-9,600 mc; includes frequency meter; Pulse
245.	TS-254/AP, (WADC) (Power Meter)	1-632 mw; 23,520-24,480 mc; Pulse
246.	AN/URM-21, (RADC) (Bridge Summation)	0-100 mw avg; 2-20 w peak; 4,000-10,000 mc; Pulse
247.	AN/URM-24, (RADC) (Bridge Summation)	0-5 w avg; 0.1-10 kw peak; 4,000-10,000 mc; Pulse

APPENDIX I

WORKING AND SUPPLEMENTARY DATA

APPENDIX I

WORKING AND SUPPLEMENTARY DATA

The nomograms included in this section are intended to be used to simplify mathematics, as well as an aid in understanding the various factors that may affect radio propagation.

NOMOGRAM A. Path attenuation. To be used in determining free space attenuation between identical antennas. (Chapter 2.)

NOMOGRAM B. Attenuation due to rainfall. This nomogram is designed to be used in determining attenuation due to rainfall where the maximum rate of rainfall is known. Results are based on the assumption that such a rate of rainfall is constant over the entire radio path. Since such a condition is extremely unlikely, results obtained from the nomogram represent the worst possible condition that may be expected. (Chapter 4.)

NOMOGRAM C. Reliability vs. fade margin. This nomogram will give the percentage of outage time due to fading, where the fade margin is known. It is based on conditions found on the North American continent, and may not be satisfactory for locations with unusual fade conditions. (Chapter 5.)

NOMOGRAM D. Characteristic impedance, open wire lines. Self-explanatory. (Chapter 6.)

NOMOGRAM E. Characteristic impedance, coaxial lines. Self-explanatory. (Chapter 6.)

NOMOGRAM F. Input reactance. "Wavelength of the transmission line" in this nomogram refers only to the fractional part of a wavelength, and not to the total number of wavelengths. Since the nomogram reads only to 1/4 wavelengths, treat shorted half-wavelength lines as open-ended quarter-wavelength lines, and open half-wavelength lines as shorted quarter-wavelength lines. (Chapter 6.)

NOMOGRAM G. Line matching impedance. While this nomogram applies equally well to coaxial cable and open wire line, it is most useful for open wire line. Having determined the necessary impedance of the matching section, Nomogram D may be used to determine the wire size and spacing of the matching section. (Chapter 6.)

NOMOGRAM H. Fresnel-zone clearance. This nomogram gives the required clearance above an obstruction. The profile of the radio path will give the actual clearance, which should be approximately the same as the value given in the nomogram. (Chapter 14.)

NOMOGRAM J. Radio and optical line-of-sight. Self-explanatory. (Chapter 14.)

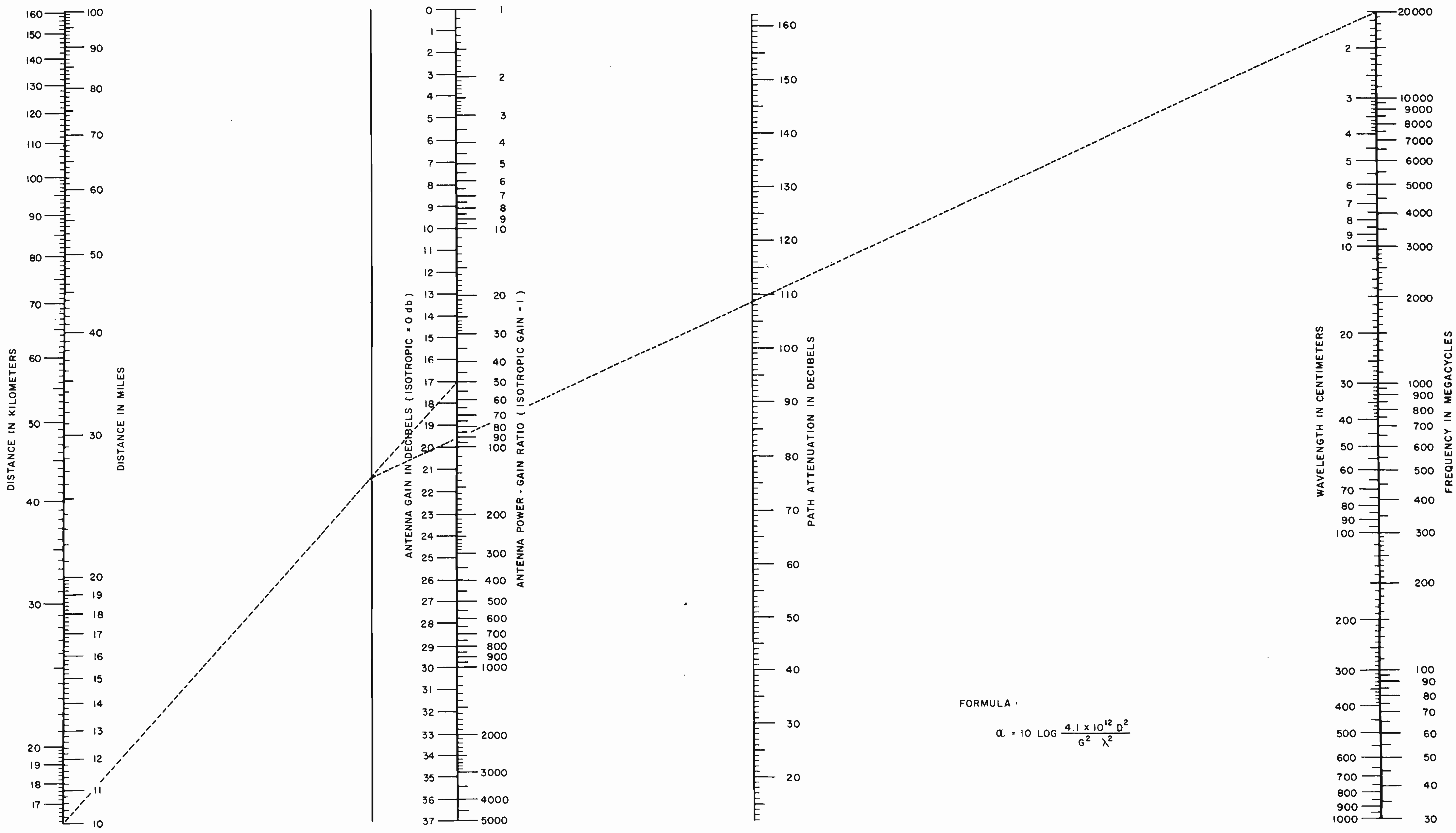
NOMOGRAM K. Parabolic antenna gain. Nomogram to be used in selecting parabolic reflectors where the use of a large reflector will give sufficient field strength at the receiver. This nomogram compensates for effective antenna area. (Chapter 7.)

NOMOGRAM L. Path attenuation between parabolic antennas. Free-space path attenuation under varied conditions of path length, frequency, and parabolic reflector size. (Chapter 7.)

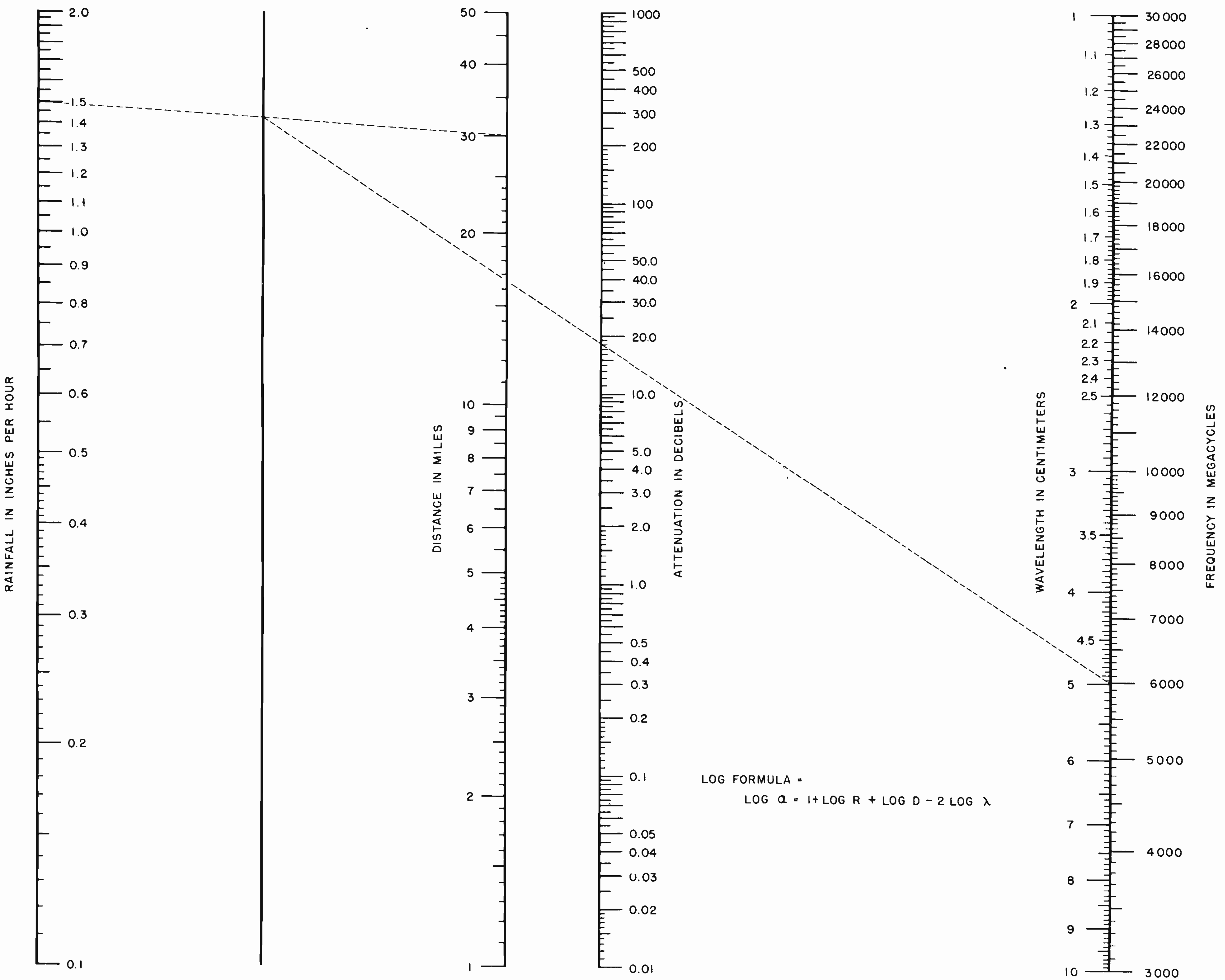
NOMOGRAM M. Increment to be added to plane earth profile for curved earth profile. Supplies correction figures to enable plotting a normal curved earth profile on rectangular coordinate paper. (Chapter 14.)

NOMOGRAM N. Wind pressure vs. velocity. Use this nomogram to determine wind loading in pounds per square inch on antenna towers. Most useful in areas where wind of hurricane velocity may occur. (Chapter 14.)

NOMOGRAM O. Conversion of dbm to microvolts. Self-explanatory.



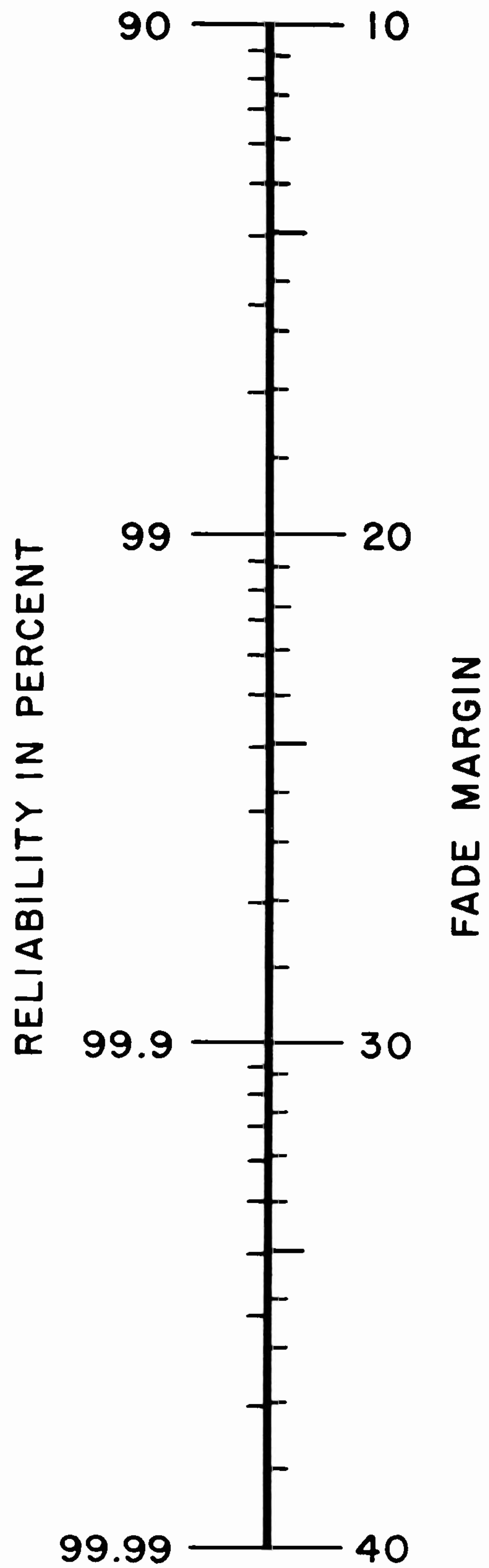
Nomogram A. Path Attenuation



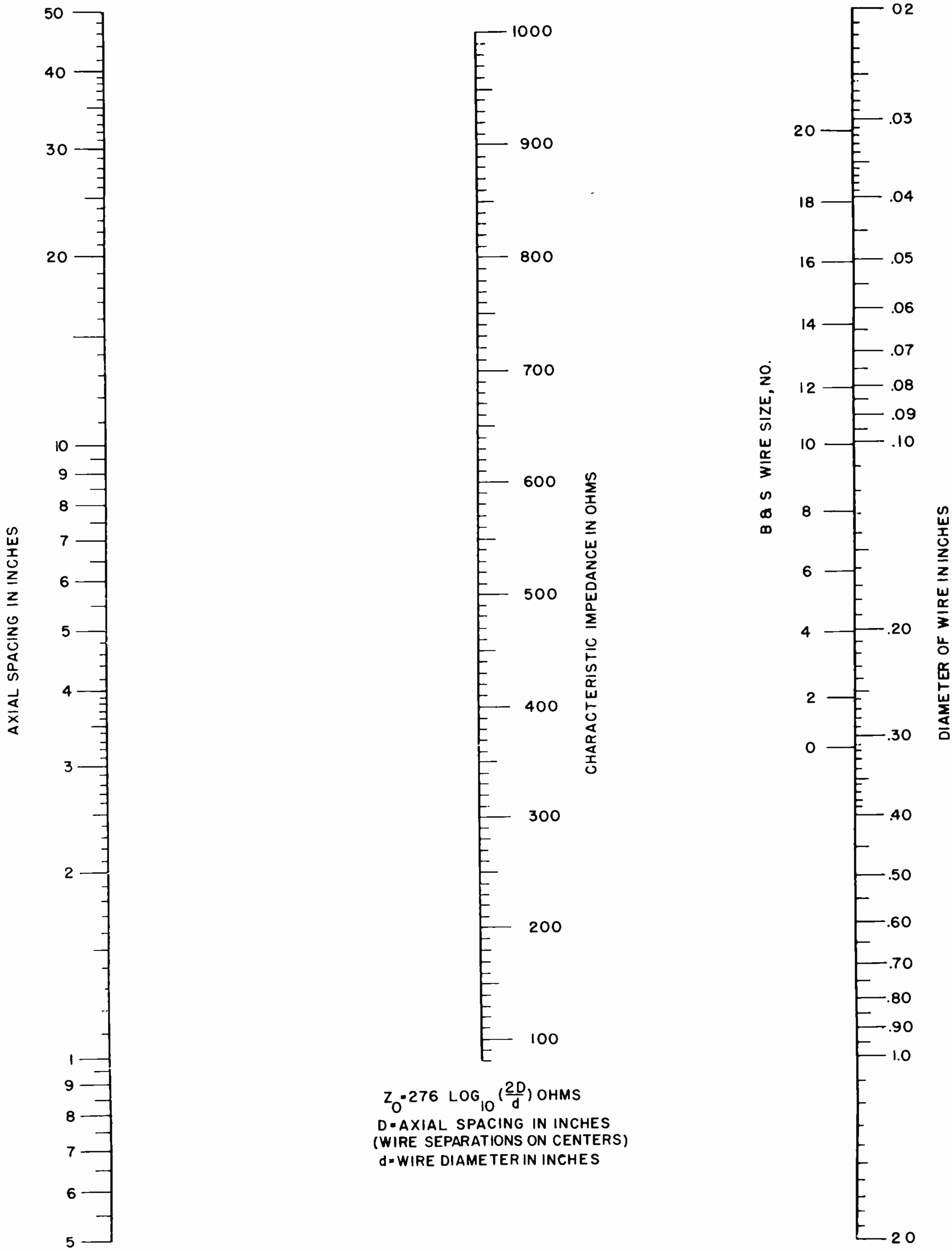
LOG FORMULA =

$\log \alpha = 1 + \log R + \log D - 2 \log \lambda$

Nomogram B. Attenuation Due to Rainfall



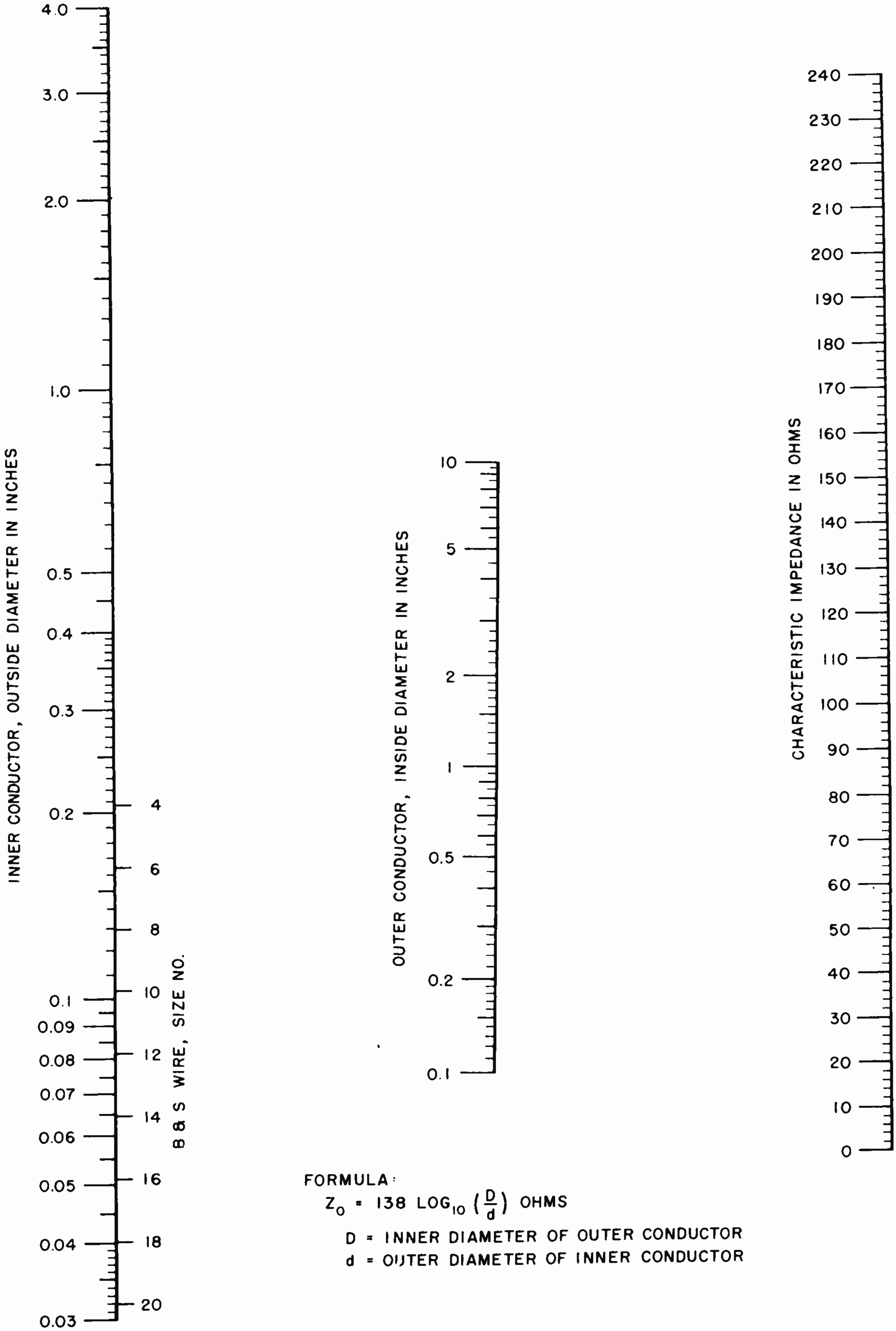
Nomogram C. Reliability



$Z_0 = 276 \text{ LOG}_{10} \left(\frac{2D}{d}\right)$ OHMS

D = AXIAL SPACING IN INCHES
(WIRE SEPARATIONS ON CENTERS)
d = WIRE DIAMETER IN INCHES

Nomogram D. Characteristic Impedance of Open Wire Lines



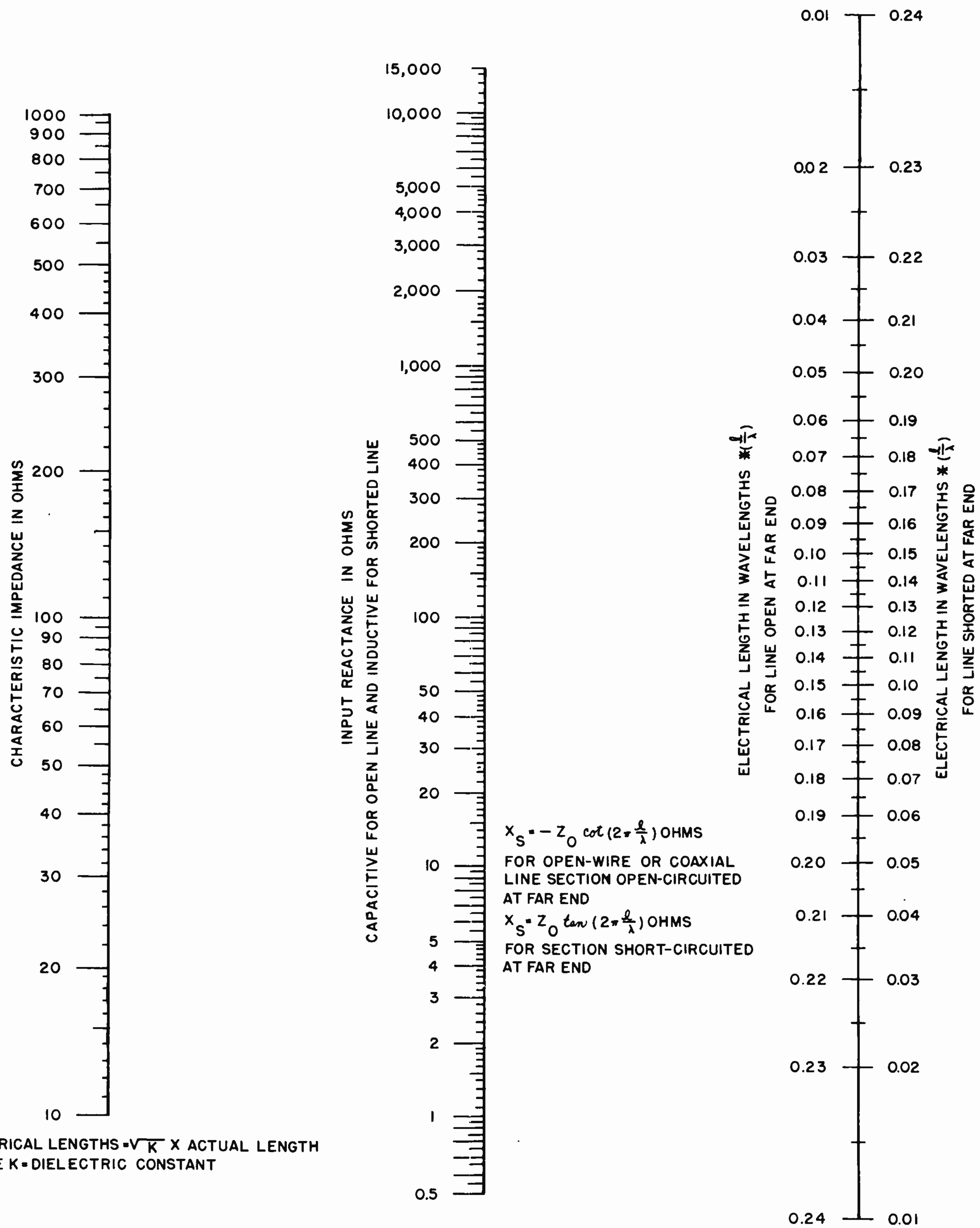
FORMULA:

$Z_0 = 138 \ LOG_{10} \left(\frac{D}{d}\right)$ OHMS

D = INNER DIAMETER OF OUTER CONDUCTOR

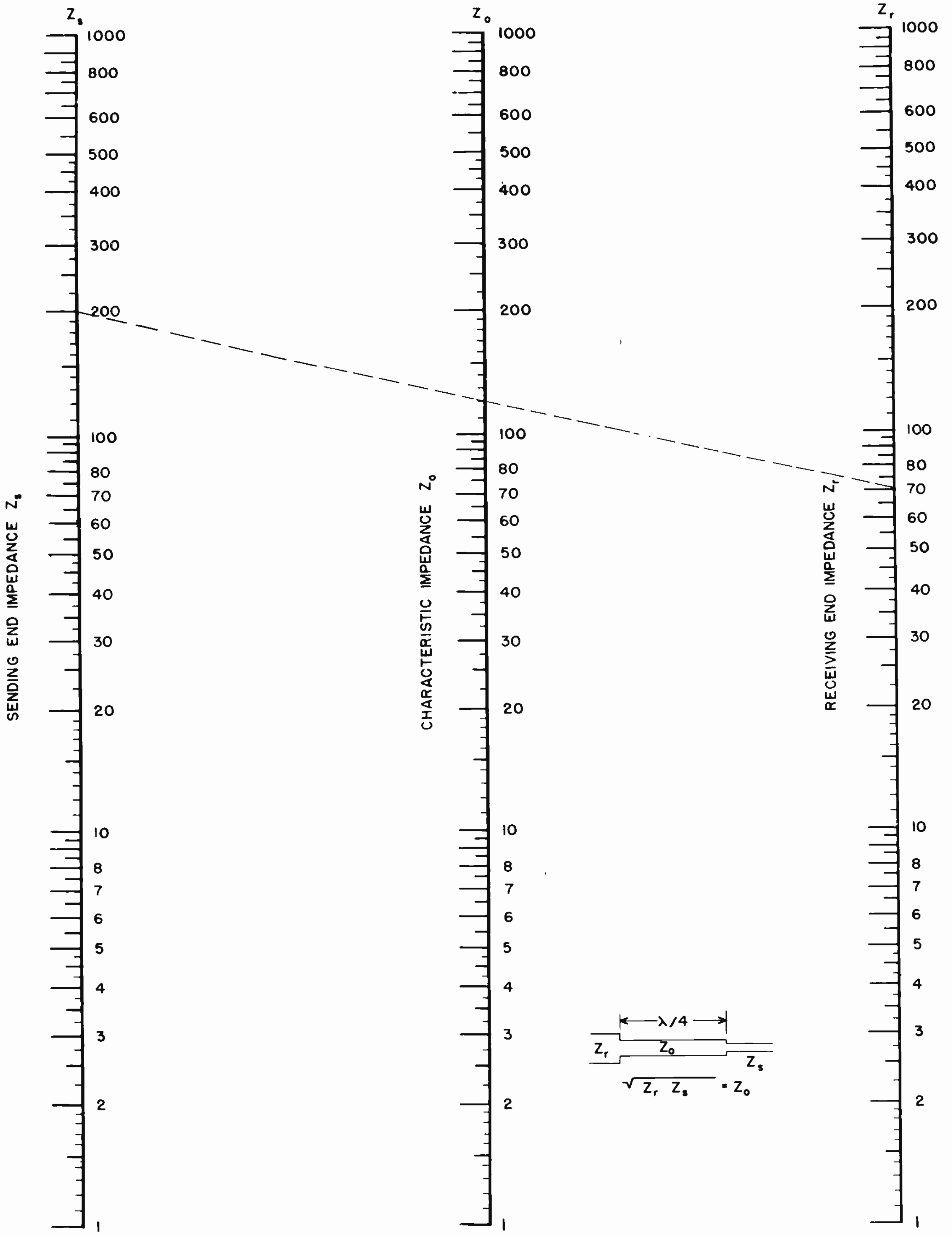
d = OUTER DIAMETER OF INNER CONDUCTOR

Nomogram E. Characteristic Impedance of Air-Filled Coaxial Lines

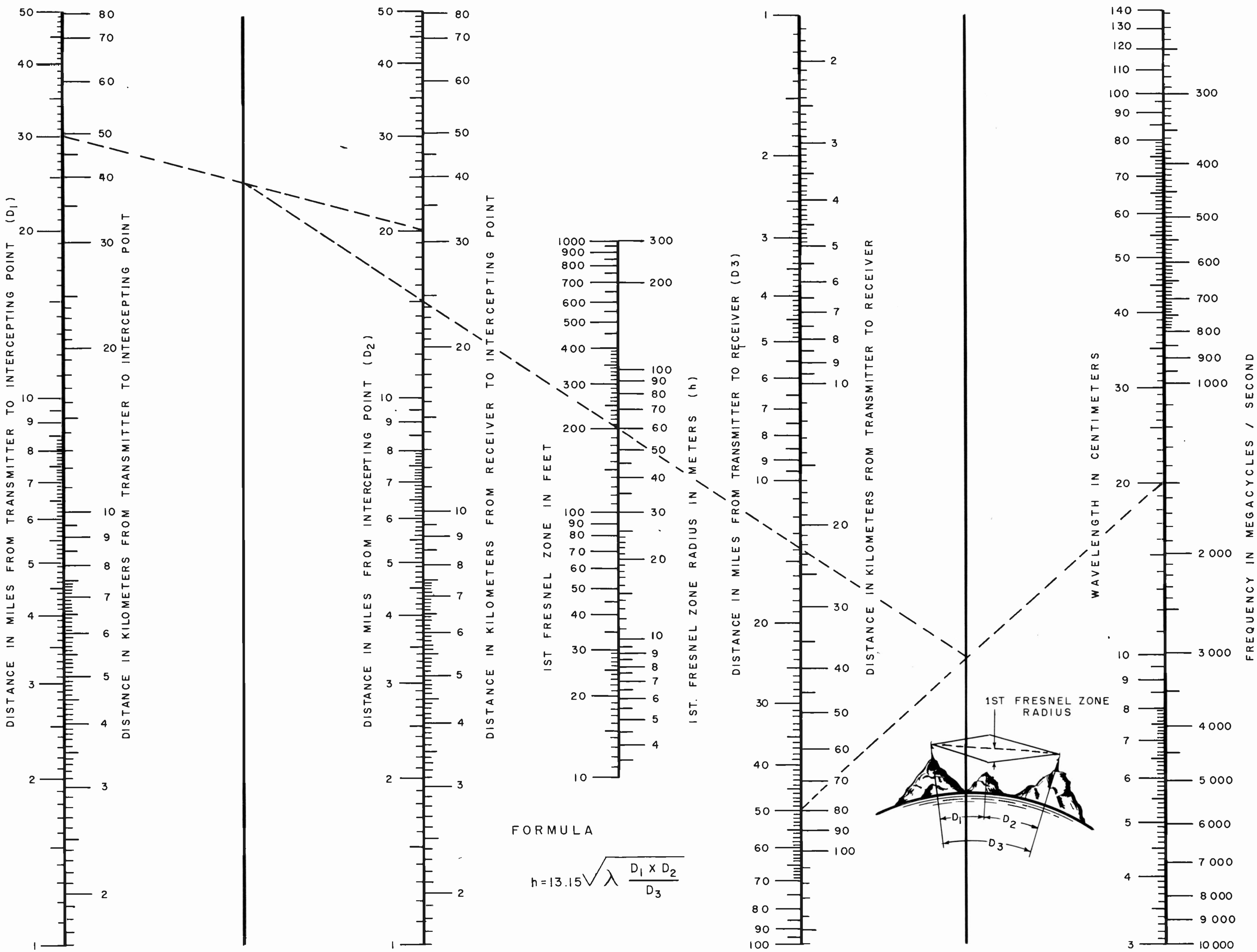


*ELECTRICAL LENGTHS = $\sqrt{K}$ X ACTUAL LENGTH
WHERE K = DIELECTRIC CONSTANT

Nomogram F. Input Reactance for Open Wire and Coaxial Lines

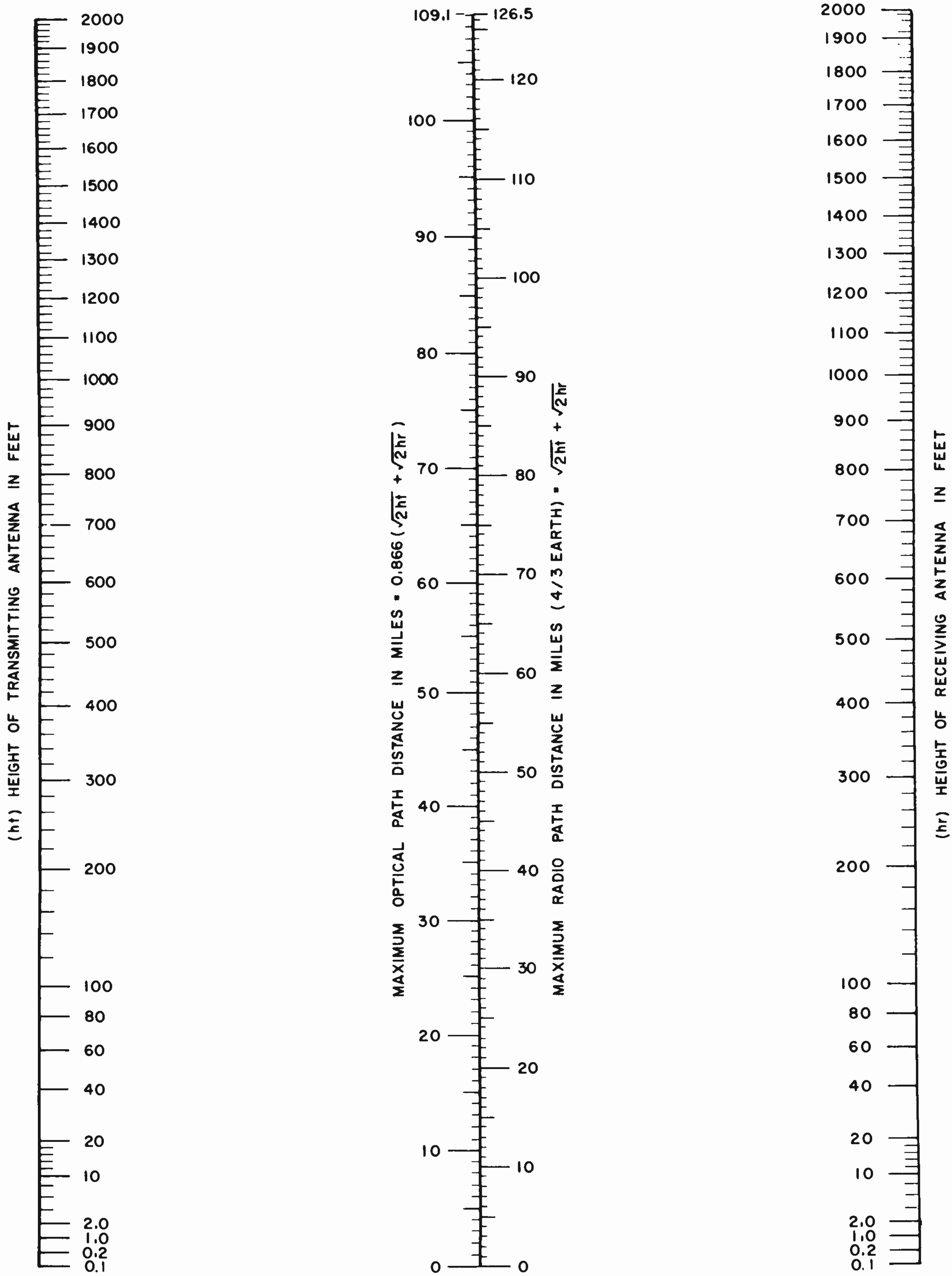


Nomogram G. Characteristic Impedance of a Quarter Wave Section Used as an Impedance Transformer

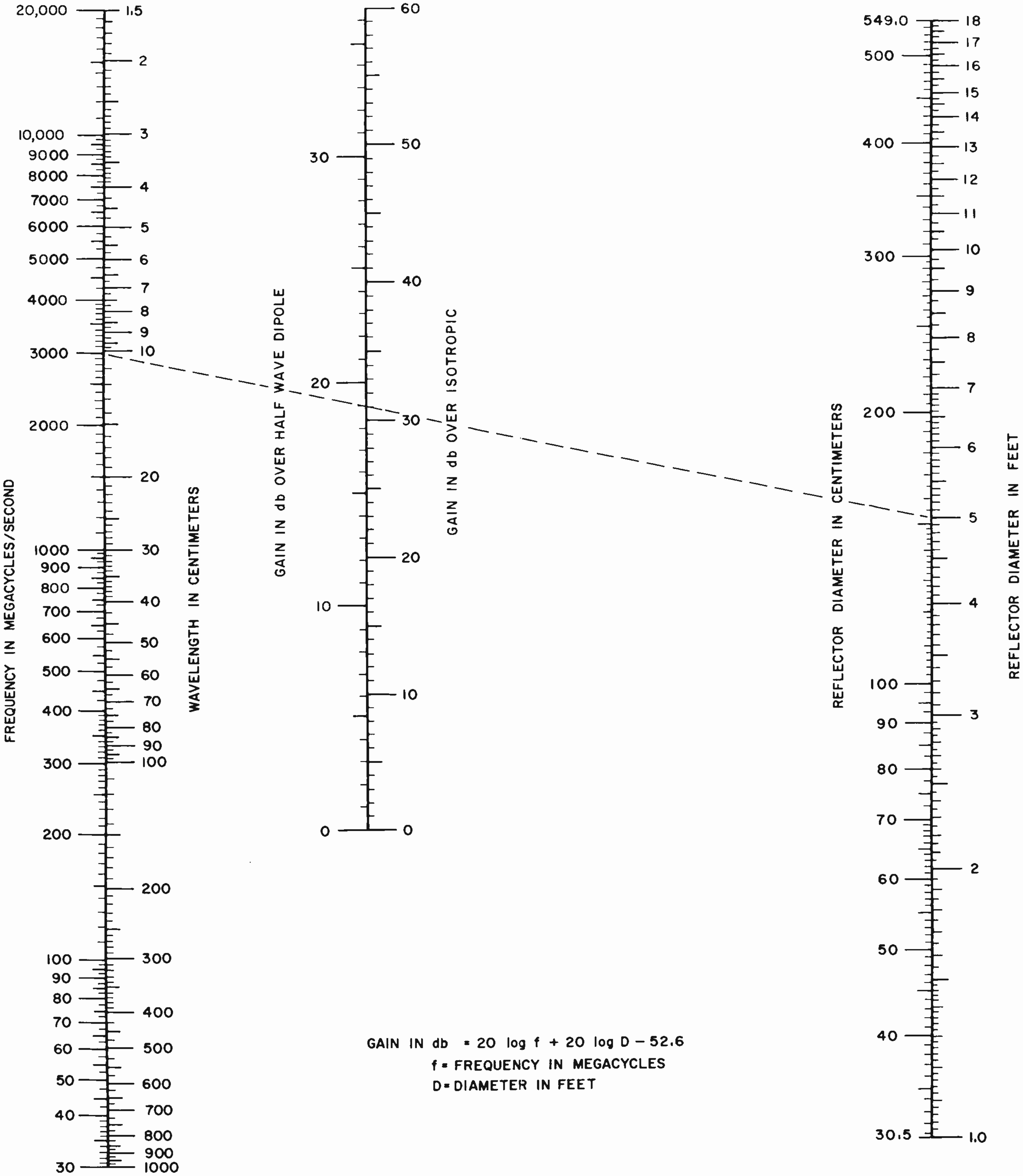


$$h = 13.15\sqrt{\lambda \frac{D_1 \times D_2}{D_3}}$$

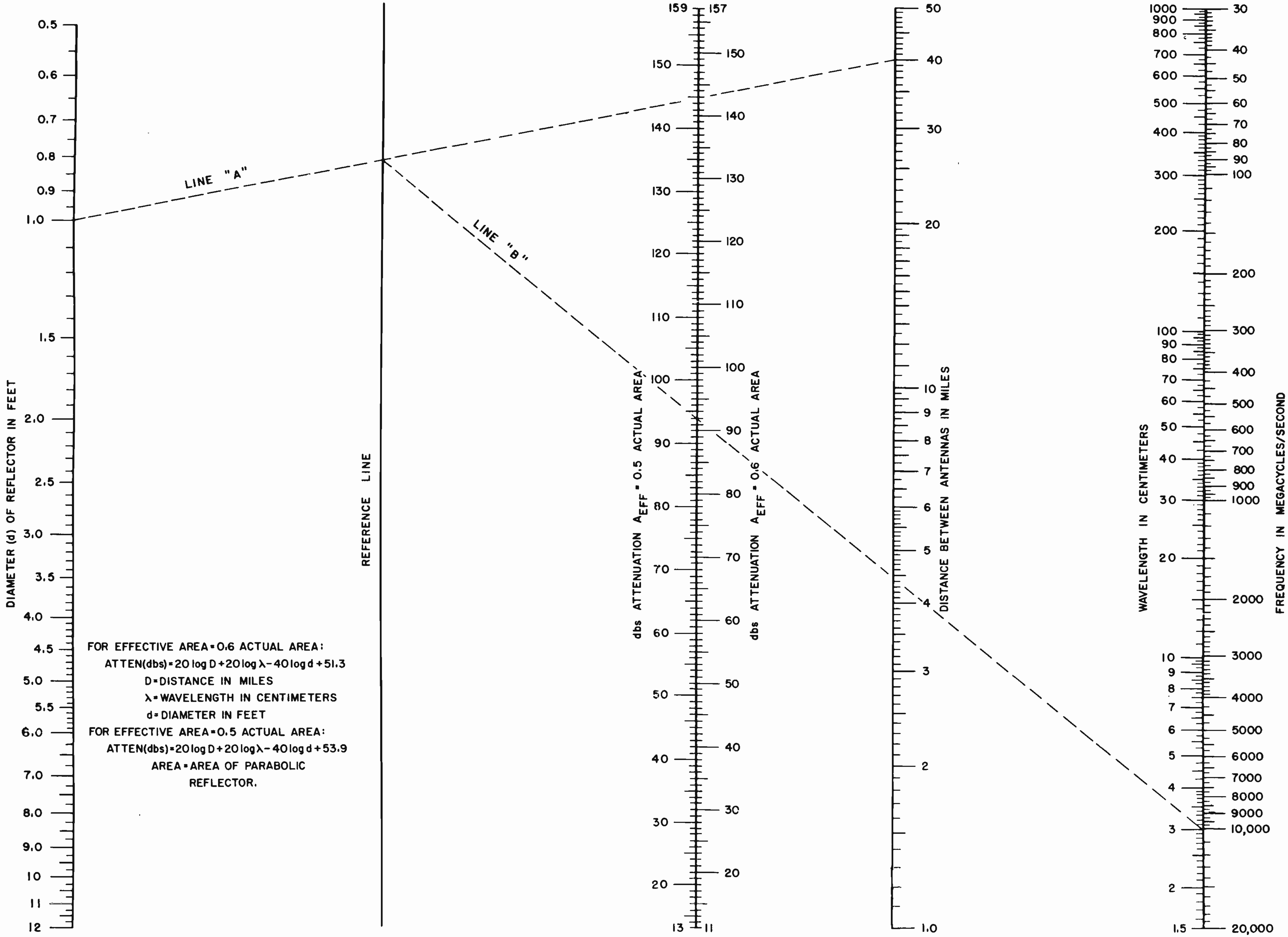
Nomogram H. First Fresnel Zone Clearance



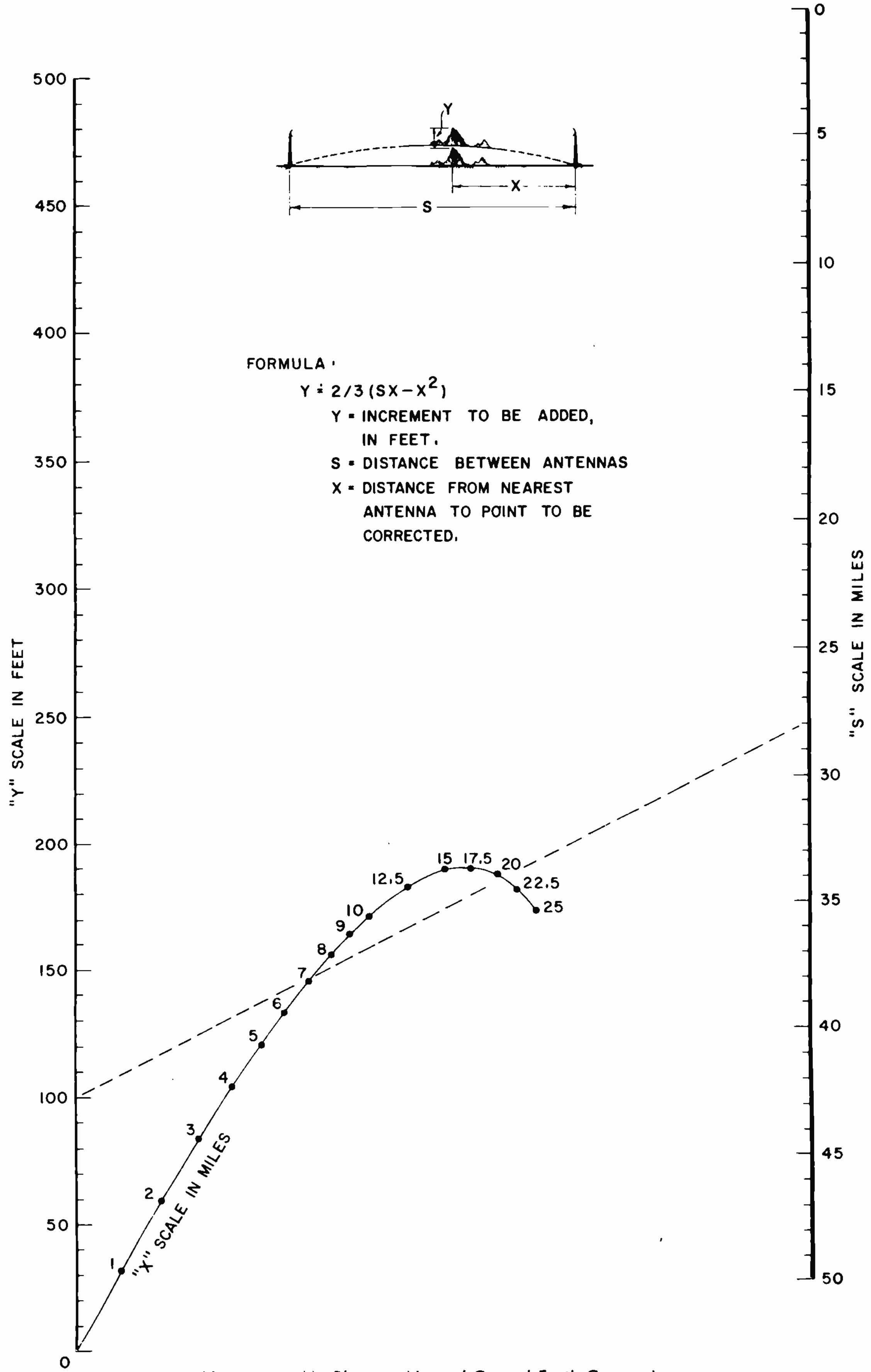
Nomogram J. Radio & Optical Line-of-Sight Distances



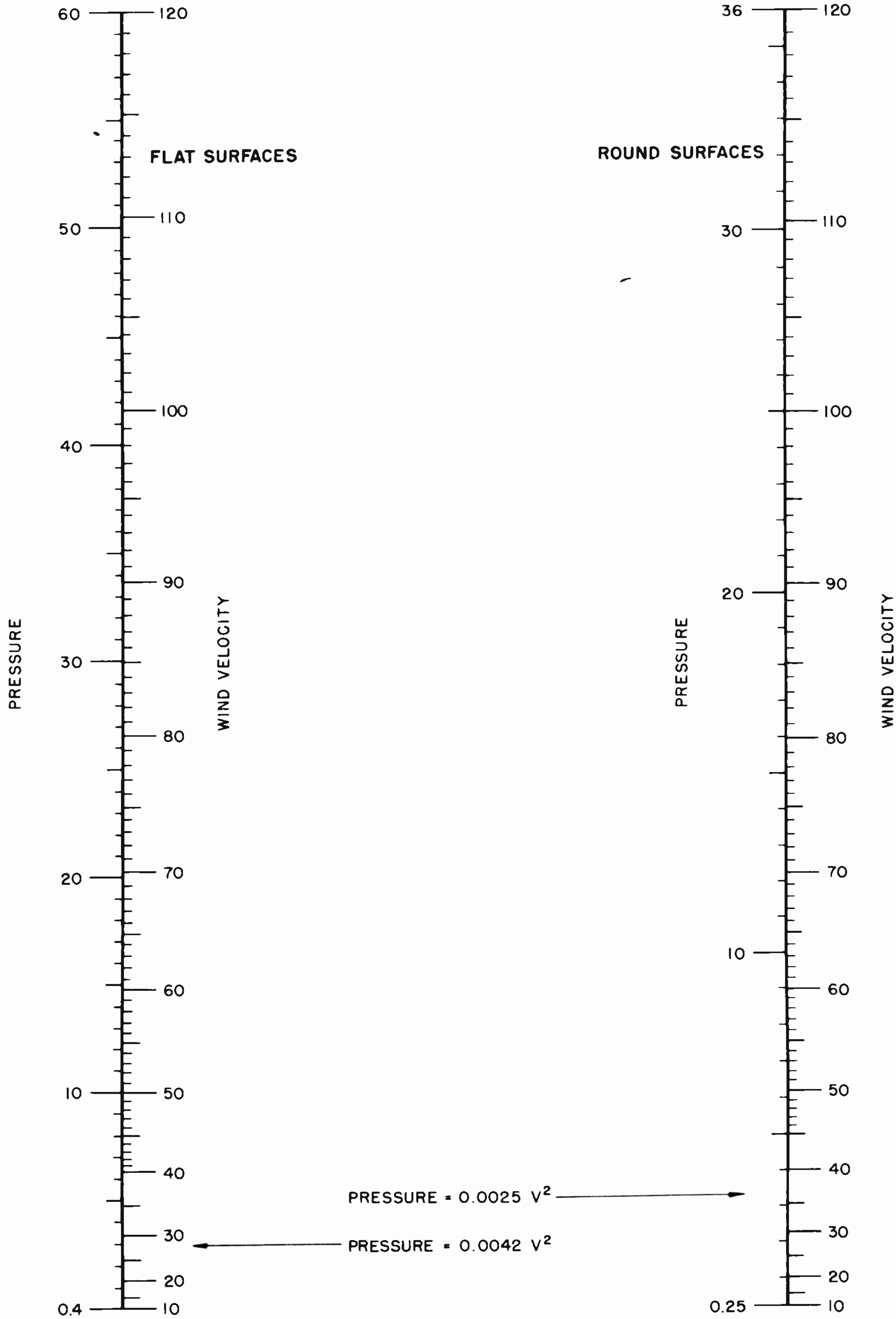
Nomogram K. Gain of Parabolic Antennas



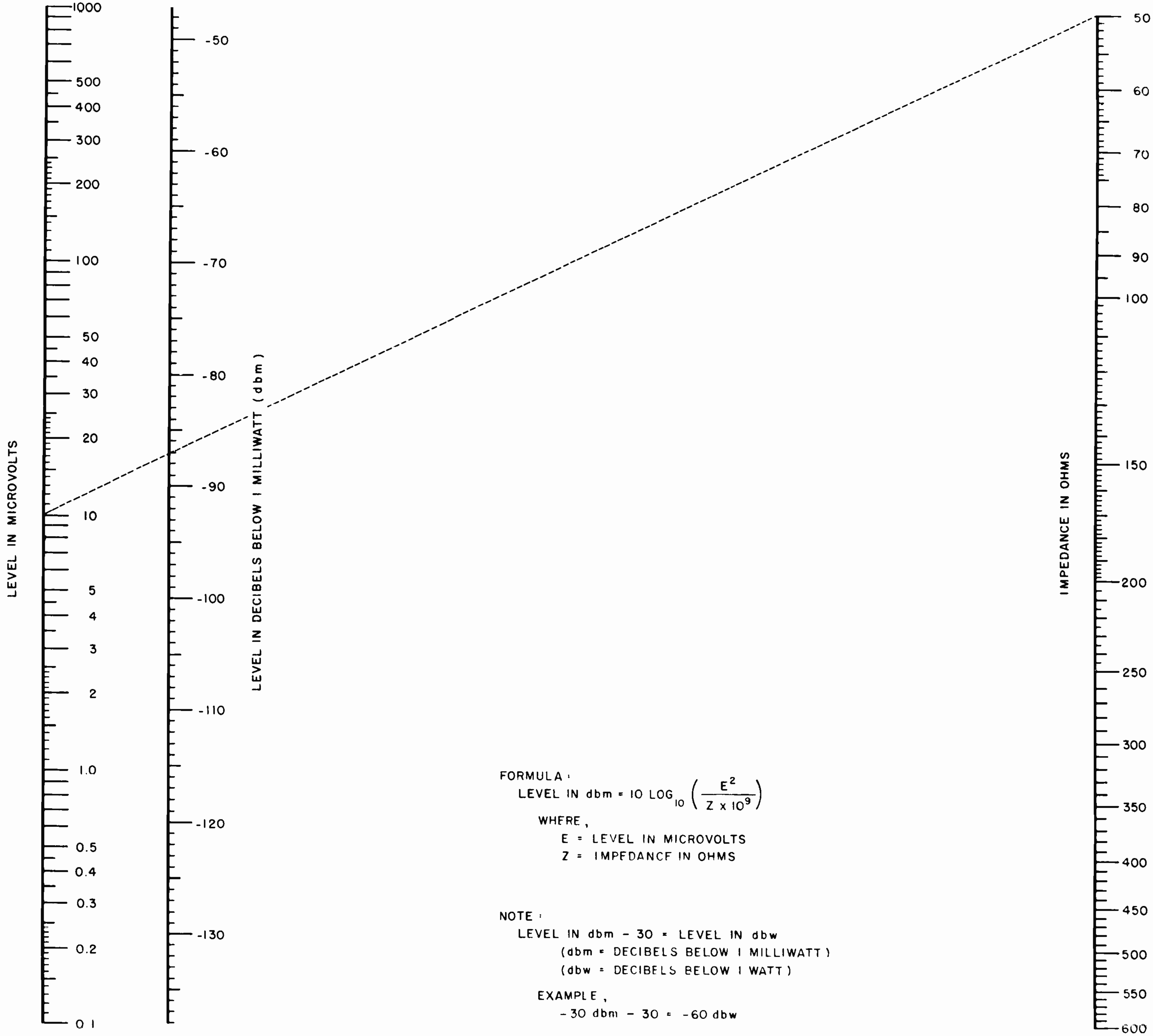
Nomogram L. Attenuation Due to Distance and Wavelength



Nomogram M. Plane to Normal Curved Earth Conversion



Nomogram N. Pressure vs Wind Velocity



FORMULA:

$$\text{LEVEL IN dbm} = 10 \log_{10}\left(\frac{E^2}{Z \times 10^9}\right)$$

WHERE,

E = LEVEL IN MICROVOLTS

Z = IMPEDANCE IN OHMS

NOTE:

LEVEL IN dbm − 30 = LEVEL IN dbw

(dbm = DECIBELS BELOW 1 MILLIWATT)

(dbw = DECIBELS BELOW 1 WATT)

EXAMPLE,

−30 dbm − 30 = −60 dbw

Nomogram O. Dbm Conversion to Microvolts

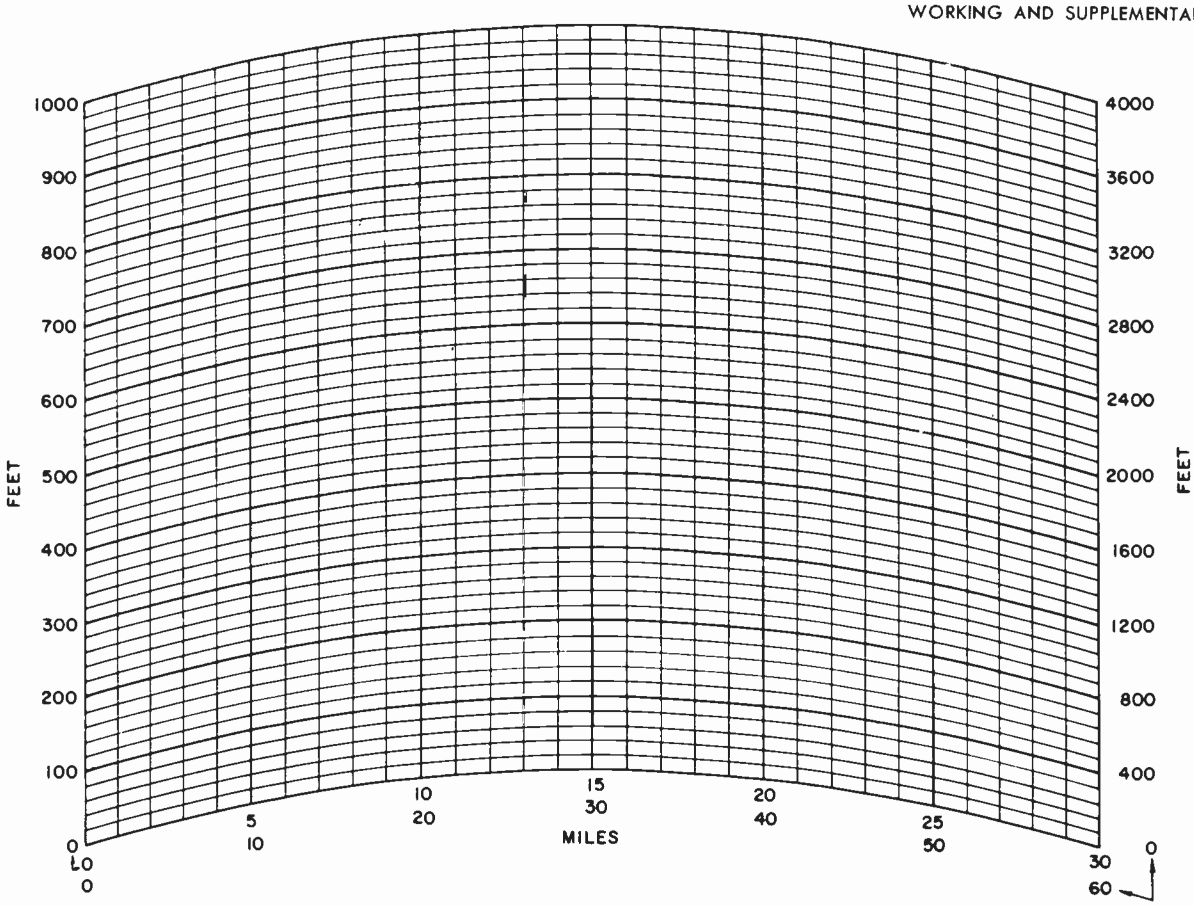


Chart P. 4/3 Earth's Profile

FREQUENCY RANGE: MEGACYCLES		30 - 300 MC	300 - 3,000 MC	3,000 - 30,000 MC	30,000 - 300,000 MC
WAVELENGTH: CENTIMETERS		1,000 - 100 CM	100 - 10 CM	10 - 1.0 CM	1.0 - 0.1 CM
BANDS			P L S	X K	Q V
DESIGNATION		VERY-HIGH FREQUENCY VHF	ULTRA-HIGH FREQUENCY UHF	SUPER-HIGH FREQUENCY SHF	EXTREMELY - HIGH FREQUENCY EHF
CHARACTERISTICS	PROPAGATION LIMITATIONS	4/3 EARTH RADIUS - SOME SPORADIC E AND ATMOSPHERIC REFLECTION	4/3 EARTH RADIUS - SPORADIC ATMOSPHERIC REFLECTION	4/3 EARTH RADIUS - SOME ATMOSPHERIC ABSORPTION	4/3 EARTH RADIUS - EXTREME ATMOSPHERIC ABSORPTION
	PREDOMINANT NOISE	MAN - MADE NATURAL	RECEIVER - NOISE MAN - MADE	RECEIVER - NOISE	RECEIVER - NOISE
	R.F. SOURCE	VARIOUS	TRIODE	MAGNETRONS KLYSTRONS	MAGNETRONS KLYSTRONS
	TRANSMISSION LINES	WIRE COAXIAL	COAXIAL WAVEGUIDE	WAVEGUIDE COAXIAL	WAVEGUIDE
	DRIVEN ANTENNA	DIPOLE	DIPOLE	DIPOLE HORN	HORN
APPLICATIONS		COMMUNICATIONS NAVIGATION TELEVISION CONTROL RADAR: SEARCH E.W. I.F.F.	COMMUNICATIONS NAVIGATION TELEVISION CONTROL RADAR: SEARCH E.W. GUNLAYING I.F.F.	MICROWAVE RELAYS NAVIGATION CONTROL RADAR: SEARCH BOMBING E.W. GUNLAYING	EXPERIMENTAL (TACTICAL RADAR, ETC)

Chart Q. High Frequency Spectrum

GLOSSARY

ABSORPTION LOSS. That part of the transmission loss due to the dissipation or conversion of electrical energy into other forms of energy (e.g., heat), either within the medium or attendant upon a reflection.

ABSORPTION MODULATION. A system for producing amplitude modulation of the output of a radio transmitter by means of a variable impedance device inserted in or coupled to the output circuit.

AMPLITUDE MODULATION (AM). Modulation in which the amplitude of a carrier is the characteristic varied.

ANTENNA. A means of radiating or receiving radio waves.

ANTENNA ARRAY. A system of antennas coupled together for the purpose of obtaining directional effects.

ANTENNA RESISTANCE. The quotient of the power supplied to the entire antenna circuit by the square of the effective antenna current referred to a specified point.

NOTE. Antenna resistance is made up of such components as radiation resistance, ground resistance, radio-frequency resistance of conductors in the antenna circuit, and equivalent resistance due to corona, eddy currents, insulator leakage, and dielectric power loss.

APERTURE (of a unidirectional antenna). That portion of a plane surface near the antenna, perpendicular to the direction of maximum radiation, through which the major part of the radiation passes.

ATMOSPHERIC DUCT. An almost horizontal layer in the troposphere, extending from the level of a local minimum of the modified refractive index as a function of height, down to a level where the minimum value is again encountered, or down to the earth's surface if the minimum value is not again encountered.

ATTENUATION. Of a quantity associated wit ah traveling wave in a homogeneous medium, the decrease with distance in the direction of propagation.

NOTE. In a diverging wave, attenuation includes the effect of divergence.

ATTENTUATION CONSTANT. For a traveling plane wave at a given frequency, the rate of exponential decrease of the amplitude of a field component (or of the voltage or current) in the direction of the propagation, in nepers or decibels per unit length.

AUTOMATIC FREQUENCY CONTROL (AFC). An arrangement whereby the frequency of an oscillator is automatically maintained within specified limits.

AUTOMATIC GAIN CONTROL (AGC). A circuit arrangement which adjusts the gain in a specified manner in response to changes in input.

BALANCED MODULATOR. A modulator, specifically a push-pull circuit, in which the carrier and modulating signal are so introduced that after modulation takes place the output contains the two sidebands without the carrier.

BANDWIDTH (of a wave). The least frequency interval outsibe of which the power spectrum of a time-varying quantity is everywhere less than some specified fraction of its value at a reference frequency.

CAUTION. This definition permits the spectrum to be less than the specified fraction within the interval.

NOTE. Unless otherwise stated, the reference frequency is that at which the spectrum has its maximum value.

(of a device). The range of frequencies within which performance, with respect to some characteristic, falls within specific limits.

BASEBAND. The frequencies containing the intelligence used to modulate the carrier of a radio relay system.

BINARY NUMBER SYSTEM. A number system which uses two symbols (usually denoted by "0" and "1") and has two as its base, just as the decimal system uses ten symbols ("0, 1, . . . , 9") and the base ten.

BOLOMETER. A specially constructed resistor which has a positive temperature coefficient and which is used for power measurements.

CARRIER WAVE. A wave generated at a point in the transmitting system and modulated by the signal.

CAVITY RESONATOR. A space normally bounded by an electrically conducting surface in which oscillating electromagnetic energy is stored, and whose resonant frequency is determined by the geometry of the enclosure.

CHANNELING. The utilization of a modulation-frequency band for the simultaneous transmission from two or more communication channels in which the separation therebetween is accomplished by the use of carriers or subcarriers, each in a different discrete frequency band forming a subdivision of the main band.

NOTE. This covers a special case of multiplex transmissions.

CHARACTERISTIC IMPEDANCE (Z_o). The ratio of the voltage to the current at every point along a transmission line on which there are no standing waves.

CLIPPER. A transducer which gives output only when the input exceeds a critical value.

CLIPPER-LIMITER. A transducer which gives output only when the input lies above a critical value and a constant output for all inputs above a second higher critical value.

NOTE. This is sometimes called an amplitude gate, or slicer.

COEFFICIENT OF REFLECTION. The square root of the ratio of the reflected power leaving a reflecting surface to the power incident to the same surface.

CONDUCTION CURRENT. The power flow parallel to the direction of propagation expressed in watts per square meter.

CONDUCTIVITY. A measure of the ability of a material to act as a path for electron flow. It is the reciprocal of resistivity and is expressed in mhos/meter.

CONVERSION TRANSCONDUCTANCE (of a heterodyne conversion transducer). The quotient of the magnitude of the desired output-frequency component of current by the magnitude of the input-frequency component of voltage when the impedance of the output external termination is negligible for all of the frequencies which may affect the result.

NOTE. Unless otherwise stated, the term refers to the cases in which the input-frequency voltage is of infinitesimal magnitude. All direct electrode voltages and the magnitude of the local-oscillator voltage must be specified, fixed values.

CONVERSION VOLTAGE GAIN (of a conversion transducer). The ratio of (1) the magnitude of the output-frequency voltage across the output termination, with the transducer inserted between the input-frequency generator and the output termination, to (2) the magnitude of the input-frequency voltage across the input termination of the transducer.

CONVERTER. A conversion transducer in which the output frequency is the sum or difference of the input frequency and an integral multiple of the local oscillator frequency.

NOTE. The frequency and voltage or power of the local oscillator are parameters of the conversion transducer. Ordinarily, the output signal amplitude is a linear function of the input signal amplitude over its useful operating range.

CRITICAL FREQUENCY. The limiting frequency below which a magneto-ionic wave component is reflected by, and above which it penetrates through, an ionospheric layer at vertical incidence.

CROSS-MODULATION. Modulation of a desired signal by an undesired signal.

CROSSTALK. The sound heard in a receiver associated with a given telephone channel resulting from telephone currents in another telephone channel.

NOTE. In practice, crosstalk may be measured either by the loudness of the overheard sounds or by the magnitude of the coupling between the disturbed and disturbing channels. In the latter case, to specify the loudness of the overheard sounds, the volume in the disturbing channel must also be given.

DECIBEL. The decibel is one-tenth of a bel, the number of decibels denoting the ratio of the two amounts of power being ten times the logarithm to the base 10 of this ratio. The abbreviation db is commonly used for the term decibel.

NOTE. With P_1 and P_2 designating two amounts of power and the number of decibels denoting their ratio,

$$n = 10 \log_{10} (P_1 / P_2) \text{ decibels.}$$

When the conditions are such that ratios of currents or ratios of voltage (or analogous quantities in other fields) are the square roots of the corresponding power ratios, the number of decibels by which the corresponding powers differ is expressed by the following equations:

$$n = 20 \log_{10} (I_1/I_2) \text{ decibels}$$

$$n = 20 \log_{10} (V_1/V_2) \text{ decibels}$$

where I_1/I_2 and V_1/V_2 are the given current and voltage ratios, respectively. By extension, these relations between numbers of decibels and ratios of currents or voltages are sometimes applied where these ratios are not the square roots of the corresponding power ratios; to avoid confusion, such usage should be accomplished by a specific statement of this application.

DEMODULATION. The process of recovering the modulating wave from a modulated carrier.

DEVIATION RATIO. In a frequency-modulation system, the ratio of the maximum frequency deviation to the maximum modulating frequency of the system.

DEW POINT. The temperature at which the water vapor in the air begins to condense.

DIFFRACTION. The bending of a wave into the region behind an obstacle.

DIPOLE ANTENNA. A straight radiator, usually fed in the center, and producing a maximum of radiation in the plane normal to its axis. The length specified is the overall length.

NOTE. Common usage in microwave antennas considers a dipole to be a metal radiating structure which supports a line current distribution similar to that of a thin straight wire, a half wavelength long, so energized that the current has two nodes, one at each of the far ends.

DIRECT WAVE. A wave that is propagated directly through space.

DIRECTION OF PROPAGATION. At any point in a homogeneous, isotropic medium, the direction of time average energy flow.

DIRECTIONAL ANTENNA. An antenna having the property of radiating or receiving radio waves more effectively in some directions than others.

DIRECTIVE GAIN. In a given direction, 4π times the ratio of the radiation intensity in that direction to the total power radiated by the antenna.

DISCRIMINATOR. A device in which amplitude variations are derived in response to frequency variations.

DISPLACEMENT CURRENT. The current at right angles to the direction of propagation determined by the rate at which the field energy changes.

DISTORTION. An undesired change in wave form.

DISTRIBUTED CAPACITANCE. The capacitance that exists between the turns in a coil or choke, or between adjacent conductors or circuits, as distinguished from the capacitance which is concentrated in a capacitor.

DISTRIBUTED INDUCTANCE. The inductance that exists along the entire length of a conductor, as distinguished from the self-inductance which is concentrated in a coil.

DUPLEX OPERATION. The operation of associated transmitting and receiving apparatus in which the processes of transmission and reception are concurrent.

EFFECTIVE AREA. The square of the wavelength multiplied by the power gain (or directive gain) in that direction, and divided by 4π.

NOTE. When power gain is used, the effective area is that for power reception; when directive gain is used, the effective area is that for directivity.

EFFECTIVE RADIUS OF THE EARTH. An effective value for the radius of the earth, which is used in place of the geometrical radius to correct for atmospheric refraction when the index of refraction in the atmosphere changes, linearly with height.

NOTE. Under conditions of standard refraction the effective radius of the earth is 8.5×10^6 meters, or 4/3 the geometrical radius.

ELECTRIC FIELD STRENGTH. The magnitude of the electric field vector.

ELECTRIC FIELD VECTOR. At a point in an electric field, the force on a stationary positive charge per unit charge.

NOTE. This may be measured either in newtons per coulomb or in volts per meter. This term is sometimes called the electric field intensity but such use of the word intensity is deprecated in favor of field strength since intensity connotes power in optics and radiation.

ELLIPTICALLY POLARIZED WAVE. An electromagnetic wave for which the electric and/or the magnetic field vector at a point describes an ellipse.

NOTE. This term is usually applied to transverse waves.

FADING. The variation of radio field strength caused by changes in the transmission medium with time.

FRAUNHOFER REGION. That region of the field in which the energy flow from an antenna proceeds essentially as though coming from a point source located in the vicinity of the antenna.

Note. If the antenna has a well-defined aperture D in a given aspect, the Fraunhofer region in that aspect is commonly taken to exist at distances greater than $2D^2/\gamma$ from the aperture, γ being the wavelength.

FREQUENCY DEVIATION. In frequency modulation, the peak difference between the instantaneous frequency of the modulated wave and the carrier frequency.

FREQUENCY-DIVISION MULTIPLEX. The process or device in which each modulating wave modulates a separate subcarrier and the subcarriers are spaced in frequency.

Note. Frequency division permits the transmission of two or more signals over a common path by using different frequency bands for the transmission of the intelligence of each message signal.

FREQUENCY MODULATION (FM). Angle modulation of a sine-wave carrier in which the instantaneous frequency of the modulated wave differs from the carrier frequency by an amount proportional to the instantaneous value of the modulating wave.

Note. Combinations of phase and frequency modulation are commonly referred to as "frequency modulation".

FRESNEL REGION. The region between the antenna and the Fraunhofer region.

Note. If the antenna has a well-defined aperture D in a given aspect, the Fresnel region in that aspect is commonly taken to extend a distance $2D^2/\lambda$ in that aspect, λ being the wavelength.

FRESNEL ZONES. Zones of wave reinforcement and destructive interference caused by interaction of direct waves and those waves reflected from the earth.

GROUND WAVE. A radio wave that is propagated over the earth and is ordinarily affected by the presence of the ground and the troposphere. The ground wave includes all components of a radio wave over the earth except ionospheric and tropospheric waves.

Note. The ground wave is refracted because of variations in the dielectric constant of the troposphere including the condition known as surface duct.

GROUNDED-CATHODE AMPLIFIER. An electron-tube amplifier with the cathode at ground potential at the operating frequency, with input applied between the control grid and ground, and the output load connected between plate and ground. (This is the conventional amplifier circuit.)

GROUNDED-GRID AMPLIFIER. An electron-tube amplifier circuit in which the control grid is at ground potential at the operating frequency, with input applied between cathode and ground, and output load connected between plate and ground. The grid-to-plate impedance of the tube is in parallel with the load instead of acting as a feedback path.

GROUP VELOCITY. Of a traveling plane wave, the velocity of propagation of the envelope of a wave occupying a frequency band over which the envelope delay is approximately constant. It is equal to the reciprocal of the rate of change of phase constant with angular frequency.

Note. Group velocity differs from phase velocity in a medium in which the phase velocity varies with frequency.

HALF-POWER WIDTH OF A RADIATION LOBE. In a plane containing the direction of the maximum of the lobe, the full angle between the two directions in that plane about the maximum in which the radiation intensity is one-half the maximum value of the lobe.

HARMONIC. A sinusoidal component of a periodic wave or quantity having a frequency which is an integral multiple of the fundamental frequency. For example, a component the frequency of which is twice the fundamental frequency is called the second harmonic.

HETERODYNE. To beat or mix two frequencies in a non-linear component so as to produce different frequencies from those introduced.

HORIZONTALLY POLARIZED WAVE. A linearly polarized wave whose electric field vector is horizontal.

IMAGE FREQUENCY. In heterodyne frequency converters in which one of the two sidebands produced by beating is selected; an undesired input frequency capable of producing the selected frequency by the same process.

NOTE. The word "image" implies the mirror-like symmetry of signal and image frequencies about the beating oscillator frequency or the intermediate frequency, whichever is higher.

IMAGE RATIO. The ratio of (1) the field strength at the image frequency to (2) the field strength at the desired frequency, each field being applied in turn, under specified conditions, to produce equal outputs.

IMPULSE NOISE. Noise characterized by transient disturbances separated in time by quiescent intervals. The frequency spectrum of these disturbances must be substantially uniform over the useful pass band of the transmission system.

INCIDENT WAVE. In a medium of certain propagation characteristics, a wave which impinges on a discontinuity or medium of different propagation characteristics.

INTERELECTRODE TRANSIT TIME. The time required for an electron to traverse the distance between two electrodes.

INTERFERENCE. In a signal transmission system either extraneous power which tends to interfere with the reception of the desired signals, or the disturbance of signals which results.

INTERMEDIATE FREQUENCY (I-F). The frequency in superheterodyne reception resulting from a frequency conversion before demodulation.

INTERMEDIATE SUBCARRIER. A carrier which may be modulated by one or more subcarriers and which is used as a modulating wave to modulate a carrier or another intermediate subcarrier.

IONOSPHERE. The part of the earth's outer atmosphere where ions and electrons are present in quantities sufficient to affect the propagation of radio waves.

NOTE. According to current opinion, the lowest level is approximately 50 kilometers above the earth's surface.

ISOTROPIC ANTENNA (UNIPOLE). A hypothetical antenna radiating or receiving equally in all directions. A pulsating sphere is a unipole for sound waves. In the case of electromagnetic waves unipoles do not exist physically but represent convenient reference antennas for expressing directive properties of actual antennas.

KELVIN SCALE (ABSOLUTE SCALE). A temperature scale using the same divisions as the centrigrade scale, but with the zero point established at absolute zero ($\cong -273°C$), theoretically the lowest possible temperature.

LEAKAGE RADIATION. In a transmitting system, radiation from anything other than the intended radiating system.

LIMITER. A transducer whose output is constant for all inputs above a critical value.

NOTE. A limiter may be used to remove amplitude modulation while transmitting angle modulation.

LINEARLY POLARIZED WAVE. At a point in a homogeneous, isotropic medium, a transverse electromagnetic wave whose electric field vector at all times lies along a fixed line.

LOWEST USEFUL HIGH FREQUENCY. The lowest high frequency effective at a specified time for ionospheric propagation of radio waves between two specified points.

NOTE. This is determined by factors such as absorption, transmitter power, antenna gain, receiver characteristics, type of service, and noise conditions.

MAGNETIC FIELD. A state of the medium in which moving electrified bodies are subject to forces by virtue of both their electrifications and motion.

MAGNETO-IONIC WAVE COMPONENT. Either of the two elliptically polarized wave components into which a linearly polarized wave incident on the ionosphere is separated because of the earth's magnetic field.

MAXIMUM USABLE FREQUENCY. The upper limit of the frequencies that can be used at a specified time for radio transmission between two points and involving propagation by reflection from the regular ionized layers of the ionosphere.

NOTE. Higher frequencies may be transmitted by sporadic and scattered reflections.

MILLIBAR. A unit of pressure equal to 1000 dynes per square centimeter. Standard atmospheric pressure at sea level is equal to 1013 millibars.

MODIFIED INDEX OF REFRACTION. In the troposphere, the index of refraction at any height increased by h/a, where h is the height above sea level and a is the mean geometrical radius of the earth. When the index of refraction in the troposphere is horizontally stratified, propagation over a hypothetical flat earth through an atmosphere with the modified index of refraction is substantially equivalent to propagation over a curved earth through the real atmosphere.

MODULATING SIGNAL (MODULATING WAVE). A wave which causes a variation of some characteristic of a carrier.

MODULATION. The process or result of the process whereby some characteristic of one wave is varied in accordance with another wave.

MODULATION INDEX. For a sinusoidal modulating wave, the ratio of the frequency deviation to the frequency of the modulating wave.

MULTIPATH TRANSMISSION. The propagation phenomenon which results in signals reaching the radio receiving antenna by two or more paths, usually having both amplitude and phase differences between each path.

MULTIPLE MODULATION. A succession of processes of modulation in which the modulated wave from one process becomes the modulating wave for the next.

NOTE. In designating multiple-modulation systems by their letter symbols, the processes are listed in the order in which the signal intelligence encounters them. For example, PPM-AM means a system in which one or more signals are used to position-modulate their respective pulse subcarriers which are spaced in time and are used to amplitude-modulate a carrier.

MULTIPLEX RADIO TRANSMISSION. The simultaneous transmission of two or more signals using a common carrier wave.

NOISE. Any extraneous electrical disturbance tending to interfere with the normal reception of a transmitted signal.

NOISE FIGURE. Of a linear system at a selected input frequency, the ratio of (1) the total noise power per unit bandwidth (at a corresponding output frequency) available at the output terminals, to (2) the portion thereof engendered at the input frequency by the input termination, whose noise temperature is standard (290°K) at all frequencies. (See NOISE TEMPERATURE.)

NOTE 1. For heterodyne systems there will be, in principle, more than one output frequency corresponding to a single input frequency, and vice-versa; for each pair of corresponding frequencies a noise factor is defined.

NOTE 2. The phrase, "available at the output terminals", may be replaced by "delivered by the system into an output termination", without changing the sense of the definition.

NOISE TEMPERATURE. At a pair of terminals and at a specific frequency, the temperature of a passive system having an available noise power per unit bandwidth equal to that of the actual terminals.

NOISE TEMPERATURE (STANDARD). The standard reference temperature T_o for noise measurements is taken as 290°K.

NOTE. $K\ T_o/e = 0.0250$ volt where e is the electron charge and K is Boltzmann's constant.

OPTICAL HORIZON. The locus of points at which a straight line from the given point becomes tangential to the earth's surface.

OPTIMUM WORKING FREQUENCY. The most effective frequency at a specified time for ionspheric propagation of radio waves between two specified points.

NOTE. In predictions of useful frequencies the optimum working frequency is commonly taken as 15 percent below the monthly median value of the maximum usable frequency, for the specified time and path.

PATH ATTENUATION. The power loss between transmitter and receiver due to all causes. It is equal to $10 \log_{10} P_t/P_r$ where P_t is the power radiated from the transmitting antenna and P_r is the power available at the output terminals of the receiving antenna and is expressed in decibels.

PERMEABILITY. A measure of the ability of a material to act as a path for magnetic lines of force.

PHASE CONSTANT. For a traveling plane wave at a given frequency, the rate of linear increase of phase lag of a field component (for the voltage or current) in the direction of propagation, in radians per unit length.

PHASE MODULATION (PM). Angle modulation in which the angle of a sine-wave carrier is caused to depart from the carrier angle by an amount proportional to the instantaneous value of the modulating wave.

NOTE. Combinations of phase and frequency modulation are commonly referred to as "frequency modulation".

PHASE VELOCITY. Of a traveling plane wave at a single frequency, the velocity of an equiphase surface along the wave normal.

PLANE WAVE. A wave whose equiphase surfaces form a family of parallel planes.

POLAR DIAGRAMS. A system of coordinates in which a point is determined by the length and the angle of a line connecting the center of the diagram and the point.

POWER GAIN. In a given direction, 4π times the ratio of the radiation intensity in that direction to the total power delivered to the antenna.

PROPAGATION CONSTANT. For a traveling plane wave at a given frequency, the complex quantity whose real part is the attenuation constant in nepers per unit length and whose imaginary part is the phase constant in radians per unit length.

PULLING. A change in the resonant frequency of a circuit due to a change in the load.

PULSE. A single disturbance characterized by the rise and decay in time or space or both of a quantity whose value is normally constant.

NOTE. In these definitions an r-f carrier, amplitude-modulated by a pulse, is not considered to be a pulse.

PULSE AMPLITUDE MODULATION (PAM) Modulation in which the modulating wave is caused to amplitude-modulate a pulse carrier.

PULSE CARRIER. A carrier consisting of a series of pulses.

NOTE. Usually pulse carriers are employed as subcarriers.

PULSE CODE.

(1) A pulse train modulated so as to represent information.

(2) Loosely, a code consisting of pulses, such as Morse Code, Baudot Code, Binary Code.

PULSE-CODE MODULATION (PCM). Modulation which involves a pulse code.

NOTE. This is a generic term, and additional specification is required for a specific purpose.

PULSE-DURATION MODULATION (PDM). Pulse-time modulation in which the value of each instantaneous sample of the modulating wave is caused to modulate the duration of the pulse.

NOTE. The terms "pulse-width modulation" and "pulse-length modulation" also have been used to designate this system of modulation.

NOTE. In pulse-duration modulation, the modulating wave may vary the time of occurrence of the leading edge, the trailing edge, or both edges of the pulse.

PULSE-POSITION MODULATION (PPM). Pulse-time modulation in which the value of each instantaneous sample of a modulating wave is caused to modulate the position in time of a pulse.

PULSE SHAPER. Any transducer used for changing one or more characteristics of a pulse.

NOTE. This term includes pulse regenerators.

PULSE-TIME MODULATION (PTM). Modulation in which the values of instantaneous samples of the modulating wave are called to modulate the time of occurrence of some characteristic of a pulse carrier.

NOTE. Pulse-duration modulation and pulse-position modulation are particular forms of pulse-time modulation.

PUSH-PUSH CIRCUIT. A circuit employing two similar tubes with grids connected in phase opposition and plates in parallel to a common load, and usually used as a frequency multiplier to emphasize even-order harmonics.

PUSHING. A change in the resonant frequency of a circuit due to changes in the applied voltages.

Q. The figure of merit of efficiency of a circuit or coil. Numerically it is the ratio of the inductive reactance to the resistance of the circuit or coil.

RADIATION EFFICIENCY. The ratio of the power radiated to the total power supplied to the antenna at a given frequency.

RADIATION INTENSITY. In a given direction, the power radiated from an antenna per unit solid angle in that direction.

RADIATION RESISTANCE. The quotient of the power radiated by an antenna divided by the square of the antenna current referred to a specific point.

RADIO-FREQUENCY PULSE. A radio-frequency carrier amplitude-modulated by a pulse. The amplitude of the modulated carrier is zero before and after the pulse.

NOTE. Coherence of the carrier (with itself) is not implied.

RADIO HORIZON. The locus of points at which direct rays from the transmitter become tangential to the earth's surface.

NOTE. On a spherical surface the horizon is a circle. The distance to the horizon is affected by atmospheric refraction.

RANDOM (OR FLUCTUATION) NOISE. Noise characterized by a large number of overlapping transient disturbances occurring at random.

REACTANCE MODULATOR. A device, used for the purpose of modulation, whose reactance may be varied in accordance with the instantaneous amplitude of the modulating electromotive force applied thereto. This is normally an electron-tube circuit and is commonly used to effect phase or frequency modulation.

REFLECTION. That phenomenon which causes a wave which strikes a medium of different characteristics to be returned into the original medium with the angles of incidence and of reflection equal and lying in the same plane.

REFRACTION. That phenomenon which causes a wave which enters another medium obliquely to undergo an abrupt change in direction if the velocity of the wave in the second medium is different from that in the first.

REFRACTED WAVE. That part of an incident wave which travels from one medium into a second medium.

REFRACTIVE INDEX. Of a wave transmission medium, the ratio of the phase velocity in free space to that in the medium.

REFRACTIVE MODULUS. In the troposphere, the excess over unity of the modified index of refraction, expressed in millionths. It is represented by M and is given by the equation

$$M = (\eta + h/a - 1)\, 10^6,$$

where η is the index of refraction at a height h above sea level, and a is the radius of the earth.

RELATIVE DIELECTRIC CONSTANT. The ratio of the dielectric constant of a material to the dielectric constant of a vacuum.

SCATTERING. When radio waves encounter matter, a disordered change in the direction of propagation of the waves.

SELECTIVE FADING. Fading in which the variation of radio field intensity is not the same at all frequencies in the frequency band of the received wave.

SELECTIVITY (of a receiver). That characteristic which determines the extent to which the receiver is capable of differentiating between the desired signal and disturbances of other frequencies.

SENSITIVITY. The least signal input capable of causing an output signal having desired characteristics.

SIDEBANDS. (1) The frequency bands on both sides of the carrier frequency within which fall the frequencies of the wave produced by the process of modulation. (2) The wave components lying within such bands.

NOTE. In the process of amplitude modulation with a sine-wave carrier, the upper sideband includes the sum (carrier plus modulating) frequencies; the lower sideband includes the difference (carrier minus modulating) frequencies.

SIGNAL-TO-NOISE RATIO. The ratio of the value of the signal to that of the noise.

NOTE. This ratio is usually in terms of peak values in the case of impulse noise and in terms of the root-mean-square values in the case of the random noise.

NOTE. Where there is a possiblity of ambiguity, suitable definitions of the signal and noise should be associated with the term; as, for example: peak-signal to peak-noise ratio; root-mean-square signal to root-mean-square noise ratio; peak-to-peak signal to peak-to-peak noise ratio, etc.

NOTE. This ratio is often expressed in decibels.

NOTE. This ratio may be a function of the bandwidth of the transmission system.

SIMPLEX OPERATION OF A RADIO SYSTEM. A method of operation in which communication between two stations takes place in one direction at a time.

NOTE. This includes ordinary transmit-receive operation, press-to-talk operation, voice-operated carrier, and other forms of manual or automatic switching from transmit to receive.

SINGLE-SIDEBAND MODULATION (SS). Modulation whereby the spectrum of the modulating wave is translated in frequency by a specified amount either with or without inversion.

SINGLE-SIDEBAND TRANSMISSION. That method of operation in which one sideband is transmitted and the other sideband is suppressed.

The carrier wave may be either transmitted or suppressed.

SPHERICAL WAVE. A wave whose equiphase surfaces form a family of concentric spheres.

STANDARD REFRACTION. The refraction which would occur in an idealized atmosphere in which the index of refraction decreases uniformly with height at a rate of 39 x 10^{-6} per kilometer.

NOTE. Standard refraction may be included in ground wave calculations by use of an effective earth radius of 8.5 x 10^6 meters, or 4/3 the geometrical radius of the earth.

STANDING WAVE. A wave in which, for any component of the field, the ratio of its instantaneous value at one point to that at any other point does not vary with time.

SUBCARRIER. A carrier which is applied as a modulating wave to modulate another carrier or an intermediate subcarrier.

SURFACE DUCT. An atmospheric duct for which the lower boundary is the surface of the earth.

TENTH-POWER WIDTH. In a plane containing the direction of the maximum of a lobe, the full angle between the two directions in that plane about the maximum in which the radiation intensity is one-tenth the maximum value of the lobe.

THERMISTOR. A specially constructed resistor which has a negative temperature coefficient, used for power measurements.

TIME-DIVISION MULTIPLEX. The process or device in which each modulating wave modulates a separate pulse subcarrier, the pulse subcarriers being spaced in time so that no two pulses occupy the same time interval.

NOTE. Time division permits the transmission of two or more signals over a common path by using different time intervals for the transmission of the intelligence of each message signal.

TRANSDUCER. A device by means of which energy can flow from one or more transmission systems to one or more other transmission systems.

NOTE. The energy transmitted by these systems may be of any form (for example, it may be electric, mechanical, or acoustical), and it may be of the same form or different forms in the various input and output systems.

TRANSVERSE ELECTRIC WAVE. In a homogeneous isotropic medium, an electro-magnetic wave in which the electric field vector is everywhere perpendicular to the direction of propagation.

NOTE. This is abbreviated "TE Wave".

TRANSVERSE ELECTROMAGNETIC WAVE. In a homogeneous isotropic medium, an electromagnetic wave in which both the electric and magnetic field vectors are everywhere perpendicular to the direction of propagation.

NOTE. This is abbreviated "TEM Wave".

TRANSVERSE MAGNETIC WAVE. In a homogeneous isotropic medium, an electro-magnetic wave in which the magnetic field vector is everywhere perpendicular to the direction of propagation.

NOTE. This is abbreviated "TM Wave".

TRAVELING PLANE WAVE. A plane wave each of whose frequency components has an exponential variation of amplitude and a linear variation of phase in the direction of propagation.

TROPOSPHERE. That part of the earth's atmosphere in which temperature generally decreases with altitude, clouds form, and convection is active.

NOTE. Experiments indicate that the troposphere occupies the space above the earth's surface to a height of about 10 kilometers.

TROPOSPHERIC WAVE. A radio wave that is propagated by reflection from a place of abrupt change in the dielectric constant or its gradient in the troposphere.

NOTE. In some cases the ground wave may be so altered that new components appear to arise from reflections in regions of rapidly changing dielectric constants; when those components are distinguishable from the other components, they are called tropospheric waves.

VECTOR. A quantity which has both magnitude and direction or an arrow drawn in the direction and whose length is proportional to the magnitude of the quantity.

VELOCITY-MODULATED OSCILLATOR (KLYSTRON). An electron-tube in which the velocity of an electron stream is varied (velocity-modulated in passing through a resonant cavity called a buncher. Energy is extracted from the bunched electron stream at a higher energy level in passing through a second cavity resonator called the catcher. Oscillations are sustained by coupling energy from the catcher cavity back to the buncher cavity.

VERTICALLY POLARIZED WAVE. A linearly polarized wave whose magnetic field vector is horizontal.

VIDEO. A term pertaining to the bandwidth and spectrum position of the signal resulting from television scanning.

NOTE. In current usage, video means a bandwidth in the order of megacycles per second, and a spectrum position that goes with a d-c carrier.

WAVEGUIDE. A system of material boundaries capable of guiding waves.

WAVE INTERFERENCE. The variation of wave amplitude with distance or time, caused by the superposition of two or more waves.

NOTE. As most commonly used, the term refers to the interference of waves of the same or nearly the same frequency.

WAVELENGTH. In a periodic wave, the distance between points of corresponding phase of two consecutive cycles. The wavelength λ is related to the phase velocity, ν, and the frequency, f, by $\lambda = \nu/f$.

MKS SYMBOLS (METERS, KILOGRAMS, SECONDS)

QUANTITY		MKS
Voltage	—	V, volts
Current	—	I, amperes
Resistance	—	R, ohms
Power	—	P, watts
Capacitance	—	C, farads
Inductance	—	L, Henrys

Characteristic Impedance — Z_o, ohms

Reactance	—	X, ohms X_L = Inductive X_c = Capacitive
Impedance	—	Z, ohms
Admittance	—	Y, mhos
Conductance	—	G, mhos
Conductivity	—	σ, mhos/meter
Permeability	—	μ, henrys/meter for free space μ_o = 1.257 x 10^{-6} henrys/m

Dielectric constant — κ, farads/meter
κ = actual dielectric constant
κ_o = dielectric constant of a vacuum
$\kappa_r = \kappa/\kappa_o$ = relative dielectric constant
κ_o = 8.854 x 10^{-12} farads/m

Magnetic field intensity—H, ampere turns/meter or amp/meter

Electric field intensity—E, volts/meter

Poynting Power Flow—p, watts/meter^2

Total current density—J, amp/meter^2

Propagation constant—Γ, hyperbolic radians/meter

Attenuation constant—α, nepers/meter or db/meter

Phase constant—ρ, circular radians/meter

Wavelength—λ, meters or centimeters

Guide wavelength—λ_g, meters

Cut-off wavelength—λ_c meters

Frequency—f, cps, mc, etc.

Velocity of propagation—v, meters/sec

Index of refraction—η, numeric

Refractive Modulus—M, numeric

Reflection coefficient—ρ, numeric

Base of Natural log—ϵ, 2.7182818, or $\log_\epsilon$ N = 2.302585 $\log_{10}$N

Imaginary unit—j = $\sqrt{-1}$

Angular velocity—ω, circular radians = 2πf where f is in cps

Quality, or figure of Merit—Q = $\omega\kappa/\sigma$ for a wave medium Q = X/R for a circuit

INDEX