UNCLASSIFIED

USAEPG-SIG 960-67

10

INSTRUCTION MANUAL

TROPOSPHERIC SCATTER

PRINCIPLES AND APPLICATIONS

TASK 45-56-0024

U.S. ARMY ELECTRONIC PROVING GROUND
FORT HUACHUCA, ARIZONA
MARCH 1960

TROPOSPHERIC SCATTER, PRINCIPLES AND APPLICATIONS

Task 45-56-0024

Prepared By

The Collins Radio Company

Under Contract Number
DA-36-039-SC-67491

for

Signal Communications Department
U. S. Army Electronic Proving Ground
Fort Huachuca, Arizona
March 1960

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont)

TABLE OF CONTENTS (Cont)

TABLE OF CONTENTS (Cont)

TABLE OF CONTENTS (Cont)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Cont)

LIST OF ILLUSTRATIONS (Cont)

LIST OF ILLUSTRATIONS (Cont)

LIST OF TABLES

INTRODUCTION

Radio engineers have searched for many years for a persistent mode of propagation to supplement services ordinarily satisfied by wire lines and cables. Until recently, microwave transmission offered the only solution. Microwave relay transmission, however, imposes line-of-sight limitations and the necessity for repeater stations in the longer circuits. Tropospheric scatter equipment provides reliable multichannel transmission for path lengths up to 300 miles. Recent developments in mobile tropospheric scatter equipment have made this mode of propagation especially adaptable to military use.

However, it has been recognized that one limiting factor in the use of tropospheric scatter equipment is the ability and training of the personnel responsible for equipment operation and maintenance. To exploit this new technique successfully in pratical applications, fundamental knowledge and experience are imperative.

It is the purpose of this manual to give the basic principles of tropospheric scatter and a description of typical equipment used for this type of communication. A prerequisite to understanding these basic concepts is a knowledge of electricity and radio and some mathematics. If you are not too well grounded in these subjects, you may have to consult other information sources before proceeding with a study of this manual. A review of logarithms and a discussion of decibels are included in the first chapter. The information included in this manual is general enough to apply to all FM tropospheric scatter equipment.

CHAPTER 1
LOGARITHMS AND DECIBELS

1.1 LOGARITHMS.

The logarithm of a quantity is the exponent, or the power, to which a given number, called the base, must be raised to equal that quantity. To illustrate, in the quantity $3^2 = 9$, the exponent, 2, is called the logarithm of 9 to the base 3. This relation is usually written $\log_3 9 = 2$. Any positive number greater than 1 might serve as a base. Two numbers have been selected, resulting in two systems of logarithms. One base, 2.718, usually indicated by the Greek letter epsilon (ϵ), is used in the "natural" logarithm system. The other base is 10; it is used in the " common" system of logarithms. In the common system, the base 10 is usually omitted in the logarithmic expression. Thus $\log 1000 = 3$ is usually written $\log_{10} 1000 = 3$. In the natural system, the base (ϵ) may be written in.

In the common system, logarithms that are exact powers of 10 are integers. Thus,

$$\begin{aligned}
\log 100 &= 2, \text{ since } 10^2 = 100\\
\log 1000 &= 3, \text{ since } 10^3 = 1000\\
\log 10000 &= 4, \text{ since } 10^4 = 10000\\
\log 0.1 &= -1, \text{ since } 0.1 = 10^{-1}\\
\log 0.01 &= -2, \text{ since } 0.01 = 10^{-2}\\
\log 1 &= 0, \text{ since } 1 = 10^0
\end{aligned}$$

For numbers not exact powers of ten, the logarithm consists of two parts, a whole number and a decimal part. The whole number is called the characteristic and the decimal part the mantissa. Thus, for example, $\log 595 = 2.7745$ (in words, the logarithm of 595 is 2.7745), the characteristic is 2, and the mantissa is .7745. The characteristic is found by inspection and the mantissa from logarithmic tables.

1.1.1 CHARACTERISTIC.

The characteristic can be determined by the following rules:

a. The characteristic of the logarithm of a number greater than 1 is positive and is one less than the number of digits to the left of the decimal point. For example, in the case of the log 595, the characteristic is 2, and for the log of 59.5 the characteristic is 1.

b. The characteristic of the logarithm of a number less than 1 is negative, and is equal to one more than the number of zeros immediately to the right of the decimal point. For example, for the log of .0595, the characteristic is -2, and for log .00595, the characteristic is -3.

When the characteristic is negative, do not put the minus sign in front of the logarithm, since it applies only to the characteristic and not to the entire logarithm. Instead, add 10 to the negative characteristic and indicate the subtraction of 10 at the end of the logarithm. Thus, the characteristic -2 is written 8. (mantissa) -10, and the characteristic -3 is written 7. (matissa) -10. Another method of indicating that the characteristic is negative is to place the minus sign above the characteristic. For example: $\bar{2}$. (mantissa), and $\bar{3}$. (mantissa).

1.1.2 MANTISSA. Find the mantissa from tables of logarithms. Numbers which have the same figures in the same order and differ only in the position of the decimal point have the same mantissa in their logarithms. For example, the mantissa of 595 is .7745; the mantissa of 59.5 is also .7745.

1.1.3 TO FIND THE LOGARITHM OF A NUMBER.

a. Determine by inspection the characteristic of the number.

b. Find the mantissa from the tables. The mantissa of the number is independent of the position of the decimal point, so you can disregard the decimal point in the number when finding the mantissa. The mantissas in the table are the decimal part of the logarithm and therefore should be preceded by the decimal point.

In four-place logarithm tables, the first column in the table contains the first two digits of the numbers whose mantissas are given in the table, and the top row contains the third digit. Thus, to find the mantissa of 595, find 59 in the left-hand column and 5 at the top. In the column under 5, and opposite 59, is 7745, the mantissa. The logarithm of 595 is then 2.7745.

c. To find the logarithm of a quantity with more than three digits, use the process called interpolation. Suppose you want to find the logarithm of 5956. The tables do not give the mantissa for 5956. However, they give the mantissas for 5950 and 5960. (The mantissa for 5950 is the same as that for 595. Likewise, the mantissa for 5960 is the same as that for 596). Since 5956 lies between 5950 and 5960, its mantissa must lie between the mantissas for these two numbers. By arranging the mantissas in the following tabular form,

Mantissa of 5960 = .7752
Mantissa for 5956 = ?
Mantissa for 5950 = .7745

you can see that 5956 is 6/10 of the way between 5950 and 5960. Obviously, the mantissa of 5956 must be 6/10 of the way between .7745 and .7752. Since the difference between the two is .0007 and since 6/10 of .0007 is .00042, add .00042 to .7745 (the mantissa of 5950). The result, .77492, is the mantissa of 5956. Therefore, the logarithm of 5956 = 3.77492.

1.1.4 ANTILOGARITHMS. The number corresponding to a given logarithm is called the antilogarithm of the number. It is written "antilog" or "$\log^{-1}$." To find the antilog reverse the process for finding logarithms. Examples:

a. To find the antilog ($\log^{-1}$) of 1.8987, first look in the logarithm table and locate the mantissa .8987. It is in line with the number 79 and under the column headed 2. Thus the number corresponding with the mantissa .8987 has the digits 792.

To determine the location of the decimal point, reverse the rule for finding a characteristic. If the characteristic were zero, the decimal point would be placed after the first digit (7.92). Since the characteristic is 1, count two places to the right and place the decimal point after the 9. Thus the antilog of 1.8987 is 79.2.

b. Find the antilog ($\log^{-1}$) of 2.4325. The tables do not show the mantissa .4325; therefore you must interpolate. The mantissa .4325 lies between the two mantissas .4330 and .4314. These mantissas correspond to 271 and 270, respectively. The difference between .4330 and .4314 is .0016, and the difference between .4325 (the given mantissa) and the mantissa for 270 is .0009. Then find the number corresponding to .4325 by adding $\frac{.0009}{.0016}$ of 1, or .563 to 270, thus giving the sum 270563. Since the given characteristic is 2, the antilog of 2.4325 is 270.563.

1.1.5 COMPUTATIONS WITH LOGARITHMS. To multiply two quantities, add their logarithms and find the antilog of the result. Example:

Find the product of 6952 and 437.

Solution: Log (6952 x 437) = log 6952 + log 437.

Log 6952 = 3.8421

Log 437 = 2.6405

Adding, log 6952 + log 437 = 6.4826

Find the antilog 6.4826

The antilog 6.4826 = 3,038,000.

Actual multiplication of 6952 by 437 would give the result 3,038,024. This indicates an error of 24 in over 3,000,000 or 0.0008 of one per cent. The error is due to the fact that the logarithm tables go to four places only. Greater accuracy would result in using five-place tables. In general, though, accuracy obtained with four-place tables is sufficient.

To divide two quantities, subtract the logarithm of the divisor from the logarithm of the dividend and find the antilog of the result. Example:

Find the quotient of 6952 divided by 437.

Solution: Log (6952 - 437) = log 6952 - log 437

Log 6952 = 3.8421

Log 437 = 2.6405

Subtracting, log 6952 - log 437 = 1.2016

Find antilog 1.2016

Antilog 1.2016 = 15.908

To raise a quantity to any power, multiply the logarithm of the quantity by the exponent, or the power, and find the antilog of the result. Examples:

a. Find the value of $(5.2)^6$

Solution: $\text{Log } (5.2)^6 = 6 \times \log (5.2)$

$\text{Log } (5.2) = 0.716$

$\text{Log } (5.2)^6 = 6 \times (0.716) = 4.296$

Antilog 4.296 = 19768

Therefore, $(5.2)^6 = 19768$

b. Find the value of $(3.7)^{1/5}$. (Exponent is fractional)

Solution: $\text{Log } (3.7)^{1/5} = \frac{1}{5} \log (3.7)$

$\text{Log } (3.7) = 0.5682$

Then $\log (3.7)^{1/5} = \frac{1}{5} (0.5682)\ 0.1136$

Antilog 0.1136 = 1.2991

Therefore, $(3.7)^{1/5} = 1.2991$

c. Find the value of $(45.6)^{-3}$ (Exponent is negative)

Solution: $\text{Log } (45.6)^{-3} = -3 \log (45.6)$

$\text{Log } 45.6 = 1.659$

Then, $\log (45.6)^{-3} = -3\ (1.659) = -4.977$

Since logarithm tables list only positive values of mantissa, change -4.977 to 5.023 (or 5.023 -10) by subtracting -4.977 from 10.

The antilog of 5.023 = 0.000010544.

Therefore, $(45.6)^{-3} = .000010544$

To find the root of a quantity obtain the logarithm of the quantity, divide it by the index of the root, and find the antilog of the result. Examples:

d. Find the value of $\sqrt[3]{1.572}$ or (1.572) .

Solution: $\text{Log } \sqrt[3]{1.572} = \frac{1}{3} \log 1.572$

$\text{Log } 1.572 = 0.19646$

Then $\log \sqrt[3]{1.572} = \frac{1}{3} (0.19646) = 0.06549$

Antilog 0.06549 = 1.1611

Therefore, $\sqrt[3]{1.572} = 1.1611$

1.2 DECIBELS.

The decibel (abbreviated db) is a logarithmic unit used in communications work to express power ratios. By definition, $\text{db} = 10 \log \frac{P_2}{P_1}$. The sign associated with the number of decibels indicates which power is greater; thus a negative sign means P_2 is less than P_1 . Two examples showing application of this equation are given below.

a. Express the gain in db of a final amplifier with an input of 0.1 watt (P_1) and an output of 1000 watts (P_2).

$$\text{db} = 10 \log \frac{P_2}{P_1}$$

$$= 10 \log \frac{1000}{0.1}$$

$$= 10 \log 10{,}000 = 10 \times 4 = 40 \text{ db}$$

(Note that log 10,000 = 4.)

b. Express the loss in db of a circuit with a 10-watt input (P_1) and a 5-watt output (P_2).

$$\text{db} = 10 \log \frac{5}{10} = 10 \log \frac{1}{2}$$

$$= 10 (\log 1 - \log 2)$$

$$= 10 (-.3) = -3 \text{ db}$$

Figure 1-1 shows the relationship between power ratios and decibels for power ratios up to 1000. The number of decibels corresponding to ratios outside the range of figure 1-1 can be obtained by separating the ratio into two factors, finding the db value for each factor, and then adding the two results. For example: convert a power ratio of 1200:1 to db. This equals the product of 600:1 and 2:1. 600:1 equals 27.8 db; 2:1 equals 3 db. Therefore, a power ratio of 1200:1 equals 27.8 db + 3 db or 30.8 db.

It will be convenient to remember a few of the power ratios for circuit gains expressed in decibels:

10 db -- a power ratio of 10

20 db -- a power ratio of 100

30 db -- a power ratio of 1000

40 db -- a power ratio of 10,000

60 db -- a power ratio of 1,000,000

For small power differences:

1 db -- a power ratio of 5/4

3 db -- a power ratio of 2

7 db -- a power ratio of 5

If only these values are remembered, other values can be estimated. These values are not exact but are accurate enough for most calculations.

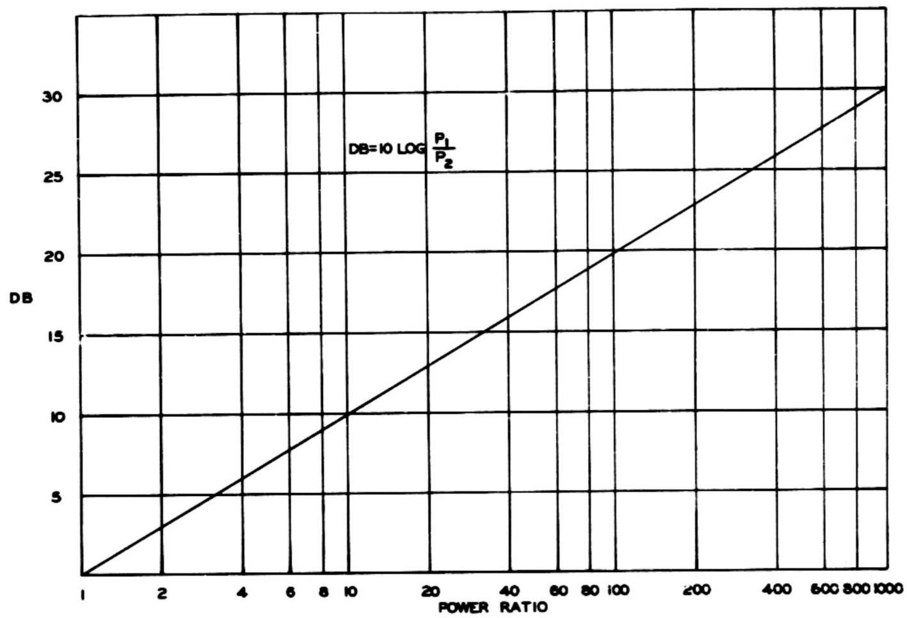


Figure 1-1. Relationship Between Power Ratio and Decibels

The ratio of voltages at two points in a circuit can also be expressed in decibels. Since $P = E \times I = \frac{E^2}{R^2}$, the ratio $\frac{P_2}{P_1}$ is equal to $\frac{E_2^2}{E_1^2}$, provided that R is the same in each case. If the resistance is the same, the ratio in decibels is:

$$db = 10 \log \frac{E_2^2}{E_1^2} = 20 \log \frac{E_2}{E_1}$$

The same equation holds for a-c circuits, provided that the impedance across which E_1 and E_2 are measured are equal. The expression for decibel gain of two-power magnitudes is not affected by the impedances.

The practical value of the decibel arises from its logarithmic nature. This permits the enormous ranges of power involved in communication work to be expressed in decibels without using inconveniently large numbers, while at the same time permitting small ratios to be conveniently expressed. For example, 1 db represents a power ratio of approximately 1.25, while 60 db represents a power ratio of 1,000,000 to 1.

The logarithmic character of a decibel also makes it possible to express the over-all gain of a circuit as the sum of the decibel gains or losses of the different parts of the circuit in cascade. An example of the use of this characteristic of decibels is given in figure 1-2, which shows the major losses and gains in a typical communications circuit. The circuit gain in db at various points in the circuit are given along the top of the figure. These db values are obtained by

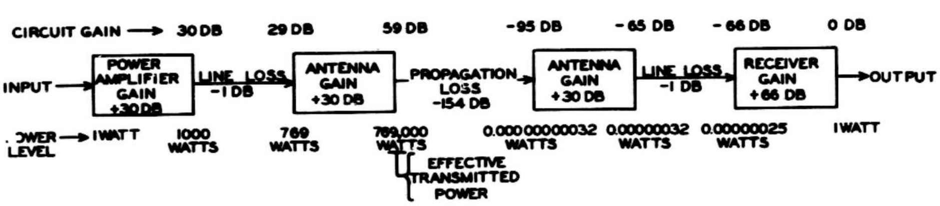


Figure 1-2. Summation of Gain and Losses in Typical Communication Circuit

adding the gains or losses algebraically as shown. Corresponding power levels, assuming 1-watt input to the power amplifier, are shown along the bottom of the figure.

Since decibels refer to ratios, they may only be used as a measure of absolute magnitude when the reference level is known. For example, to state a power level as 20 db would tell nothing unless it was stated that the level was 20 db above or below some given reference. Common references used in communications are dbw and dbm. Power levels expressed in dbw are referred to a level of 1 watt

$$\left(\text{dbw} = 10 \log \frac{P_x}{1}\text{watt}\right) .$$

Thus, a level expressed as 10 dbw means that this level is 10 db above one watt (10 watts); and -10 dbw in a power level 10 db below a watt (0.1 watt). In figure 1-2, all of the circuit gain values shown could be expressed in dbw, where the reference level would be the 1-watt input. For example, the power amplifier output is 30 dbw, and input to the receiver is -66 dbw. Power levels expressed in dbm are referred to reference level of 1 milliwatt

$$\left(\text{dbm} = 10 \log \frac{P_x}{0.001}\right) .$$

The same power level can be expressed in dbw or dbm. For example: assume a power level of 100 watts.

$$\text{dbw} = 10 \log \frac{100}{1} = 10 \log 100 = 20$$

$$\text{dbm} = 10 \log \frac{100}{.001} = 10 \log 100{,}000 = 50$$

As can be seen by reference to the expressions for dbw and dbm, a power level of one dbw is 1000 times larger than a power level of one dbm. This ratio expressed in decibels is:

$$\text{db} = 10 \log \frac{\text{dbw (1)}}{\text{dbm (.001)}} = 10 \log 1000 = 30$$

Therefore, to convert dbw to dbm add 30; to convert dbm to dbw subtract 30.

Other reference levels are used for certain noise measurements in telephone circuits. For example, when noise measurements are being made, the symbol db RN means decibels above reference noise. This is a unit which shows the relationship between the interfering effect of a noise frequency, or band of noise frequencies, and a fixed amount of noise power commonly called reference noise. A tone of 1000 cps with a power level of -90 dbm was selected as the reference noise power. This value was selected for reference because noise at this level in a telephone circuit appeared to have negligible interfering effect.

The reference level db RN was originally determined by measuring the noise in telephone circuits which used Type 144 Handsets. Later an improved type of handset (Western Electric Company, Type F1A) came into general use. This required an adjustment of the reference noise level and a different unit. The reference level was changed to -85 dbm, and the new unit was called dba. This is an abbreviation for db RN adjusted.

CHAPTER 2
RADIO WAVE PROPAGATION

2.1 INTRODUCTION.

The principles involved in tropospheric scatter propagation are a logical development from the principles and theories for all other types of radio wave propagation. Therefore, in order to understand the mechanics of tropospheric scatter propagation, it is necessary to know certain basic fundamentals of radio wave propagation. This chapter provides you with this required background information and serves as an introduction to chapter 3, where the details of tropospheric scatter are discussed.

2.2 GENERAL CHARACTERISTICS OF RADIO WAVES.

Radio waves are a form of electromagnetic radiations similar to light and heat. They differ from these other radiations in the manner in which they are generated and detected and in their frequency range which is from approximately 10 kc to 30,000 mc. This spectrum is divided into various bands of frequencies as shown in the following table.

TABLE 2-1. RADIO FREQUENCY SPECTRUM

FREQUENCY	DESCRIPTION	ABBREVIATION
10 kc - 30 kc	very-low-frequency	vlf
30 kc - 300 kc	low-frequency	lf
300 kc - 3 mc	medium-frequency	mf
3 mc - 30 mc	high-frequency	hf
30 mc - 300 mc	very-high-frequency	vhf
300 mc - 3000 mc	ultrahigh-frequency	uhf
3000 mc - 30,000 mc	superhigh-frequency	shf

Radio waves travel at the same velocity as light waves which in free space is approximately 186,000 miles per second or 300,000,000 meters per second. The wavelength of any radio wave can be found by the formula: $\lambda = \frac{c}{f}$ where λ is the wavelength, f the frequency in cycles, and c the velocity in meters or miles per second.

A radio wave consists of traveling electric and magnetic fields, with the energy evenly divided between the two types of fields. The lines of force of these fields are at right angles to each other, in a plane perpendicular to the direction of travel. The polarization of the radio wave is determined by the direction of the lines of force in the electric field as shown in figure 2-1. If the direction of the electric field is perpendicular to the earth, the wave is vertically polarized. If the direction of the electric field is parallel to the earth, the wave is horizontally polarized.

Radio waves can be reflected, refracted, and diffracted in a manner similar to light and heat waves. They may be reflected from any sharply defined discontinuity of suitable characteristics and dimensions encountered in the medium. The wave is not reflected from a single point on a surface, but rather from an area. The size of the area required for reflection depends on the wavelength and angle of incidence. When a wave is reflected from a plane surface, a phase shift occurs, as shown in figure 2-2. The amount of phase shift depends on the polarization of the wave and the angle of incidence.

As in the case of light, a radio wave is bent when it moves from one medium into another in which the velocity of propagation is different from that of first medium. This bending is called refraction. The

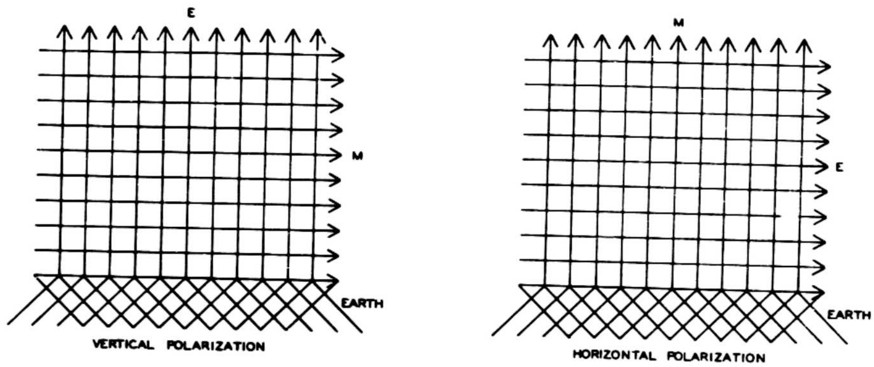


Figure 2-1. Representation of Magnetic and Electric Fields in Wave Front

bending is always toward that medium in which the velocity is the least. If the wave front is traveling obliquely away from the earth, and it encounters a medium with a higher velocity of propagation, the part of the wave that enters the new medium travels faster. This tends to swing the wave front around or refract it in such a manner that is is directed back to earth.

A radio wave is also bent when it passes the edge of an object. This bending, called diffraction, results in a change of direction of part of the energy from a straight or line-of-sight path. This makes it possible to receive energy at some distance below the summit of an obstruction or around its edges.

The wave which is transmitted from an antenna can be considered to have three major components. One component is the ground wave which is that part of the total radiated energy that travels along the ground and follows the curvature of the earth. It is directly affected by the presence of the earth and its surface features. Another component is the tropospheric wave. This is that part of the total radiation that undergoes refraction, reflection or scattering from regions in the troposphere. Another component is the ionospheric wave which is radiated

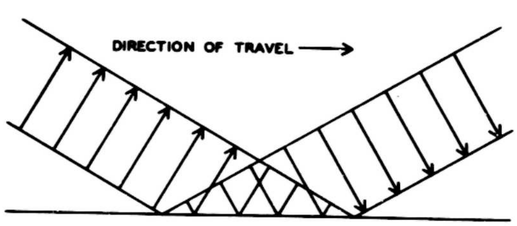


Figure 2-2. Reflection of Electromagnetic Energy from A Plane Surface

in an upward direction and returned to the earth at some distant location due to reflection, refraction or scattering from the ionosphere.

All of these waves are affected by the condition of the earth's surface and by the atmosphere. However, in order to introduce fundamental concepts of propagation in a simple manner and to obtain a formula that is used in figuring losses for all types of propagation, we begin by considering the properties of transmission in free space. This is transmission between two antennas so isolated in space that no objects exert measurable influence on the transmission.

2.3 TRANSMISSION IN FREE SPACE.

There is a certain amount of attenuation or loss of energy for radio signals transmitted in free space. This loss is due to the spreading of energy over a greater area as the transmission distance is increased. The loss is directly related to the frequency and transmission distance. The formula for free-space loss is:

$$L_{FS} = 37 + 20 \log F + 20 \log D$$

where L_{FS} is ratio of transmit power to receive power in decibels,

F is frequency in mc,

D is transmit distance in miles.

This is an important formula, and an understanding of it is required for understanding all other propagation losses. In the following paragraphs, this formula is developed and explained.

The circumstances required for true free space would be realized only if the transmit and receive antennas were isolated in unbounded empty space.

However they are realized well enough for practical purposes if the following conditions are fulfilled:

a. No large obstacles intervene between the antennas along an optical line of sight.

b. No alternate transmission path can be followed by a substantial fraction of the radiated energy.

c. The intervening atmosphere has a constant index of refraction so that no bending of the wave occurs at the frequency used.

d. The intervening atmosphere does not absorb energy from the wave at the frequency used.

If these conditions are fulfilled, the transmitted wave has spherical wavefronts. These spherical wavefronts spread so that the intensity of radiated energy varies inversely as the square of the distance. The intensity of energy is the power per unit of area on the spherical front. This relationship between intensity of energy and distance is illustrated in figure 2-3 which shows the pattern of radiated energy from an isotropic antenna in free space. An isotropic antenna radiates energy uniformly in all directions or receives energy equally well from all directions. Referring to figure 2-3, assume a given unit of area (A) on the surface of a sphere at a distance of D_1 from the isotropic antenna. The total area of the sphere at this distance is $4\pi D_1^2$. Since power is uniformly distributed over the entire area of the sphere, the fraction of total power in the area A is equal to $\frac{A}{4\pi D_1^2}$. Now increase the distance to D_2 and consider the intensity of energy on the same area A on the sphere at this distance. The area of the sphere has increased, and therefore area A has become a smaller fraction of the total area. Thus, the fraction of total power incident to A is decreased.

The power flow through a unit area at a distance D from an iostropic antenna is found by dividing the total radiated power by $4\pi D^2$. However, if a directive antenna is used, the energy is concentrated in certain directions, and there is not uniform distribution over the sphere. In this case, the power flow through a unit area at a given distance differs by factor G from that which would be produced by an isotropic antenna. This factor G is called the gain of the antenna. The greater the concentration of energy in a given direction, the greater the gain will be in that direction. By definition, an isotropic antenna has a gain of 1 in all directions. A directive antenna has a gain greater than 1 in some directions and less than 1 in other directions. However, the total gain taken over the entire sphere must be 1.

The formula for antenna gain is: $G = \frac{4\pi A}{\lambda^2}$ where A is the area of the antenna aperture, and λ is the wavelength of the transmited signal. The units for A and λ must be the same. This expression is derived as follows: Using an aperture with dimension X in both directions, the angular width of the beam determined by diffraction is about $\frac{\lambda}{X}$ radians. The radiated power is then concentrated in a solid angle beam of $\frac{\lambda^2}{X^2}$. An

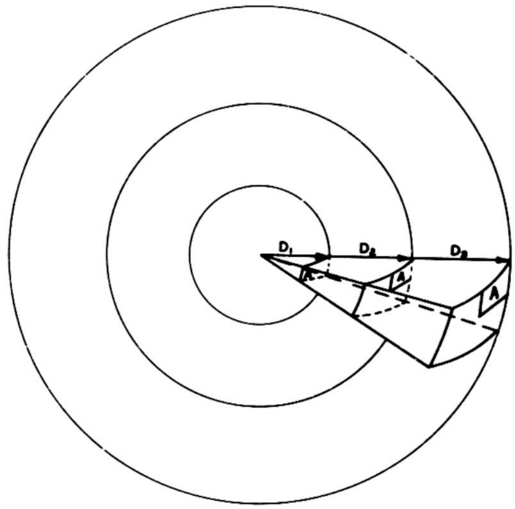


Figure 2-3. Free-Space Pattern of Isotropic Antenna Showing Change in Intensity of Radiated Energy with Distance

isotropic antenna should spread the same power over a solid angle of 4π. Therefore, the gain in concentration of energy is equal to: $\frac{4\pi}{\frac{\lambda^2}{X^2}} = \frac{4\pi X^2}{\lambda^2}$ since X^2 equals the area of the aperture A, $G = \frac{4\pi A}{\lambda^2}$.

Before an equation for attenuation of radio waves in free space can be derived, one other factor must be defined. This is called the effective receiving cross section of an antenna. The effective cross section is equal to the total signal power available at the antenna terminals divided by the power density (power per unit area) of the incident wave. This, in most cases, is different from the actual physical area of the antenna. The effective cross section is a quantity which tells the effectiveness of the antenna in capturing the power in the incident wave. If all of the energy incident on the aperture A is absorbed, then the effective cross section is equal to the area of the aperture. The formula for effective cross section:

$$A_R = \frac{G\lambda^2}{4\pi}$$

All of the discussion so far concerning free-space propagation can be summarized by providing the equation for power received over a free-space circuit. This equation is: $P_R = P_T\left(\frac{G_T A_R}{4\pi D^2}\right)$

where: P_R is the total power at the output terminals of the receiver antenna.

P_T is the power input to the transmit antenna
A_R is the effective cross section of the receive antenna
G_T is the gain of the transmit antenna
D is the distance between antennas.

This equation shows the inverse relationship between the received power and distance. It also shows that the received power is directly dependent upon the amount of power transmitted, the gain of transmit antenna, and the effective cross section of the receive antenna.

Assuming isotropic antennas, $G_T = 1$ and $A_R = \frac{\lambda}{4\pi}$,

therefore, for isotropic antennas in a free-space circuit, the expression for total received power is:

$$P_R = P_T\left(\frac{\frac{\lambda^2}{4\pi}}{4\pi D^2}\right) = P_T\left(\frac{\lambda^2}{4\pi}\right)\left(\frac{1}{4\pi D^2}\right) = P_T\left(\frac{\lambda^2}{16\pi^2 D^2}\right)$$

Free-space loss or attenuation is the difference between the input power to the transmit antenna and the output from the receive antenna $P_T - P_R$ assuming isotropic antennas at each end of the circuit. In decibels, this free-space loss is equal to $10 \log \frac{P_T}{P_R}$. Using the expression given above for received power with isotropic antennas, the formula for free-space loss is developed as follows:

$$\frac{P_T}{P_R} = \frac{16\pi^2 D^2}{\lambda^2}$$

For F in megacycles and C in miles per second:

$$\lambda = \frac{C}{F} = \frac{186000 \times 10^{-6}}{F} = \frac{0.186}{F}$$

Therefore: $\frac{P_T}{P_R} = \frac{16\pi^2 D^2}{\left(\frac{.186}{F}\right)^2} = \frac{16\pi^2 D^2 F^2}{(.186)^2}$

In decibels free-space loss = L_{FS}

$$= 10 \log \left(\frac{16\pi^2 D^2 F^2}{.186^2}\right)$$

$= 10 \log 16 + 20 \log \pi + 20 \log D + 20 \log F - 20 \log 0.186$

The constants $10 \log 16 + 20 \log \pi - 20 \log 0.186$ equal 37. Therefore, the final expression for free-space loss in db with F in megacycles and D in miles is: $L_{FS} = 37 + 20 \log D + 20 \log F$.

To see how this expression is used, try a simple example. Find the free-space loss over a circuit with a distance of 100 miles between transmit and receive antennas and a transmit frequency of 1000 mc.

$L_{FS} = 37 + 20 \log 100 + 20 \log 1000$

$L_{FS} = 37 + (20 \times 2) + (20 \times 3)$

$L_{FS} = 37 + 40 + 60 = 137$ db

This means that the output from the receive antenna will be down 137 db from the power input to the transmit antenna because of free-space loss alone assuming isotropic antenna; or expressed another way, the received power is approximately 0.00000000000001 of the transmit power input.

It can be seen from this one example that free-space loss introduces a substantial attenuation to the transmitted signal. This is the basic loss which occurs for all types of radio transmissions. For line-of-sight circuits, where the conditions for free-space propagation are closely approximated, the total loss can be considered to be the free-space loss. However, for long distance communications, where either ground-wave, sky-wave, or scatter propagation is used, other losses are introduced by the effects of the earth and atmosphere. Each of these losses must be added to the free-space loss to find the total attenuation to the transmitted signal.

2.4 IONOSPHERIC PROPAGATION.

2.4.1 THE IONOSPHERE.

A basic understanding of the ionosphere, what it is, and how it is formed, is required to understand how it affects radio wave propagation. The following paragraphs provide a description of the ionosphere.

The ionosphere is the area of the atmosphere in which the gas atoms are charged by ultraviolet light from the sun and by meteor activity. These charged atoms are called ions, and the process by which they are charged is called ionization.

This ionization occurs as follows. When a high-energy electromagnetic wave, such as ultraviolet light, hits an atom, it is capable of knocking an electron completely out of an atom. When this occurs, a positively charged atom, called a positive ion, remains in space along with the negatively charged free electron. The electron has absorbed energy from the incident wave. The rate of ion and free-electron formation depends upon the density of the atmosphere and the intensity of the ultraviolet light. As the ultraviolet wave produces positive ions and free electrons, its intensity is decreased because of the absorption of energy by the free electrons. Therefore, the ionized region will tend to form in a layer for a given frequency of ultraviolet light. When the wave first hits the atmosphere it has high energy, but the atmosphere is not dense, therefore little ionization will occur. As the wave passes through the atmosphere, the density increases, but the energy level of the wave decreases. An ionization layer is formed where the combined effect of atmospheric density and wave energy is maximum.

Since there are different ultraviolet wave frequencies, several ionized layers are formed, as shown in figure 2-4. The lower frequency ultraviolet waves tend to produce a higher ionic layer and lose all of their energy at the high altitude. The higher frequency waves tend to penetrate deeper before producing appreciable ionization.

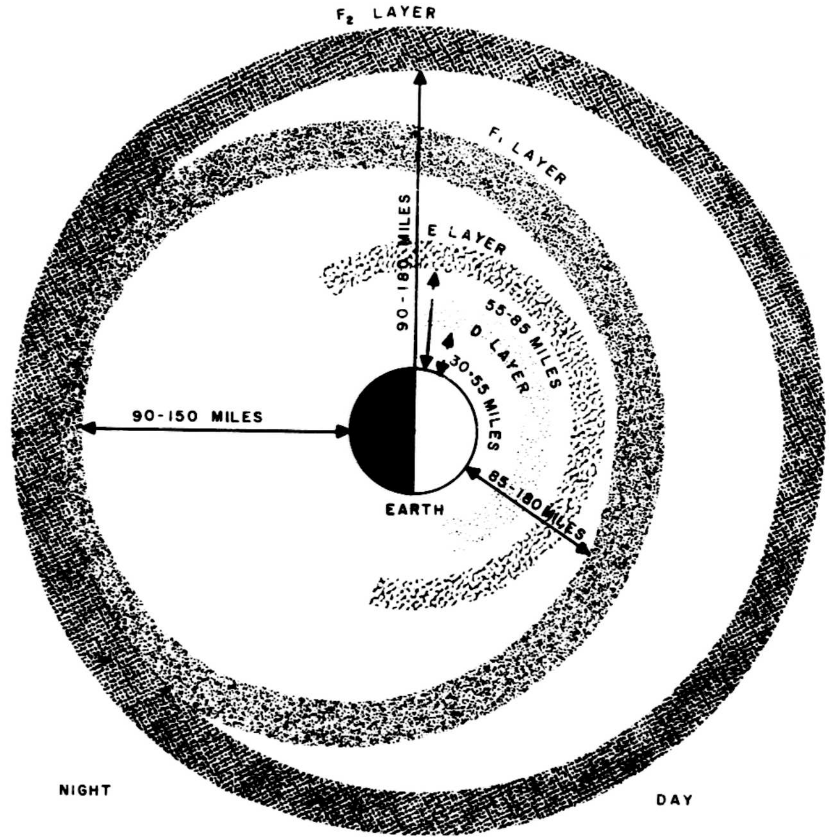


Figure 2-4. Structure of the Ionosphere

Figure 2-4 shows that there are four distinct layers designated D, E, F1, and F2 in the order of their height from the earth's surface. The height, thickness, and intensity of ionization for each of the layers is determined by transmitting r-f pulses vertically into the atmosphere and then receiving the reflected pulse. The echo time indicates height of the ionospheric layer, and the strength of the received signal indicates thickness of the layer. When pulses of various r-f frequencies are transmitted, there will be found a frequency above which the vertical sky wave will not be reflected back to earth. This frequency, called the critical frequency, indicates the extent of ionization. The higher the critical frequency, the greater the ionization.

Since the ionospheric layers are caused chiefly by ultraviolet light emitted from the sun, their height and thickness change with the season and time of day. The D layer exists only in the daytime. Very little sky-wave reflection is obtained from this layer, and there is a very pronounced absorption effect at frequencies below 2 mc. The E layer exists only during daylight hours at a height of about 55 to 85 miles. The F1 layer exists at a height between 85 and 155 miles during daylight hours. When the sun sets, the F1 layer merges with the next higher ionic layer, the F2 layer. The F2 layer is the most useful ionic layer for sky-wave transmission because it exists during the night as well as the day. This layer is between 90 and 150 miles above the earth during the night for all seasons. During the day in the summer, it is between 90 and 180 miles high. This variation in height is accounted for by the effect of solar heat which increases the height of the layer and decreases its ion density during the summer. The reduction of solar heat in the late afternoon causes the layer to descend.

Besides the seasonal and daily changes in the ionosphere, there are also other variations which occur. There is a noted increase in ionization with an increase of sunspot activity. These sunspots produce vast amounts of ultraviolet energy. Therefore, the greater the number of sunspots, the greater the amount of ionization. Observers of solar activity over the past 100 years have confirmed that sunspot activity is cyclic, with the cycle repeating every 11.1 years. There are variations within this cycle and variations from cycle to cycle, which make it necessary to know the predicted sunspot number for a given time in order to determine the probability of sky-wave communication.

Another change in the ionosphere is called sudden ionospheric disturbance (SID). When a SID occurs, daytime communication by high-frequency sky-wave propagation is made impossible by great absorption in the ionosphere. This condition usually starts very suddenly and may decrease very gradually. The condition may last from a few minutes to several hours. A SID is accompanied by a sudden increase in the ion density of the highly absorptive D region, as well as an increase in the moderately absorptive E layer. A SID is apparently the result of some eruption on the sun.

Magnetic storms also cause changes in the ionosphere which affect the sky-wave propagation. Magnetic storms are apparently caused by particle radiation from the sun, with the radiated particles being deflected by the earth's magnetic field. For this reason, the effects of a magnetic storm are most severe in the two geomagnetic poles and result in a marked increase in absorption in the polar regions.

Another phenomenon associated with the ionosphere is called the sporadic E layer. This occurs roughly at the height of the E layer at irregular times and locations, both day and night. The critical frequency of the sporadic E layer may be much higher than that of the regular E layer, and the reflection may be very effective at frequencies up to 60 mc.

Meteors may also cause an increase in ionization as they drop through the atmosphere. A meteor, as it enters the earth's surface, heats by friction against the atmosphere and leaves a trail of ionized atmosphere. During the time that it takes the ionized molecules to recombine, a small ionized cloud exists. If this cloud is in the signal path, it may cause an increase in received signal strength because of increased reflections from the ionosphere.

2.4.2 REFRACTION IN THE IONOSPHERE.

When a radio wave is transmitted into the ionosphere, the wave is refracted. This refraction or bending is caused by a change in refractive index of the medium as the wave passes from the nonionized atmosphere into an ionized region. The amount of bending depends upon the electron density of the ionized region and the frequency of the transmitted wave. The effect of each of these factors is discussed in the following paragraphs.

Each ionized layer consists of a central region of relatively dense ionization which tapers off in intensity both above and below the maximum area. As a wave passes into a region of increasing electron density, the velocity of the upper part of the wave increases. This causes the wave to bend back toward the earth. If the wave enters into a region of decreasing electron density, the velocity of the upper part of the wave decreases, and the wave is bent away from the earth. The amount of bending in both cases depends upon the variation in electron density with height in the ionized region. If there is an appreciable change in ionization within a short interval of travel, the wave will be bent so rapidly that reflection may take place. This rapid change in ionization required for a reflection must take place within a difference in height comparable to a wavelength. Therefore, ionospheric reflection is more apt to occur at longer wavelengths (lower frequencies). If there is a gradual change in electron density, the wave is bent more slowly as shown in figure 2-5.

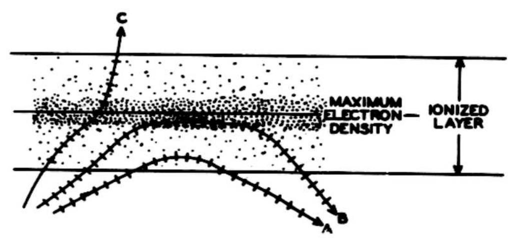


Figure 2-5. Change in Bending of Waves in Ionized Layer with Change in Angle of Incidence

The amount of bending required to return the wave at a given frequency to earth depends on the angle at which the wave enters the ionized region (incident angle). This relationship is shown in figure 2-5. Wave A approaches the ionized region at a small angle, and only a slight amount of bending is required to return the wave to earth. Wave B approaches at a larger angle and penetrates deeper into the ionized region. A longer path is required for bending because variation in density with height is slight. Wave C approaches at almost vertical incidence, and the ionized region is unable to return the wave to earth.

The amount of bending which occurs for a given incident angle and ionization condition depends on the frequency of the wave. For a given electron density, the degree of refraction becomes less as the wavelength becomes shorter (as the frequency increases). The bending, therefore, becomes less as the frequency increases. As the frequency is raised, the bending becomes too slight to bring the wave back to earth even though the incident angle may be very small.

The amount of bending, then, that is required to return a wave to earth is dependent upon the incident angle, and the amount of bending which will occur for a given ionization condition decreases as the frequency is increased. The amount of bending required to return a wave to earth is, of course, maximum with a vertical incident angle. If transmission is at vertical incidence and continues to increase the frequency, the lower frequencies will be returned, but eventually a point will be reached where the signals are not returned. The highest frequency that will be returned at vertical incidence is called the critical frequency. This will vary with the seasons, time of day or any other effects which cause the density of the ionosphere to change. The critical frequency is greater in the daytime than at night.

As the incident angle is decreased, the highest frequency which will be returned is increased. The factors which determine the actual incident angle which can be used for a communication circuit are the height of the ionized layer used for refraction and the distance between the two ends of the circuit. The maximum frequency which will be refracted back for a given distance of transmission is called the maximum useful frequency (MUF). This is always greater than the critical frequency.

Experience has shown that the MUF may vary during the operating period because of changes in the ionosphere. Therefore an optimum working frequency (OWF) slightly lower than the MUF is used so that variations in the ionosphere will have less effect on the communication circuit. If the frequency selected for use is too high, the required refraction will occur at a lower incident angle, and the signal will be transmitted beyond the receiving site. If the frequency is too low, the attenuation in the D layer will be increased.

2.4.3 ABSORPTION IN THE IONOSPHERE.

As the radio wave passes into the ionosphere, it interchanges energy with free electrons and ions. If the ions do not collide with gas molecules or other ions, all of the energy is reconverted back into electromagnetic energy, and the wave continues to be propagated with no change in intensity. However, if the ions do collide with other particles, they dissipate the energy which they have acquired from the wave. This results in absorption of the energy from the wave. Since absorption of energy is dependent upon collision of particles, the greater the density, the greater the probability of collisions, and therefore the greater the absorption. The higher density D and E layers provide the greatest absorption for the ionospheric wave.

Because the amount of attenuation to the sky wave depends upon the density of the ionosphere which varies with seasonal and daily conditions, it is impossible to express a concrete relationship between distance and field strength for ionospheric propagation. Under favorable conditions, only free-space attenuation will occur. Under other conditions, the absorption is so great that communicating over any distance is difficult. Due to variations in the atmosphere, the sky-wave intensity varies from minute to minute, month to month, and year to year.

2.4.4 FADING.

Signals received over an ionospheric path may vary in intensity over short periods of time. There are three major reasons for this fading. When the wave is reflected from surfaces in the ionosphere or the earth's surface, random variations in polarization of the wave may occur. These will cause changes in received signal level. Fading may also occur if the operating frequency is too close to the MUF. If this is the case, any slight change in the ionosphere will cause a change in signal level. The major reason for fading on ionospheric circuits is multipath propagation. This is described in the following paragraphs.

Figure 2-6 shows the various paths a signal can travel between two sites in a typical circuit. One signal from the transmitter may follow the path XYZ, which is the basic ground wave. Another signal follows the path XEA. It is reflected from the E layer and received at A, but not at Z. Another path is XFZFA which results from a higher transmission angle and two reflections from the F layers. At point Z , the received signal is the resultant of the ground wave and the sky wave. If these two waves are received in phase, a stronger signal results. If they are out of phase, they will produce a weak signal. Small changes in the path may change the phase relationship of the two signals, periodically causing fading. The same adding of various signal components occurs at point A. At this point, the double-skip F layer reflection signal may be in or out of phase with the signal arriving from the E layer.

Fading resulting from multipath propagation is frequency sensitive since each frequency arrives at the receive point over a different path. When a wide

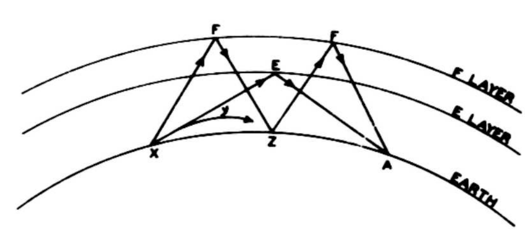


Figure 2-6. Multipath Transmission Which Could Result in Fading

band of frequencies is transmitted, all of the frequencies will not encounter the same amount of fade. This is called selective fading. When selective fading occurs, all of the frequencies in the transmitted band may not be propagated in the same relative amplitude and phases they had at the transmitter. This may cause severe distortion of the signal and limit the total bandwidth which can be transmitted.

2.4.5 IONOSPHERIC SCATTER.

In explaining refraction and describing the ionosphere, it has been assumed that the density variation and variation in refractive index within a layer is gradual and uniform. This picture of the ionosphere provides a very satisfactory explanation for most radio wave phenomena associated with the ionosphere. However, there are also turbulent and irregular variations in the ion density and refractive index of the ionosphere. These irregular variations result in the scattering of the radio wave. Figure 2-7 illustrates how scattering can occur. When the wave front encounters a sudden change in ion density, irregular variations in the wave front result.

Scattering of radio energy by the ionosphere is similar to the scattering of light by small water droplets in a cloud or fog. When a beam of light shines through the

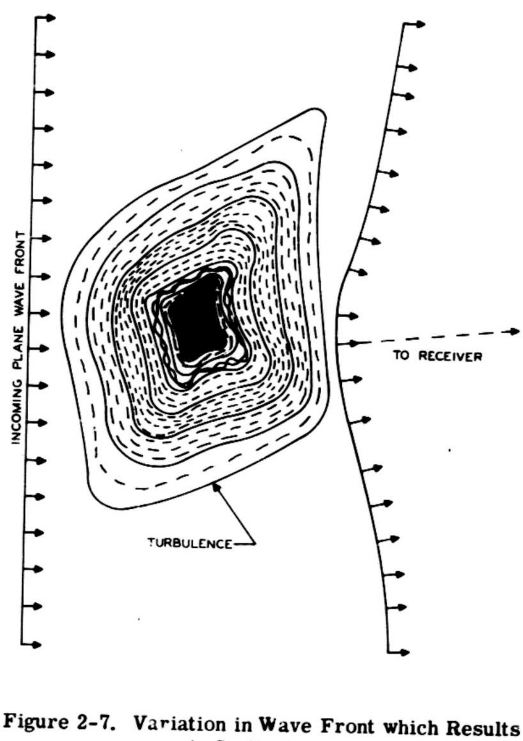


Figure 2-7. Variation in Wave Front which Results in Scattering

mist, moving points of light are seen in the beam. These are caused by water droplets which scatter light to the observer.

This scattering takes place from the E layer of the ionosphere. Under suitable conditions, this type of propagation may be used for large incident angles with frequencies as high as 100 mc. Ordinarily, radio waves with frequencies up to about 25 mc will pass directly through the E layers and are refracted from the F layer. However, the scattering process causes some of this energy to be returned from the E layer to the receiving antenna.

Only a very small percentage of the total energy is returned. If the scattering region is within visible range of the transmitting and receiving antennas, the total loss is free-space loss plus a scatter loss which is dependent upon the size and strength of the irregular variation in the medium, and the angular change in direction of the wave front. This scatter loss is large, being in the range of 60 to 100 or more db. Since the scatter loss is so large, this form of propagation requires the use of high-power transmitters and highly directional antennas.

Signals received over an ionospheric scatter circuit are weak, but do not show the extreme changes in level which are sometimes encountered with other types of propagation. The field strength reaches a minimum value in a daily cycle at about 2000 local time. There is, however, no nighttime disappearance of the signal as in the case of regular E-layer propagation. There is also an annual cycle with minimum field strength during the spring and fall seasons. The received signal is also characterized by rapid fading, punctuated by strong bursts evidently associated with ionized meteor trails.

The principal factors that limit the utilization of ionospheric scatter are frequency and distance. The useful frequency range for ionospheric scatter extends from about 20 to 60 mc. The region in which the scattering occurs is approximately fixed at a height of about 60 miles. This determines the optimum range of distance which is about 500 to 1100 miles.

2.4.6 SUMMARY.

The ionospheric wave is that part of the total radiated energy which is propagated upward and returned to the earth by reflection, refraction or scattering in the ionosphere. The ionosphere is not a constant region, but changes in height, thickness, and density with the time of day, the seasons, and cyclic sunspot variations in the sun. These changes affect the received signal level. The ionospheric method of propagation is also frequency sensitive. Low frequencies suffer high absorption, and high frequencies penetrate the ionosphere and are not refracted. The maximum useful frequency is the highest frequency which will be returned to the receiving site over the operating circuit. A frequency slightly less than the maximum usable frequency is normally selected for use on a circuit. This is called the optimum working frequency. If it is too high, the circuit will be unreliable, since the received level will be too dependent on changes in the ionosphere. If the selected frequency is too low, the signal will suffer high absorption in the ionosphere.

2.5 GROUND WAVE PROPAGATION.

2.5.1 DEFINITION.

The ground wave is that part of the total radiated energy which is propagated at a low angle from the antenna and is therefore directly affected by the presence of the earth and its surface features. The ground wave can be considered to be composed of two major components: a surface wave, and a space wave. The space wave is composed of a direct wave and a reflected wave. These various waves and their relationship are shown in figure 2-8.

The transmission characteristics of the ground wave are determined by the frequency of the wave, the earth's surface conditions, and conditions of the lower atmosphere. The effects of each of these conditions on each of the components of the ground wave are discussed in the following paragraphs.

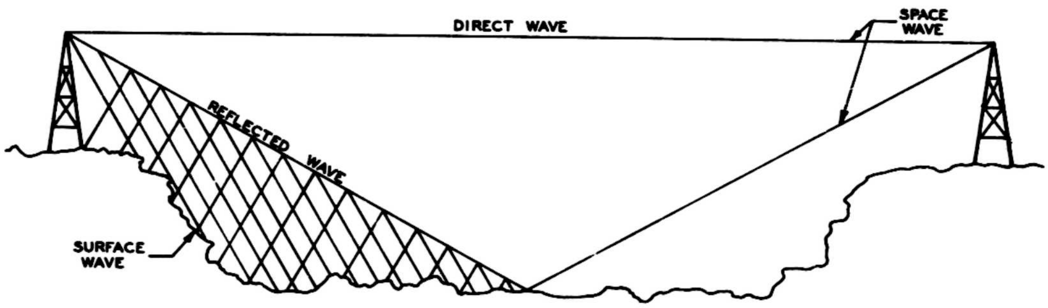


Figure 2-8. Components of the Ground Wave

2.5.2 SURFACE WAVE.

The surface wave is an earth-guided wave which glides over the surface of the earth. It must be vertically polarized or the entire field would be shorted to the earth. As the wave moves over the surface, it is accompanied by charges induced in the earth's surface. The charges move with the wave to make up a current. Since the earth offers resistance to the flow of this current, energy is dissipated. This energy is absorbed from the surface wave, and the portion of the wave in contact with the earth is continuously wiped out. This is replenished by diffraction of energy downward from the portions of the ground wave immediately above the earth. This produces a slight forward tilt in the wave front as shown in figure 2-9.

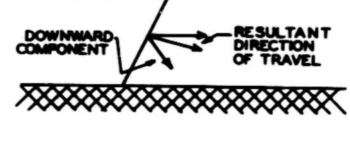


Figure 2-9. Movement of Surface Wave Over Earth's Surface

Attenuation of the surface wave due to absorption depends upon the condition of the earth's surface and the transmitted frequency. The table below gives the relative conductivity of various earth surfaces.

SURFACE	RELATIVE CONDUCTIVITY
Sea water	Good
Flat, loamy soil	Fair
Large bodies of fresh water	Fair
Rocky terrain	Poor
Desert	Poor
Jungle	Unusable

From this table, it is apparent that the best surface wave transmission occurs over sea water and that the highest degree of attenuation is found over jungle areas. Attenuation over all types of terrain increases rapidly as the frequency is increased. Extremely high losses make it impractical to use the surface wave for long-distance transmissions with frequencies above the broadcast band.

Since fading is due primarily to atmospheric fluctuations, there is no fading associated with surface wave propagation. The received signal level is also independent of seasonal and day and night effects.

2.5.3 SPACE WAVE.

When the transmit antenna is elevated, a space wave is propagated in addition to the surface wave. This wave consists of two rays--direct and ground reflected--as shown on figure 2-8.

The characteristics of the space wave are dependent upon the distance from the transmit antenna. There are two definite regions with different characteristics. These are called the line-of-sight region and the diffraction region as shown on figure 2-10.

Figure 2-10. Two Regions in Space Wave Propagation

The line-of-sight distance, which is the straight line distance from the antenna to the horizon, is, of course, dependent upon antenna height and the curvature of the earth. The radio horizon is slightly beyond the actual horizon. This extension of the horizon distance results from refraction in the earth's atmosphere. The refractive index decreases with height above the earth's surface because of changes in humidity, pressure, and temperature. As a radio wave passes from a region of high density to a region of low density, it is bent back toward the earth. This bending effectively decreases the earth's curvature and extends the radio horizon. The effect varies with atmospheric and climatic conditions. However, for average conditions, the direct path distance of a radio wave can be plotted as a straight line if the earth's radius is increased by a factor of 4/3 (1.33). This factor may vary from 1.1 in cold, dry climates to 1.6 in hot, humid climates. Unless otherwise specified, the value is assumed to be 1.33. The actual radius of the earth is 3960 miles. The effective radius for calculating radio horizon is 5280 miles (3960 x 4/3).

Within the line-of-sight distance, the direct and reflected rays suffer little attenuation other than that caused by spreading, and therefore each has a field strength inversely proportional to distance. The total received signal is the sum of the field strength of the two rays. If the two rays are close to 180° out of phase, the received signal level will be very low. If they are in phase, the received signal level will be high. Therefore, the line-of-sight field can vary from a very low value to a value equal to approximately twice the free-space field of the direct ray. The actual resultant field strength depends upon the transmission path length.

Since conditions in the atmosphere are constantly changing, the extent to which the direct ray is bent will also vary. This changes the relative path lengths of the direct and reflected rays, and thus causes a change in phase relationship between the two waves. The received signal level varies or fades as the phase relationship of the two rays changes.

The signal level does not drop off abruptly when the radio horizon is reached. The direct ray is diffracted as it passes over the earth's surface at the horizon, and a diffraction field results in the shadow area. The field intensity drops very rapidly with increased distance into the shadow area. The field intensity also decreases with an increase in frequency. Attenuations of approximately 1 db per mile are typical for this shadow region.

The diffraction field is supplemented some 20 or 30 miles beyond the horizon by another component which is easily distinguished by its characteristic fading. The strength of this field is not nearly as sensitive to frequency change as the diffraction field. This supplementary field is generally called the tropospheric scatter field. It is introduced in the following paragraphs and discussed in detail in section 3 of this manual.

2.6 TROPOSPHERIC SCATTER PROPAGATION.

In previous paragraphs, the portion of radiated energy which is acted upon by the ionosphere and returned to earth and the portion of radiated energy which is propagated along the earth's surface are considered. In the following paragraphs, that part of the total radiated energy which undergoes reflection and refraction in regions of abrupt changes of refractive index in the tropospheric will be discussed.

2.6.1 APPLICATION OF TROPOSPHERIC SCATTER PROPAGATION.

When a space wave is transmitted, it undergoes approximately free-space attenuation within the line-of-sight distance. When the horizon is reached, the wave is diffracted and follows the earth's curvature. The rate of attenuation increases very rapidly for distances beyond the horizon, and signals in the diffraction field are extremely weak. Tropospheric scatter provides a usable signal at distances beyond where the diffraction field drops below a usable level.

The scattered wave is able to reach beyond the diffraction field because of the height at which the scattering takes place. The tropospheric region which contributes most strongly to the scatter field lies near midpath and just above the horizon rays of the antennas.

Tropospheric scatter propagation is used for point-to-point communications. A correctly designed tropospheric scatter circuit will provide highly reliable multichannel service for distances from 50 miles to 300 miles. A tropospheric scatter circuit is not affected by atmospheric and auroral disturbances. The usable frequency range extends from approximately 100 mc to 10,000 mc.

The following paragraphs present a description of the troposphere and tell how tropospheric scattering occurs.

2.6.2 CHARACTERISTICS OF THE TROPOSPHERE.

The troposphere is the lowest region of the atmosphere, extending from the ground to a height of slightly over six miles. Virtually all weather phenomena occur in this region of the atmosphere. There is practically no ionization of air molecules in the troposphere. Generally, the troposphere is characterized by a steady decrease of temperature and pressure with height.

The index of refraction in the troposphere is a function of various meteorological variables, such as the amount of water vapor in the air, air temperature, and air pressure. The refractive index decreases gradually with height as the density of the air decreases. However, the change is not uniform. Because of the uneven heating of the earth's surface, the air is in constant motion. This motion causes small turbulences, or eddies, to be formed. These turbulences are quite similar to the whirlpools in a rapidly moving stream of water. The turbulence is most intense near the earth's surface and gradually diminishes with elevation.

For frequencies up to about 30 mc, the wavelength is large compared to the size of the turbulences, and therefore they have little effect on the transmitted signal. However, as the frequency is increased, these local deviations become increasingly important. They are responsible for tropospheric scatter transmission.

2.6.3 TROPOSPHERIC SCATTERING.

As a wave front passing through the troposphere encounters a turbulence, a small amount of energy is scattered away from the incident beam. The scatter effect is the same as if each turbulence received the signal and reradiated it. This is similar to the ionospheric scatter shown in figure 2-7. The total received signal is an accumulation of the energy received from each one of the turbulences.

The word "scatter" implies that the spreading of energy is equally probable in all directions. However, the direction of energy distribution in tropospheric scatter propagation differs only slightly from the direction of the path of the main wave front. The scattering occurs chiefly in the forward direction, and therefore the term "forward scatter" is sometimes used when talking about tropospheric scatter.

The magnitude of the received signal depends on the number of turbulences illuminated in the common volume between the transmitting and receiving antenna beams. This volume is called the scatter volume. The scatter volume is a function of the scatter angle formed by the rays of the transmitting and receiving antennas. The scatter volume and scatter angle are shown in figure 2-11. As the scatter angle is increased, the amount of received scattered energy decreases very rapidly. This is because as the angle is increased, the amount of scattering required to return the signal to the receive antenna is also increased.

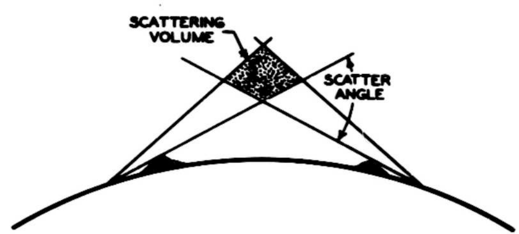


Figure 2-11. Geometry of Tropospheric Scattering

The amount of received energy decreases as the height of the scatter volume is increased. There are two reasons for this. One reason is that the scatter angle increases as the height is increased; the other reason is that the amount of turbulence decreases with height. As the circuit distance is increased, the height of the scatter volume is also increased. Therefore, the received signal level decreases as the circuit distance is increased.

Since tropospheric scatter depends on turbulences in the atmosphere, changes in atmospheric conditions will affect the received signal level. Daily and seasonal variations are noted. These changes are called long term fading. Besides the long-term fading, the tropospheric scatter signal is also characterized by very rapid fading. This fading is caused by propagation over the multiple paths from the scatter volume. The signal level at any one time is the sum of all of the signals received from each of the turbulences in the volume. Since the turbulent condition is constantly changing, the path lengths and individual signal levels are also changing. This results in a rapidly changing signal. Although the signal level is constantly changing, the signal is persistent, and no complete fade out occurs.

Another characteristic of a tropospheric scatter signal is its relatively low level. The scatter volume can be pictured as a relay station, located above the horizon, receiving the transmitted energy and reradiating it to some point beyond the line-of-sight distance. Since most of the transmitted energy is not reradiated to the receiver, the efficiency is very low, and the signal level at the final receive point is low. To compensate for the low efficiency in the scatter volume, the incident power must be high. This is done by using high-power transmitters and high-gain antennas which concentrate the transmitted power into a beam. This increases the intensity of energy on each turbulence in the volume. The receiver must also be very sensitive to detect the low-level signals.

Several factors determine the frequency range most suitable for tropospheric scatter. For frequencies below 30 mc, the troposphere appears to be uniform, and scattering does not occur. Also, at these lower frequencies, it is difficult to construct the required high-gain antennas. As the frequency is increased, the amount of scattering loss and the free-space loss increases. Above 10,000 mc, the wave is greatly affected by atmospheric conditions. Therefore, there seems to be some optimum frequency for tropospheric scatter. This will be discussed in more detail in chapter 3.

This chapter has provided an introduction for tropospheric scatter. A detailed description of the characteristics of tropospheric scatter is provided in the next chapter.

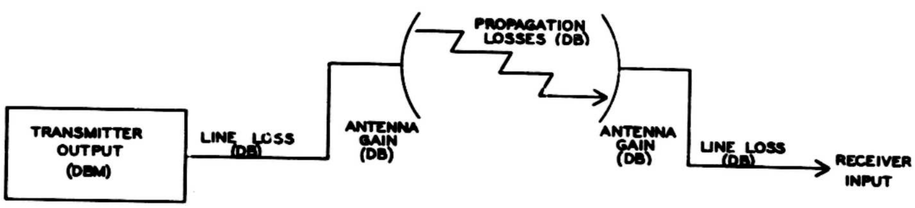
CHAPTER 3
PRINCIPLES OF TROPOSPHERIC SCATTER

3.1 INTRODUCTION.

In chapter 2, the basic fundamentals of radio propagation were discussed. In this chapter, these fundamentals will be enlarged upon and applied in a discussion of the principles of tropospheric scatter. An understanding of these principles is a requirement for understanding the design and operation of tropospheric scatter equipment.

Basically, tropospheric scatter propagation is beyond line-of-sight propagation which uses energy received from a common scatter volume. The transmit antenna illuminates a certain volume in the troposphere. The receive antenna is aimed at this same volume and receives energy scattered from the volume. The antennas are set for point-to-point communication between two sites.

In order to predict performance of a circuit between the two sites, it is necessary to know what the expected received signal level will be. The signal level at the receiver input terminals is the difference between the effective transmitted power and the propagation losses, as shown in figure 3-1. These factors are discussed in the following paragraphs.



RECEIVER INPUT SIGNAL LEVEL = EFFECTIVE TRANSMITTED POWER - PROPAGATION LOSSES
EFFECTIVE TRANSMITTED POWER = TRANSMITTER OUTPUT + TRANSMIT ANTENNA GAIN + RECEIVE ANTENNA GAIN - LINE LOSSES

Figure 3-1 Losses and Gains in a Tropospheric Scatter Circuit

3.2 PROPAGATION LOSSES.

The total propagation loss for a circuit depends on the distance between the antennas, the frequency used, and the type of terrain between the antennas. The effect of each of these factors is described in the following paragraphs.

To begin to figure the losses in a tropospheric scatter circuit, assume a free-space path with a smooth, spherical earth and with isotropic antennas being used. In this case, the circuit loss is given by the expression:

$$L_{FS} = 20 \log F + 20 \log D + 37$$

where: L_{FS} is ratio of transmit power to receive power in decibels
F is frequency in megacycles
D is path length in miles.

However, with a scatter circuit, the path is not free space. It is broken in the middle by the scatter volume which adds another loss to the free-space loss. The sum of the two losses is called the basic propagation loss (BPL). This is the loss figured for a tropospheric scatter circuit if you assume that isotropic antennas are used and that the surface of the earth between the two antennas is perfectly smooth. Any change from these conditions will change the total propagation loss. The expression for BPL is:

$$\begin{aligned} BPL &= \text{free-space loss} + \text{scatter loss} \\ &= 20 \log F + 20 \log D + 37 + \text{scatter loss.} \end{aligned}$$

Therefore, to figure the BPL for a circuit, we have to know the frequency, the circuit distance, and the scatter loss.

The scatter loss has been derived experimentally by measuring losses over tropospheric scatter circuits and then comparing these losses with free-space losses. Figure 3-2 compares scatter loss to free-space loss at 800 mc for distances up to 400 miles.

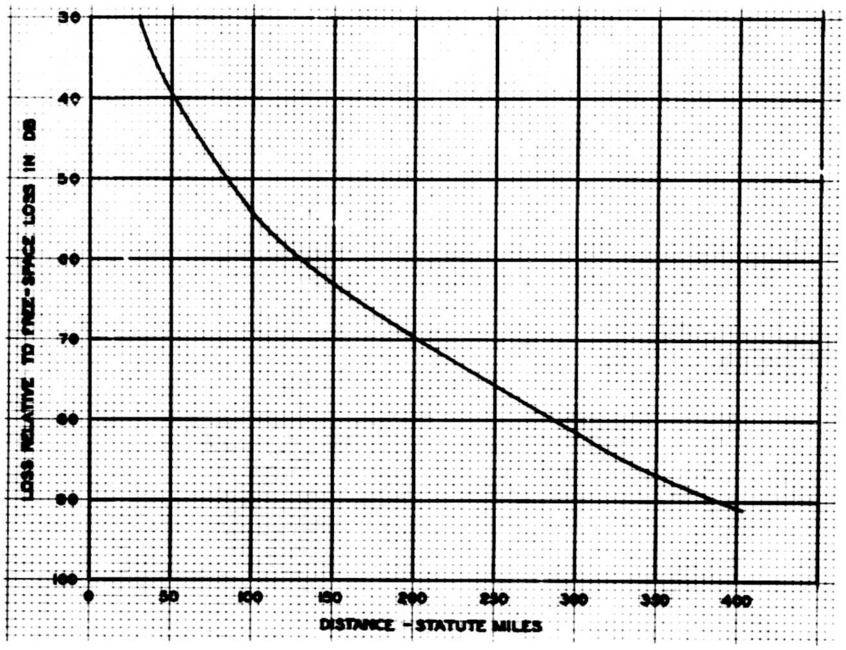


Figure 3-2. Scattering Loss Relative to Free-Space Loss at 800 Megacycles

The curve of figure 3-2 shows the rapid increase of scatter loss with distance. For example, at 80 miles, the scatter loss at 800 mc is 49 db above free-space loss. At 160 miles, the scatter loss increases to 65 db. This is an increase of 16 db as the distance is doubled. In comparison, free-space loss increases with distance according to the expression 20 log D. This means that each time the distance is doubled, free-space loss increases 6 db ($20 \log \frac{D_2}{D_1} = 20 \log 2 = 20 \times 0.3 = 6$).

The rapid increases of scatter loss with distance can be explained by referring again to the path geometry for a tropospheric scatter circuit (figure 2-10). As the circuit distance is increased, the scatter angle and the height of the scatter volume are also increased. The fraction of total radiated energy scattered to the receive antenna decreases very rapidly as these factors are increased.

Figure 3-2 could be used for calculating the basic propagation loss if the operating frequency is 800 mc. For example, assume a circuit distance of 100 miles. The free-space loss is:

$$L_{FS} = 20 \log 800 + 20 \log 100 + 37$$
$$= 20\,(2.9) + 20\,(2) + 37$$
$$= 58 + 40 + 37 = 135 \text{ db.}$$

The scatter loss for 100 miles, obtained from figure 3-2, is 54 db. Therefore, the basic propagation loss for 800 mc at 100 miles is 135.0 + 54 = 189.0 db.

If scatter loss were constant with frequency, figure 3-2 could be used to determine scatter loss for all frequencies. However, experimental results have shown that scatter loss increases with frequency at a rate of approximately 10 log F. This, added to the frequency dependence factor of 20 log F for free-space loss, gives a total frequency dependence of 30 log F for basic propagation loss. This means that each time the frequency is doubled, the basic propagation loss will increase 9 db ($30 \log \frac{F_2}{F_1} = 30 \log 2 = 30 \times 0.3 = 9$).

Figure 3-3 shows the basic propagation loss at 1000 mc for distances up to 400 miles. This curve is based on experimental data obtained over circuits of various path lengths. Since basic propagation loss is directly proportional to frequency, the loss for any other frequency can be determined by correcting the loss shown in figure 3-3 according to the expression $30 \log \frac{F_2}{F_1}$. Figure 3-4 shows the required frequency corrections relative to 1000 mc.

The basic propagation loss for a given frequency is obtained by adding the correction factor, shown in

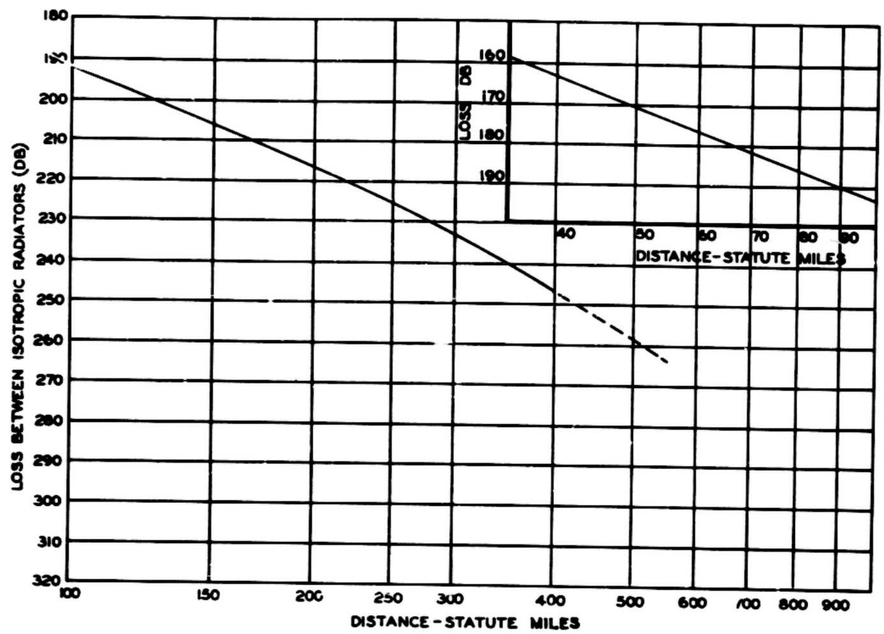


Figure 3-3. Basic Propagation Loss at 1000 Megacycles

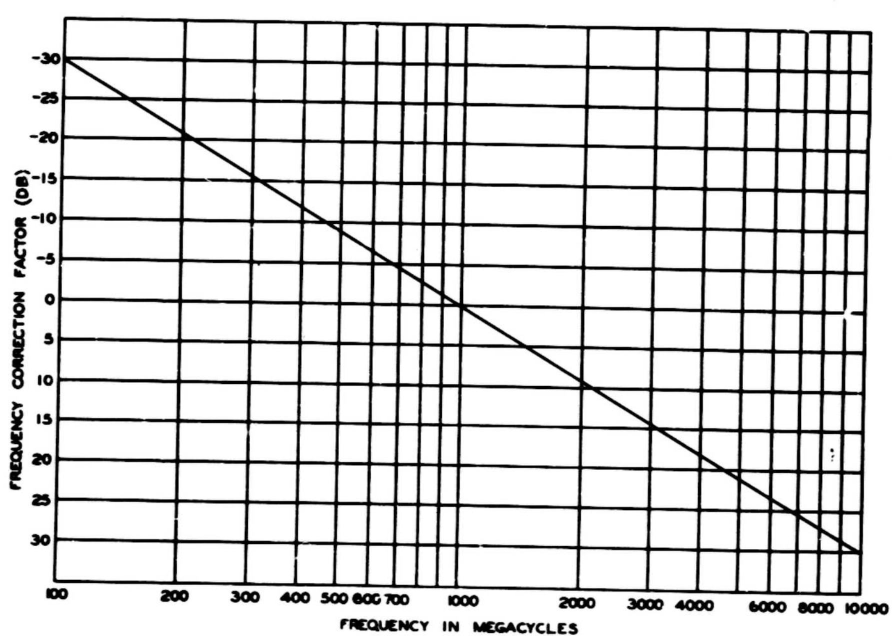


Figure 3-4. Basic Propagation Loss Relative to the Loss at 1000 Megacycles

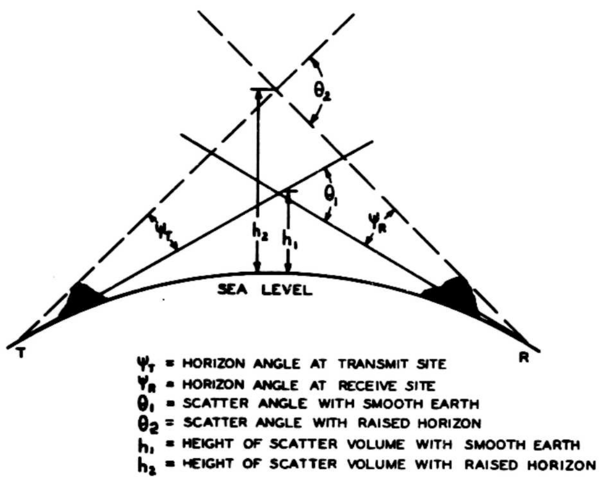


Figure 3-5. Effect of Positive Horizon Angle on Height of Scatter Volume and Size of Scatter Angle

figure 3-4, to the loss for 1000 mc, given in figure 3-3. As an example, find the basic propagation loss for a circuit 200 miles in length with an operating frequency of 750 mc. From figure 3-3, the BPL for 1000 mc is 214 db. Since the loss increases with frequency, we know that the loss at 750 mc will be less than 214 db. Figure 3-4 provides a correction factor of -4 for 750 mc. This added to 214 gives a basic propagation loss of 210 db.

Try another example with a frequency higher than 1000 mc. Find the basic propagation loss for a

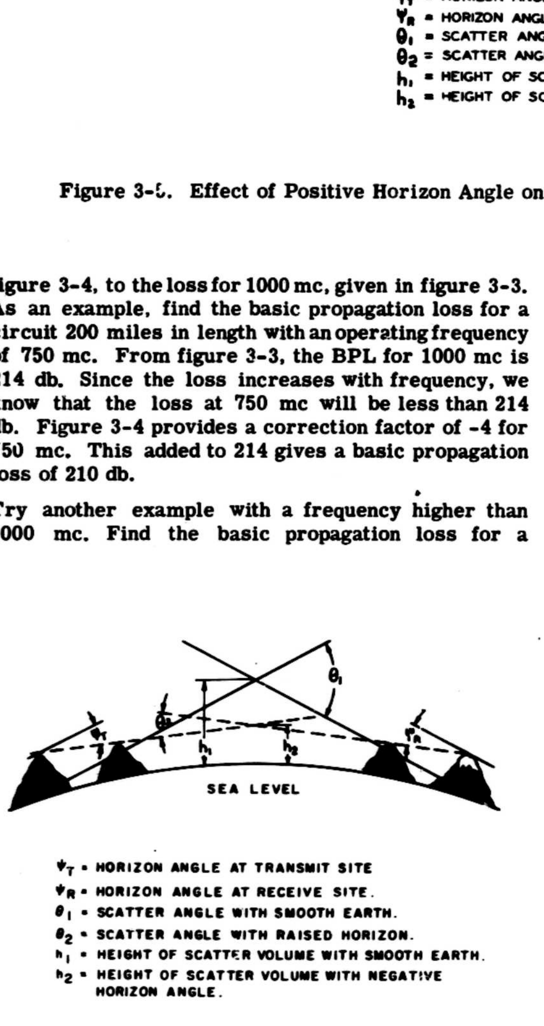


Figure 3-6. Effect of Negative Horizon Angle on Height of Scatter Volume and Size of Scatter Angle

circuit 200 miles in length with an operating frequency of 1500 mc. The BPL for 1000 mc at 200 miles is 214 db. The frequency correction factor for 1500 mc is +5 db. This added to 214 db gives a basic propagation loss of 219 db.

Now stop and review the meaning of basic propagation loss. In the first place, this is the loss over a tropospheric scatter circuit if a smooth earth and isotropic antennas are assumed. It is the sum of the free-space loss and the loss which occurs in the scattering process, and it is directly proportional to frequency and distance. Basic propagation loss is the basic loss for all tropospheric scatter circuits regardless of the circuit characteristics. For a given circuit, the size of antennas and the characteristics of the terrain between the antennas will introduce additional loss factors. These must be added to the BPL to obtain the total propagation loss.

The angle at which the antenna must be aimed to clear the horizon is called the horizon angle. If the surface of the earth is smooth, the antenna is aimed along the tangent to the earth's surface, and the horizon angle is zero. This is the condition assumed when figuring basic propagation loss. When there are obstructions, such as a hill or mountain in front of the antenna, it must be aimed at a higher angle to clear the horizon. This increases the scatter angle and raises the height of the scatter volume, as shown on figure 3-5. An increase in either of these circuit features increases the scatter loss for the circuit. This increase in loss must be added to the basic propagation loss.

If the antenna is located on a site higher than the terrain in the foreground, it can be aimed down from a zero horizon angle. This negative horizon effect is shown in figure 3-6. As can be seen on this figure, when the horizon angle is negative, the scatter angle and the height of the scatter volume are decreased. This results in a circuit loss which is less than the basic propagation loss.

If the antenna at one site is aimed at a certain angle to clear the horizon, the scatter angle and height of the scatter volume are increased a corresponding amount from zero angle conditions. An increase in horizon angle at the other site will also cause a corresponding increase in scatter angle and height. The total effect of horizon angles on circuit loss depends on the sum of the horizon angles.

Figure 3-7 is a series of curves which show the loss due to the sum of horizon angles for various circuit distances. These curves show that the additional loss due to elevated horizon angles is greater at the shorter distances. For example, a total horizon angle of 2 degrees will cause an additional loss of 30 db at 50 miles, 26 db at 100 miles, and 23 db at 150 miles.

The effect of horizon angles decreases with circuit length because the fractional increase in scatter angle for a given change in horizon angle is much greater at the shorter distances.

Figures 3-3, 3-4, and 3-7 can be used to determine propagation loss over any tropospheric scatter circuit if isotropic antennas are assumed. For example, assume that a circuit over a distance of 150 miles is to be set up; the operating frequency is to be 900 mc, and the sum of the horizon angles is 1.5°. The basic propagation loss obtained from figure 3-3 is 206 db. The correction factor obtained from figure 3-4 for 900 mc is -1.5 db. Therefore, the basic propagation loss for 900 mc at 150 miles is 204.5 db. The additional loss due to horizon angle sum of 1.5° at a distance of 150 miles is 18 db (from figure 3-7). This loss added to 204.5 db gives a total propagation loss of 222.5 db.

A propagation loss of 222.5 db, in terms of received signal level, means that the power at the receive antenna will be 222.5 db less than the power radiated

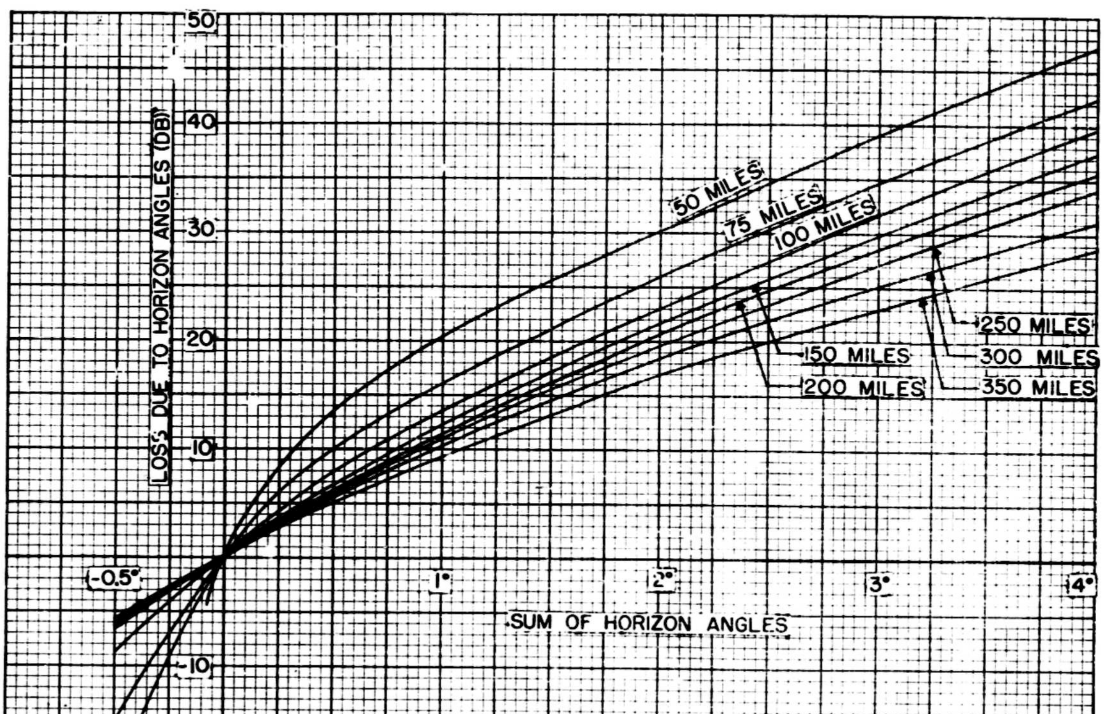


Figure 3-7. Effect of Horizon Angles on Propagation Loss

from the transmit antenna. An expression for this ratio can be derived using the basic formula for decibels:

$$db = 10 \log \frac{P_T}{P_R}$$

$$222.5 = 10 \log \frac{P_T}{P_R}$$

$$\log \frac{P_T}{P_R} = 22.25$$

$$\frac{P_T}{P_R} = 10^{22.25}$$

To see what this ratio means in terms of watts, assume that the transmit power is one watt. The received power will be only about 0.000000000000000000000001 watts.

The propagation loss encountered in a tropospheric scatter circuit results in a very weak received signal. To obtain a detectable and usable signal, the effective transmitted power must be raised. This is done by increasing the transmitter output power, and by concentrating the energy into a beam by using antennas with parabolic reflectors. The factors which determine the effective transmitted power are discussed in the following paragraphs.

3.3 EFFECTIVE TRANSMITTED POWER.

Figure 3-8 shows the effect that an increase in transmitter power and the use of high-gain antennas has upon the received level. In this example, first assume a total propagation loss of 200 db, isotropic antennas, and 1-watt transmitter output power. This results in a received signal level which is 200 db below the transmitter output power, or a received signal level of 10^{-20} watts. Now increase the transmitter output power to 1000 watts, which is a 30-db increase in power. This results in a 30-db increase in received power. The received signal level is 200 db below the 1000 watts transmitter output or 10^{-17} watts. By using an antenna with 30-db gain at each site, the received level is raised by a factor 10^6. This increases the received level to 10^{-11} watts.

Normally, all levels are expressed in dbm or dbw. Figure 3-8 shows the advantages for doing this. The numbers expressing the quantities are made easier to work with, and the various losses and gains in the circuit can be added algebraically to obtain the final received level. For example, the first received level on figure 3-8 with 1-watt transmitter output and 200 db loss is 200 db below a watt. This is the same as -200 dbw or -170 dbm. A 30-db increase in transmitter power results in received level of -170 dbw or -140 dbm, and a total antenna gain of 60 db increases the received level to -110 dbw or -80 dbm.

The total effective transmitted power is determined by adding transmitter output to the combined antenna gain, and then subtracting from this sum the line losses at each terminal. To do this, the transmitter output must be expressed in dbw or dbm. For example, a transmitter output of 1000 watts in dbw is equal to 30 dbw or 60 dbm. (The method for converting watts to dbm or dbw is given in chapter 1.)

The line loss depends on particular station installations. It is the total loss between the transmitter output and antenna at the transmit site, and between the receive antenna and receiver input at the receive site.

	RECEIVED POWER IN WATTS	DBW	DBM
RECEIVED POWER ASSUMING PROPAGATION LOSS OF 200 DB, ISOTROPIC ANTENNAS, AND 1-WATT TRANSMITTER POWER.	0.000,000,000,000,000,000,01	-200	-170
TRANSMITTER OUTPUT POWER INCREASED TO 1000 WATTS (30-DB GAIN).	→ 0.000,000,000,000,000,01	-170	-140
AN ANTENNA WITH 30-DB GAIN USED AT EACH SITE (60-DB TOTAL GAIN).	→ 0.000,000,000,01	-110	-80

Figure 3-8. Effect of Transmitter Power and Antenna Gain on Received Level

It will be recalled that in figuring the propagation losses, it was assumed that an isotropic antenna was used at each end of the circuit. Under these conditions, the energy is propagated equally in all directions. By using directional antennas which concentrate energy into a beam, the propagation loss is reduced. This has the same effect as increasing the transmitted power. Therefore, antenna gain is added to transmitter output power to obtain the effective transmitted power. The total antenna gain is the sum of the gain that each antenna has over an isotropic radiator.

Antenna gain is achieved by concentrating the transmitted power into a narrow beam with the use of parabolic reflectors. This increases the intensity of incident energy on each turbulence in the scatter volume. In general, the smaller the beam width, the greater the gain. The beam can be made smaller by either increasing the frequency or increasing the size of the reflector. Antenna gain is obtained from the formula: $G_{db} = 20 \log F + 20 \log R - 52.6$ (where F is frequency in mc, and R is the diameter of the reflector in feet).

Antenna gain can be obtained from figure 3-9 which is a nomogram drawn in accordance with the gain equation. To obtain antenna gain from this figure for any combination of frequency and reflector diameter, line a straight edge between the correct points on the frequency and reflector diameter scales. The gain is then found at the intersection of the straight edge and the gain scale. For example, assume that 15-foot reflectors are to be used with an operating frequency of 900 mc. The straight edge crosses the gain scale at 30 db. This means that the total gain of each antenna is 30 db, and the total antenna gain for the circuit is 60 db. If this frequency and antenna combination were used, the received level would be increased by a factor of 60 db (1,000,000) over that which would be obtained with isotropic radiators.

The full theoretical gain of antennas used on a tropospheric scatter circuit may not be obtained. This is because as the beam width is narrowed to achieve more gain, less of the useful scatter volume is included. The difference between the theoretical gain and the actual antenna gain realized over a circuit is called aperture-to-medium coupling loss. This loss increases with frequency and reflector diameter, since an increase in either of these factors will decrease the beam width. Aperture-to-medium coupling loss also increases with distance.

3.4 PREDICTED RECEIVED LEVEL.

If the effective transmitted power and the propagation losses for a circuit are known, the received signal level can be predicted. A typical example of the steps involved in predicting received signal level is given in figure 3-10. To determine the predicted received level, first find the total propagation loss expected over the circuit. The transmitter output power is

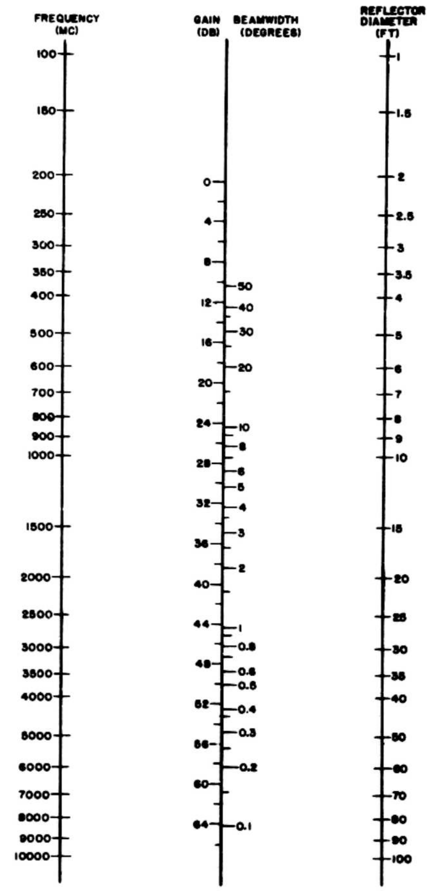


Figure 3-9. Nomogram for Estimating Antenna Gain and Beamwidth as a Function of Frequency and Antenna Diameter

expressed in dbm and the total antenna gain is added. The effective transmitted power is this sum minus the line losses which occur at each terminal. The predicted received level is then found by subtracting the effective transmitted power from the total propagation loss.

The predicted received level is that level which would be obtained for one set of conditions in the troposphere. If the tropospheric conditions did not change, the received signal level would remain constant at the predicted level. However, conditions in the troposphere are constantly changing. These changes cause the received signal level to vary about the predicted signal level. The variation in signal level is called fading. This is described in the following paragraphs.

CIRCUIT CHARACTERISTICS

DISTANCE —— 150 MILES

FREQUENCY —— 900 MC

SUM OF HORIZON ANGLES — 1.5°

ANTENNA REFLECTOR DIAMETER — 15 FT

TRANSMITTER POWER — 1000 WATTS

BASIC PROPAGATION LOSS (FROM FIGURES 3-3 AND 3-4)	-204.5 DB
LOSS DUE TO ELEVATED HORIZONS (FROM FIGURE 3-7)	- 18.0 DB
TOTAL PROPAGATION LOSS	222.5 DB
TRANSMITTER POWER	60 DBM
ANTENNA GAIN (FROM FIGURE 3-9)	60 DB
COMPOSITE LINE LOSS	-2 DB
EFFECTIVE TRANSMITTED POWER	118 DBM
PREDICTED RECEIVED LEVEL	-104.5 DBM

Figure 3-10. Method for Calculating Predicted Received Level

3.5 FADING.

If the received signal level is recorded for any given period of time, many variations in the signal level during this recording period will be noted. The signal level which is exceeded 50 percent of the recording period is called the median level for the recording period.

Assume that signal level received over a tropospheric scatter circuit for a 24-hour period is being measured. Many rapid variations in the signal level will continue throughout the observation period. Change in median level from hour-to-hour will also be noted. If the recorded signal is measured for a year, hourly, daily, and seasonal changes can be noted in the median level. The rapid variations will continue at approximately the same rate. The rapid variations in signal level are called short-term fading, and the variations noted over a longer recording period are called long-term fading. Short-term fading is caused by rapid changes in the scatter volume; long-term fading is caused by changes in the general condition of the troposphere.

Short-term fading results because the turbulences which cause the scattering are constantly changing position in the scatter volume. The received signal level is the vector sum of the signals received from the contributing turbulences. The magnitude of the received signal is dependent on the phase relationship between the individual signals and the magnitude of each signal. Because the turbulences are constantly changing, the phase and signal level from each turbulence also change. This results in a received signal level which is continuously changing.

When an infinite number of vectors arriving in random phase and amplitude relationships are combined, the magnitude of the resultant follows what is called the Rayleigh Probability Distribution. The level of the signal received over a tropospheric scatter circuit follows this distribution.

To see what this distribution of signal level means, assume that the received signal level for one hour has been measured. During this measuring period, the signal level which is exceeded say 90, 70, 50, 30, and 10 percent of the time also has been determined. These are called percentile levels. Special test equipment is necessary to determine these percentile levels.

Assume that the percentile levels were as follows:

LEVEL	PERCENTAGE OF TIME LEVEL WAS EXCEEDED
-88 dbm	90%
-83 dbm	70%
-80 dbm	50%
-77.5 dbm	30%
-74.5 dbm	10%

This means, for example, that 90 percent of the measuring period the signal level was above -88 dbm, and 10% of the measuring period the signal was below this level. Since the signal level, -80 dbm, was exceeded half of the time, this is the median level for the measuring period. Drawing a graph which shows the signal level versus the time percentages gives the percentile distribution of the received signal.

Figure 3-11 shows the signal level distribution for a tropospheric scatter circuit where the signal level follows the Rayleigh distribution. The levels relative to the median level are given on the vertical axis, and the time percentages are given on the horizontal axis. From this curve it can be seen that on a theoretical scatter circuit with Rayleigh distribution, the median level is exceeded by 5.5 db 10 percent of the time, and by 3.0 db 30 percent of the time. For 70 percent of the time, the signal level does not drop more than 3 db below the median; and for 90 percent of the time, the signal level does not drop more than 8.0 db below the median.

The difference between the 10 and 90 percentile levels is called the fade range of the signal. This is a measure of the amount of swing that the signal level takes about the median. In a scatter circuit, where the signal level follows the Rayleigh distribution, the fade range is approximately 13.5 db.

In practice, the signal will never follow the Rayleigh distribution completely. Constant level components, resulting from diffraction or reflection, will change the distribution. The vector addition of a constant level component to the scatter components results in a decreased fading range. The relative strengths of the various components in the received signal are indicated by the fade range. Thus, a fade range of 13.5 db indicates the signal is pure scatter; a fade range of 8 db indicates the presence of some constant level component.

Figure 3-12 is a recording of signal level received over a tropospheric scatter circuit. This recording shows the effects of fast fading. It has been labeled to further illustrate the meaning of percentile levels, median level, and fade range. The median level, which is that level exceeded 50 percent of the time, is -80 dbm. The 10 percentile level is 5.5 db above this, or -74.5 dbm; the 90 percentile level is 8 db below the median, or -88 dbm. The fade range, which

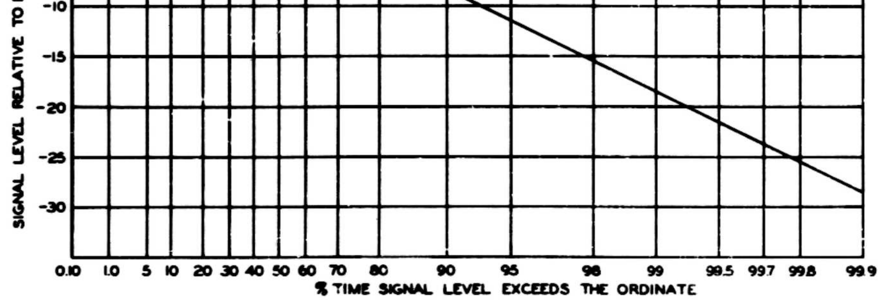


Figure 3-11. Short-Term Fading Characteristics

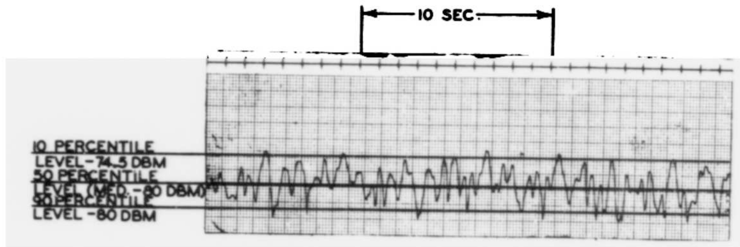


Figure 3-12. Typical Signal Level Recording Showing Short-Term Fading

is the difference between the 10 and 90 percentile levels, is 13.5 db. Only a few peaks exceed the 10 percentile level, and only a small portion of the signal drops below the 90 percentile level.

The rate at which the signal level changes is called the fade rate. This is directly dependent on transmission frequency and on wind velocities across the path. The fade rate at 1000 mc is approximately one cycle per second.

Superimposed on the short-term fading of the scatter signal is the long-term fading which is caused by daily and seasonal changes in the atmosphere. Figure 3-13 is a recording of a signal received over a tropospheric scatter circuit. This recording shows both the short-term fading and the long-term variation in median level.

The effects of long-term fading may be analyzed by plotting a distribution of hourly median levels. Figure 3-14 shows the distribution for various circuit lengths. The vertical axis gives signal level relative to the yearly median level, and the horizontal axis gives the percentage of time that a given level is exceeded.

The curves on figure 3-14 show that long-term fading range decreases with distance. This becomes apparent when the fade ranges (difference between 10 and 90 percentile levels) for the various circuit lengths are compared. The fade range for a 50-mile circuit is 30 db; the fade range for a 300-mile circuit is 11.2 db.

There are two reasons for the decrease in long-term fading range with distance. One reason is that any change in atmospheric conditions is much more apt to change propagation conditions on a short circuit than on a long circuit. Another reason is the increase in height of the scatter volume with an increase in circuit length. Most of the atmospheric changes which cause long-term fading occur in the lower part of the troposphere. As the circuit distance is increased, the scatter volume is elevated to a less changeable region.

The short-term percentile distribution shown in figure 3-11, and the long-term percentile distribution shown in figure 3-14 can be combined to obtain the composite fading characteristics of a scatter circuit. The composite fading characteristic is used to determine circuit reliability.

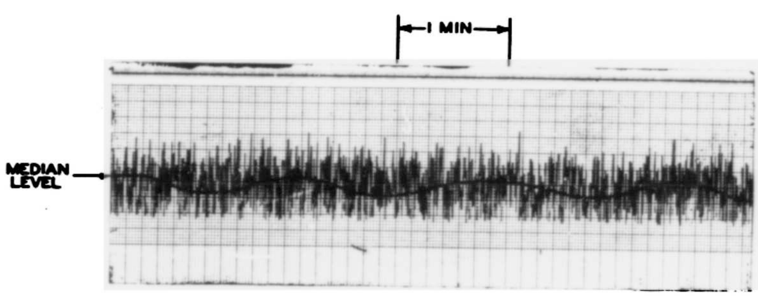


Figure 3-13. Typical Recording Showing Long-Term and Short-Term Fading

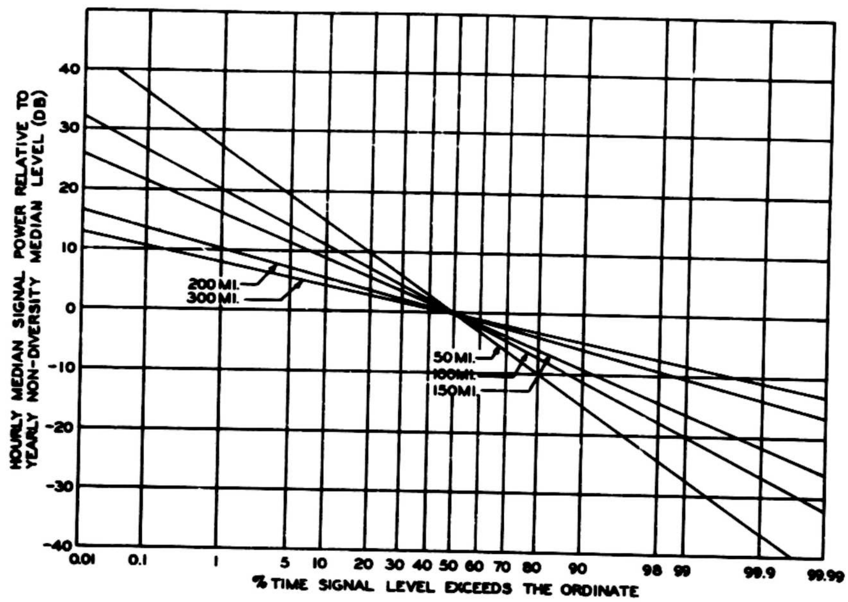


Figure 3-14. Long-Term Distribution of Hourly Medians

3.6 RELIABILITY.

The reliability of a circuit depends on the percentage of time that the received signal level exceeds a minimum usable level. This minimum usable level is determined by the noise characteristics of the receivers used in the circuit. If the receiver noise level is low, the received signal can drop to a low level and still be usable. If the receiver noise level is high, the minimum usable level is raised.

The percentage of time that the minimum usable level is exceeded is determined by the combined effects of the long-term and short-term fading on the circuit. If the combined fading effect results in a signal which is below the minimum usable level a high percentage of the time, the circuit reliability will be low.

The percentage of reliability for a circuit can be predicted if the maximum allowable fade for the circuit and the composite fading characteristic are known. The maximum allowable fade is the difference between the predicted received median level and the minimum usable level. This difference is called the fade margin. The composite fading characteristics determine what percentage of the time the fade margin is exceeded. Figure 3-15 shows the composite fading characteristic in terms of reliability versus fade margin. Since the long-term fading range varies inversely with distance, a family of curves showing reliability for various distances is provided.

Figure 3-15 shows that the same percentage of reliability requires a greater fade margin on a short circuit than on a longer circuit. For example, a reliability of 99 percent requires a fade margin of about 28 db on a 50-mile circuit, and about 20 db on a 300-mile circuit.

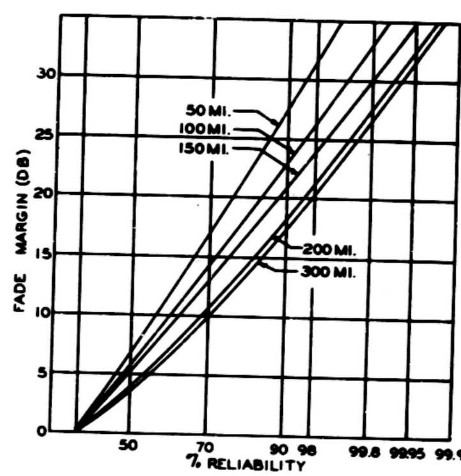


Figure 3-15. Reliability as a Function of Fade Margin for a Nondiversity System

The reliability percentage of a circuit can be increased by making the fade margin greater. This can be done by increasing the predicted received median level, or by decreasing the receiver noise power. The predicted received median is the difference between the propagation losses and the effective transmitted power. Since the propagation losses for a given circuit are fixed, the predicted receiver median level can only be increased by increasing transmitter output power or antenna gain. The receiver design features which affect the noise level are discussed in chapter 8.

The reliability of a circuit can also be increased by reducing the effects of fading. Long-term fading results from general changes in the atmosphere, and therefore the effects of this fading cannot be reduced. However, short-term fading is determined by the instantaneous arrangement of turbulences in the scatter volume. This instantaneous arrangement appears different at two separated receive points. The short-term fading characteristics of the two signals are different. When one signal is at minimum, it is possible that the other signal will be at maximum. If a receive system is used that continuously selects the strongest of the received signals, the reliability will be increased. This increases reliability by increasing the probability that the selected received signal will be above the minimum usable level.

The system which provides two or more inputs with independent short-term fading characteristics to a receiver is called diversity reception The terminology of dual-diversity, triple-diversity, and quadruple-diversity indicates the number of independent received signals. Any technique that provides two or more received signals with independent short-term fading characteristics can be used. These include employing spaced antennas (space diversity) and dual-transmission frequencies (frequency-diversity).

When space diversity is used, if the spacing between the antennas is slight, the difference in the signal levels will be very small. As the spacing is increased, the signal level difference becomes greater until a distance is reached where the signal characteristics at each antenna are not the same. The signals fade independently; when one signal is up, the other signal level may be down. The spacing required for this condition to occur is called diversity spacing. When frequency diversity is used, there will be a certain minimum frequency difference which provides diversity reception.

The greater the number of independent signals received, the greater the reliability. The reason for this is shown in figure 3-16. As the degree of diversity is increased, the fading range is decreased. For example, in a nondiversity system, the difference between the 10 and 90 percentile levels is 13.5 db. In a dual-diversity system, this difference decreases to 8.5 db; and in a quadruple-diversity system, the difference is only 5 db. At the same time that the fade range is decreased, the median level for the system is increased.

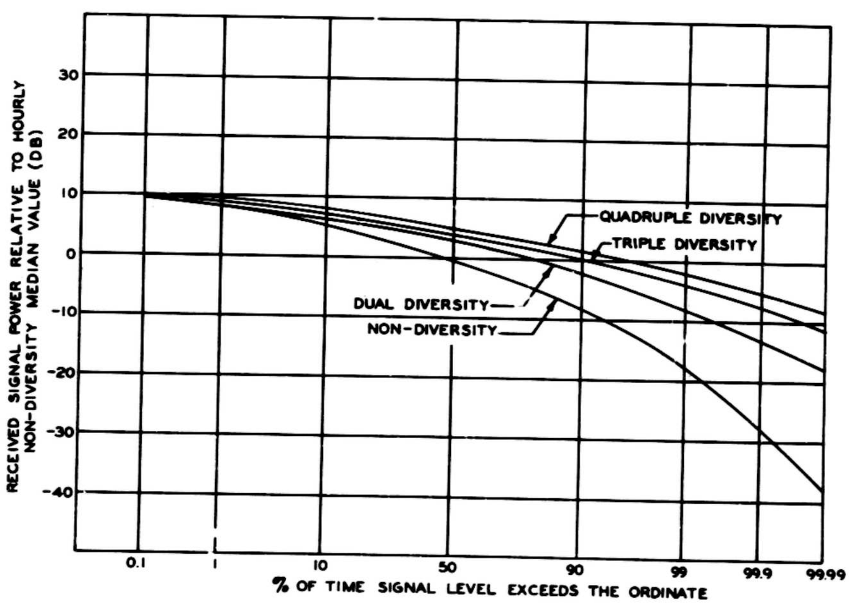


Figure 3-16. Effects of Short-Term Fading on Diversity and Nondiversity Circuits

Figure 3-17 shows reliability percentages as a function of fade margin for a dual-diversity system. A comparison of figures 3-17 and 3-15 shows the increase in reliability obtained with dual diversity. In a nondiversity system, a fade margin of 21 db provides a reliability of 98 percent on a 150-mile circuit. With dual-diversity, this same fade margin provides a reliability of approximately 99.2 percent on a 150-mile circuit. To attempt to obtain this same increase in reliability on a nondiversity circuit, the fade margin would have to be increased by about 3 db. This would require a 3-db increase in effective transmitted power, or a 3-db decrease in receiver noise power. Greater increases in reliability can be obtained with triple or quadruple diversity.

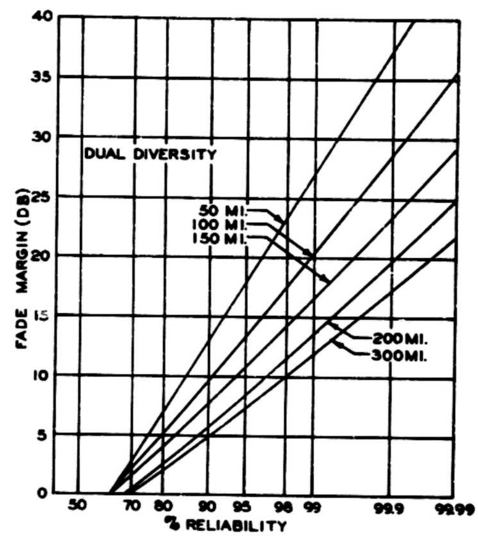


Figure 3-17. Reliability as a Function of Fade Margin for a Dual-Diversity System

3.7 TYPICAL EQUIPMENT ARRANGEMENT.

The characteristics of tropospheric scatter propagation require the use of high-power transmitters, high-gain antennas, and selective receivers. The antennas and receivers are arranged for diversity reception to provide an increase in circuit reliability.

A tropospheric scatter terminal consists of equipment necessary to transmit simultaneously in a given direction and receive from that same direction. This is called duplex operation. The antennas are set for point-to-point communication between two terminals. If communication is required over a distance which exceeds the practical one-hop distance, a system of repeater stations must be used, as shown in figure 3-18. Each repeater station consists of two terminals. Signals received from one direction by one terminal are transferred to the other terminal for retransmission in the same direction. With this system, full-duplex operation is possible from one end station to the other. The end stations in a system of stations consist of only one terminal.

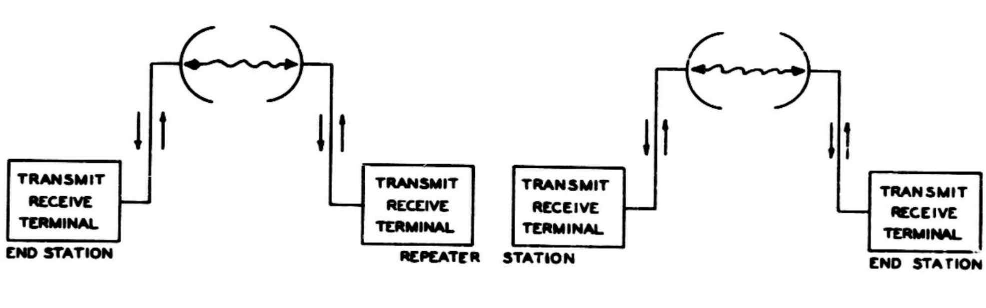


Figure 3-18. Multihop Tropospheric Scatter System

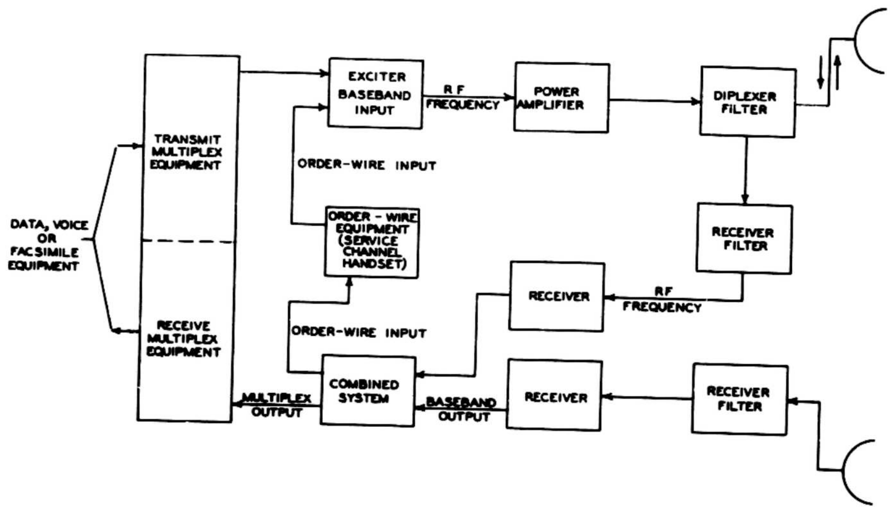


Figure 3-19. Typical Receiver-Transmit Tropospheric Scatter Terminal with Dual-Diversity Reception

Figure 3-19 is a general block diagram of a typical tropospheric scatter terminal. Two antennas, spaced for dual-diversity reception, are used. One of these antennas is used for both transmitting and receiving. Simultaneous transmission on this antenna is made possible by a diplexer filter which prevents interference between the transmit and receive sections of the terminal. For proper operation of this filter, the transmit and receive frequencies must be separated by at least 10 percent of the highest frequency.

The transmit section of a terminal converts baseband modulating frequencies into an r-f output at the terminal transmit frequency. The baseband modulating frequencies normally consist of multiplex and order-wire signals, as shown in figure 3-20.

The multiplex input is obtained from the transmit multiplexing equipment. This equipment combines information input from data, voice, or facsimile equipment into 4-kc voice channels. Each voice channel is then raised in frequency and placed in a particular location in the band of frequencies used for the multiplex input. The total bandwidth of the multiplex input depends on the number of voice channels. Figure 3-20 shows a typical baseband spectrum arrangement for a 12-channel system.

The order wire portion of the baseband contains an engineering service channel and any other information which can be carried in the spectrum allocation. Normally, the order wire is used for communicating between stations without interfering with normal circuit traffic.

The baseband modulation inputs are applied to an exciter which converts these inputs to the transmit frequency. Low-level output from the exciter is amplified to the required high-power output by a power amplifier. This output is applied through a diplexer filter to the antenna.

The receive section of a terminal receives signals at the terminal receive frequency and detects the baseband signals. Two independent antenna-receiver combinations are used for dual-diversity reception. The baseband outputs from each receiver are combined to provide an output signal which is better than the signal from either of the individual receivers. The received baseband is divided into the multiplex and order wire portions. The order wire is applied to the order wire equipment. The receive multiplex equipment separates the multiplex bands into the original voice channels and demodulates the voice channels to obtain the transmitted information signals.

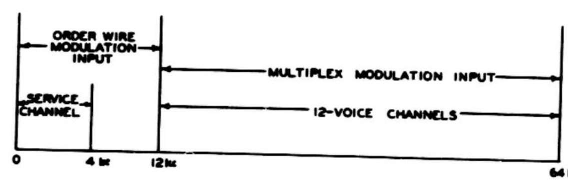


Figure 3-20. Typical Baseband Spectrum for 12-Channel System

CHAPTER 4
MODULATION TECHNIQUES

4.1 INTRODUCTION.

The purpose of any communications circuit is to transmit information of some type between two points. Very often, the information itself is not in a form capable of being transmitted over the intervening medium. The process by which the information signals are changed into a form which can be transmitted is called modulation. At the receive end of the circuit, the original information signals are regained by a process called demodulation.

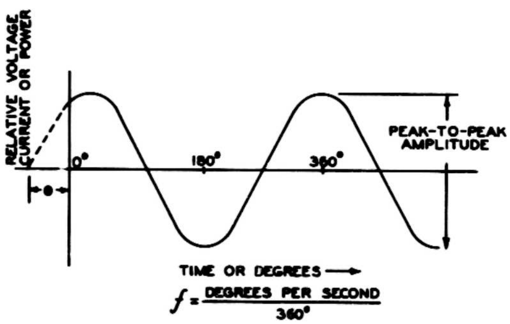


Figure 4-1. Characteristics of the Carrier

In radio circuits, modulation is accomplished by varying one of the characteristics of the r-f output in accordance with the information to be transmitted. The r-f output is called the carrier since it essentially carries the information across the intervening medium to the receive site. Figure 4-1 illustrates the three characteristics of the r-f wave which can be made to vary. Amplitude modulation (AM) results when the level of the r-f wave is made to vary in accordance with the instantaneous amplitude of the information signals. Frequency modulation (FM) results when the number of cycles in the r-f carrier vary in a given period of time. However, if the number of cycles in a given period of time is held constant and the phase (θ) of the r-f carrier is changed, then phase modulation is attained.

The performance of any communications circuit depends, in part, on the characteristics of the modulation used. Frequency modulation and single-sideband modulation, which is a form of amplitude modulation, appear to have the greatest promise for tropospheric scatter circuits. This chapter gives you the basic principles of each type of modulation.

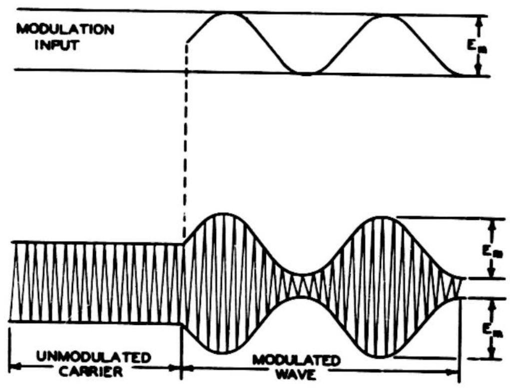


Figure 4-2. Amplitude Modulation

4.2 SINGLE SIDEBAND MODULATION.

4.2.1 GENERAL DESCRIPTION OF AMPLITUDE MODULATION.

Amplitude modulation is the process by which the magnitude of a carrier wave is varied according to the instantaneous amplitude of the modulation input. Figure 4-2 is a representation of an amplitude-modulated wave. From this drawing, you can see that the amplitude of an amplitude-modulated wave is high during the positive peaks of the signal voltage and near zero during the negative peaks. The rate of amplitude change is determined by the frequency of the modulation input. The extent of the amplitude change is determined by the amplitude of the modulation input.

4.2.2 SPECTRUM OF AMPLITUDE-MODULATED WAVE.

When a carrier is amplitude modulated with a single modulating frequency, the resultant wave is composed of the following frequencies:

a. The carrier frequency.
b. An upper-side frequency which is equal to the carrier plus the modulating frequency.
c. A lower-side frequency which is equal to the carrier minus the modulating frequency.

Figure 4-3 shows the frequencies which appear in the modulator when a carrier frequency of 500 kc is modulated by a 1-kc tone. These frequencies

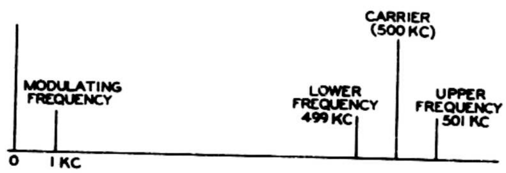


Figure 4-3. Spectrum of Amplitude-Modulated Wave with Single-Tone Modulation

are the 500-kc carrier, the 1-kc modulating frequency, the 501-kc upper-side frequency, and the 499-kc lower-side frequency. The r-f frequencies are within the range of the transmitter circuit, and can be acted upon by tuned circuits. Therefore, they are the ones to be transmitted.

The transmitted signal does not contain the modulating frequency. Because of this, simple low-pass filters which are commonly used to separate frequencies are not capable of separating the signal from the carrier. Instead, receivers use a detector stage which generates or reproduces the signal frequency when the carrier and side frequencies are applied to it.

Up to this point, we have considered only a single-frequency modulating input. However, in practice, the modulating input is a complex wave composed of a number of frequencies. This causes the upper and lower frequencies to become bands of frequencies rather than single frequencies.

As an example, consider what happens when a voice signal is used to modulate a 15-mc r-f frequency. An intelligible voice signal contains audio frequencies over the range of 300- to 3000-cycles per second (cps). When this audio signal is mixed with the 15-mc r-f frequency, upper and lower sidebands are formed, as shown on figure 4-4. The frequency range of the upper sideband corresponds to the sum of the 15-mc carrier frequency and the two extremes of the audio signal. The resultant sum frequencies cover the range of 15,000,300 to 15,003,000 cps.

The upper frequency of the lower sideband is 15,000,000 - 300 or 14,999,700 cps. The lower frequency of the lower sideband is 15,000,000 - 3000 or 14,997,000 cps. Thus, the original voice band is inverted. That is, the lower frequency of the audio input has become the upper frequency of the lower sideband, and vice versa. Note that the carrier frequency is not included in either sideband.

The modulated wave covers a total band of frequencies equal to the difference in frequency between the extremes of the two sidebands. In the example shown in figure 4-4, the total bandwidth is 6 kc. The power relationship between the carrier and the sidebands depends on the extent that the carrier is modulated. For 100% modulation, half of the power is represented in the carrier. Since the carrier does not contain any information, the power in the carrier is wasted, and only the power in the sidebands is usable.

Each sideband contains all the information necessary for communication. Therefore, power and frequency space can be saved by transmitting only one sideband. This is called single sideband transmission (SSB).

4.2.3 METHODS OF SIDEBAND COMMUNICATION.

Several methods of sideband communications are possible. The term SSB generally refers to a method which is more accurately termed single-sideband suppressed carrier. In this method, only one sideband is transmitted, and the carrier is suppressed to the point of nonexistence. To demodulate the single-sideband signal requires conversion of the signal with a locally generated signal close to the proper frequency, but with no phase relationship required. In the single-sideband pilot-carrier system, only one sideband is transmitted, but a low-level carrier of sufficient amplitude for reception is also transmitted. To demodulate this signal, the pilot carrier is separated from the sideband in the receiver, then amplified and used as the conversion frequency to

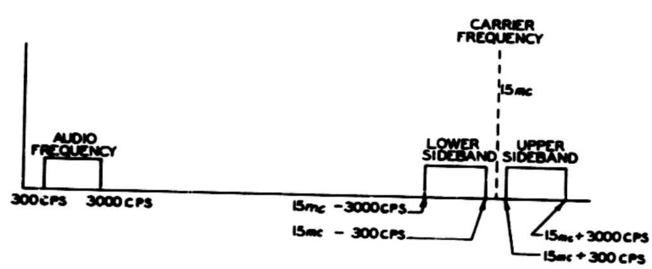


Figure 4-4. Location of Sidebands with Voice Modulation

demodulate the sideband signal. In another method, the pilot carrier is used for automatic frequency control of the receiver. In the double-sideband (DSB) system, both the upper and lower sidebands of the signal are transmitted with the carrier suppressed to the point of nonexistence. To demodulate the double sideband requires insertion of a locally generated carrier of both the proper frequency and the proper phase. This system depends upon an automatic frequency and phase control, derived from double-sideband signal, for control of the locally generated carrier. In the single-sideband controlled-carrier system, only one sideband is transmitted, but a carrier which varies inversely with the signal level is also transmitted. This allows an appreciable average carrier level for automatic frequency control without reducing the sideband power below the full transmitter rating.

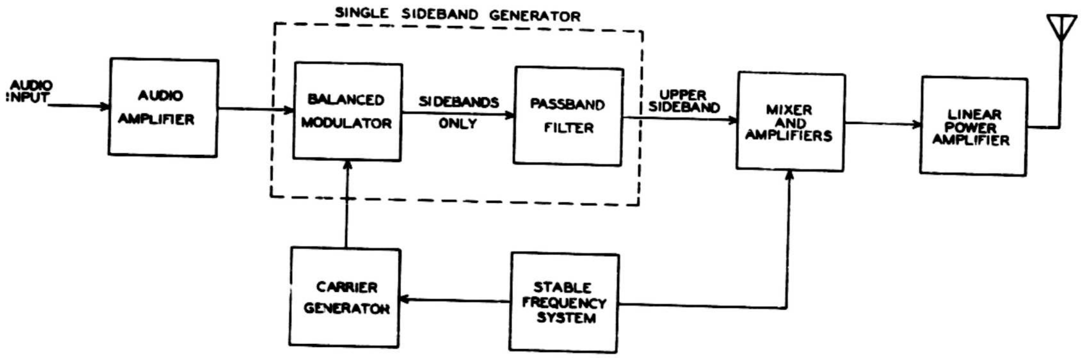


Figure 4-5. Single-Sideband Transmitting System

4.2.4 SINGLE-SIDEBAND TRANSMITTING SYSTEM.

4.2.4.1 GENERAL CHARACTERISTICS.

Figure 4-5 shows the basic functional units of an SSB transmitter. An SSB transmitter generates a radio-frequency sideband from an audio-input signal, translates this r-f sideband to the final-output frequency, and amplifies the signal to the required-output level. Five main-design considerations for a single-sideband transmitter are as follows:

a. Suppression of the carrier.
b. Elimination of one sideband.
c. Linear operation of the entire transmitter.
d. Elimination of unwanted frequency radiations.
e. High-frequency stability.

4.2.4.2 AUDIO INPUT CIRCUITS.

The audio amplifier, shown in figure 4-5, is of conventional design. Audio filtering is not required because highly selective filtering takes place in the SSB generator. The input signal may be any desired intelligence signal and may cover all or any part of the frequency range between 100 and 6000 cps. The upper limit of the audio input is determined by the channel bandwidth and the upper cutoff frequency of filtering in the SSB generator. The lower limit of the input audio signal is determined by the lower cutoff frequency of filtering in the SSB generator.

4.2.4.3 SINGLE-SIDEBAND GENERATOR.

The actual generation of the single-sideband signal is perhaps the most important part of an SSB transmitter. An SSB generator generates the desired sideband and suppresses the carrier and other sideband.

An SSB generator includes a balanced-modulator circuit for carrier suppression. In a balanced-modulator circuit, the carrier is introduced in such a way that current at the carrier frequency cancels out in the output circuit. Figure 4-6 is a simplified schematic of a balanced modulator. The audio input is applied in push-pull, and the r-f carrier is applied in parallel to two tubes. The two outputs are combined in the center-tapped primary of an output transformer. Since the carrier is applied in phase to the two tubes, the carrier current from one tube is balanced or canceled out by the carrier current from the other tube. Therefore, with no modulation input, there will be no output. When push-pull modulation input is applied, the circuit is balanced. The modulation is not balanced for the sidebands, and they appear in the output.

The amount of carrier suppression obtained is dependent upon the matching of the two tubes and their

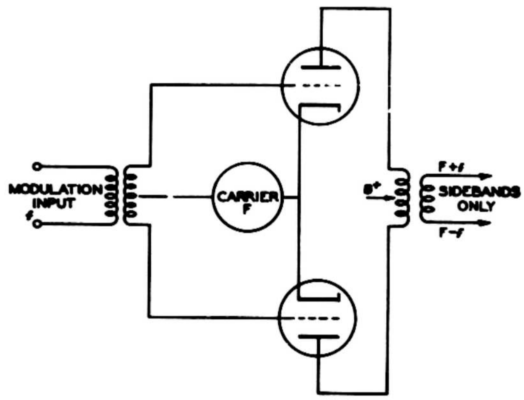


Figure 4-6. Simplified Schematic Diagram of Balanced Modulator

associated circuits. Normally, two tubes of the same characteristics can be adjusted to give at least 30 db of carrier suppression without further filtering. Balance is difficult to maintain since tube characteristics change with age and supply voltage variations. Since in suppressed-carrier single-sideband transmission it is desirable to suppress the carrier at least 40 db, added carrier suppression is provided by filters following the balanced modulator.

It is the property of all amplitude modulators that the output consists of a pair of sidebands located on each side of the carrier frequency. Since the objective is to transmit only a single sideband, a means must be found to select the desired sideband and suppress the undesired sideband. This may be accomplished through the use of either a filter or phase-shift system. The two systems are shown in figure 4-7.

The filter system (figure 4-7A) uses a band-pass filter which has sufficient selectivity to pass one

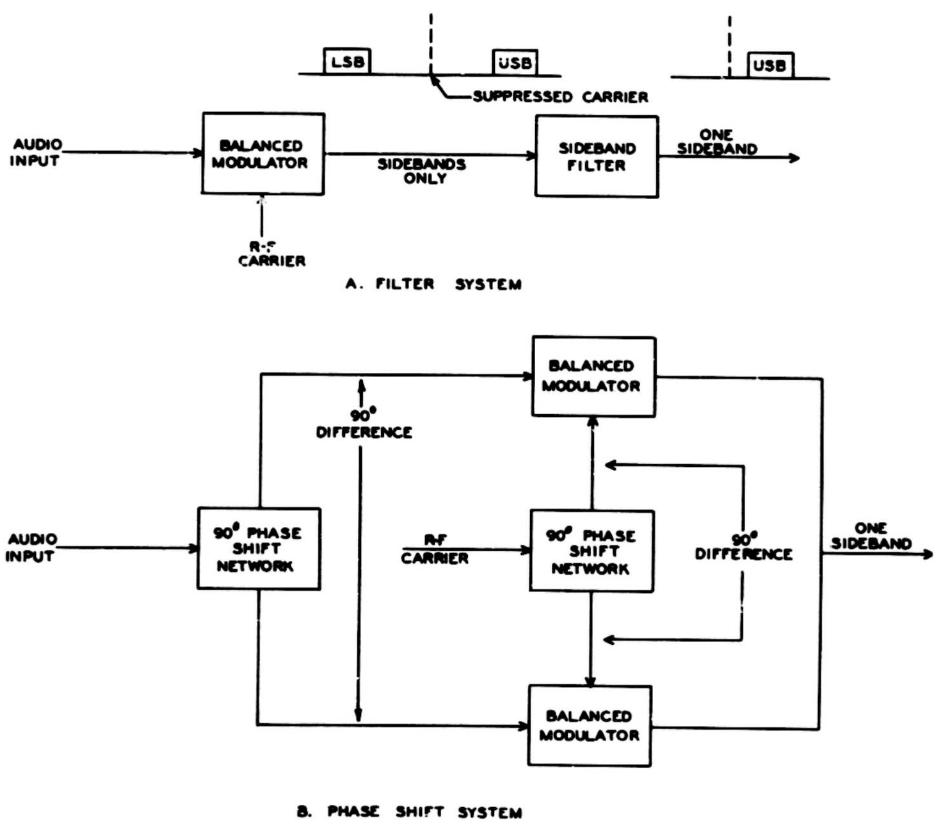


Figure 4-7. Two Systems for Generating Single-Sideband Signals

sideband and reject the other. This filter is required to do several things:

a. It should pass the desired sideband.

b. It should limit the bandwidth of the desired sideband to that required for an intelligible communication circuit.

c. It should provide adequate suppression to the undesired sideband.

d. It should provide some attenuation to the carrier frequency.

Filters having the necessary characteristics are normally constructed for relatively low frequencies, (below 500 kc). However, mechanical or metal plate filters have been built to operate up to 600 kilocycles, and crystal filters have been made to work at frequencies as high as 5 megacycles. A frequency conversion system must be used to translate the SSB output to the desired operating frequency.

Removing the unwanted sideband through the use of a selective filter has the advantage of simplicity and good stability. The stability of sideband suppression is determined by the stability of the elements used in constructing the sideband filter.

In the phase-shift method (figure 4-7B), two balanced modulators are used. As shown in the diagram, the audio signal is split into two components that are identical except for a phase difference of 90 degrees. The r-f carrier, which may be at the output frequency, is also split into two separate components having a 90-degree phase difference. One r-f and one audio component are combined in each of the two balanced modulators. The carrier is suppressed in each modulator. The relative phases of the sidebands are such that one sideband is balanced out, and the other is reinforced in the combined output.

Since no selective filter is required in the phase-shift system, it is possible to generate the single sideband at the operating frequency with no frequency conversion being required. The amount of suppression is dependent on the accuracy with which the undesired sideband components are canceled. To obtain complete cancellation, it is necessary to maintain accurately the phase shifts and amplitudes of the signals in each balanced modulator. These requirements place a very stiff specification on the phase shift and amplitude control properties of the circuit. The circuit is also somewhat more complex since two modulators are required.

4.2.4.4 TRANSLATION OF THE OPERATING FREQUENCY.

The single-sideband signal is translated to the operating frequency by the use of one or more frequency changers (mixers). These frequency changers perform their function through the modulation process which is identical to that used to generate the sideband signal. The sideband signal modulates a high-frequency carrier which has a frequency such that the upper or lower sideband is on the desired operating frequency. As a result of this modulation process, the sideband signal will be shifted to a new frequency that is either the sum of the carrier and sideband frequency or the difference between the carrier and sideband frequency. One of these frequencies is selected by tuned circuits. If the lower sideband of the translation-modulation process is selected, an inversion of the sideband signal occurs. That is, an upper sideband will be converted into a lower sideband.

Another important consideration in the frequency-translation process is the frequency accuracy and stability of the carrier frequency used in the modulation process. Any error in the carrier frequency is passed on to the sideband signal in exact proportion.

The frequency conversions required to produce the desired operating frequency also produce other higher order frequencies. However, the undesired difference frequency or the undesired sum frequency, along with the higher order frequencies, are attenuated by interstage tuned circuits. The problem of attenuating undesired frequencies is also simplified by using two stages of frequency translation.

4.2.4.5 LINEAR AMPLIFIERS.

The single-sideband signal is amplified to the desired output level by use of linear amplifiers. Linear amplifiers are required for SSB because it is essential that the plate output r-f signal be a replica of the grid input signal. Any nonlinear operation results in mixing between the frequencies of the input signal. This mixing produces distortion in the output. Distortion in the linear amplifiers is kept low by choice of tubes, operating conditions, and use of r-f feedback circuits.

4.2.4.6 STABLE FREQUENCY SYSTEM.

The frequency accuracy requirements for single-sideband communications are very precise when compared with most other communications systems. This is because the frequency relationship between the suppressed carrier and the transmitted sideband must be maintained both in the transmitting and receiving system. A very stable oscillator is used to generate a frequency standard for the original carrier and the insertion frequencies in the translator stages. The standard frequency is normally obtained from a crystal oscillator with the crystal housed in an oven. Stable thermal control of the oven is necessary.

4.2.5 SINGLE-SIDEBAND RECEIVING SYSTEM.

The operation of a single-sideband receiver is, in a limited sense, the reverse of the process carried out in a single-sideband exciter. The received single-sideband r-f signal is amplified, translated to a low i-f signal, and then converted to useful audio-frequency signal. The basic functional units of an SSB receiver are shown in figure 4-8.

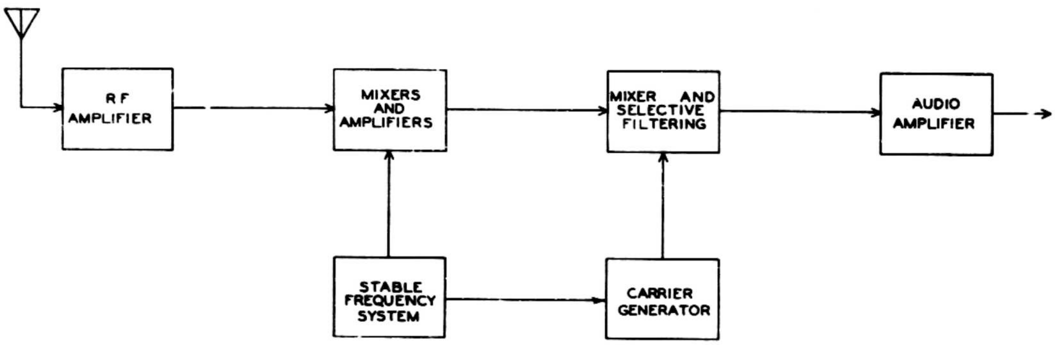


Figure 4-8. Single-Sideband Receiving System

The absence of a carrier in the received SSB signal accounts for the principal difference between SSB and AM receivers. In order to recover the intelligence from the SSB signal, it is necessary first to restore the carrier. This local carrier must have the same relationship with the sideband components as the initial carrier used in the transmitter. To achieve this, the oscillator which produces the reinserted carrier must have extremely good frequency accuracy and stability. The total frequency error of the system must be less than 100 cycles per second, or the intelligibility of the received signal will be degraded.

4.3 FREQUENCY MODULATION.

4.3.1 DIRECT AND INDIRECT FM.

Figure 4-1 shows that, beside the amplitude, the frequency and phase of the carrier can also be varied in accordance with the modulating signal. Phase and frequency modulation are sometimes classified together as angle modulation. This is because a change in the angular velocity of the carrier results from either PM or FM.

If the frequency of the carrier is changed, the instantaneous phase is also changed. Likewise, if the instantaneous phase is changed, the frequency is indirectly changed. If the phase is varied directly proportional to the modulation input, the result is called phase modulation. If the frequency is varied directly, the result is called frequency modulation. The characteristics of each of these types of modulation are described in the following paragraphs.

4.3.2 REPRESENTATION OF FM.

Figure 4-9 shows what happens when an audio input is used to frequency modulate an r-f wave. As can be seen in this figure, when there is no modulating signal, the frequency of the r-f wave remains unchanged. This frequency with no modulation is called the center frequency. As the modulating signal increases in amplitude, the frequency of the r-f increases. At instant 2 on figure 4-9, the amplitude of the modulating voltage and the frequency of the r-f wave are both maximum. As the modulating input decreases in amplitude, the frequency of the r-f wave decreases. At instant 3, the modulating voltage is zero, and the r-f is back at center frequency. As the modulating voltage goes negative, the r-f decreases below the center frequency and reaches a minimum value at instant 4. As the modulating voltage increases again, the frequency of the r-f wave increases; and at instant 5, it is again back at the center frequency. Note that only the frequency of the r-f wave is changed. The amplitude of the r-f wave is unchanged.

Thus, from figure 4-9 and the above discussion, we can see that in frequency modulation, the amount of change in r-f frequency at any instant is proportional to the instantaneous amplitude of the modulating voltage. The amount of frequency change from the center frequency is called frequency deviation.

To illustrate the change in frequency with modulating voltage, use a simple equation and some actual figures. The equation is:

$$F_i = F_c + V \times D$$

where:

F_i = instantaneous r-f frequency

F_c = center frequency

V = instantaneous value of modulating voltage

D = frequency change (deviation) per volt of modulating input.

Figure 4-10 illustrates the use of this equation with some actual figures. In this example, the center frequency, F_c, is 10 mc; and the deviation, D, is set to be 0.01 mc per volt of modulation input. The peak value of modulation input is 5 volts; therefore, the peak deviation is 5 x 0.01 mc, or 0.05 mc.

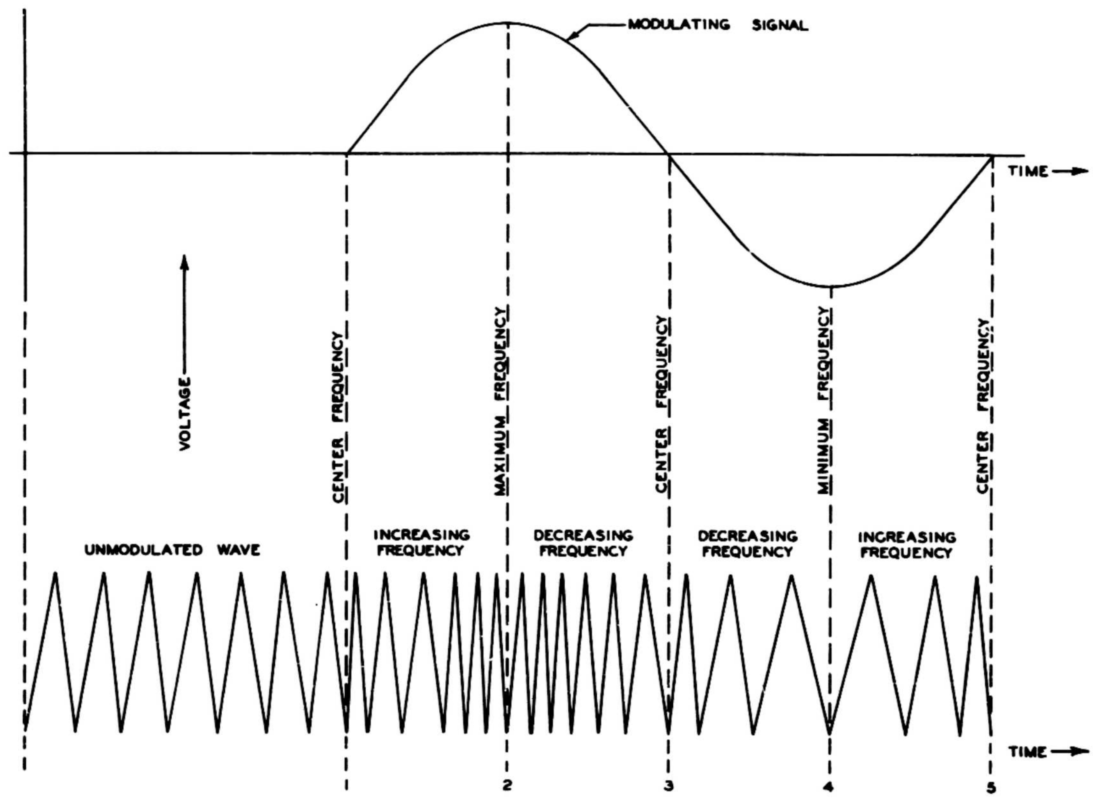


Figure 4-9. Representation of FM

The maximum instantaneous frequency, using the above equation and the figures on figure 4-10, is:

F_i = 10 + 5 x 0.01 = 10.05 mc.

The minimum instantaneous frequency occurs when the modulation input is at the lowest negative value:

F_i = 10 + (-5 x 0.01) = 10 - 0.05 = 9.95 mc.

Figure 4-10 shows the values which occur between the maximum and minimum values. The change in the frequency of the r-f wave completes one cycle of change (center frequency - maximum - center frequency - minimum - center frequency) with each cycle of modulation input. Therefore, the rate at which the frequency changes varies directly with frequency of the modulating input.

Figure 4-11 summaries how frequency deviation in an FM system is affected by modulating voltage. Figure 4-11A shows that the deviation varies directly with the modulating voltage. The amount of change in deviation with a change in modulating voltage determines the slope of the line in 4-11A.

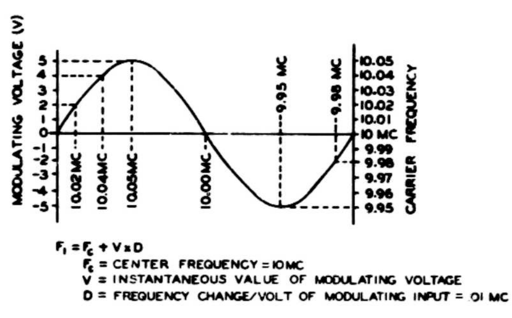


Figure 4-10. Relationship Between Change in R-F Frequency and Change in Modulating Voltage for Frequency Modulation

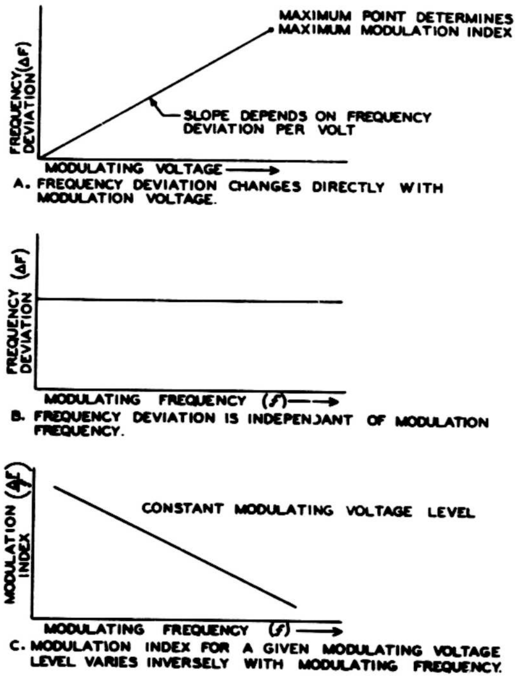


Figure 4-11. Effect of Modulating Voltage and Frequency in an FM System

This is represented by D in the equation $F_t = F_c + V \times D$. Normally, in an FM system, the deviation is limited to some value by predetermining D and limiting the amplitude of the modulation voltage.

Figure 4-11B shows that the amount of frequency deviation is not affected by the frequency of the modulating input. In a properly designed FM system, the deviation will be constant for a given modulating voltage for all frequencies in the modulating range.

The ratio of the frequency deviation (ΔF) to the rate at which the frequency is changing is called the modulation index. In an FM system, the frequency changes at a rate equal to the modulation frequency. Therefore, the modulation index (M) can be expressed as follows:

$$M = \frac{\text{Frequency Deviation}}{\text{Modulating Frequency}} = \frac{\Delta F}{f}$$

where:

ΔF is the frequency deviation, and f is the frequency of the modulation input. Since the deviation does not change with modulating frequency, the modulation index decreases as the modulating frequency increases. This relationship is shown in figure 4-11C.

Use some actual figures to see how the modulation index is determined. Assume that the deviation per volt is 0.01 mc and that modulation is occurring with a 1-kc tone with a constant peak amplitude of 2 volts. The peak frequency deviation is then 2 x 0.01 mc, or 0.02 mc, or 20 kc. The modulation index is equal to $\frac{\Delta F}{F} = \frac{20}{1} = 20$. Use a 2-kc tone with the same amplitude. The modulation index is then equal to $\frac{20}{2} = 10$. Therefore, by increasing the modulating frequency, the modulation index is decreased accordingly.

All of the discussion so far has shown FM characteristics with a single-tone modulating input. When a complex modulating signal is used, the relationship between the deviation and modulation index becomes very complex. Therefore, the modulation index is usually determined by dividing the peak frequency deviation by the highest modulating frequency. In paragraph 4.3.5, the importance of the modulation index in determining the FM spectrum and bandwidth requirements can be seen.

4.3.3 REPRESENTATION OF PM.

Phase modulation is the process of varying the phase of a carrier according to the instantaneous amplitude of the modulating input. Phase increase or decrease must always be stated in some relation to some reference. In modulation, when the phase of a signal is said to lead or lag, it is always understood that this lead or lag is with respect to the unmodulated position of the carrier.

A phase-modulated wave and its modulating audio signal are shown in figure 4-12. The unmodulated carrier, shown in the lower portion of the figure as a solid line, is used as the reference. The phase of the modulated wave, shown by the dotted line, is shifted in reference to the unmodulated carrier.

To see the effects of phase modulation on the carrier wave, trace the variation in phase with instantaneous modulating voltage, as shown on figure 4-12. As the modulating voltage increases in the positive direction, the modulated wave advances in phase by the amount indicated as a. As the audio voltage increases more, the phase angle increases to the amount indicated as b. As the instantaneous modulating voltage decreases, the phase shift follows and decreases a corresponding amount. When the modulating voltage goes negative, the modulated wave is retarded in phase. With maximum negative modulating voltage, the modulated wave is retarded by the amount shown as c. For accurate representation of the modulating voltage, the phase shift of the modulated voltage has to be proportional to the instantaneous modulating voltage.

One interesting feature of phase modulation can be noted by carefully observing figure 4-12. As the phase of the carrier is changed, the time interval for

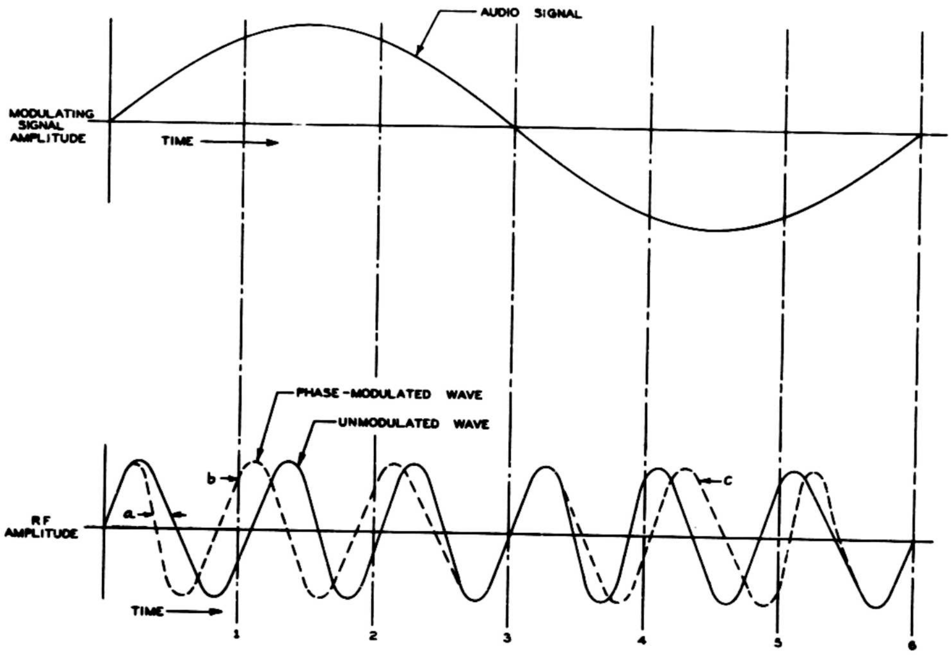


Figure 4-12. Representation of Phase Modulation

each cycle is varied. This is equivalent to changing the frequency, since frequency of a wave is the reciprocal of the time required for one cycle. Therefore, if we retard the phase, the time required for the cycle is increased, and the frequency is decreased. If we advance the phase, the time required for the cycle is decreased, thus increasing the frequency. From this example, it can be seen that phase modulation, like frequency modulation, also results in a change in frequency of the carrier.

There is, however, one notable difference between FM and PM. With frequency modulation, the frequency deviation is directly proportional to the amplitude of the modulating voltage. With phase modulation, the frequency deviation is directly proportional to the rate at which the modulating voltage is changing, as well as the amplitude of the modulating voltage. If the modulating voltage is held at some constant d-c level, the phase difference between the modulated and unmodulated carrier becomes constant, and therefore there is no frequency change.

Figure 4-13 illustrates this difference between FM and PM. Suppose that we are using a modulating voltage with the semisquare wave shape shown. If this wave is used to frequency modulate the carrier, the frequency increases with modulating voltage. The frequency is maintained at the high value during the time that the modulating voltage is constant at the maximum positive level. As the modulating voltage decreases, the frequency decreases and assumes a constant low value during the time that the modulating voltage is constant at the maximum negative level. In phase modulation, however, the frequency changes only during the time that the modulating voltage is changing. During the periods that the modulating voltage is constant, the phase relationship is constant, and there is no frequency change.

This relationship between change in transmitted frequency and rate of change in the modulating voltage is illustrated further in figure 4-14 which is a vector representation of the process. In this figure, the vector A represents the carrier. The angular velocity at which it is rotating determines the transmitted frequency. This relationship between angular velocity and frequency is given in the following expression:

$$\text{Frequency} = \frac{\text{Angular velocity (degree per second)}}{\text{No. of degrees per cycle (360°)}}$$

This shows a direct relationship between the speed of rotation of the vector and the instantaneous frequency. The dashed vectors show the position of the carrier vector for corresponding modulating voltage. The phase deviation per volt is set by system design and is constant for all frequencies in the modulating frequency range. Figure 4-14 shows that the phase deviation follows the amplitude of the modulation input and that there is one complete swing of phase deviation for each cycle of the modulation input.

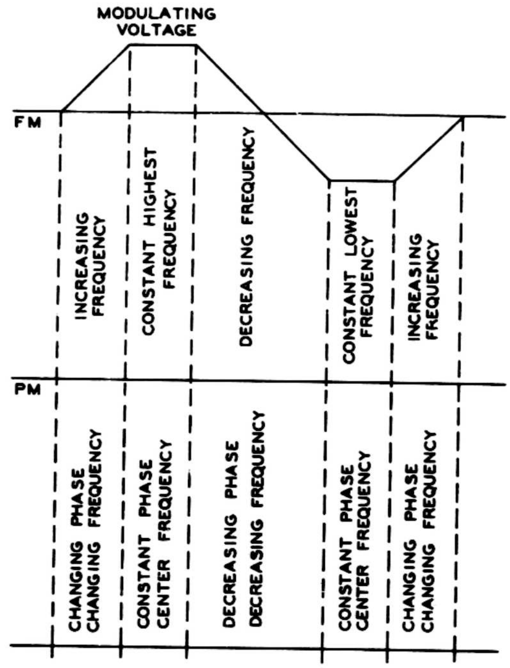


Figure 4-13. Comparison of FM and PM

Figure 4-14 also illustrates that the rate at which the phase changes, and therefore the amount of equivalent instantaneous frequency deviation, varies directly with the modulating frequency. For example,

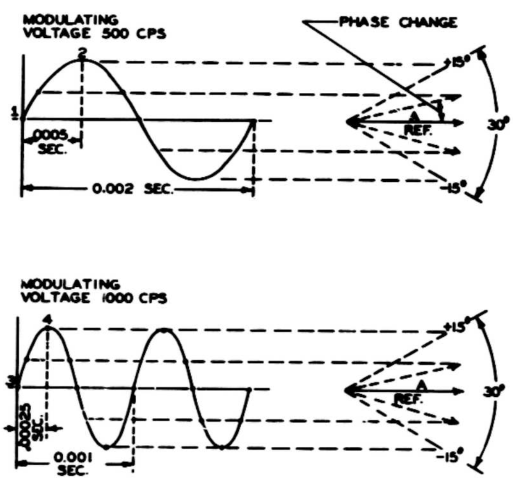


Figure 4-14. Vector Representation of Phase Modulation

assume that the modulating frequency is 500 cps and that the peak phase deviation, resulting when the modulation voltage is maximum, is 15°. This means that the vector must move 15° in the time required for the voltage to change from zero to maximum. Therefore, the total phase deviation (p_d) for each cycle of modulating voltage is 60° (0° to +15°, +15° to 0°, 0° to -15°, and -15° to 0°). The vector A in figure 4-14 must travel 60° in the time required for one cycle of the modulating input. The period or time required for one cycle, when the frequency is 500 cps, is $\frac{1}{500}$, or 0.002 seconds. The change in angular velocity of the vector is equal to the total change in degrees divided by the time it takes to make this change; or, in the example, this is $\frac{60°}{0.002}$, or 30,000° per second. Since frequency is equal to $\frac{\text{angular velocity}}{360°}$, the equivalent frequency deviation is $\frac{30{,}000}{360}$, or 83.33 cps.

The lower part of figure 4-14 shows what happens to the change in angular velocity and equivalent frequency deviation if the amplitude of the modulating signal is kept the same, but the frequency is doubled to 1000 cps. In this case, the carrier phase is changed the same amount (60°), but it goes through this phase change twice as fast. Therefore, the change in angular velocity and the equivalent frequency deviation is doubled. The period required for a 1000-cps signal is $\frac{1}{1000}$, or 0.001 seconds. The change in angular velocity is $\frac{60}{.001}$, or 60,000° per second. The equivalent frequency change is $\frac{60{,}000}{360}$, or 166.66 cps. This is twice the frequency deviation resulting when the modulating frequency is 500 cps.

The above examples show that the amount of equivalent frequency deviation in a phase-modulation system depends directly on the total amount of phase deviation and on how rapidly the phase is changed. The total amount of phase deviation is proportional to the instantaneous amplitude of the modulating signal. The rapidity of phase shift is directly proportional to the modulating frequency. Consequently, the equivalent frequency deviation in phase modulation is directly proportional to both the amplitude and frequency of the modulating signal.

This relationship between deviation and the amplitude and frequency of the modulating voltage is illustrated in figure 4-15. Figure 4-15A shows that the phase deviation varies directly with the modulating voltage. Normally, in a PM system, the phase deviation is limited to some maximum value by designing for a given deviation per volt of modulating input and limiting the amplitude of the modulating input. As long as the phase deviation is less than the maximum design value, the equipment functions properly.

Figure 4-15B shows that the frequency deviation also varies directly with the modulating frequency. The difference between FM and PM can be noted by comparing this figure with 4-11B.

The effect of modulating frequency on modulation index is shown in 4-15C. This figure shows that for a constant voltage level, a change in modulating frequency has no effect on modulation index. This can be easily explained if it is remembered that the modulation index is equal to $\frac{\Delta F}{f}$. An increase in modulating frequency f causes a corresponding increase in ΔF. Therefore, the ratio remains constant.

4.3.4 USING PHASE MODULATORS FOR INDIRECTLY PRODUCING FM.

There are several disadvantages to using phase modulation in its pure form in a communication system. One thing to remember is that when one characteristic of the carrier is changed in accordance with the intelligence signal, in order to regain this signal at the receiving end, it is necessary to detect the change in the carrier. Phase modulation is difficult to detect. Also, phase changes in the carrier, caused by propagation conditions, will introduce errors in the detected signal. Another disadvantage in PM is connected with the fact that equivalent frequency deviation in a PM system varies directly with modulating frequency. Therefore, large deviations are required for high-capacity systems where a large range of modulating frequencies are used. Paragraph 4.3.5 shows that large deviations require wide bandwidth systems with accompanying design problems.

Although phase modulation in its pure form is very seldom used in communications systems, it is used for indirectly producing frequency modulation. A direct FM system requires an oscillator which can be shifted through the required frequency deviation range. This involves an oscillator system which is relatively unstable. However, if a phase modulator is used to produce FM indirectly, a stable crystal oscillator can be used with a phase-shifting system.

A phase modulator can be made to have the same characteristics as a frequency modulator. To do this, it is necessary to make the equivalent frequency deviation independent of the modulating frequency and dependent only on change in modulating voltage. The discussion in paragraph 4.3.3 showed that the frequency deviation with phase modulation depends on two factors:

a. Amount of phase deviation - determined by the amplitude of the modulating signal.

b. Rate of phase change - determined by the frequency of the modulating signal.

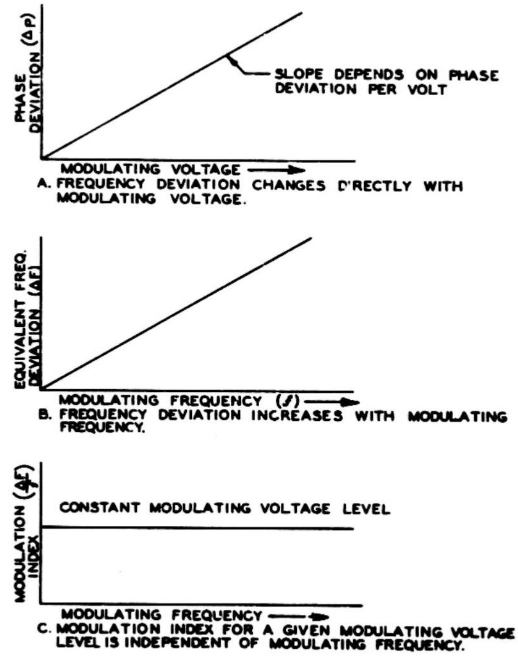


Figure 4-15. Effect of Modulating Voltage and Frequency in a PM System

Therefore, the frequency deviation can be made independent of modulating frequency if the modulating voltage is reduced in proportion to frequency. For example, assume that the modulating signal is 1000 cps with a peak amplitude of 2 volts. With this combination of voltage level and frequency, a certain frequency deviation will be obtained. Now, assume that the modulating frequency is increased to 2000

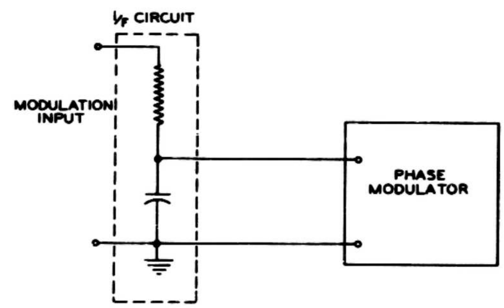


Figure 4-16. Correction Network Used with a Phase Modulator for Producing Frequency Modulation

cps, and the peak amplitude remains at 2 volts. With pure phase modulation, the frequency deviation will be doubled. However, if the peak voltage is reduced to one-half of the original value when the frequency is doubled, the frequency deviation will remain constant.

The circuit which is used to cause this reduction in level in proportion to frequency is called a 1/f network. This consists of a simple resistor-capacitor circuit shown in figure 4-16. The circuit acts as a voltage divider. The input is applied across the resistor and capacitor in series, and the output is taken from across the capacitor and applied to the phase modulator. The voltage drop across the capacitor is inversely proportional to the frequency since it is determined by the value of the capacitive reactance X_c. The reactance is determined with the following equation:

$$X_c = \frac{1}{2\pi fc}$$

where: f is the modulating frequency.

Thus, since the reactance of the capacitor varies inversely with the frequency f, the magnitude of the voltage across the capacitor also varies inversely with the frequency.

4.3.5 FM SPECTRUM AND BANDWIDTH REQUIREMENTS.

In the discussion of amplitude modulation, it was found that each audio frequency produced two side frequencies--the carrier plus the modulating frequency, and the carrier minus the modulating frequency. The bandwidth for an AM system has to be wide enough to include both of these frequencies or approximately two times the highest audio frequency. In single-sideband systems, where one of the sidebands is eliminated, the bandwidth has to be approximately equal to the highest modulating frequency.

When the carrier is frequency modulated, each modulating frequency produces a whole series of side frequencies on each side of the carrier, spaced at intervals equal to the modulating frequency. This arrangement of frequencies is called the frequency spectrum.

There is another major difference between the spectrum of an amplitude-modulated wave and that of a frequency-modulated wave. With amplitude modulation, the total transmitted power varies in accordance with the modulation level. With FM, the total transmitted power remains constant and is not determined by the modulation level. With frequency modulation, the level of modulation only changes the distribution of power in the side frequencies. The bandwidth for a given deviation has to be wide enough to include all of the side frequencies which represent a significant portion of the transmitted power.

Figure 4-17 is a large-scale reproduction of a single cycle of a frequency-modulated wave. This sketch shows the wave at an instant when the frequency is increasing and shows, in an exaggerated way, how the wave is not a pure sine wave. The changing frequency causes the time required to complete one-quarter cycle. The resultant wave is distorted, consisting of a number of sine waves superimposed. These superimposed waves are spaced at each side of the carrier at integral multiples of the modulating frequency.

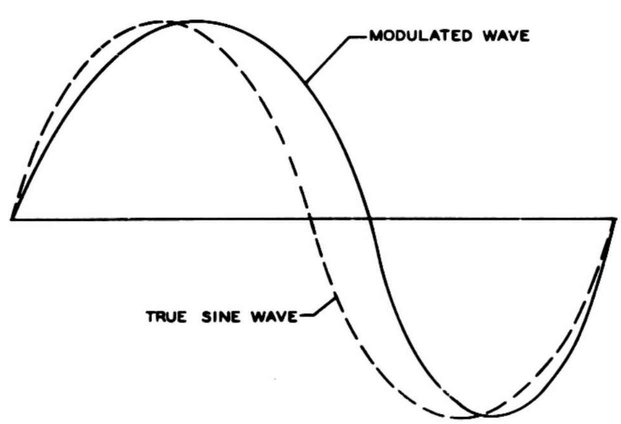


Figure 4-17. Comparison of Frequency-Modulated Wave and True Sine Wave

This frequency distribution is as follows:

1st side frequency pair = F_c ± 1 x modulating frequency

2nd side frequency pair = F_c ± 2 x modulating frequency

3rd side frequency pair = F_c ± 3 x modulating frequency

4th side frequency pair = F_c ± 4 x modulating frequency

This distribution continues for as many effective side frequencies as there are in the spectrum. Theoretically, there is an infinite number of side frequencies in an FM spectrum. However, the amplitudes of the higher level components diminish rapidly beyond a certain point. The relative amplitude of the carrier and each set of side frequencies is determined by the ratio of the frequency deviation to the modulating frequency $\left(\frac{\Delta F}{f}\right)$. This is the modulation index (M).

Figure 4-18 shows the spectral distributions for three different modulating frequencies. The center frequency and the frequency deviation are the same for each of these distributions. This figure shows that as the modulating frequency is decreased for a given frequency deviation, the number of significant side frequencies is increased. There are 11 significant side frequencies when the modulating frequency is 500 cps, and only 3 when the modulating frequency is 2000 cps. However, since the side frequencies are spaced closer together with the lower modulating frequency, the total bandwidth is less. Therefore, for a given deviation, the width of the FM spectrum increases with modulating frequency.

Figure 4-19 shows the spectral distributions for three different values of frequency deviation. The center frequency and modulating frequency are the same for

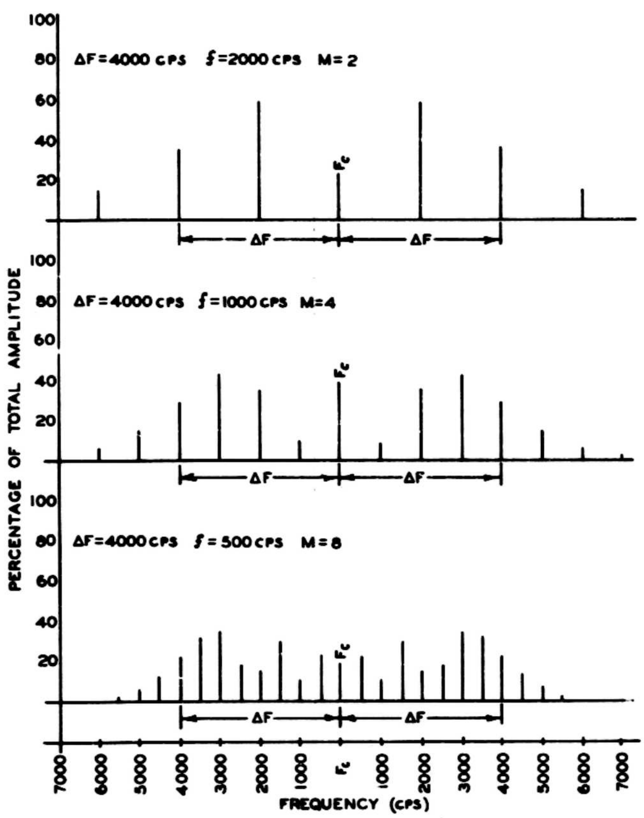


Figure 4-18. FM Spectrums for Various Modulating Frequencies with Constant Frequency Modulation

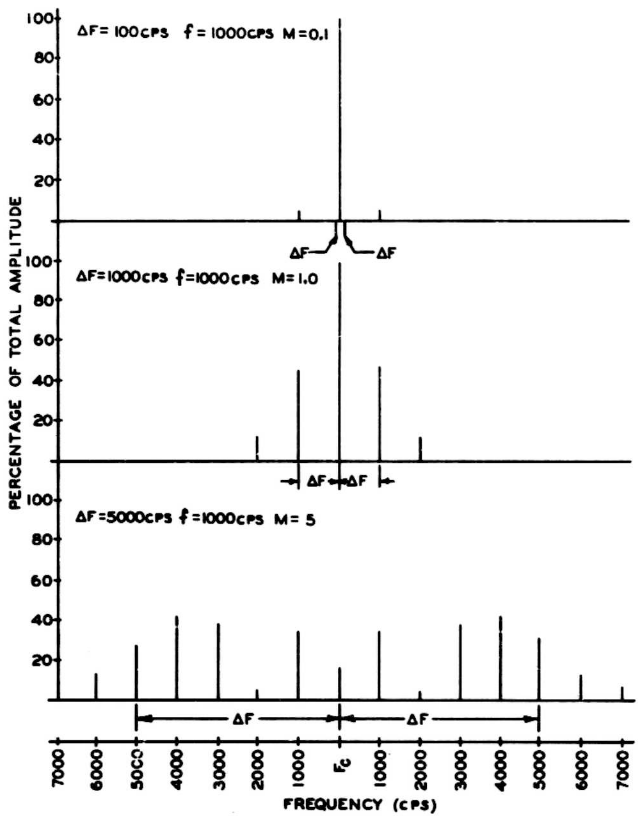


Figure 4-19. FM Spectrums for Various Frequency Deviations with Constant Modulating Frequency

each of these distributions. This figure shows that the number of significant side frequencies, and therefore the width of the spectrum, increases with an increase in frequency deviation. For example, with a ΔF = 100 cps and f = 1000 cps, there is only one significant pair of side frequencies. The width of this spectrum is identical to the AM spectrum which would result if the same modulating frequency were used. When ΔF is increased to 5000 cps and the modulating frequency is held at 1000 cps, there are six pairs of significant side frequencies. The spacing between the side frequencies remains the same since the modulating frequency has not been changed. However, the width of the total spectrum is increased since the number of pairs is increased.

From figures 4-18 and 4-19, it can be seen that the characteristics of the FM spectrum are determined both by the modulating frequency and the frequency deviation. The modulating frequency determines the spacing between the side frequencies; and the deviation, which is set by the modulating voltage level determines the number of significant side frequencies for a given modulating frequency.

The power relationship between the carrier and the side frequencies is determined by the modulation index M. This is shown in both figures 4-18 and 4-19. For example, in the top section of figure 4-19, where M = 0.1, almost all of the transmitted power is represented in the carrier. When M is increased to one, more side frequencies appear, and the amplitude of the carrier is decreased. From figures 4-18 and 4-19, the spectrums for various values of modulation index, from 0.1 to 8, can be determined.

There is a cyclic variation in the relative amplitudes of the various sidebands and the carrier as the modulation index is changed. This variation is shown for the carrier and the first three side frequencies in figure 4-20. With no modulation voltage applied, the modulation index is zero, and all of the power is in the carrier. As the modulation index is increased, the carrier power is decreased, and the power of the carrier is placed in the sidebands. When the modulation index reaches the value of 2.4, all of the power is represented in the sidebands. This is called <u>carrier disappearance.</u>

Carrier disappearance is useful for checking the deviation of an FM transmitter. For example, to operate with a frequency deviation of 28.8 kc, the first carrier disappearance will occur when:

$$2.4 = \frac{\Delta F}{f} = \frac{28.8 \text{ kc}}{f}$$

To solve this equation for the modulating frequency f:

$$f = \frac{28.8 \text{ kc}}{2.4} = 12 \text{ kc}$$

If this frequency is used as the modulating frequency, and the modulating level is adjusted, the desired deviation will be obtained at the point where the carrier component first disappears.

Figure 4-20 shows that as the modulation index is increased above 2.4, the phase of the carrier is reversed, and the level returns to zero when the modulation index is 5.5. This cyclic variation continues as the modulation index is increased with carrier disappearance at the various values noted on figure 4-20. Each of the side frequencies vary in a similar manner and disappear at various modulation indexes. The total power, which is the sum of the power in the carrier and the sidebands, remains constant for all values of modulation index.

For every value of modulation index, the amplitude of the carrier and each set of side frequencies is determined. When these waves, specified by the modulation index, are added together in correct amplitude and phase relationships, the correct FM wave, constant in amplitude and varying only in frequency, will result. To do this, the bandwidth has to be wide enough to include the significant side frequencies. For purposes of calculating bandwidth, any side frequency which has an amplitude greater than 1 percent of the total amplitude is considered significant. Table 4-1 shows the distribution of power in the spectrum for various values of modulation index. Table 4-1 does not take into account the differences in phase of the various components, and therefore the percentages do not add up to 100. The table gives the percentage of total power represented in each significant side frequency. It shows that the number of significant side frequencies increases directly with the modulating index. The required bandwidth is twice the number of significant side frequencies times the modulating frequency. This is because the significant side frequencies are spaced at intervals equal to the modulating frequency on each side of the carrier.

Try some numerical examples to illustrate the significance of table 4-1. Assume that the modulating frequency is 1 kc and the frequency deviation is 5 kc. The modulation index is equal to $\frac{\Delta F}{f}$, or $\frac{5}{1} = 5$.

Table 4-1 shows that when the modulation index is 5, there are eight significant side frequencies on each side of the carrier. The required bandwidth is 16 f, or 16 kc. Suppose the deviation is increased to 10 kc. With a modulating frequency of 1 kc, the modulation index increases to 10. Table 4-1 shows that there are fourteen significant side frequencies for this modulation index, and the required bandwidth is 28 f; or, in our example, this would be 28 kc. From these examples, we can see that for a given

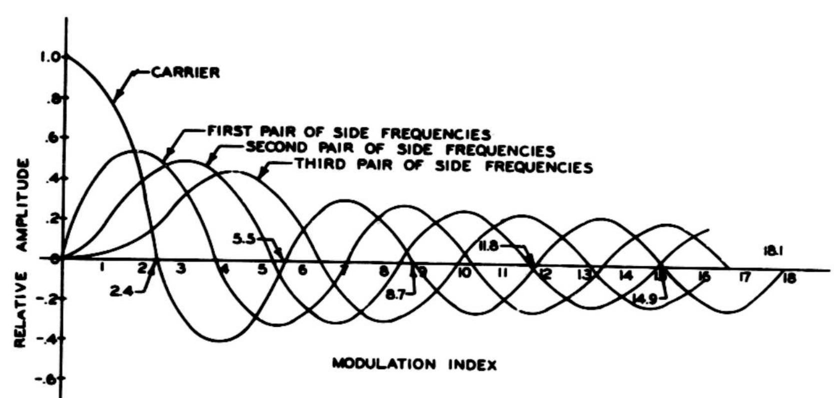


Figure 4-20. Variation of Amplitude of Carrier and First Three Pairs of Side Frequencies with Modulation Index

modulating frequency, the required bandwidth increases directly with deviation. If the bandwidth is not made wide enough to include all significant sidebands, distortion results. An equation used for estimating the bandwidth required for negligible distortion in an FM system is:

$$\text{Bandwidth} = 4\Delta F$$

Stated another way, in a system with a given bandwidth, the maximum allowable deviation in one-fourth the bandwidth. For example: assume that the system bandwidth is 240 kc. The maximum allowable deviation for this system is $\frac{240}{4} = 60$ kc. If the deviation is allowed to extend an appreciable amount beyond this value, distortion will result.

TABLE 4-1. DISTRIBUTION OF POWER IN THE FM SPECTRUM

$\frac{\Delta F}{f}$	CARRIER AND SIGNIFICANT SIDE FREQUENCIES EXPRESSED IN PERCENT OF TOTAL CARRIER LEVEL															REQUIRED BANDWIDTH
	F_c	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
0.1	99.75	4.99														2 f
0.5	93.85	24.23	3.10													4 f
1.0	76.52	44.01	11.49	1.96												6 f
2.0	22.39	57.67	35.28	12.89	3.40											8 f
3.0	26.01	33.91	48.61	30.91	13.20	4.30	1.14									12 f
4.0	39.71	6.60	36.41	43.02	28.11	13.21	4.91	1.52								14 f
5.0	17.76	32.76	4.66	36.48	39.12	26.11	13.10	5.34	1.84							16 f
6.0	15.06	27.67	24.29	11.48	35.76	36.21	24.58	12.96	5.65	2.12						18 f
7.0	30.01	0.50	30.14	16.76	15.78	34.79	33.92	23.36	12.80	5.90	2.30	0.80				22 f
8.0	17.17	23.46	11.30	29.11	10.54	18.58	33.76	32.06	22.35	12.63	6.10	2.60	0.96			24 f
9.0	9.03	24.53	14.48	18.10	26.55	5.50	20.43	32.75	30.51	21.49	12.47	6.20	2.73	1.10		26 f
10.0	25.59	4.35	25.46	5.83	21.96	23.41	1.45	21.67	31.79	29.19	20.75	2.31	6.34	2.90	1.20	28 f

In a system where a large range of modulating frequencies is used, the modulation index and corresponding bandwidth requirements will vary considerably from the lowest to the highest modulating frequency. For example: suppose that a system has been set for operation with a peak deviation of ±120 kc and that the modulating frequency range is from 12 kc to 60 kc. The modulation index with peak deviation will vary from 10, at 12 kc, to 2, at 60 kc. From table 4-1, the bandwidth required for 12 kc with a modulation index of 10 is 28 f, or 28 x 12 = 336 kc. The bandwidth required for 60 kc with a modulation index of 2 is 8 f, or 480 kc. For a given peak deviation, the greatest bandwidth is required for the highest frequency. For this reason, in systems where a wide range of modulating frequencies is used, the modulation index is expressed as the ratio of peak deviation to the highest modulating frequency. It is then called the <u>deviation ratio</u> for the system. The deviation represents the maximum modulating conditions for the system.

CHAPTER 5
NOISE AND INTERFERENCE

5.1 INTRODUCTION.

Interference from other communication systems and noise are two factors which limit the useful range of all radio equipment. This chapter provides a description of the various types of noise and tells how noise affects the received signal. The performance of AM and FM receivers in the presence of noise and interference is discussed. Special circuits which are used to offset the effect of noise in an FM system are described.

5.2 SOURCES OF NOISE.

The output of a receiver with no signal present is not zero, but fluctuates around some average value. This average output with no signal is called the output noise level of the receiver. There are various types of receiver noise, but they can all be considered to arise from one of the following sources: (1) noise generated in some external source and picked up by the receive antenna; (2) thermal noise generated in the antenna; and (3) noise generated within the receiver because of thermal agitation and other tube noise effects. Each of these sources is considered in the following paragraphs.

5.2.1 EXTERNAL NOISE.

There are three major types of noise resulting from sources external to the receive system. The first of these is atmospheric noise which results chiefly from lightning discharges in thunderstorms. Noise arising from this source is the principal limitation of radio service on lower frequencies. At frequencies above 30 mc, atmospheric noise falls to levels generally lower than the level of noise generated in the receiving system.

Another type of noise generated external to the receiver system is cosmic and solar noise received from the sun and certain constellations. The amount of noise received from these sources is dependent on the direction of propagation and antenna gain. Under certain propagation conditions with high-gain antennas and very low-noise receivers, cosmic and solar noise could be the limiting factor in receiver operation.

The third type of noise generated external to the receiver system is man-made noise. This is noise arising from such sources as: ignition systems, electric motors and generators, high-tension line leakage, and industrial heating generators. This noise is, of course, greatest in densely populated and industrial areas. Man-made noise is usually of no importance at the frequency range of tropospheric scatter equipment.

5.2.2 THERMAL NOISE.

Another major source of receiver noise is thermal or heat noise. Thermal noise is developed across any type of resistor because of the movement of free electrons within the resistor. The free electrons in any substance are never stationary, but are being continuously moved about because of the heat energy applied to the substance. As each electron moves, the charge concentration within the material is changed. Each change in charge concentration generates a small potential difference across the terminals. The additive effect of all the potential differences thus created, is a large number of sinusoidal voltages with a completely random distribution in both frequency and phase. This random distribution of sinusoidal voltages is called thermal noises. The antenna connected across the receiver input terminals acts as a source of thermal noise.

The amount of thermal noise generated is directly dependent upon temperature. This is because as the temperature increases, the movement of electrons increases, thus creating more and greater disturbances.

Thermal noise is equally distributed across the r-f spectrum. This means that the amount of noise for a given bandwidth will be the same regardless of the frequency. This relationship is shown in figure 5-1. Suppose that the receiver bandwidth is 0.5 mc. This bandwidth can slide along the noise spectrum and the noise power will remain constant. The amount of noise accepted by the receiver is directly proportional to its bandwidth. For example in figure 5-1 if the bandwidth is increased from 0.5 mc to 1.0 mc, the amount of noise admitted to the receiver is doubled. For a given temperature then, the average

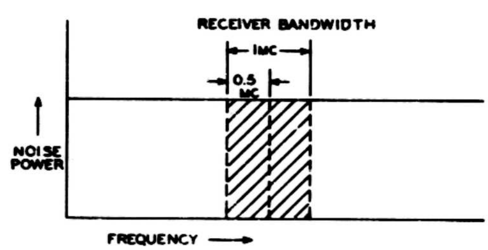


Figure 5-1. Relationship between Thermal Noise Frequency

level of thermal noise is constant across the spectrum, and the wider the bandwidth, the greater the amount of noise admitted to the receiver.

If we assume a temperature of 80°F, the amount of noise per kc of bandwidth is 4×10^{-15} watts. This relationship can be expressed with the following equation:

$$N = (4 \times 10^{-15})\ BW$$

where:

N = noise power in watts
BW = bandwidth in kilocycles.

It is convenient to express this relationship with the noise power expressed in dbm. This allows work with larger numbers and logarithmic relationships. Since powers expressed in dbm are referenced to a milliwatt, the expression for noise is changed to provide noise power in milliwatts:

$$N = (4 \times 10^{-15})\ BW$$

$$N\ (DBM) = 10 \log 4 + 10 \log 10^{-15} + 10 \log BW$$

$$= 10\ (.6) + (-150) + 10 \log BW$$

$$= -144 + 10 \log BW$$

where:

BW = bandwidth in kilocycles

N = noise power expressed in dbm.

To see how a change in bandwidth affects noise power, substitute some actual values in the equation. First assume a bandwidth of 1 kc. The noise power in dbm is then:

$$N = -144 + 10 \log BW$$

$$= -144 + 10 \log 1$$

$$= -144 + (10)\ (0) = -144 \text{ dbm}$$

Now, increase the bandwidth to 10 kc; the noise power in dbm is then:

$$N = -144 + 10 \log 10$$

$$N = -144 + (10)\ (1) = -134 \text{ dbm}$$

Figure 5-2 shows the change in noise power with change in bandwidth.

5.2.3 RECEIVER NOISE.

The other major source of noise is the receiver itself. Thermal and other types of emission noise

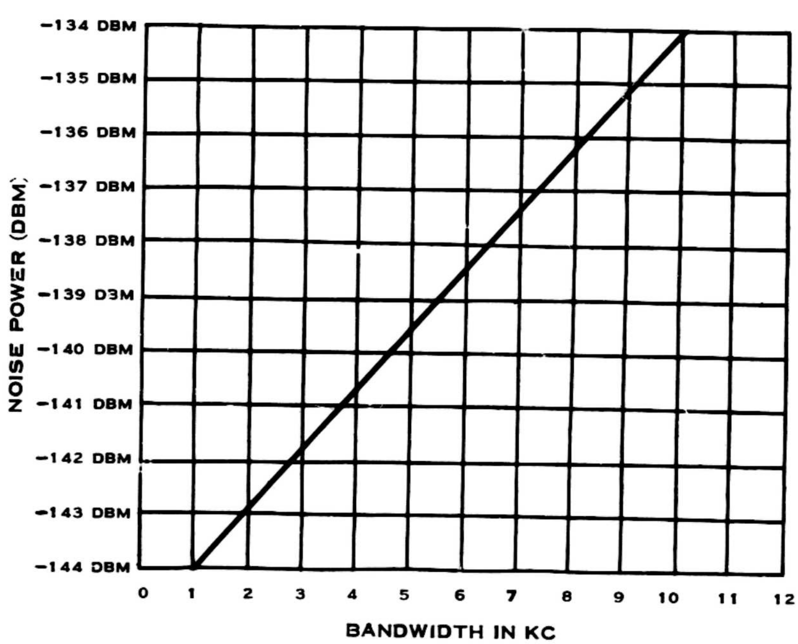


Figure 5-2. Change in Noise Power with Bandwidth

originate from each stage and add to the input noise. The term receiver noise figure is used to define the actual amount of noise generated in the receiver. The expression for noise figure is:

$$\text{Noise figure (F)} = \frac{\text{Actual Noise Output}}{\text{Ideal Noise Output}}$$

The noise accepted by a receiver will be amplified along with the input signal. Therefore, the output noise level is equal to the input noise times the over-all receiver gain plus the noise generated in the receiver:

$$N\ (\text{out}) = N\ (\text{in}) \times \text{Gain} + \text{Noise Generated in Receiver}$$

An ideal receiver has gain with no noise added, and therefore the ideal noise output in the expression for noise figure is equal to the receiver gain times the input noise:

$$\text{Noise Figure (F)} = \frac{\text{Actual Noise Output}}{\text{Gain x Input Noise}}$$

In the uhf range of frequencies used for tropospheric scatter, the greatest percentage of receiver input noise is the antenna thermal noise. Therefore, when determining receiver noise figure, it is assumed that the only input noise is noise arising from this source, and that the ideal noise output is equal to receiver gain times the antenna thermal noise. Thus, the noise figure of an actual receiver is:

$$F = \frac{\text{Actual Noise Output}}{\text{(Receiver Gain) (Antenna Thermal Noise)}}$$

Since some noise is added in an actual receiver, the noise figure is larger than unity. A low ratio indicates a low-noise receiver. A receiver with no noise would have a noise figure of 1. Typical uhf receivers now in use have a maximum ratio of approximately 6.3.

This ratio is usually expressed in db. When this is done, the noise figure can be added directly to the input noise to obtain the total noise in the receiver output. For example, the maximum ratio of 6.3 would be equivalent to a noise figure of 8 db. This tells us that the actual noise output is 8 db higher than the noise output of an ideal receiver.

5.3 OUTPUT SIGNAL-TO-NOISE RATIO.

Normally, when predicting performance of a tropospheric scatter circuit, it is necessary to know what the signal-to-noise ratio in the receiver output will be for predicted propataion conditions, receiver bandwidth, and receiver noise figure. This ratio is:

$$\frac{S\ (\text{out})}{N\ (\text{out})}$$

From the expression for noise figure, the actual noise output is equal to:

N_{out} = (Antenna Thermal Noise) (Gain) (F), where F = receiver noise figure.

The output signal is equal to the input signal times the receiver gain:

$$(S_{in} \times \text{Gain})$$

Therefore, the output signal-to-noise ratio is:

$$\frac{(S_{in})\ (\text{Gain})}{\text{(Antenna Thermal Noise) (Gain) (F)}}$$

The gain expression appears in both the numerator and denominator and cancels out so that the signal-to-noise ratio becomes:

$$\frac{S_{out}}{N_{out}} = \frac{S_{in}}{\text{(Antenna Thermal Noise) (F)}}$$

The higher this ratio is for a given signal, the greater will be the ability of the receiver to receive weak signals. From this equation, it can be seen that the ability of the receiver to receive weak signals is determined by the antenna thermal noise and the noise figure. An increase in either of these factors causes a corresponding increase in the minimum detectable signal. In the discussion of thermal noise, it was shown that the noise level admitted to the receiver increases directly with bandwidth. Therefore, an increase in bandwidth raises the minimum detectable signal level.

The chief source of receiver noise, and therefore the principal determinant of the receiver noise figure, is the first stage of the receiver. This is because any noise created in this stage will be amplified by all of the other stages in the receiver.

The equation for signal-to-noise ratio can be used to set up an expression for some minimum required signal. Let R equal the required ratio so that:

$$R = \frac{S_{in}}{\text{(Antenna Thermal Noise) F}}$$

The S_{in} required for this minimum ratio is then equal to:

$$S_{in} = \text{(Antenna Thermal Noise) (F) (R)}$$

If this equation is converted to a logarithmic relationship, S_{in} can be expressed in dbm and the expression already given for antenna thermal noise in dbm (-144 + 10 log BW) can be used.

The minimum detectable signal is then:

$$S_{in} = -144 + 10 \log BW + F + R$$

where:

S_{in} = minimum signal in dbm

BW = receiver bandwidth in kilocycles

F = receiver noise figure expressed in db

R = minimum signal-to-noise ratio expressed in db.

To see how this equation is used, try a few examples. Suppose that a system is being operated with a bandwidth of 250 kc; the receiver noise figure is 8 db, and the required signal-to-noise ratio is 10 db. The minimum detectable signal level is:

$$S_{in} = -144 + 10 \log (250) + 8 + 10$$

$$S_{in} = -144 + 10 \times 2.4 + 18$$

$$= -144 + 24 + 18 = -102 \text{ dbm}$$

Now suppose that all other conditions are kept the same, but the bandwidth is doubled to 500 kc. The minimum signal level is increased as follows:

$$S_{in} = -144 + 10 \log (500) + 18$$

$$S_{in} = -144 + 10 \times 2.7 + 18$$

$$= -144 + 27 + 18 = -99 \text{ dbm.}$$

The minimum detectable signal level is increased 3 db when the bandwidth is doubled. There is a direct increase in the minimum signal level when the noise figure is increased.

Since the noise level in the carrier output is constant for a given bandwidth and noise figure, the output signal-to-noise ratio increases with an increase in input signal level. Figure 5-3 shows the relationship between input signal level and signal-to-noise ratio for AM and FM systems. This is a typical representation, with the average FM power equal to the peak AM power and all other circuit parameters equal. These plots show the advantage of each system. In an AM system, the signal-to-noise is proportional to input signal level for all values of input level. In an FM system, the signal-to-noise is proportional to input signal level down to a certain signal level called the threshold level. At this point, the signal-to-noise ratio decreases very rapidly with a further decrease in signal level. However, for all levels above threshold, the FM system produces a higher signal-to-noise ratio for a given signal level than the AM system. This increase in signal-to-noise ratio is called FM improvement and is one of the most valuable properties of an FM system.

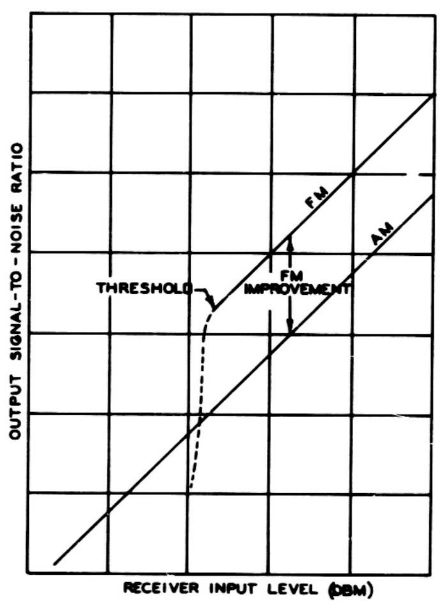


Figure 5-3. Change in Signal-to-Noise Ratio at Receiver Output with Change in Input Signal Level for AM and FM Systems

5.4 FM IMPROVEMENT.

FM improvement can be explained by comparing the effect of noise in AM and FM systems. Noise in an AM system causes amplitude disturbances which appear in the receiver output. Noise in an FM system causes amplitude disturbance, and it also produces variations in the frequency swing of the signal, resulting in a noise frequency-modulation disturbance. In an FM receiver, the amplitude disturbances are removed by the receiver-detection system which is insensitive to amplitude variations. The disturbances caused by frequency deviation, however, will be detected and be present in the receiver output. However, the effect of this frequency deviation caused by noise can be made small if the deviation caused by the signal is large in comparison. This noise reduction effect holds when the signal is appreciably stronger than noise, that is, when the signal level is above the threshold value, shown in figure 5-3.

Noise suppression and threshold in an FM system can be explained by using a vector representation given in figure 5-4. In this figure, S represents the frequency-modulated signal. It is assumed to rotate at an angular velocity equal to 360° x F, where F is the instantaneous frequency. The S vector is alternately advanced and retarded by an angle equal to θ, where θ is the angle relative to the unmodulated carrier. The size of θ is determined by the modulation index. The noise power can be considered to be a vector rotating at the end of the signal vector, causing amplitude and angular changes. The amplitude variations are eliminated in the receiver detection system. In A of figure 5-4, the maximum angular variation caused by the noise is X. From this figure, it can be seen that the greater the deviation due to signal is made, the less significant is the angular variation due to noise. The FM improvement increases with the frequency deviation.

There is, however, a limit to the amount of noise reduction which can be obtained by increasing the deviation due to signal level. In chapter 4, it was found that the required bandwidth increases with an increase in deviation. In this chapter, it has been seen that the noise power increases with bandwidth. Therefore, as the deviation is increased for greater

FM improvement, the bandwidth must be increased accordingly. This increases the noise level and raises the threshold level of the receiver. A compromise must be made between FM improvement and threshold level.

Figure 5-4B shows the effect of high level noise. As the noise level is increased, the angular variation caused by it becomes a high percentage of the total angular deviation caused by the signal. Finally, a point is reached where the deviation is being controlled by the noise rather than by the signal. This point is called the threshold level. As the noise peaks exceed the signal peaks, the signal-to-noise ratio at the FM receiver output deteriorates very rapidly.

5.5 RECEIVER THRESHOLD.

Normally, in an FM system, the threshold level is defined as the received input power which produces approximately a 10-db rms signal-to-noise ratio. This threshold level can be obtained from this equation:

S (threshold) = -144 + 10 log BW + F + R

BW = receiver bandwidth in kilocycles

F = receiver noise figure expressed in db

R = S/N ratio required for threshold expressed in db for FM systems
R = 10 db.

Therefore:

S (threshold) = 144 + 10 log BW + F + 10 =
-134 + 10 log BW + F

The 10-db signal-to-noise ratio required for FM threshold can be explained by referring to figure 5-5. The detector system in an FM receiver is controlled by peak voltage. Thus, if the noise peaks exceed the signal peaks, the detector will follow noise input rather than signal. Therefore, a peak signal voltage must, at least, equal peak noise or the signal-to-noise ratio deteriorates rapidly. In determining signal-to-noise ratios, rms values are used, since these are the values which will be measured. Peak noise is defined as the voltage level which is exceeded one percent of the time. In a random distribution, such as represented by noise, the ratio of the peak value to rms value is 13 db. The ratio of peak to rms for the sinusoidal signal is 3 db. Therefore, the ratio of the rms value of signal to the rms value of noise at threshold, where the noise peaks equal signal peaks, is 10 db.

5.6 DE-EMPHASIS AND PRE-EMPHASIS CIRCUITS.

In figure 5-4, the noise voltage super-imposed on the signal actually causes phase modulation of the signal. Deviation in phase modulation increases with frequency as shown in chapter 4. Therefore, although the noise spectrum is flat across the receiver bandwidth, the interference caused by noise increases with frequency. This would result in an increased effect of noise and a lower signal-to-noise ratio at the high end of the signal spectrum.

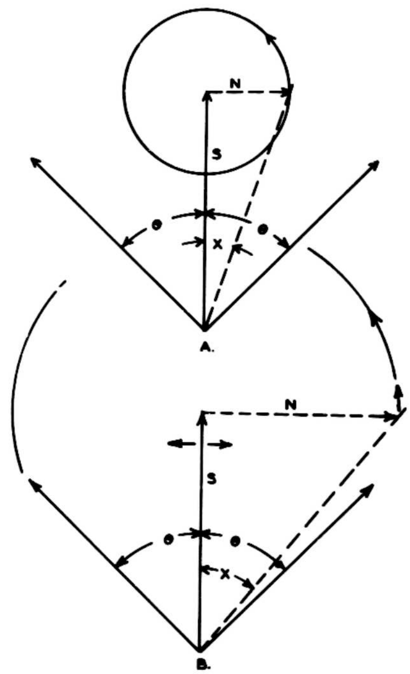


Figure 5-4. Vector Representation of Signal and Noise in an FM System

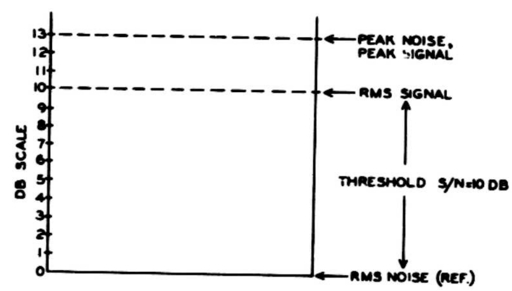


Figure 5-5. Relationship between Peak and RMS Value for Noise and Signal at Threshold

This increased effect of noise with frequency can be counteracted by using what is called a de-emphasis circuit in the receiver. This circuit is shown in figure 5-6. It is an RC circuit with the input and output connections arranged to provide a 1/f circuit. As shown in figure 4-6, the output is taken from across the capacitor. A decrease in capacitive reactance with frequency causes a decrease in output voltage.

Since both noise and signal voltages are applied through the de-emphasis circuit, the signal voltage will also be decreased with frequency. To counteract this effect, the amplitude of the higher frequencies is increased before modulation at the transmitter. The circuit for doing this is called a pre-emphasis circuit. This circuit and the variation in amplitude of modulating signal with frequency before and after pre-emphasis is shown in figure 5-6. The pre-emphasis circuit, like the de-emphasis circuit, is also an RC circuit. However, in this case, the output voltage is taken from across the resistor so that the output level increases with frequency. This arrangement is called an f network. If the circuit provided true f relationship for all frequencies, there would be very low output at the low end of the signal spectrum. This would result in extremely low deviations at the low frequencies. Therefore, most pre-emphasis circuits provide an increase in output level with frequency only for the high-frequency end of the spectrum. This change in characteristics at the high frequencies is provided by using the proper time constant (RC) in the pre-emphasis circuit. The characteristics of the pre-emphasis and de-emphasis circuits used in the same system must match.

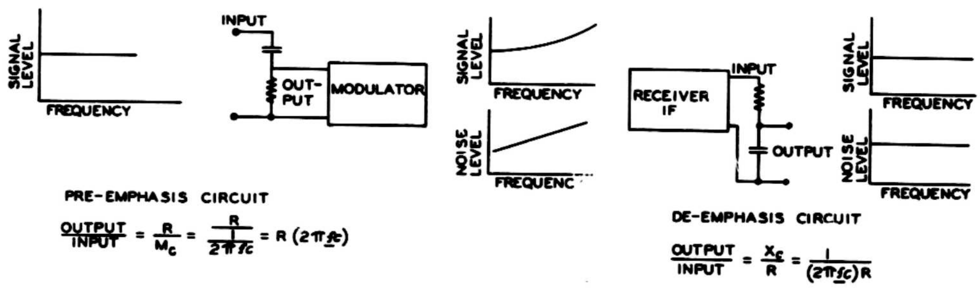


Figure 5-6. Pre-Emphasis and De-Emphasis Circuits

CHAPTER 6
FM EXCITERS

In chapter 4, the basic principles of frequency modulation were discussed. In this chapter, it will be seen how these principles are applied in the design of FM exciters. The operation of each of the major components of an FM exciter is described. This information is essential for understanding the basic adjustments and maintenance procedures which are required for an FM exciter. A description and explanation of the basic adjustments are also included in this chapter.

6.1 PURPOSE OF EXCITER AND DESIGN REQUIREMENTS FOR TROPOSPHERIC SCATTER.

The FM exciter in a tropospheric scatter system provides a modulated signal at the transmit frequency for excitation of the associated power amplifier or for direct feed to an antenna. The r-f carrier is modulated in accordance with baseband input from the multiplex and order-wire equipment.

An FM exciter designed for use with tropospheric scatter equipment must meet certain performance requirements. These requirements are listed and described below.

6.1.1 OUTPUT FREQUENCY.

The exciter must provide an input to the power amplifier at the final transmit frequency, since normally the power amplifier is only a straight amplifier with no frequency translation facilities. The exciter must be capable of being tuned to any selected frequency in the transmit frequency range. One reason for this tuning range requirement is that provision must be made for selection of proper transmit frequency to avoid interference with and from other services. Also, the various transmitters operating in the same tropospheric scatter circuit must be operated at separated frequencies.

6.1.2 FREQUENCY STABILITY.

The output frequency must remain within a small percentage of the selected value. This is because once a transmit frequency has been selected, all of the tuned circuits in the exciter, power amplifier, and associated receiver are tuned for proper operation at the frequency. If the frequency is allowed to drift, distortion and high losses will result. Also, the greater the instability of the exciter, the wider must be the bandwidth of the associated receiver. This is because the bandwidth must be broad enough to compensate for drift in carrier frequency. The increase in receiver bandwidth to compensate for exciter instability reduces the sensitivity of the receiver.

The frequency stability requirement is significant because it determines the type of modulating system which can be used. To meet the stability requirements, a system of indirectly creating frequency modulation from a phase modulator is often used. The phase modulated is driven by a highly stable crystal oscillator. Another method which could be used is a direct FM modulator with automatic frequency control.

6.1.3 MAXIMUM DEVIATION.

It was stated in chapter 5 that the amount of FM improvement in the receiver output signal-to-noise ratio varies directly with the deviation. However, the receiver bandwidth requirements, and thus the receiver threshold level, increase as the deviation is increased. Therefore, the maximum deviation to be used for a system presents a compromise between the desired FM improvement factor and the received signal power necessary to exceed threshold. The exciter used in the system must be capable of producing the specified deviation.

6.1.4 DISTORTION.

FM exciters used for tropospheric scatter communications must meet low distortion requirements while producing the required deviation. This becomes difficult when phase modulators are used to produce large deviations because one characteristic of phase modulation is that the percentage of distortion increases very rapidly for large phase deviations.

The relationship between modulating frequency, phase deviation, and equivalent frequency deviation is:

$$\Delta\theta = 57.3\frac{\Delta F}{f} \text{ or } F = \frac{f\Delta\theta}{57.3}$$

where:

$\Delta\theta$ is phase deviation in degrees

ΔF is equivalent frequency deviation

f is modulating frequency

ΔF and f must be expressed in same units.

For large frequency deviations, the phase deviation must be increased accordingly. As the phase deviation is increased, the distortion also increases. This difficulty of increased distortion with large deviations is overcome by allowing only a small phase deviation in the modulator, and then passing the modulator output through circuits which multiply the frequency deviation to the required value.

From the expression, $\Delta F = \frac{f\Delta\theta}{57.3}$, it can be seen that for a given ΔF, the product $f\Delta\theta$ must be constant.

Therefore, the greatest phase deviation is required at the lowest modulating frequencies. For example: for a fixed F of ±200 kc and modulating frequency range of 4 to 50 kc, the following required phase deviations are at each end of the modulating frequency range.

$$200 = \frac{4(\Delta\theta)}{57.3} \qquad \Delta\theta = 2865°$$

$$200 = \frac{50(\Delta\theta)}{57.3} \qquad \Delta\theta = 229.2°$$

Since the greatest phase deviation is required at the lower modulating frequencies, distortion in the phase modulator is most serious at these frequencies. Therefore, the phase modulator is designed to meet low distortion requirements at the low end of the modulating frequency range.

In addition to modulator distortion, there is also the problem of distortion in the circuits which follow the modulator. To pass the FM signal properly, the tuned circuits must be broad enough to pass all of the significant side frequencies. If the bandwidth is too narrow, distortion, due to phase nonlinearity, results. In chapter 4 (table 4-1) it was seen that the required bandwidth for a given deviation increases with modulating frequency. Therefore, distortion arising in the tuned circuits is most serious at the highest modulating frequencies.

6.1.5 SPURIOUS OUTPUT.

Any r-f signals from the exciter other than the assigned carrier frequency are called spurious outputs. To prevent internal interference and interference with other systems, these spurious outputs must be kept at a low level. This is usually done by using tuned circuits which filter out r-f frequencies other than the carrier. However, the tuned circuits must be broad enough to pass all of the significant side frequencies. This means that sufficient tuned circuits with proper selectivity characteristics must be provided to accomplish the desired suppression.

6.1.6 INCIDENTAL FM.

Theoretically, the deviation produced in an FM exciter should drop to zero when no modulating voltage is applied. However, in practice, there is a certain amount of deviation present even with no modulating voltage. This is called incidental FM and is usually caused by ripple in the power supply, short term instability of the modulated oscillator, microphonics in low-level tubes due to blower or other vibrations, or, in some cases, faulty tuning of the stages which follow the modulator. Incidental FM results in noise which appears in the receiver detected output and, therefore, should be kept at a very low level. The amount of incidental FM is usually stated as the ratio of the maximum deviation to the deviation resulting from incidental FM. This can be stated in db, where: Incidental FM (db) = 20 log (Peak Deviation-Incidental FM Deviation). This is a voltage relationship, since voltage at the receiver output is proportional to deviation.

6.1.7 OUTPUT POWER.

Unlike the usual practice in AM transmitters, modulation in an FM exciter is usually accomplished at a low-power level. Output from the modulator must be amplified to provide sufficient drive for the associated power amplifier. In some cases, the exciter output may be applied directly to the antenna for low-power line-of-sight circuits. The output level must be adjustable to provide for various line losses in different installations and drive power requirements of the particular power amplifier with which it is to be used.

6.1.8 RESPONSE CHARACTERISTICS FOR MODULATING FREQUENCY RANGE.

An FM exciter used for tropospheric scatter must be capable of operating with a wide range of modulating frequencies, depending on the channel capacity of the system. To be a true FM system, the frequency deviation must be independent of frequency throughout this range. A phase modulator, with no corrective network ahead of the modulator, produces a frequency deviation which increases in proportion to the modulating frequency. To convert the output of the phase modulator to FM, the modulating input must be fed through a network which provides an input level to the modulator, which is inversely proportional to the frequency. This network, called a 1/f circuit, is described in chapter 4.

In chapter 5 in the discussion of noise, intereference caused by noise is greatest at the high end of the modulating frequency range. To compensate for this noise characteristic, a de-emphasis circuit is used in the receiver. This is also a 1/f circuit which provides a decreased level at the high end of the band. Since the high signal frequencies are also attenuated in the de-emphasis circuit, a circuit must be provided in the exciter to pre-emphasize the high frequencies.

The required 1/f and pre-emphasis circuits are usually combined in the exciter. It is not feasible to design a combined circuit which will operate satisfactorily for all the modulating frequencies required by the various channel capacities of a system. Therefore, in some exciters, several combined circuits are provided, each designed for operation in a particular modulating frequency range.

6.2 BLOCK DIAGRAM OF TYPICAL FM EXCITER.

Figure 6-1 is a block diagram of a typical FM exciter used for tropospheric scatter communications. The frequency scheme is shown for an output frequency in the range of 755 to 985 mc.

The exciter may be divided into three basic sections. The modulator section produces a signal modulated in accordance with the baseband input with the required maximum deviation. The UHF injection section generates a frequency which, when mixed with output from the modulator, produces the required output frequency. The third section amplifies output from the mixer to the required output level. Both the modulator section and the UHF injection section use a crystal oscillator. The exciter stability is a function of both of the oscillators.

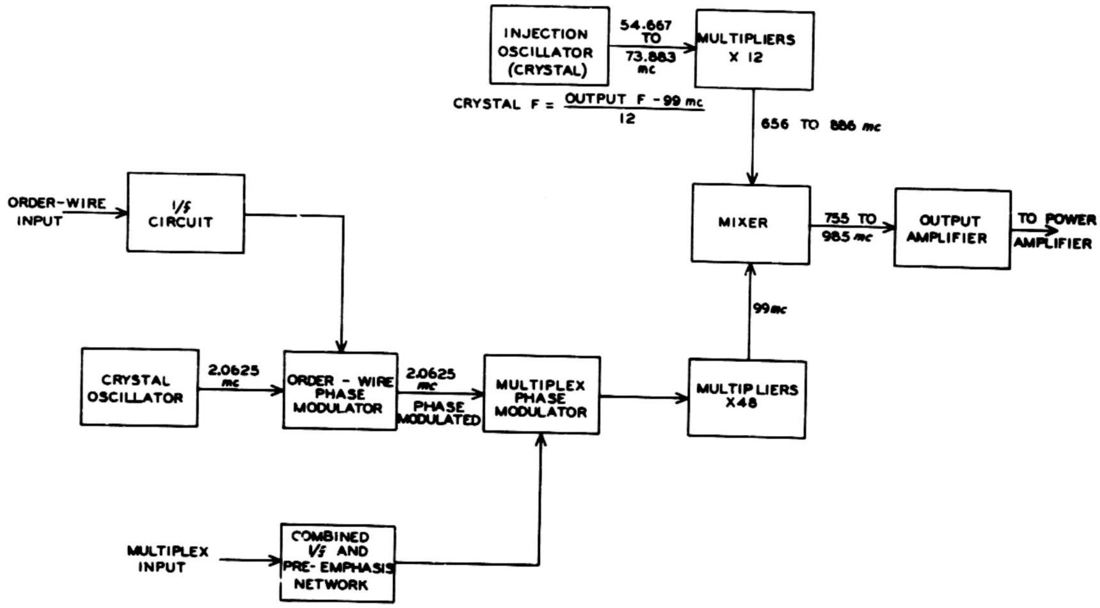


Figure 6-1. Block Diagram of Typical FM Exciter Used for Tropospheric Scatter Communications

Output from the crystal oscillator in the modulator section is phase modulated by order wire and multiplex inputs. Figure 3-20 shows a typical spectrum arrangement for these inputs. The multiplex input is applied through a combined 1/f and pre-emphasis network which provides for correct response through the range of multiplex frequencies. The order wire input is applied through a 1/f network which results in an FM characteristic in the order wire range.

The distortion created in a phase modulator increases very rapidly for large phase deviations; therefore, the original phase deviation is kept very small. The required deviation is obtained by following the modulator with multiplier stages. Output from the phase modulator is at the frequency of the crystal oscillator with a very small amount of deviation. When this signal is passed through a multiplier stage, the deviation and center frequency are multiplied by the same factor.

As shown on figure 6-1, output from the phase modulator is first passed through a series of multipliers which provide a total multiplication factor of eight. This results in a frequency of 16.5 mc, which is 8 times the original frequency of the crystal oscillator, and a deviation 8 times that produced in the phase modulator. The 16.5-mc signal is then applied through a series of multipliers which multiply the center frequency and deviation by a factor of 6. This results in a carrier frequency which is 48 times that f the original frequency, and a deviation which is 48 times that produced by the phase modulator.

The amount of multiplication required in the modulator section is determined by the ratio of maximum deviation required to the maximum deviation possible in the modulator for a specified distortion level. For example: assume that the maximum deviation required is ±12 kc at the lowest modulating frequency of 4 kc. This makes a deviation ratio (M) of 3. Also, suppose that the distortion requirements set a maximum phase deviation of 3.6 degrees at the modulator. We can determine the corresponding deviation ratio from this expression:

$$\Delta\theta = 57.3\left(\frac{\Delta F}{f}\right)$$

$$\Delta\theta = 57.3\ M$$

$$M = \frac{3.6}{57.3} = 0.063$$

Therefore, the required multiplication is $\frac{3}{0.063} = 47.6$, or the closest whole number factor is 48. This is the multiplication factor shown in figure 6-1.

After the required multiplication factor is determined, the next consideration is the frequency of the oscillator which drives the phase modulator. The ratio of the modulator carrier frequency to highest modulating frequency must be high. In figure 6-1, notice that output from the multiplier stages in the modulator section is applied to a mixer, along with output from the UHF injection section. The mixer output is the sum of the two input frequencies. The combination of the two frequencies must be arranged so both input frequencies and other combinations of the various harmonics will be outside the pass band of the tuned circuits in the mixer output. After the required mixer input frequencies have been set, the frequency of the crystal oscillator can be determined by dividing the required mixer input by the required multiplication factor. For example: in figure 6-1, the required mixer input is 99 mc, and the required multiplication factor is 48. Therefore, the oscillator frequency is $\frac{99 \text{ mc}}{48} = 2.0625$ mc.

Another consideration in the selection of the output frequency from the modulator section is that it is very advantageous to have this frequency equal to the i-f frequency in the associated receiver. When these frequencies are the same, back-to-back testing at a terminal is made easier.

The exciter output frequency is set by selecting the correct crystal for the injection oscillator, and then tuning the multiplier stages which follow the oscillator. The range of oscillator frequencies is placed where good performance is obtained from crystal oscillators. The multiplication factor is required to raise the oscillator frequency to the required mixer input. Since the output from the modulator section is fixed, the output from the UHF injection section is equal to: output frequency - modulator section output frequency. The crystal frequency of the carrier oscillator is equal to this difference divided by the multiplication factor following the oscillator. For example, in figure 6-1, the crystal frequency is equal to:

$$\frac{\text{Output Frequency} - 99 \text{ mc}}{12}$$

6.3 CIRCUITS MODIFYING AUDIO RESPONSE OF PHASE MODULATORS.

The audio response of the transmit and receive equipment is a very important characteristic of the communications system. The response is modified to produce a signal-to-noise ratio which is approximately constant across the receiver baseband output. The noise which must be considered, when modifying the response for constant signal-to-noise ratio, consists essentially of two types. One type of noise is incidental FM which originates in the transmit equipment and generally has the greatest effect at the low end of the baseband spectrum. The other type is receiver thermal noise. The effect of this noise increases with frequency. Therefore, the noise level encountered across the range of modulating frequencies is not constant, and the signal level response must be modified accordingly.

The signal voltage output of the receiver is a function of the deviation produced in the exciter. Therefore the desired signal voltage response of the system can be obtained by modifying the audio response of the phase modulators in the exciter. In chapter 4, it was shown that a phase modulator produces an equivalent frequency deviation that increases with an increase in modulating frequency. This unmodified response of a phase modulator, shown in figure 6-2, results in a receiver signal output which increases directly with frequency. The desired signal response is attained by changing the modulating voltage input level to the modulator in accordance with the modulating frequency.

Typical tropospheric scatter systems operate with a wide range of modulating frequencies. A typical spectrum arrangement is shown in figure 3-20. The order wire input usually extends from 250 cps to 4 kc, 8 kc or 12 kc. The multiplex input extends from the upper limit of the order wire to an upper frequency limit determined by the channel capacity of the system. Since it is difficult to design one modulator to operate properly throughout the entire range, separate modulators are sometimes provided for the order wire and multiplex inputs. Each modulator is preceded by an

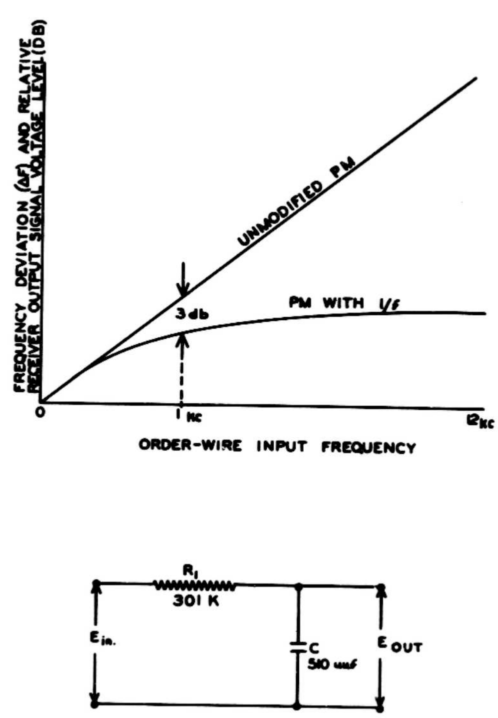


Figure 6-2. Response of Order Wire Modulator with Typical 1/f Circuit

RC circuit which modifies the response of the modulator to provide the desired response for the particular range of requencies.

The response of the order wire phase modulator is modified to produce frequency deviation independent of modulating frequency across the major portion of the order wire spectrum. This is done with a 1/f network, which is an RC circuit used to obtain a modulating voltage inversely proportional to frequency. A typical 1/f circuit and resultant phase modulator response are shown in figure 6-2. The input is applied across a resistor and capacitor in series, and the output is taken from across the capacitor. To see how this circuit modifies the response of the phase modulator, first derive the expression for the ratio of input voltage to output voltage:

$$\frac{E_{in}}{E_{out}} = \frac{IZ_{in}}{IZ_{out}} = \frac{\sqrt{R^2 + X_c^2}}{X_c}$$

At low frequencies, X_c is much greater than R, and the output voltage is essentially equal to the input voltage:

$$\frac{E_{in}}{E_{out}} = \frac{\sqrt{(X_c)^2}}{X_c} = 1$$

Therefore, for the low frequencies, the response of the phase modulator is unaltered as shown in figure 6-2. As the frequency increases, X_c decreases and approaches the value of R. The network begins to act as a voltage divider, resulting in the 1/f response characteristic shown in figure 6-2. The response curve bends away from the natural response of the phase modulator and assumes the 1/f characteristic at the frequency where $R_1 = X_c$. At this point, the output voltage is 3 db below the input voltage:

$$\frac{E_{in}}{E_{out}} = \frac{\sqrt{R^2 + X_c^2}}{X_c} \qquad X_c = R$$

$$\frac{E_{in}}{E_{out}} = \frac{\sqrt{2\,X_c^2}}{X_c} = \sqrt{2}$$

$$\frac{E_{in}}{E_{out}}\ (\text{db}) = 20 \log \sqrt{2}$$

$$= 10 \log 2 = 10 \times 0.3 = 3\ \text{db}$$

The expression for the frequency where the 1/f characteristic starts is derived as follows:

$$R_1 = X_c = \frac{1}{2\pi fC}$$

$$(R_1)\,(2\pi fC) = 1$$

$$f = \frac{1}{2\pi CR} = \frac{0.16}{RC}$$

where: f is in cycles, R in ohms, and C in farads. In the circuit shown in figure 6-2, the frequency is:

$$f = \frac{0.16}{(301 \times 10^3)(510 \times 10^{-12})} = 1050\ \text{cps}$$

As the frequency increases, X_c becomes much smaller than R. The ratio of E_{in} to E_{out} can then be simplified as follows:

$$\frac{E_{in}}{E_{out}} = \frac{\sqrt{R^2 + X_c^2}}{X_c}$$

$$\frac{E_{in}}{E_{out}} = \frac{\sqrt{R^2}}{X_c} = \frac{R}{X_c} = (R)(2\pi fC)$$

This ratio can be expressed in db:

$$\frac{E_{in}}{E_{out}} = 20 \log (R)(2\pi fC)$$

Since all the other parameters for a given circuit are constant except f, this can be simplified for our discussion to:

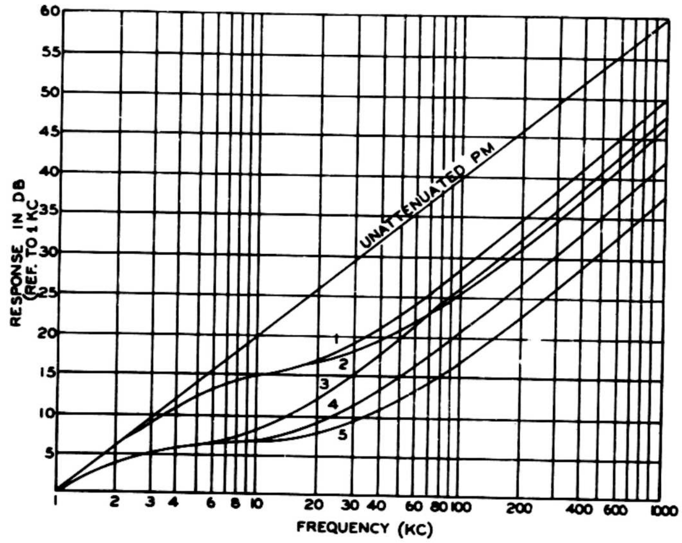
$$\frac{E_{in}}{E_{out}} = K + 20 \log f$$

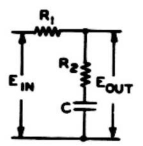
where K is constant, depending on the component values in the circuit. We can see that after X_c decreases below the value of R, the output voltage from the 1/f network decreases at a rate of 20 log f, or 6 db per octave. This is the rate of decrease required to cancel the normal increase in deviation in a phase modulator to produce frequency deviation independent of modulating frequency as shown in figure 6-2.

The multiplex phase modulator operates over a much wider range of frequencies than the order wire modulator. The response of the multiplex modulator has to be altered to compensate for the increase in noise power with frequency across the multiplex range. The natural response of a phase modulator would provide the required increase in signal level. However, if this response were used, deviation and corresponding signal level would be too low at the low end of the multiplex range. Therefore, the response of the multiplex modulator is altered with a combined 1/f and pre-emphasis circuit, which provides deviation independent of modulating frequency at the low end of the band, and deviation which increases with modulating frequency for the higher modulating frequencies.

Figure 6-3 shows a typical combined 1/f and pre-emphasis circuit with response curves for various component values. This circuit is similar to the 1/f circuit used with the order wire modulator except that a resistor (R_2) is added in series with the capacitor in the output circuit.

To see how the combined 1/f and pre-emphasis circuit provides the required response, first notice how circuit operation changes with frequency. The values of circuit components are selected so that R_2 is much smaller than R_1, and X_c is much larger than $R_1 + R_2$ at the low end of the band. Therefore, for the low frequencies, $E_{out} = E_{in}$, and there is little change in the response of the modulator. This condition continues until X_c approaches the value of R .





CURVE	R_1	R_2	C	1/f FREQ.	PRE-EMPHASIS FREQ. KC	MULTIPLEX CHANNELS	MULTIPLEX FREQ. RANGE
1	100 K	36.4 K	220UUF	7.26 KC	20.0	12	12-60 KC
2	100 K	22.9 K	220UUF	7.26 KC	32.5	24	12-108 KC
3	300 K	82.5 K	220UUF	2.42 KC	8.81	4	4-20 KC
4	300 K	36.4K	220UUF	2.42 KC	20.0	12	4-52 KC
5	300 K	22.3K	220UUF	2.42 KC	32.5	24	4-100KC

Figure 6-3. Response of Phase Modulator with Typical Combined 1/f and Pre-Emphasis Circuits

The circuit then begins to function as a 1/f network to produce frequency deviation independent of modulating frequency as shown in figure 6-3. The frequency at which the output voltage is 3 db below the input voltage is determined by the same equation given for the order wire 1/f circuit:

$$f = \frac{0.16}{R\,C}$$

Frequencies for various values of R_1 and C are given in the table in figure 6-3.

As the frequency increases, X_c continues to decrease, and the circuit continues to function as a 1/f network. This continues until the value of X_c approaches the value of R_2. At .his frequency, the response characteristic is changed to produce frequency deviation which increases with frequency. This is called the pre-emphasis part of the response curve. The frequency at which the response curve assumes the pre-emphasis characteristic can be determined by finding the R_2C time constant and using the expression:

$$f = \frac{160}{R_2C}$$

where f is in kilocycles, and R_2C is in microseconds.

The value of R_2 and C are usually expressed as time constant in microseconds rather than in particular values of R_1 and C. This makes it possible to match the pre-emphasis circuit to a corresponding de-emphasis circuit in the receiver.

Figure 6-3 shows some typical response curves with various values of R_1 and R_2. The curves show frequency deviation relative to the deviation which would result with no network preceding the modulator. These response curves illustrate that the frequency at which the response departs from the natural phase modulator response and assumes a 1/f characteristic is determined by the value of R_1C. The frequency at which the pre-emphasis characteristic starts is determined by the value of R_2C.

The curves, shown in figure 6-3, are the response curves resulting from selecting the values of R_1 and R_2 for the proper operation with various baseband frequency arrangements. The frequency at which the pre-emphasis occurs is selected to correspond with the multiplex channel capacity. For example, curve 4 on figure 6-3 shows the response characteristic for a 12-channel systen starting at 4 kc. The value selected for R_2 is 36.4K. This makes a pre-emphasis time constant of eight microseconds and sets the frequency at 20 kc. Curve 1 shows the response characteristic for a 12-channel system starting at 12 kc For comparison with a larger capacity system, look at curve 2 which shows the response for a 24-channel system. The value of R_2 is decreased to 22.9K to make a pre-emphasis time constant of five microseconds and a frequency of 32.5 kc.

6.4 ORDER-WIRE MODULATOR.

Output from the modulator crystal oscillator is applied to an order wire modulator, where it is phase modulated in accordance with the order wire modulation input.

To produce a frequency-modulated wave with a phase modulator, several requirements must be met. First, the amplitude of the modulating voltage must be reduced in amplitude in proportion to the frequency. This is done by passing the modulating voltage through a 1/f network which precedes the modulator. A second requirement is that the phase deviation produced in the modulator must be made to vary linearly with the amplitude of the modulating input. Also, in the process of causing phase deviation, any amplitude change in the r-f signal must be held to a minimum. If these requirements are fulfilled, the equivalent frequency deviation will be independent of modulating frequency and dependent only on the amplitude of the modulating voltage.

Figure 6-4 shows a simple circuit for producing phase modulation. The vector diagrams in this figure show the various voltages in the circuit, with the r-f input voltage (V_1) used as a reference. This r-f input is applied through a circuit which shifts the input through a fixed angle θ. Output from the phase-shift circuit is passed through variable resistor R_2. This results in a voltage V_2, which is displaced from V_1 at the angle θ. The amplitude of V_2 depends on the setting of R_2. Another path for the input voltage V_1 is directly through variable resistor R_1. Output from this path is V_3 which is in phase with V_1 and at some value less than V_1, depending on the setting of R_1. The variable resistors are mechanically linked so that as one is increased, the other is decreased. This means that as V_2 is increased, V_3 is decreased accordingly, and vice versa.

The output voltage V_R is the vector sum of V_3 and V_2. This vector relationship is shown in figure 6-4 for two settings of R_1 and R_2. When the resistors are set in the position shown for diagram A, V_2 is increased and V_3 is decreased. This results in the resultant V_R displaced at angle X from the input voltage. When the resistors are set in the opposite direction, V_2 is decreased and V_3 is increased. This causes a change in the angle X. These vector diagrams show how the phase angle is changed with only an insignificant change in the amplitude of the resultant vector V_R. If resistors R_1 and R_2 could be made to move in accordance with the amplitude of the modulating voltage, phase modulation would result.

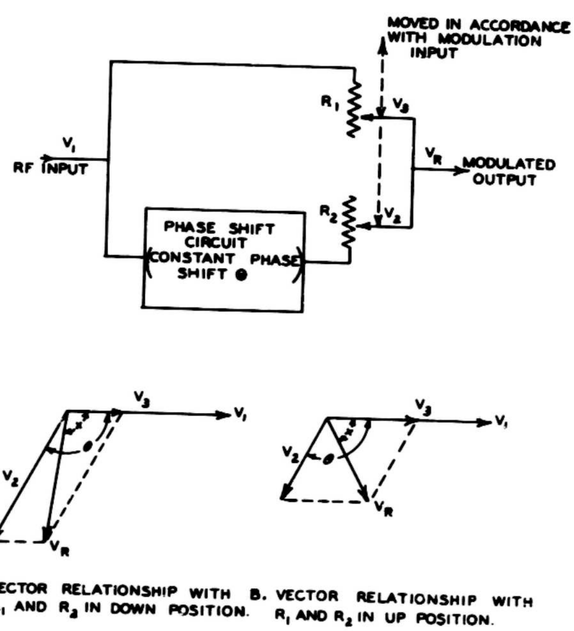


Figure 6-4. Method of Obtaining Phase Modulation with Phase-Shifting Network

In practice, the variable resistors are replaced with a vacuum-tube circuit shown in figure 6-5. In this diagram, the r-f input voltage V_1 is again used as reference. One path for the signal voltage is through a phase-shifting network, C_1 and R_1, and then from grid to plate in the tube. The phase-shifting network shifts the voltage V_4 at the grid by something less than 90°. As V_4 is amplified in the tube, it is shifted 180°, resulting in the voltage V_2 at the plate. The second path for the signal V_1 is directly to the plate through R_4. This voltage V_3 is in phase with V_1. The voltage V_3 and V_2 add to produce the modulator output.

As order wire input is applied, the bias on the tube varies at the audio rate. As the grid goes positive, there is less voltage drop from grid to plate, and the voltage V_2 increases. At the same time, current through the tube increases, causing an increase in voltage drop across R_4 and a reduction in V_3. The vector sum of the voltages V_2 and V_3 is essentially the same as with no modulation, but the resultant is shifted in phase. As the modulation input causes the grid to go negative, voltage V_2 decreases and voltage V_3 increases. The vector sum of V_2 and V_3 is again essentially the same as with no modulation, but the phase shift is in the opposite direction as that caused by a positive modulating signal. The circuit causes the phase of the r-f output to vary at an audio rate with no significant change in amplitude. Any amplitude variations are reduced by shaping circuits which follow the modulation.

The variable resistor in the cathode circuit (R_5) is used to adjust the bias for operation on the linear portion of the tube characteristic curve. If the bias is not adjusted properly, the resultant phase deviation will not vary linearly with the amplitude of the modulating voltage, and distortion will result. Capacitor C_4 in the plate-tank circuit is adjusted for maximum output.

6.5 MULTIPLEX MODULATORS.

The multiplex modulator must operate with a much wider range of modulation frequency than the order wire modulator. The linearity and distortion requirements are also more stringent for the multiplex modulator. Also, the multiplex modulator superimposes multiplex modulation on the crystal oscillator output which is already modulated by the order wire input. For these reasons, the multiplex modulator is usually more complex. Two common types of multiplex phase modulators, used in tropospheric scatter exciters, are the pulse phase modulator and the vector modulator. These are described in the following paragraphs.

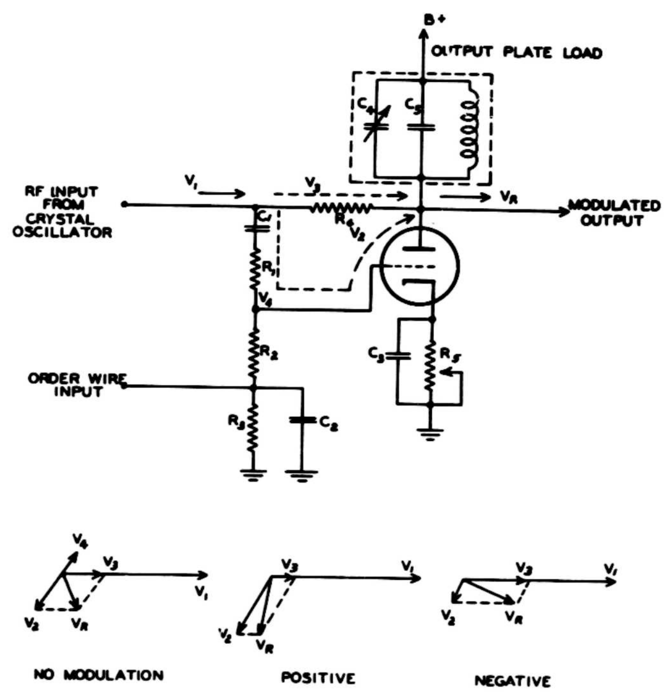


Figure 6-5. Typical Order Wire Modulator Circuit

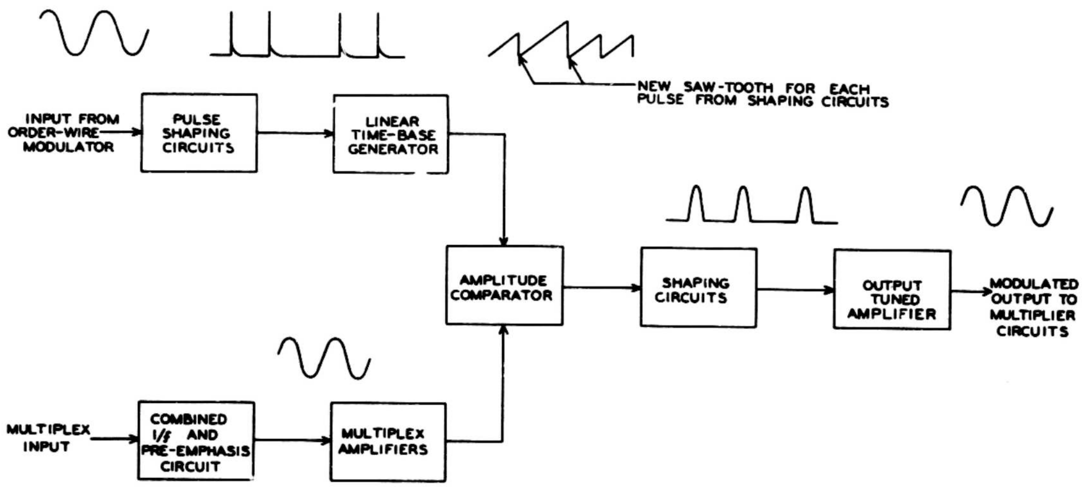


Figure 6-6. Block Diagram of Pulse Phase Modulator

6.5.1 PULSE PHASE-MODULATOR.

Figure 6-6 is a block diagram of a typical pulse phase modulator. Input from the order wire modulator is passed through shaping circuits which produce positive pulses. These positive pulses trigger a linear time base generator which generates a series of very linear saw-tooth waves. The phase of the saw-tooth waves are varied in accordance with the phase deviation in the order wire modulator output. The saw-toothed waves are then directly coupled to the grid of an amplitude comparator which is biased so that the conduction begins about halfway up the sawtooth. A pulse is generated each time the tube conducts. The multiplex modulation input varies the bias so that the conduction point is moved up and down the saw-tooth. Thus, the time position of the pulses is varied to produce phase modulation. The pulses are then used to excite a tank circuit in the output tuned amplifier stage. Since the pulses are varied in time-position, phase modulation of the carrier results when the pulses are caused to ring in the tank circuit.

First, look at a typical circuit to see how the phase of the saw-tooth wave is made to vary with the order wire input. Figure 6-7 shows a typical saw-tooth generator circuit and the waveforms at various points in the circuit. Output from the order wire modulator is passed through clipper and pulse shaper circuits which form a positive pulse for each cycle of the r-f input. The spacing between the pulses is varied in accordance with the modulation information as shown in figure 6-7.

The saw-tooth generator circuit produces one saw-tooth wave for each positive pulse. This saw-tooth wave is the voltage developed across capacitor C_3 as it charges through tube V_3 and resistors R_4 and R_3. When a positive pulse is applied, C_3 discharges through trigger tube V_1. The remainder of the saw-tooth generator circuit provides a constant current source for capacitor C_3.

When a positive pulse is applied to the grid of V_1, the tube conducts heavily. Grid current charges capacitor C_1 to a high negative potential which cuts off V_1. Capacitor C_1 must discharge through R_1. Since the time constant of the discharge circuit is large, V_1 is held at cutoff until the next positive pulse is applied.

When the trigger tube is cut off, capacitor C_3 begins to charge through R_4, R_3, and diode V_3. The saw-tooth waveform, which develops across C_3, is fed to the amplitude comparator. This voltage increases linearly until the next positive pulse is applied to the trigger tube. At this time, capacitor C_3 discharges very rapidly through the tube. This results in a saw-tooth wave which coincides with the positive pulses as shown in figure 6-7. When order wire modulation is applied, the position of the pulses is shifted. This causes a corresponding phase shift in the saw-tooth waves.

The saw-tooth wave developed across C_3 is also applied to the grid of V_2. This causes an increase in voltage developed across cathode resistor R_2 which is fed through C_2 to the tap on potentiometer R_3. Because of the size of C_2, this voltage increases linearly with the saw-tooth voltage. Therefore, the voltage across R_4 and the lower part of potentiometer R_3 is maintained substantially constant. This results in a constant current source in the charge path of capacitor C.

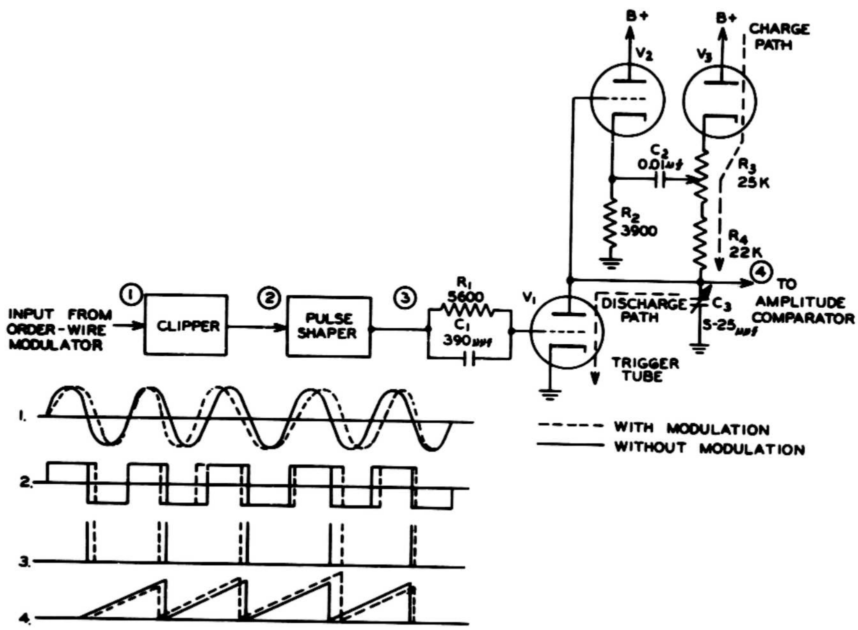


Figure 6-7. Linear Saw-Tooth Generator

The slope of the saw-tooth wave can be varied by adjusting C_3. Linearity of the saw-tooth can be adjusted with potentiometer R_3.

Multiplex modulation information is added to the saw-tooth wave in the amplitude comparator circuit shown in figure 6-8. Input from the saw-tooth generator is applied to the plate of diode V_1, and multiplex modulation input is applied to the cathode along with B+ bias voltage. This voltage biases the tube so that the tube will not conduct until the signal voltage applied exceeds the bias voltage. Figure 6-8 shows the waveforms at various points in the circuit. With no multiplex modulation applied, the tube always conducts at the same position of the saw-tooth waveform. Multiplex modulation applied to the cathode changes the biasing point to vary effectively the tube firing point on the saw-tooth wave. The bias adjustment R_2 is set for a biasing point that is on the most linear part of the saw-tooth waveform.

Each time the amplitude comparator fires, a pulse is generated. The length of time between pulses is varied with the modulating signal. The order wire signal changes the phase of the saw-tooth wave applied to the plate, and the multiplex signal slides the firing point of the amplitude comparator along the saw-tooth wave. Thus, the spacing between the pulses is varied by both the order wire and multiplex modulation inputs.

The pulses are applied to an output amplifier stage. The plate circuit of this stage is tuned to the modulator crystal oscillator frequency. Essentially one sine wave is generated in the tuned circuit for each pulse output from the amplitude comparator; but, since the pulses vary in time due to the modulating information, the signal output is a phase-modulated sinusoidal waveform.

6.5.2 VECTOR MODULATOR.

Another type of circuit used for multiplex modulation is the vector modulator shown in figure 6-9. This circuit is similar in operation to the order wire modulator shown in figure 6-5. Phase modulation is obtained by combining two voltages which are separated in phase by a fixed angle and varied in amplitude in accordance with the modulating input.

The r-f input to be modulated is applied to the cathodes of two amplifier stages. The input to one cathode is shifted a fixed angle from the other input. The plates of the two amplifier stages are connected together to make the modulator output the resultant of the two amplifier outputs.

The gain of the two amplifiers is made to vary proportionately with the modulating input. As the gain of one stage increases, the gain of the other stage decreases the same amount. This is done by passing the modulating input through a phase-splitting circuit which

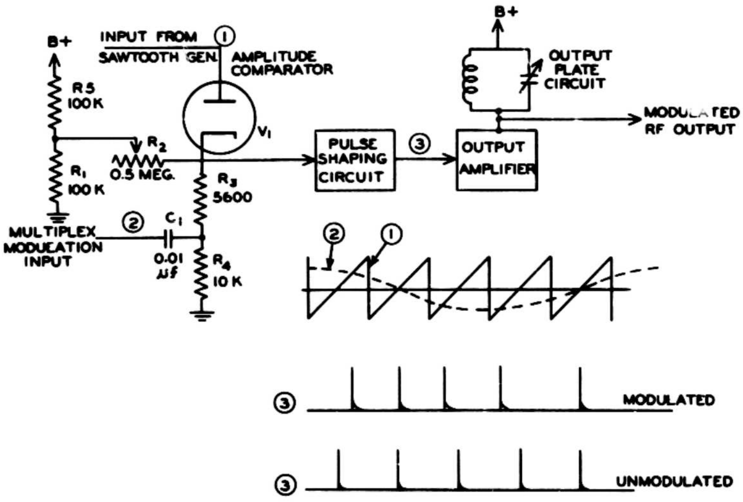


Figure 6-8. Amplitude Comparator and Output Amplifier for Pulse Phase Modulator

provides two inputs, 180° out-of-phase, for the grids of the two amplifiers.

Figure 6-9 shows the vector relationship of the various voltages in the circuit. The output of one amplifier (V_4) is in phase with the r-f input voltage (V_1) and the other output (V_5) is shifted in phase by some fixed angle θ produced in the phase-shifting network. With no modulation input, the gain of the two stages are equal, producing equal outputs. This produces a

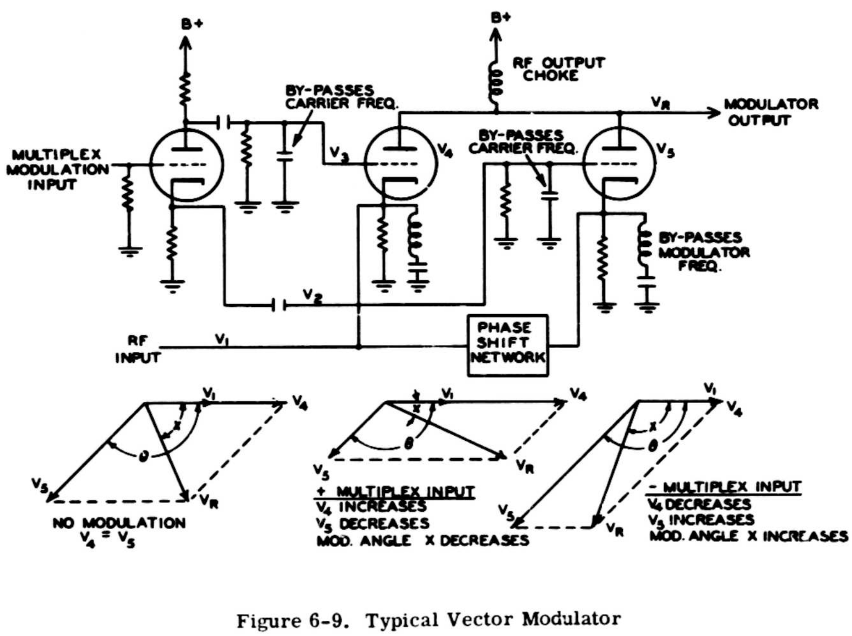


Figure 6-9. Typical Vector Modulator

resultant V_R displaced some angle X from the r-f input. When the modulation input goes positive, the first amplifier gain increases, and the second amplifier gain decreases. This increases V_4 and decreases V_5, thus causing a decrease in the angle X. When the modulation input goes negative, V_4 decreases and V_5 increases. This causes an increase in the angle X. Thus, this circuit produces a phase change which follows the amplitude of the modulation input. By correct design of the phase networks and by selection of proper tube characteristics, it is possible to produce very linear phase modulation with this circuit. Amplitude variations, which are produced in the modulating process, are eliminated by following the modulator with limiter stages.

6.6 OSCILLATORS.

Oscillators are used in the exciter to generate r-f input for the modulator, and to generate an r-f signal which is mixed with the modulator output to produce the required output frequency. In addition to these r-f oscillators, the exciter may also include an audio oscillator for generating a fault-tone signal which is included in the order wire portion of the baseband. Typical oscillator circuits used in the exciter are described in the following paragraphs.

6.6.1 TYPICAL CRYSTAL OSCILLATOR CIRCUITS.

The frequency stability of an oscillator can be made very high by replacing the usual resonant circuit that controls the frequency with a crystal. Normally, exciters used for tropospheric scatter must use crystal oscillators to fulfill the stability requirements.

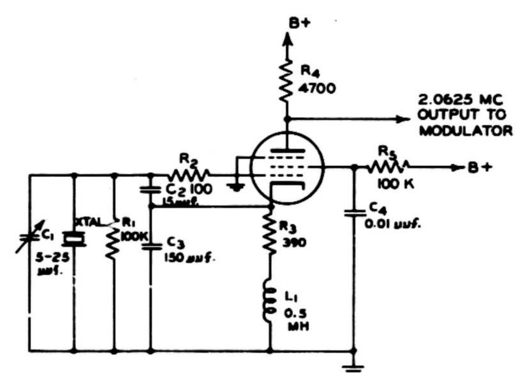


Figure 6-10. Crystal-Controlled Oscillator Used to Provide Modulator Input

Figure 6-10 shows a crystal oscillator circuit used in tropospheric scatter exciters to provide modulator input. This crystal oscillator circuit, called a Pierce oscillator, provides excellent stability, requires no tuned circuits, and may be used over a wide frequency range without change in circuit value. The crystal is operated at its fundamental frequency.

Although a pentode is used, the oscillator functions as a triode oscillator. The screen acts as the oscillator plate with feedback applied from the screen to the grid circuit through interelectrode capacitance. Output is taken from the plate circuit which is shielded from the oscillator circuit with the grounded suppressor. This circuit arrangement isolates the oscillator proper from the load and, therefore, provides for greater oscillator stability.

Variable capacitor C_1, connected across the crystal, is used as a trimmer adjustment to compensate for changes in crystal frequency which may occur as the crystal ages. This capacitor is also used to compensate for manufacturer's tolerance and circuit effects on crystal frequency.

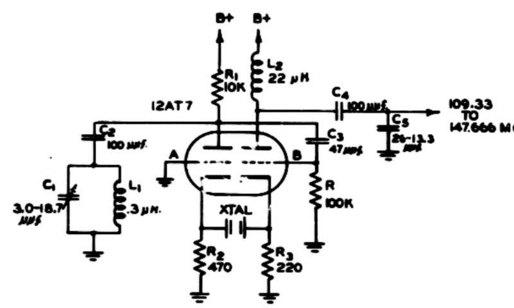


Figure 6-11. Crystal-Controlled Oscillator Used to Provide Injection Signal

Another crystal oscillator circuit, used in tropospheric scatter circuits, is shown in figure 6-11. This circuit, called a Butler oscillator, consists of a grounded grid amplifier, followed by a cathode follower. Feedback voltage is applied from the cathode follower back to the amplifier through a crystal. This circuit is especially suitable for overtone crystals because the crystal current is low. The circuit operates the crystal at series resonance. Normally, the miniature-type twin-triodes are used.

Section A, in figure 6-11, is the grounded grid amplifier stage, and section B is the cathode follower. The plate tank of section A (C_1 and L_1) provides a low impedance to ground for all frequencies except the resonant frequency. It is tuned to this resonant frquency by adjusting C_1.

Output from the grounded grid section (A) is coupled to the cathode follower section (B) through C_3. Voltage developed across cathode resistor R_3 is fed back to the grounded-grid section through a crystal selected to produce the required oscillator fundamental. This feedback voltage causes the circuit to oscillate at the crystal frequency. No trimmer is required for the crystal because fine frequency adjustments may be made with C_1 in the tank circuit.

The circuit, shown in figure 6-11, can be used as an oscillator doubler by tuning the output to the second harmonic of the fundamental frequency. This is done by adjusting C_5. This output tuning has no effect on the oscillator fundamental frequency. The circuit shown is capable of operation over a wide range of frequencies. The desired frequency is obtained by selection of the proper crystal and tuning C_1 and C_5 for the required output.

6.6.2 PHASE-SHIFT OSCILLATORS.

Many oscillators use resistance-capacitance networks for providing positive feedback coupling between the input and output circuits. Such oscillators are called resistance-capacitance, or simply RC oscillators.

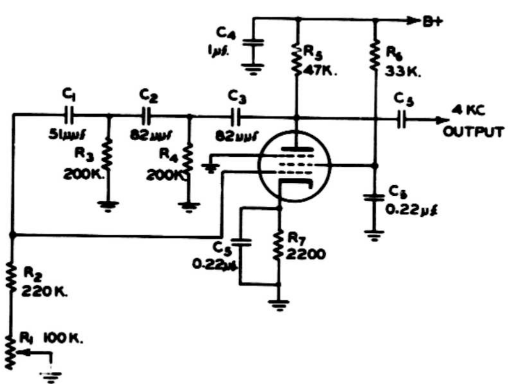


Figure 6-12. Typical Phase-Shift Oscillator

One type of RC oscillator, shown in figure 6-12, is a phase-shift oscillator. This type of oscillator is used to provide a fault-tone signal in the order wire frequency range. It is especially suitable for operation at a fixed audio frequency.

The phase-shift oscillator circuit consists of a single amplifier tube and a phase-shifting feedback circuit. The phase-shift network must provide a 180° phase shift for the signal fed from the plate to the grid. This function is performed by three resistance-capacitance sections. Each section provides a phase shift of approximately 60°. The oscillator can be adjusted to the desired frequency by adjusting R_1 for the required 180° phase shift.

6.7 MULTIPLIER STAGES.

The deviation produced in the phase modulator must be held low to reduce distortion. Therefore, the modulator must be followed by multiplier stages to result in the required output deviation. Multiplier stages offer a means for obtaining a greater amount of frequency deviation without causing distortion in the modulator.

A multiplier (or harmonic generator) is a nonlinear amplifier in which the plate tank circuit is tuned to a harmonic of the input frequency. In a multiplier, the grid is allowed to swing sufficiently positive to allow saturation current to flow through the tube. This produces harmonics in the plate output. The required harmonic is selected by tuning a tank circuit in the plate output. Figure 6-13 shows a typical multiplier. This circuit is a doubler stage, with the output tuned to twice the input frequency.

The power output from a multiplier is less than what would be obtained if the circuit were an ordinary amplifier. Because of this increase in power loss with increase in multiplication factor, the required multiplication is usually accomplished with a series of multipliers rather than in one multiplier stage. To compensate for power loss and to provide sufficient

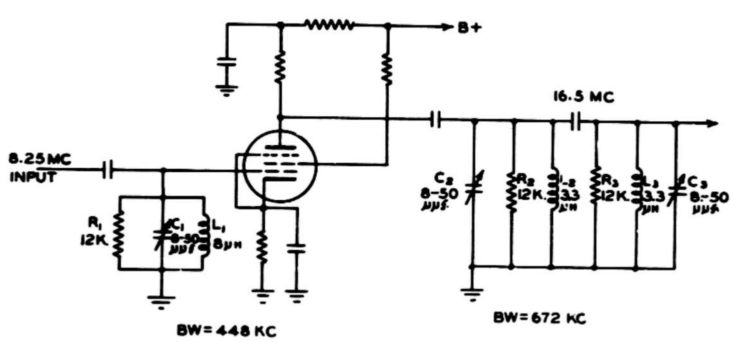
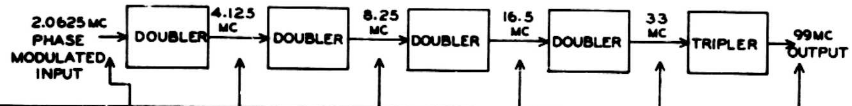


Figure 6-13. Doubler Circuit



DEVIATION	7 KC	14 KC	28 KC	56 KC	112 KC	336 KC
(x) MOD. INDEX	0.0625	0.1250	0.250	0.500	1.000	3.000
(x) MAX. BW REQ.	224 KC	224 KC	224 KC	448 KC	772 KC	1,344 MC

(x) MODULATION INDEX AND MAXIMUM BANDWIDTH REQUIREMENT FIGURED FOR HIGHEST MODULATING FREQUENCY AT 112 KC

Figure 6-14. Typical Multiplier Arrangement for Exciter

grid input for the next stage, multipliers are usually followed by an amplifier. These amplifiers have the added effect of providing the proper bandwidth characteristics.

Figure 6-14 shows a typical multiplier arrangement for an exciter. In this particular arrangement, the total multiplication factor is 48 (2.0625 mc to 99 mc) provided by four doubler stages and a tripler stage.

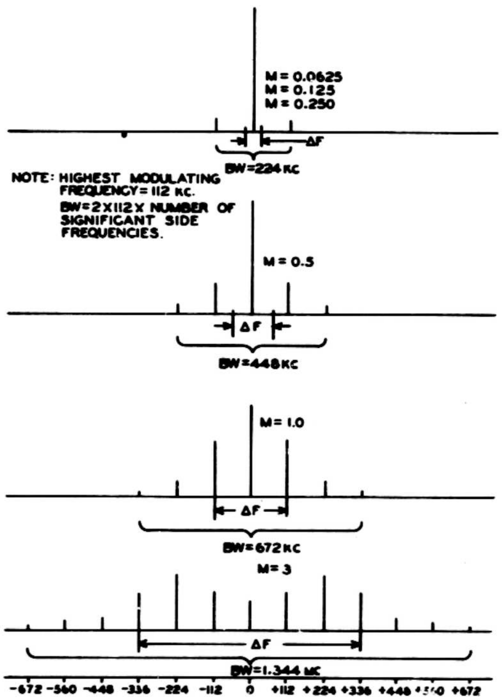


Figure 6-15. Change in Bandwidth Requirements as Modulation Index is Increased in Multiplier Stages

Figure 6-14 also shows that as the carrier frequency is multiplied, the modulation index is multiplied by the same factor. This results in the increased bandwidth requirements shown in the figure.

The modulation index is increased by the same factor as the carrier frequency. Look at the third doubler stage shown in figure 6-13. In this case, the input frequency is 8.25 mc, and the deviation of the input signal is 28 kc. This means that the frequency swings in the input from 8.25 mc - 28 kc to 8.25 mc + 28 kc. Since the output is tuned to twice the input, each of these frequencies will be doubled to produce the following extremes in frequency: 16.5 mc - 56 kc to 16.5 mc + 56 kc. Thus, the deviation has been doubled along with the carrier frequency. However, the modulating frequency or the rate at which the carrier frequency is deviated remains unchanged. Therefore, the modulation index, which is equal to $\frac{\Delta F}{\text{Mod. Freq.}}$, is increased by the same factor as the carrier frequency. This increase is shown in figure 6-14 for each multiplier stage.

An increase in modulation index causes a corresponding increase in bandwidth requirements of the output stage. Table 4-1, in chapter 4, shows the relationship between modulation index and bandwidth. As the modulation index is increased, the number of significant side frequencies is increased. The total bandwidth is equal to the produce of twice the number of significant side frequencies times the highest audio frequency. The multiplier arrangement in figure 6-14 shows the bandwidth requirement for the modulation index in each stage with a highest modulating frequency of 112 kc. If these stages have the required bandwidth for 112 kc, they will operate satisfactorily for all modulating frequencies less than 112 kc. Figure 6-15 shows why the bandwidth must be increased as the modulation index is increased. Insufficient bandwidth or faulty tuning will distort to the spectrum. This can cause incidental FM and other types of distortion in the output signal.

6.8 MIXERS.

After the required deviation is obtained with multipliers, it is then necessary to increase the carrier frequency to the required frequency without changing the deviation. This is done by mixing the modulator multiplier output with an injection frequency produced in a separate crystal oscillator circuit.

Figure 6-16 shows a typical mixer circuit. The tube is a grounded grid amplifier with the two frequencies to be mixed, both applied to the cathode. In order that the desired mixing action take place, it is necessary that the plate current versus voltage characteristic have some nonlinearity or curvature. The components of the plate current will be the d-c, the two input frequencies, the sum of the two input frequencies, the difference between the two input frequencies, and various harmonics and other spurious products. The sum frequency is selected by tuning the output tank circuit. Normally, the difference between the two frequencies is so great that the desired frequency is easily selected, and the unwanted frequencies rejected with the output tuned circuit. The mixer is followed with tuned amplifier stages which provide additional rejection.

A typical mixer frequency scheme is shown on the exciter block diagram (figure 6-1). Output from the injection oscillator section and the output from the modulator multipliers are both applied to the mixer which selects the sum frequency. The desired output frequency is obtained by selection of the correct crystal for the injection oscillator, and then tuning the multiplier stage which follows the oscillator. The frequency of the input from the modulator multipliers is fixed. This arrangement eliminates the necessity of tuning the more critical modulator circuits when the transmit frequency is changed.

6.9 OUTPUT AMPLIFIERS.

The mixer circuit is followed by a series of output amplifiers which amplify the selected mixer output to the output level required for power amplifier drive or for direct application to the antenna. These amplifiers must operate in the uhf output frequency range. Amplifier circuits operating in this range present several problems involving tube construction and external circuits connected to the tube. A discussion of these problems and how they affect amplifier design is included in the following paragraphs.

6.9.1 TUBE CONSTRUCTION PROBLEMS.

The tube construction problems which limit operating frequency of the usual or conventional amplifier are the interelectrode capacitances in the tube, the inductances of the leads, and the transit time.

At low frequencies, the interelectrode capacitances in a vacuum tube form reactances which are large enough not to cause any serious trouble. However, as frequencies increase, the reactance of these capacitances becomes small enough to materially affect the performance of a circuit.

Since interelectrode capacitances are effectively in parallel with the tuned circuit, they affect the frequency at which the tuned circuit resonates. Not only does interelectrode capacitance limit the frequency by establishing a minimum capacity below which it is impossible to go, but it also varies with the applied voltages and with the loading of the tuned circuits. This causes frequency instability, particularly when the interelectrode capacitance forms a large part of the tuning capacitance.

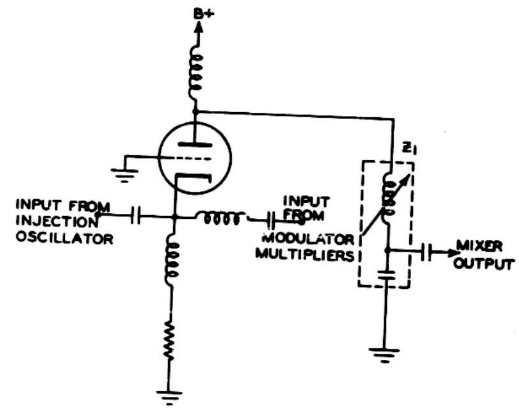


Figure 6-16. Mixer Circuit

Another frequency limiting factor within a tube is the inductance of the leads to each tube element. They may represent a major portion of the inductance of the tuned circuit and limit the frequency by setting a minimum limit on the inductance.

A third limitation imposed by tube construction is transit time, which is the time required for electrons to travel from cathode to plate. Typical transit time of conventional tubes is approximately 0.001 microsecond. At low frequencies, this time is a small part of the period required for one cycle, and therefore has little effect on circuit operation. As the frequency becomes higher, the period decreases to a value comparable to the transit time. For example: at 100 mc, 0.001 microsecond is 1/10 of the period for one cycle; at 1000 mc, the period is equal to the transit time. At frequencies where transit time and the period of the signal are comparable, transit time can seriously affect circuit operation. It can cause loading of the input grid circuit, stretched-out plate current pulses for class C amplifiers, and bombardment of the cathode. All of these effects result in inefficient circuit operation.

The tube limitation problems are overcome by using tubes which are properly designed for operation in the uhf range. In these tubes, the electrodes are closely spaced and constructed in a manner to reduce transit time. These tubes are also constructed with leads having virtually no inductance. Conventional pin connectors are replaced by discs or flanges which are

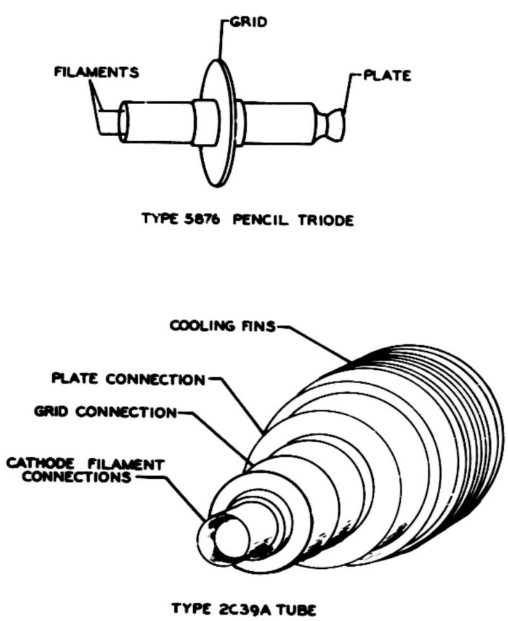


Figure 6-17. Typical Tubes Designed for UHF Operation

sealed to the tube envelope. A typical example of a tube designed for uhf operation is the pencil triode, Type 5876, shown in figure 6-17. Another tube type is the 2C39A shown also in figure 6-17. This tube is expecially designed for operation in coaxial line circuits. Continuous multipoint connections can be made from coaxial line element to the tube elements through sets of flexible fingers. Cooling fins are attached to the plate terminal.

6.9.2 EXTERNAL CIRCUIT PROBLEMS.

Inductances and capacitances in the tuned circuits for uhf amplifiers must be very small. Conventional inductances and capacitors cannot be used. Figure 6-18 shows a special type of tank circuit for uhf circuits called a Hubbard tank. The slotted ring serves as a tunable inductance. It is tuned by a contact spring which is attached to a tuning shaft. As the shaft is turned to a given frequency setting, the contact spring is moved along the slotted ring to select the proper amount of inductance. The inductance is resonated by a very small capacitor which is formed by a Teflon sheet located between the solid output ring and the end plate. These circuits function at low-power levels; however, since the elements are spaced so close together, the power handling capabilities are limited.

Coaxial resonator circuits are normally used for the high-power output amplifiers in the exciters. Figure 6-19 is a cross sectional view of a typical coaxial resonator used with a Type 2C39A tube. The plate, grid, and cathode connections are arranged so that the

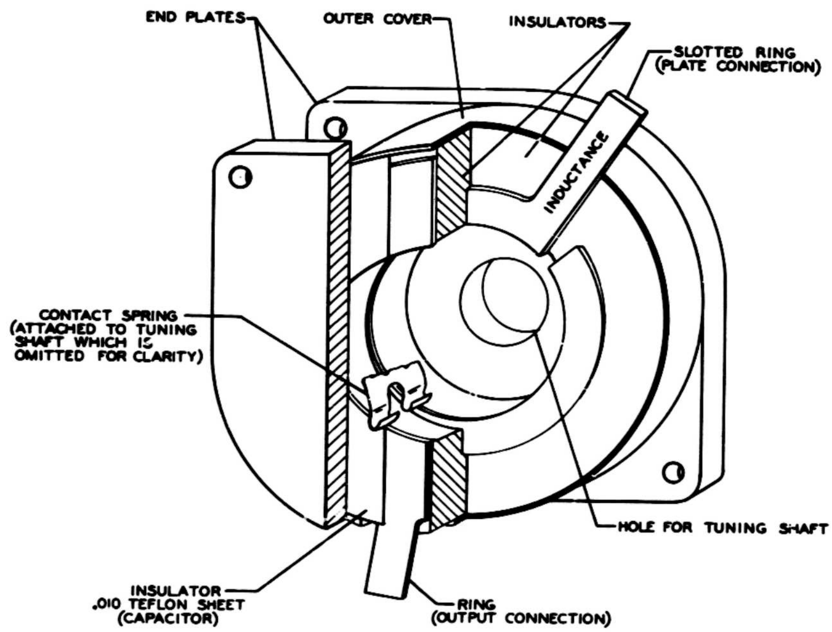


Figure 6-18. A Special Tank Circuit Designed for UHF Application

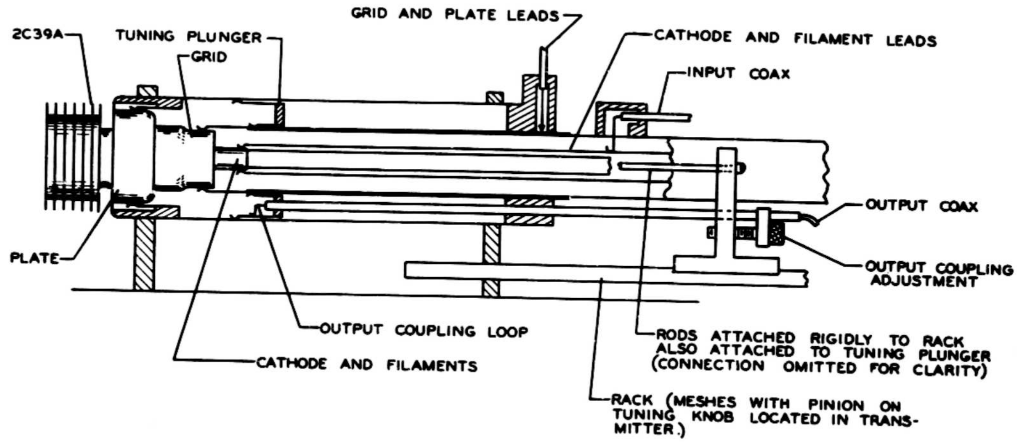


Figure 6-19. Cross Sectional View of Coaxial Resonator

tube may be mounted directly in the tuning assembly. The coaxial resonator consists of three cylindrical conductors. Each cylinder makes contact with one of the tube elements. The cylinders have the required characteristic impedance and line length to resonate at the required frequency. The cylinders are short-circuited by movable plungers to adjust the electrical length. The characteristic impedance is fixed by initial design. The frequency of operation is determined largely by the plunger setting in the output circuit. The tuning of the grid circuit is less sharp. Coaxial resonators must be carefully tuned because change in operation conditions resulting from a shift in frequency may produce damaging tube currents. For this reason, preliminary tuning should be conducted at reduced plate voltages.

Coaxial resonators have several characteristics which make them specially valuable for use with uhf amplifiers. They are self-enclosed and therefore do not radiate energy. Many tubes operating in the uhf range are of circular symmetrical design and hence are suited for use with coaxial resonators. Also, a linear relation exists between plunger displacement and resonant wavelength.

6.9.3 TYPICAL AMPLIFIER CIRCUITS.

Figure 6-20 shows two typical amplifier circuits. The circuit for the low-power amplifier shows an equivalent circuit for the Hubbard tank. The high-power amplifier shows the equivalent circuit for a coaxial resonator.

Grounded grid amplifiers are used. The input is applied between cathode and grid, and the output is taken from the plate and grid. The grid is the common element between the input and output circuits and acts as a shield between plate and cathode. This reduces the plate-cathode capacitance to a low value, thus reducing the feedback that causes oscillation.

6.10 EXCITER ADJUSTMENT PROCEDURES.

The detailed test and adjustment procedures depend on the particular type of exciter. However, the basic procedures are the same for all FM exciters used for tropospheric scatter communications. These basic procedures are described in the following paragraphs.

6.10.1 TUNING AND OUTPUT POWER ADJUSTMENTS.

Before any adjustments are made, the exciter output should be terminated in the proper load. This lowers the Q of the output tank circuits to prevent excessive plate currents in the output tubes. Having the output properly loaded also provides for more accurate tuning. Another precaution is that all preliminary tuning should be done with low-plate voltage applied in the output stage. Figure 6-21 shows the tuning adjustments to be made.

The first step in setting the exciter for operation at a given transmit frequency is selection of the correct crystal for the uhf injection oscillator. The crystal frequency is determined by the multiplier and mixer arrangement which follows the crystal oscillator.

After the correct crystal is inserted, the oscillator and multiplier stages are tuned. Usually, d-c test jacks are associated with each tuning adjustment in the oscillator and multiplier stages. These stages are tuned for maximum indication on a d-c vtvm connected to the test jack. After the oscillator and multiplier stages have been tuned, the tuning should be checked by noting the dial markings on the adjustments. They should be at approximately the transmit frequency.

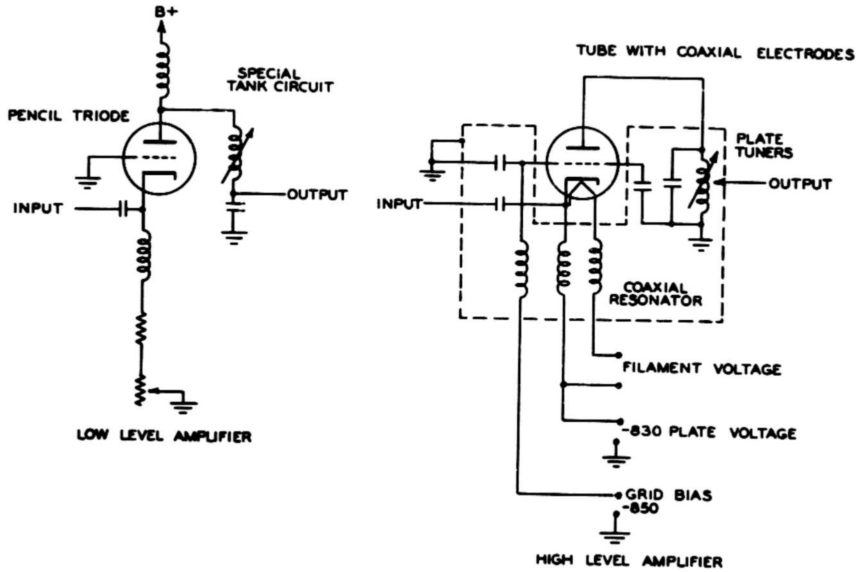


Figure 6-20. Two Typical Output Amplifier Circuits

There is a level adjustment for each of the mixer injections. These adjustments should be made very carefully. If the mixer injection levels are not correct, improper mixer action will result.

Normally, the modulator oscillator and the modulator multipliers require no tuning. The only adjustment required for this injection signal is the level adjustment. This is made by terminating the output in an

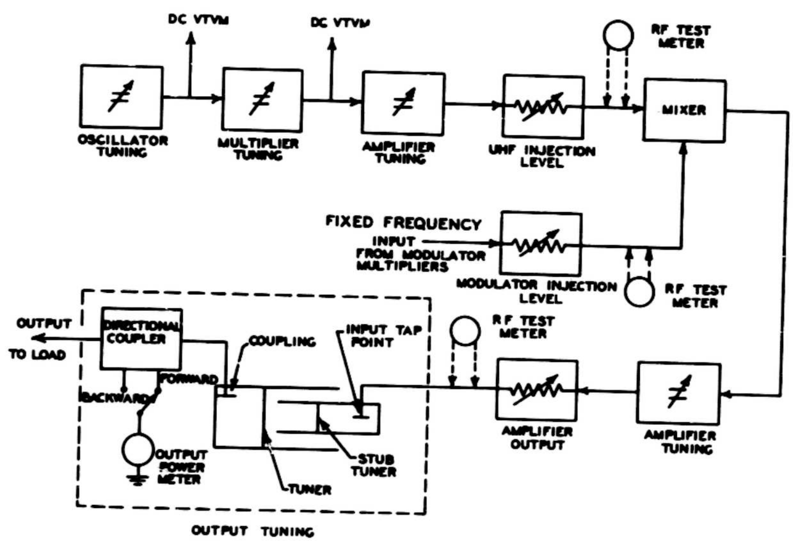


Figure 6-21. Typical Exciter Tuning Adjustment

r-f test meter which provides the correct output impedance, and making the level adjustment for a specified output for proper mixer operation. After this adjustment is made, the injection signal is reconnected to the mixer, and the amplifier tuning and level adjustment procedures for the uhf injection are performed.

The level adjustment for the uhf injection is first set for maximum injection. The final amplifier in the uhf injection circuit is then tuned for maximum indication on the r-f test meter connected in the mixer input. After this tuning is completed, the r-f test meter is connected to the output of the low-level amplifiers in the mixer output. These stages are then tuned for maximum output with the output adjustment set for maximum. After the tuning is completed, the injection level is reduced, and the mixer input is retuned for maximum output. The injection level is reduced before retuning the input because with maximum injection level, the mixer operates at saturation, and the input tuning has little effect on the output. After the input is returned for maximum, the injection level is set for maximum output.

In the exciter shown in block diagram form in figure 6-1, it is possible to tune the amplifiers for operation at the uhf injection frequency instead of at the desired output frequency. This possible faulty tuning can be checked by disconnecting the uhf injection from the mixer. The output indication should then be zero.

After the injection circuits and the low-level amplifiers are adjusted, the final amplifiers should be tuned. The preliminary tuning is done with low plate voltage. The exciter output power amplifier is usually used as the tuning indicator when tuning the output stage. After the stages are tuned, the amplifier output adjustment is then adjusted for the desired output level.

6.10.2 DEVIATION ADJUSTMENTS.

The frequency deviation requirements increase with system channel capacity. As the deviation increases, the bandwidth requirements of the system also increase. The required bandwidth depends on the allowable distortion in the system, but normally for multichannel systems, the relationship between bandwidth and deviation is:

$$BW = 4\Delta F$$

where: BW = system bandwidth

ΔF = peak deviation of exciter

(BW and ΔF are in same units).

Typical deviation and bandwidth requirements for various multiplex channel capacities are given in the following table:

CHANNEL CAPACITY	PEAK DEVIATION	BANDWIDTH
4	65 kc	260 kc
12	150 kc	600 kc
24	336 kc	1.5 mc

Normally, the receivers operating in the system have the minimum bandwidth required for the system channel capacity. If the deviation in the exciter output is allowed to exceed the peak deviation corresponding to the receiver bandwidth, distortion results in the receiver. Therefore, the exciter deviation adjustments must be made very carefully.

Deviation adjustments are provided in both the order wire and multiplex modulator circuits as shown in figure 6-22. These adjustments set the relationship

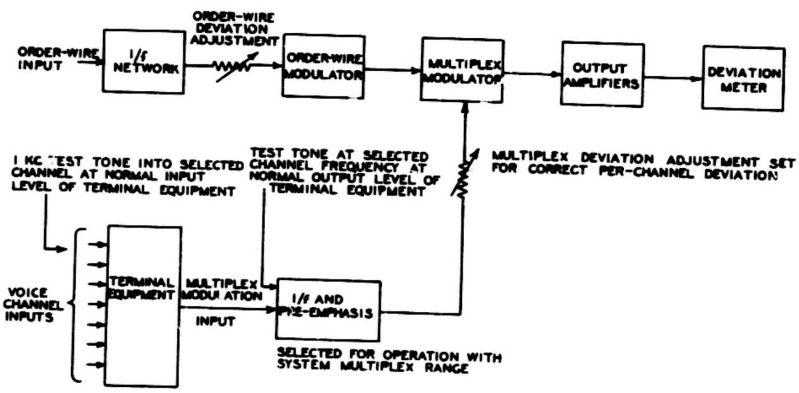


Figure 6-22. Deviation Adjustments

between modulation input level and frequency deviation in the output. The modulation input is set at a predetermined level and frequency. The corresponding adjustment is then set for the correct deviation.

The order wire adjustment is normally set at a point which provides adequate order wire service without causing interference in the multiplex channels. A 1000-cycle tone is applied to the order wire amplifier at a specified level, and then the adjustment is made until the required deviation is obtained. These values depend on the multiplex channel capacity of the system and the multiplex range of frequencies. Typical system values are given in the following table:

MULTIPLEX CHANNEL CAPACITY	MULTIPLEX RANGE	ORDER WIRE INPUT LEVEL	ORDER WIRE DEVIATION (PEAK)
4 channel	4-20 kc	-16 dbm	5 kc
12 channel	4-52 kc	-16 dbm	7 kc
12 channel	12-60 kc	-16 dbm	6 kc
24 channel	4-108 kc	-16 dbm	13 kc

To understand the details of the multiplex deviation adjustment, first look at the nature of the multiplex modulation input. It consists of a particular number of voice channels multiplexed into one modulation input. The input signal level per channel is determined by the characteristics of the multiplex equipment operating with the exciter. The signal level in each channel contributes to the output deviation. The problem in setting the multiplex deviation is to determine the deviation which can be allowed for each channel without exceeding the allowed peak deviation.

If the channel inputs were constant level signals, and if all the channels were used all of the time, the deviation per channel would be simply equal to the peak deviation divided by the number of channels. However, with a number of voice inputs and with a varying number of active channels, there is not a direct relationship between peak deviation and deviation per channel. The problem becomes one of the statistical properties of voice fluctuations and channel activity in a normal communications system. An extensive study of these properties, based on experimental investigations, has been made.[1]

Voice inputs are not at a constant level, but fluctuate about some average level. A series of experiments with various voice sources have been performed to determine the distribution of voice levels. This distribution of a normal voice input shows that the peak level is 18 db above the rms level. Peak is defined as that level which is exceeded less than one percent of the time. The peak-to-rms ratio of the multiplexed input depends on the number of voice channels being combined. The ratio decreases as the number of channels increases, and approaches 13 db as the number of channels becomes very large.

Another statistical property of the multiplex input is the percentage of time that the various channels will be used. Experiments have also been performed to determine this distribution. A communications system must supply the required channel capacity to accommodate the busy hours. Therefore, during a great percentage of the time, all channels will not be in use. In addition, voice will actually be transmitted over the channel only a fraction of the time it is in use.

The statistical properties of the voice fluctuations and channel usage are combined into a voice-loading curve shown in figure 6-23. This curve shows the ratio of the output peak deviation to the deviation which can be allotted per channel without exceeding the peak deviation more than one percent of the time. The ratio of peak deviation to deviation per channel is given in db. This can be done if one remembers that deviation is a voltage function and that when the ratio of the deviations is found, what is actually being found is the ratio of the voltages which causes the deviations.

The voice-loading curve shows that as the number of channels increases, the portion of the total deviation which can be allotted per channel decreases. For example, with 12 channels, the deviation per channel is 16 db below peak deviation. When the channel capacity is increased to 24 channels, the deviation which can be allotted to each channel decreases so that the deviation per channel is 16.8 db below peak deviation.

The relationship between per channel deviation, peak deviation, and the ratio given in figure 6-23 is:

$$R = 20 \log \frac{\Delta F}{X}$$

where: R is the ratio in db, given in figure 6-23,

1. B. O. Holbrook and J. T. Dixon, "Load Rating Theory for Multichannel Amplifiers," Bell System Technical Journal. Vol, 8, October 1939, pp. 624-644.

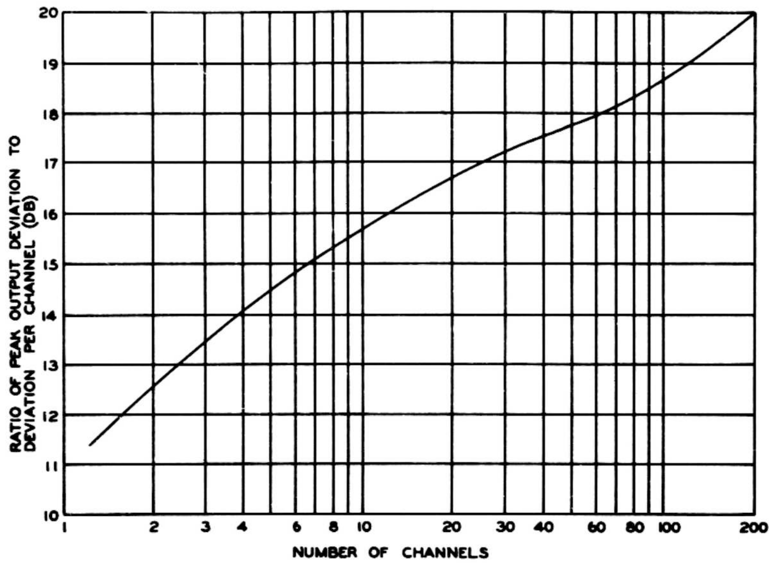


Figure 6-23. Voice Loading Curve for Multichannel Systems

ΔF is the peak deviation, and X is the deviation per channel for the system.

To see how figure 6-23 can be used to determine the multiplex deviation adjustment, use a practical example. Assume that the exciter is set for operation with 12-channel multiplex equipment which has a multiplex frequency range of 12 kc to 60 kc. The first step is to make certain that the correct combined 1/f and pre-emphasis network is used in the the modulator. Then connect a deviation meter in the exciter output and apply a test tone at the required frequency and input level at the multiplex modulation input. This test tone can either be applied at voice frequency to one of the channel inputs of the multiplex equipment, or at a channel frequency to the exciter multiplex input. In either case, the input level at the exciter should be monitored to make certain it is at the required level for operation with the particular multiplex equipment.

The next step is to determine the per-channel deviation for a 12-channel system. The bandwidth of a typical receiver used for a 12-channel system is 600 kc. Therefore, the maximum peak deviation which can be produced in the exciter without resulting in excessive distortion is $\frac{600 \text{ kc}}{4}$, or 150 kc. The per-channel deviation for a 12-channel system is 16 db below the peak deviation as illustrated in figure 6-23. Remembering that deviation is a voltage function, obtain the following equation:

$$20 \log \frac{150 \text{ kc}}{X} = 16 \quad \text{where } X = \text{deviation per channel}$$

$$20 \log (150 - \log X) = 16$$

$$\log 150 - \log X = 0.8$$

$$\log X = 2.17609 - .8 = 1.37609$$

$$X = 23.8 \text{ kc}$$

This means that with a test tone substituted for one of the channel inputs, set the multiplex deviation adjustment for 23.5-kc deviation at the exciter output.

The next step is to determine the channel to be used for making the adjustment. In a normal FM system where the deviation is independent of frequency, the per-channel deviation is the same for all channels. However, in systems where pre-emphasis is used, the deviation increases with modulating frequency. The response for the modulator used with the pre-emphasis circuit for a multiplex range of 12-60 kc is shown in curve 1 of figure 6-3. With this type of response if a high-frequency channel is used and the multiplex deviation adjustment for the calculated per-channel deviation is set, under deviation will result and the full capabilities of the system will not be used. If a low-frequency channel is used, over deviation under normal load conditions will result. This will cause excessive distortion in the receiver. To avoid these undesirable results, a channel which is closest to the mean frequency of the multiplex range is used.

The mean multiplex frequency is calculated from the following equation:

$$F_m = \sqrt{\frac{(F_h)^3 - (F_L)^3}{3(F_h - F_L)}}$$

where: F_m is the mean multiplex frequency, F_h is the upper frequency circuit, and F_L is the lower frequency limit. In the high capacity systems where F_h is much larger than F_L, the equation can be simplified to:

$$F_m = \frac{F_h}{\sqrt{3}}$$

In the system, deviation is adjusted for $F_h = 60$ kc, and F = 12 kc. Therefore, $FM = \frac{60}{\sqrt{3}} = 34.7$ kc.

If the test tone is inserted directly into the exciter, this frequency can be used. If the test tone is to be inserted at the input to the terminal equipment, determine the channel which has an output frequency closest to 34.7 kc. Then, insert a one-kc test tone at the required level into this channel.

To make certain that the procedures for determining the per-channel deviation and mean multiplex frequency are understood, go through the calculations for several other typical systems.

TABLE 6-1. TYPICAL CHARACTERISTICS OF MULTICHANNEL SYSTEMS

SYSTEM CHARACTERISTICS				RATIO OF PEAK DEVIATION TO PER CHANNEL DEVIATION (R) (Fig. 6-23)	PER CHANNEL DEVIATION (X)	MEAN MULTIPLEX FREQUENCY $F_m = \sqrt{\frac{(F_h)^3 - (F_L)^3}{3(F_h - F_L)}}$
CHANNEL CAPACITY	MULTIPLEX RANGE	BANDWIDTH	PEAK DEVIATION ΔF			
4	4-20 kc	260 kc	65 kc	13.8 db	13.3 kc	12.9 kc
12	4-52 kc	600 kc	150 kc	16.0 db	23.8 kc	30.6 kc
12	12-60 kc	600 kc	150 kc	16.0 db	23.8 kc	38.6 kc
24	4-108 kc	1.5 kc	336 kc	16.8 db	48.6 kc	63.5 kc

First, find the ratio (R) of peak deviation (ΔF) to per-channel deviation (X) given in figure 6-23 for the particular channel capacity. Using this ratio, find the per-channel deviation (X) by using the following equation:

$$\log X = \log \Delta F - \frac{R}{20}$$

For example, in the first system given in table 6-1, a ratio of 13.8 db for a 4-channel system is found. The peak deviation is 65 kc. The equation is then:

$$\log X = \log 65 - \frac{13.8}{20}$$

$$\log X = 1.813 - 0.69 = 1.123$$

$$X = 13.3 \text{ kc}$$

The mean multiplex frequency is found by using the equation given above. In the systems where the upper frequency is much higher than the lower frequency, such as is the case with the 24-channel system, the simplified expression, $F_m = \frac{F_h}{\sqrt{3}}$, can be used.

These calculations provide the mean multiplex frequency and deviation information which is required to adjust accurately the multiplex deviation.

Frequencies other than the mean multiplex frequency for a test tone input can be used, provided that the per-channel deviation is adjusted in accordance with the pre-emphasis response of the system. Therefore, if a frequency higher than the mean frequency is to be used, the per-channel deviation must be increased. If a lower frequency is to be used, the per-channel deviation must be decreased.

CHAPTER 7
POWER AMPLIFIERS

7.1 PURPOSE OF POWER AMPLIFIER AND DESIGN REQUIREMENTS FOR TROPOSPHERIC SCATTER.

The exciter in a tropospheric scatter system supplies output at the transmit frequency at a low-power level. This output must be amplified to a much higher level before being propagated over a tropospheric scatter circuit. The required amplification is provided bv the power amplifier associated with the exciter.

A tropospheric scatter terminal is normally equipped with a 1-kilowatt or 10-kilowatt power amplifier which supplies the required output across the transmit frequency range. Because the power amplifier is supplied with low level input, it must produce high gain to obtain the rated output. The output must be produced with high efficiency to reduce power supply requirements. The amplification must also be accomplished without adding noise and distortion to the signal. These stringent requirements, plus the added difficulty of operating in the uhf range at high power, make the use of conventional tubes impractical. Therefore, in most cases, klystron amplifier tubes are used.

Klystron tube operation depends on the transit time of electrons. This is the characteristic which makes conventional tubes impractical at high frequencies. With conventional tubes, the electrodes must be spaced very closely to counteract the effect of transit time. However, high power requires increased spacing between tube elements and larger parts to dissipate the power involved. Since the klystron uses the transit time effect, no compromise has to be made with

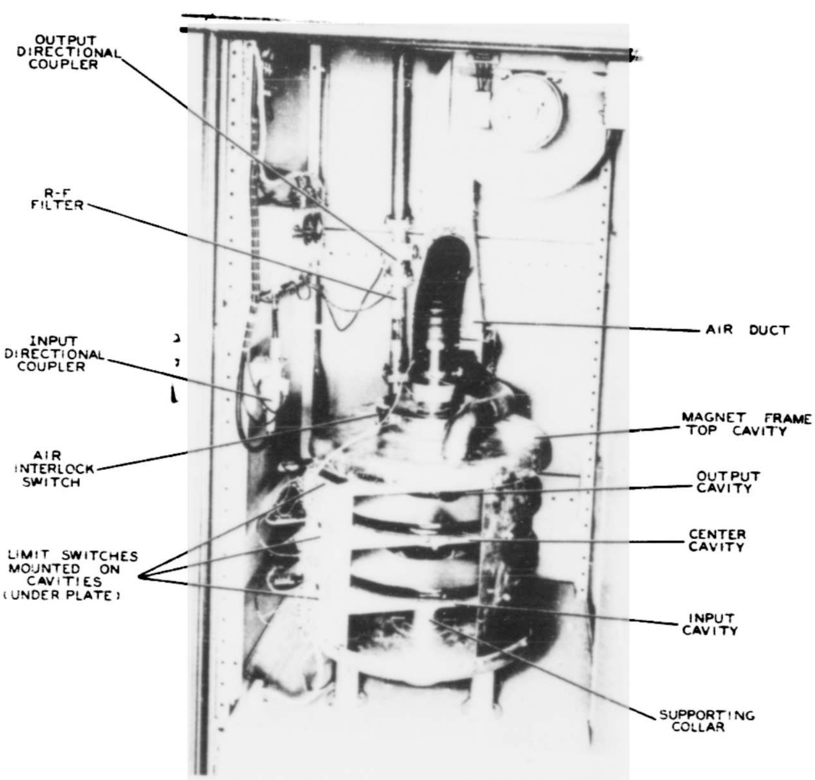


Figure 7-1a. Typical Klystron Amplifier Assembly (1000 mc)

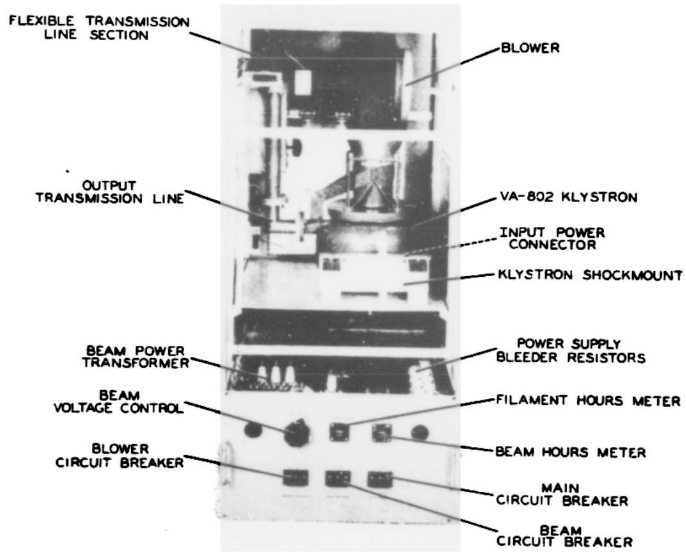


Figure 7-1b. Typical Klystron Amplifier Assembly (2000 mc)

element spacing and size. At power levels of 1 kilowatt or more, the klystron begins to exceed the usefulness of conventional tubes at frequencies about 200 to 300 megacycles. Figure 7-1 shows two typical klystron amplifier assemblies. The following paragraphs describe the major sections of a klystron and klystron operation.

7.2 DESCRIPTION OF KLYSTRON.

The klystron can be divided into three functional systems: electron gun, r-f section, and collector.

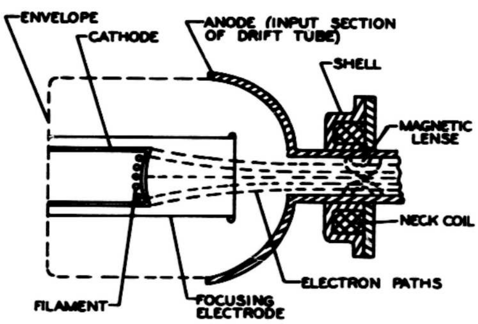


Figure 7-2. Klystron Electron Gun, Simplified Sectional View

Each of these sections is described in the following paragraphs.

7.2.1 ELECTRON GUN.

The electron gun is the source of the electron beam. It consists of a filament, cathode, focusing electrode anode (input section of drift tube), and prefocusing or neck coil. Figure 7-2 is a sectional drawing which shows the general physical arrangement of these components.

Two general types of cathodes are used, depending on the power rating of the tubes. Tubes designed for one kilowatt output normally use oxide-coated, indirectly heated cathodes. Tubes designed for 10 kilowatt output normally use pure metal cathodes which are heated by electron bombardment. Electron bombardment is accomplished by making the cathode approximately 2000 volts more positive than the filament. Electrons emitted from the filament travel at high velocities to the cathode, where they release all their kinetic energy in the form of heat.

The cathode is a concave section located inside one end of the focusing electrode which is a cylindrical metal piece. The focusing electrode is maintained at a negative potential with respect to the cathode. Both the cathode and focusing electrode are maintained at a very high negative potential with respect to the r-f section and collector. The input section of the drift tube functions as the anode. High negative potential applied to the cathode forces the electrons away from the cathode toward the anode, and the negative charge on the focusing electrode causes the electrons to move in toward the axis. As a result of the two forces acting on the electrons, they form a converging beam which focuses inside the first drift-tube section. The neck coil directs and concentrates the beam until it enters the main magnetic field. The neck coil is enclosed in a special magnetic shell containing an annular air gap. The flux outside the air gap forms a magnetic lens located on the axis of the klystron at the approximate point where the convergent paths of the electrons would focus. This magnetic lens overcomes the tendency of the electron beam to diverge and strike the drift tube wall, and it directs the beam down the center of the drift tube.

7.2.2 R-F SECTION.

The r-f section of a klystron is made up of a drift tube, and several resonant cavities and focusing magnets which surround the drift tube at intervals along its length. The drift tube is interrupted by gaps located along the length of the tube. The drift-tube gaps are surrounded by cylindrical, ceramic-envelope sections. A resonant cavity is assembled around each drift-tube gap.

The klystron requires a strong axial magnetic field to maintain and direct the electron beam throughout the length of the drift tube. This magnetic field is established by several magnets which form the magnetic assembly in which the klystron is mounted. Either fixed magnets or electromagnetic coils may be used. When coils are used, the direct currents used to energize the coils are made individually adjustable to permit variation of the field strength along the drift tube.

7.2.3 COLLECTOR.

The electron beam transfers some of its energy to the r-f circuits as it flows through the r-f section of the klystron, and it carries the balance of its energy out of the r-f section into an electrode called the collector. The collector gathers the electrons and passes them to the positive terminal of the beam supply. The large energy content of the partially spent beam is disposed of by the collector. When the electrons collide with the collector surface, all their kinetic energy is transformed into thermal energy which heats the collector. Therefore, forced air and liquid cooling must be provided for the collector.

7.3 KLYSTRON OPERATION.

At the beginning of its passage through the drift tube, the electron beam is a continuous stream of electrons moving at a constant velocity. Although the electrons are not confined to a wire, the beam is a direct current of electricity flowing through the free space enclosed by the drift tube. This direct current must be modulated in some manner so that amplification of the r-f signal can result. This is accomplished by velocity modulating the d-c electron beam as it

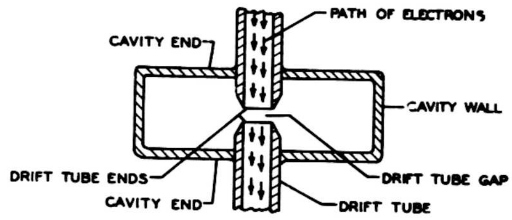


Figure 7-3. Klystron Cavity Assembly, Simplified Sectional View

passes through each gap in the drift tube. The cavities are arranged so that the ends of the drift-tube sections protrude into the cavity at opposing high-voltage points in the cavity wall (see figure 7-3). Thus, the drift-tube tips become the capacitive loading elements in the cavity, and large r-f voltages are induced across them when the cavities are excited. Electrons are deaccelerated or accelerated, depending on the field polarity, as they pass through the gap. The faster electrons begin to overtake the slower electrons as the beam travels down the axis of the tube. This causes the density of the beam to vary periodically at signal frequency.

To explain further how velocity modulation occurs, refer to figure 7-4 which is a simplified representation of klystron operation. The entire klystron can be thought of as a concentrated or focused electron beam surrounded at intervals by tuned circuits which represent the cavities in an actual assembly. The emitter is made highly negative with respect to the collector to cause electrons to flow from emitter to collector. With no r-f energy applied to the circuit, the average velocity and density of the electron beam is constant along the drift tube. However, when r-f drive power is applied to the input tuned circuit, the velocity of the electrons is changed in accordance with the r-f electric field applied to the beam. When the r-f voltage is positive, the electric field is in a direction to cause acceleration of the electrons. Conversely, when the r-f voltage is negative, the electric field causes deacceleration of the electrons.

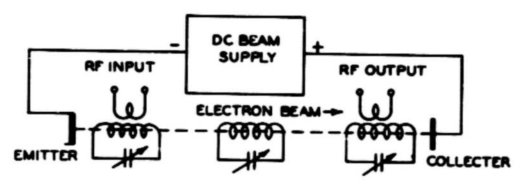


Figure 7-4. Representation of Klystron Operation

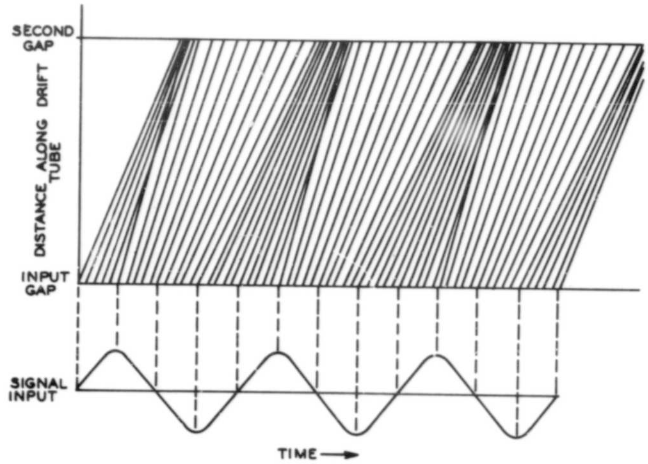


Figure 7-5. Mechanics of Velocity Modulation

The velocity modulation process is shown in figure 7-5. The vertical scale on this figure represents distance along the drift tube, and the horizontal scale represents time in relation to the period of the r-f input. The slope of each line $\left(\frac{\text{Distance}}{\text{Time}}\right)$ represents the velocity of a given electron. Electrons that pass through the input gap when the r-f input voltage is negative are slowed down and leave the resonator with lower than average velocity. The electrons leaving a half period later have higher than average velocity and so catch up in the drift space with the slower ones that left earlier. Similarly, they draw away from the slower electrons which follow. The differences in velocity cause the electrons to bunch as they travel down the tube between emitter and collector. This causes the density of the beam to vary, with maximum points occurring at certain distances from the first gap. After bunching has been achieved, the electron beam represents a current which is varying in amplitude at the input frequency.

Tuned circuits placed at optimum points along the drift tube increase the bunching effect. The variations in beam density start oscillations in the tuned circuit in the same way that plate current variations cause oscillations in the output circuit of a tuned amplifier. The tuned circuits are adjusted to provide the correct phase relationship between oscillations occurring in the tuned circuit and the variations in beam density. They must be adjusted so that the beam acts as a load, and energy is delivered to the beam to increase bunching. The cavity preceding the output cavity is normally tuned above resonance to provide an inductive circuit. This causes the current in the tuned circuit to lag the correct amount to provide proper phasing with the variations in beam density. Each tuned circuit along the drift tube increases the bunching so that well-defined bunches arrive at the output cavity.

Energy is extracted from the bunched beam by the same mechanism used to velocity modulate it. As the beam travels through the output gap, the gap polarity varies in such a way that the denser portions of the beam are deaccelerated, while the less dense portions are accelerated. As a result, there are many more electrons being made to give up energy to the circuit than there are electrons which take energy. This causes a transfer of power from the electron beam into the klystron output circuit. With proper adjustments, the amount of power required to produce the bunching is relatively small compared to the amount of energy delivered by the electron beam to the output tuned circuit. As a result, the klystron tube is a power amplifier. Energy delivered to the output cavity is coupled to the antenna circuit.

There are two principal ways in which energy can be inserted into and removed from a cavity resonator The first is by inserting a probe into the cavity. The current which flows in the probe sets up an electric field parallel to it. This field starts oscillation in the cavity. Another method uses a magnetic loop which is placed in the region where the magnetic field will be located. In the output cavity, it is very important that the device used for coupling be set for maximum coupling. If the device is set improperly, the output cavity will be unloaded. This causes an excessive build-up of energy within the output cavity, resulting in arcing and other cavity breakdown.

7.4 POWER SUPPLIES AND METERING CIRCUITS.

Figure 7-6 is a simplified schematic which shows how power amplifier voltages are applied and metered.

The power supply circuits include a filament supply and a high-voltage beam supply. A current supply for the focusing magnets is also required in assemblies where permanent magnets are not used. The positive side of the beam supply is connected to the collector which is operated at approximately ground potential. The negative side is connected to the klystron cathode. Voltage applied across the klystron amplifier is measured by a beam voltage meter as shown in figure 7-6.

Bias voltage for the klystron focusing electrode is developed across a potentiometer in the cathode circuit. This circuit arrangement makes the focusing electrode more negative than the cathode to result in focusing of the electron stream.

Since the electron-gun section of the tube (filament, cathode, and focusing electrode) is operated at very high-negative potential, it is located in an interlocked compartment to protect personnel during maintenance and adjustment. All high voltage is removed by action of interlock circuits, when the front panel to this compartment is removed. The remainder of the tube, containing the cavities and tuning adjustments, is operated at close-to-ground potential.

A separate bombarder supply is required for tubes using electron bombardment to heat the cathode. This supply is connected between the filament and cathode to place the cathode at a positive potential with respect to the filament.

Some of the electrons in the klystron beam strike the drift tube walls instead of passing on through the klystron to the collector. These are captured by the wall of the drift tube and returned to the beam supply through the body current meter. The electron beam which reaches the collector is measured by the collector current meter connected between the collector and the positive side of the beam supply. The sum of the collector current and the body current is equal to the total beam current emitted from the cathode. The collector current, expressed as a percentage of the total cathode current, is called beam transmission. The body current, expressed as a percentage of the total cathode current, is called beam loss, since electrons lost to the wall of the drift tube are lost as far as r-f power is concerned. When the klystron beam is first energized, the magnet and focusing controls are adjusted for the lowest possible body current with no drive applied.

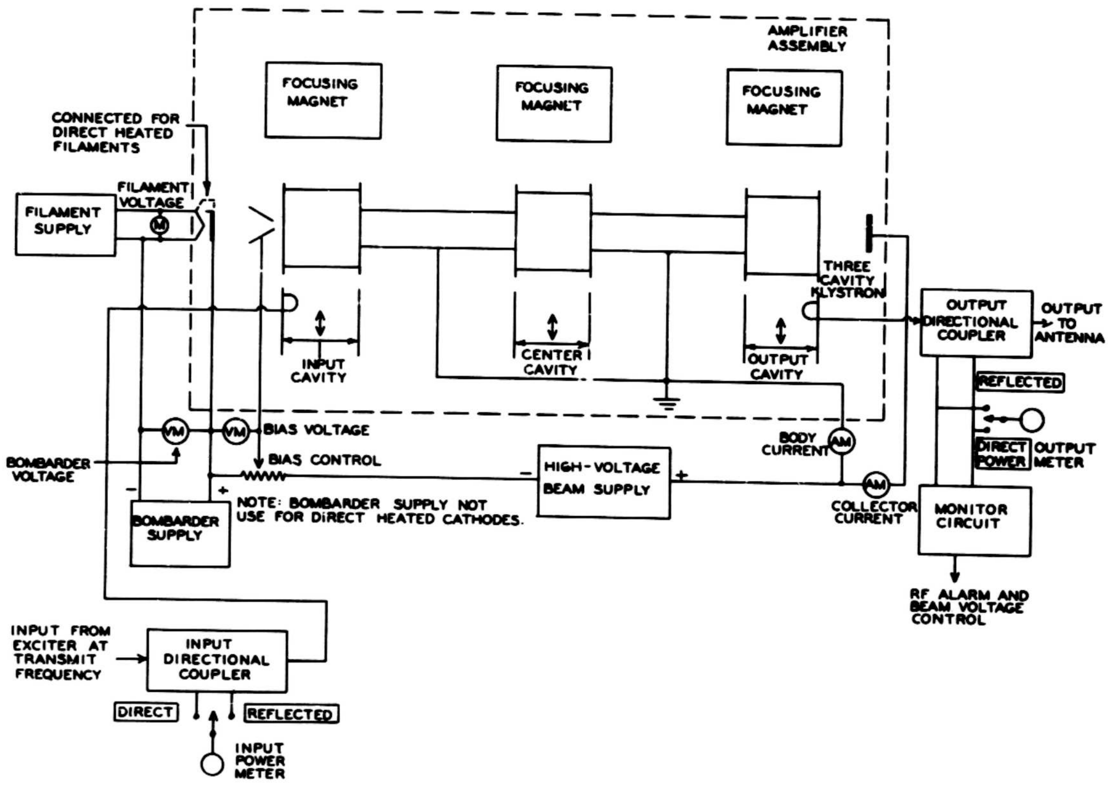


Figure 7-6. Power Supplies and Metering Circuits

Calibrated directional couplers are employed as monitoring devices on the input and output coaxial lines. A selector switch and meter associated with each coupler allow direct and reflected power to be read from both transmission lines. Output from the output directional coupler can also be connected to a monitor circuit which can be used to actuate alarm circuits if the direct power becomes too low or if the reflected power on the output line becomes too high. This circuit can also be used to cause removal of the beam voltage if the reflected power reaches a predetermined value.

7.5 CONTROL CIRCUITS.

Power amplifier control circuits must perform three functions: (1) They must supply circuit control; (2) They must provide personnel protection; and (3) They must provide equipment protection. The control circuits are sequenced and interlocked so that the filament voltage must be applied, and the cooling system must be functioning before the beam voltage can be applied. Interlock circuits break the high-voltage circuit when the access panel for the high-voltage compartment is removed.

Figure 7-7 is a simplified schematic of a typical control circuit arrangement. The control circuits are connected across 120 volts. Power application is controlled by "filament-on" and "beam-on" pushbuttons on a main control panel. Either on-button can be operated to start the power-on sequence. The filament control relay is energized when either on-button is operated. When the filament-on button is operated, the filament control relay is energized and held energized by holding contacts in parallel with the pushbutton. If the beam-on control is operated and all interlock circuits are closed, the beam-holding relay is energized. The filament control relay is then energized through contacts on the beam-holding relay which parallel the filament-on pushbutton.

After the filament control relay is energized, voltage is applied through the blower contactor to the blowers used for cabinet and collector cooling. When the cabinet blower reaches sufficient speed, the air interlock switch closes, causing the filament contactor to close. Voltage is then applied to all filament circuits.

After the klystron filament voltage has reached operating value, the filament voltage interlock relay operates. This starts the beam time delay. After a preset time delay has elapsed, the contacts in the beam contactor circuit close, readying this circuit for operation. If the beam-on pushbutton was operated to start the power-on sequences, the beam contactor circuit will be completed, and at the end of the delay period, the beam voltage will be automatically applied. If the power-on sequence was started with the filament-on pushbutton, beam voltage will be applied only after the beam-on pushbutton is operated.

Operation of the beam-off control removes the beam voltage. Operation of the filament-off control removes all tube voltages. However, the blowers

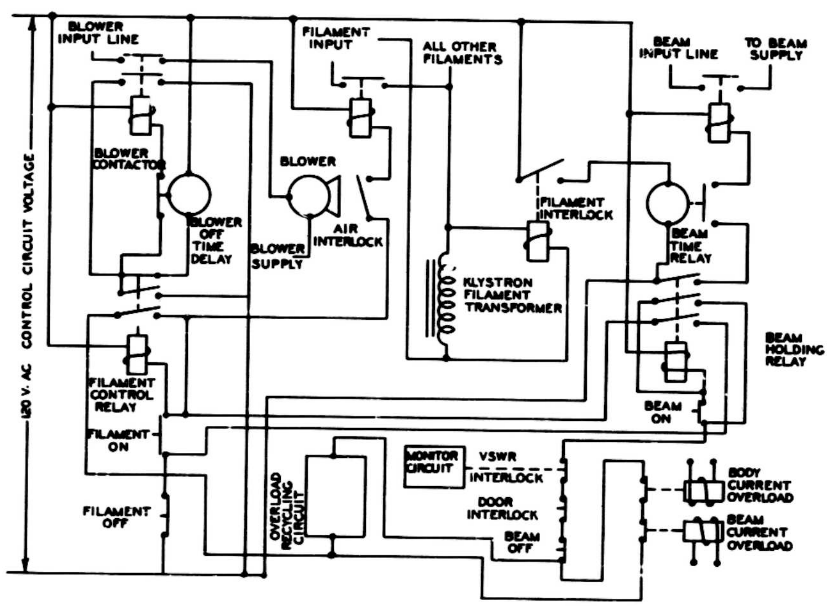


Figure 7-7. Typical Power Amplifier Control Circuit

continue to operate for a period of time controlled by the setting of the blower-off time delay circuit. This delay circuit allows the blower to operate for cooling the tube after the voltages have been removed. When the filament-off control is operated, the filament control relay is de-energized. This causes the time-delay relay to be energized. The blower contactor remains energized until its operating circuit is opened by action of the time-delay relay after the preset time period has elapsed.

The beam-on circuit is completed through normally closed contacts on body current and beam current overload relays. These relays serve as tube protection to remove the beam voltage in the event of damaging current levels. An overload recycling circuit automatically reapplies the beam voltage if a momentary overload occurs because of arcing or some other nonpersistent cause. If the overload persists through the automatic recycling period, the beam-on control must be operated to reapply beam voltage.

7.6 TESTS AND ADJUSTMENTS.

7.6.1 GENERAL TUNING PROCEDURES.

The detailed step-by-step tuning procedures will depend on the particular klystron installation. However, the general tuning procedures and order of performance are the same for all klystrons:

a. The initial tuning steps are performed with no beam voltage and r-f drive power applied to prevent damage to the tube and tuning cavities. A beam voltage control is usually provided on the main control panel. This should be initially set to the lowest setting to avoid accidental application of high beam voltage.

b. The filament voltage is applied and adjusted for the correct operating voltage. The filament voltage will normally rise slowly and lag the adjustment setting because of high starting current. Therefore, the filament voltage adjustment should be done slowly and carefully to avoid application of excessive voltage.

c. If the electromagnetic focusing is used, the magnet currents should be set to the specified values before any beam voltage is applied. This prevents excessive body current resulting from insufficient focusing. The cavities should also be set to the approximate operating frequency.

d. After these preliminary adjustments are made, the low-beam voltage is applied. The focusing magnets are then mechanically adjusted for minimum body current with no drive power applied. This adjustment obtains the best possible d-c power transmission through the tube.

e. The r-f drive power at the transmit frequency is then applied. The level should be set to the value specified for the particular klystron.

f. The input cavity is then tuned for best impedance match to the input line. This is usually done by tuning the cavity for minimum reflected power indication on the input power meter. The particular tuning instructions given for each cavity should be followed carefully to avoid tube damage and to provide for correct operation. Body current should be observed while tuning, and readjustment of the focusing magnets may be required to provide minimum body current.

g. The cavity preceding the output cavity must always be tuned to a frequency above resonance. It should be remembered that there is no load connected to this cavity to absorb energy. To protect the cavity against damage resulting from large resonant voltages and currents, it must be tuned carefully so that the electron beam acts as the load.

h. The output cavity is tuned to resonance at the transmit frequency and must be adequately loaded at all times. This is done by placing the output coupling loop in the maximum coupling position before any attempt is made to resonate the output cavity.

i. After the cavity has been tuned to resonance, the coupling is adjusted for maximum output, and then the cavity is retuned. The output coupling and output cavity tuning interact because of the change in cavity loading caused by the coupling adjustment.

j. After retuning, the body current should be adjusted, if necessary, by resetting the magnet current.

k. After the cavities have been tuned, the beam voltage is increased in steps until the required output power is obtained. The cavities and magnets are readjusted for maximum power output for each increase in beam voltage. The body current changes with an increase in beam voltage due to the change in electron velocity and bunching effects.

7.6.2 DETERMINATION OF POWER AMPLIFIER EFFICIENCY.

When adjusting the power amplifier, we are interested in obtaining the desired output with the highest obtainable efficiency figure. Power amplifier efficiency is equal to $\frac{\text{r-f power output}}{\text{d-c power input}}$. The d-c input power is equal to the product of the voltage applied across the tube (beam voltage) and the d-c beam current flowing in the tube. The beam current is the sum of the body current and the collector current. Therefore, the equation for efficiency is:

$$\text{Efficiency} = \frac{\text{Power Output}}{\text{Beam Voltage (collector current + body current)}} \times 100.$$

To see how this equation is used, substitute some typical values for one-kilowatt output. Assume that the meter indications are:

Collector current: 330 ma (0.330 amp)
Body current: 30 ma (0.03 amp)
Beam voltage: 7000 volts
Power output: 1 kw

$$\text{Efficiency} = \frac{1000}{7000\,(0.330 + 0.03)} \times 100 =$$

$$\frac{1000}{2520} \times 100 = 39.7\%$$

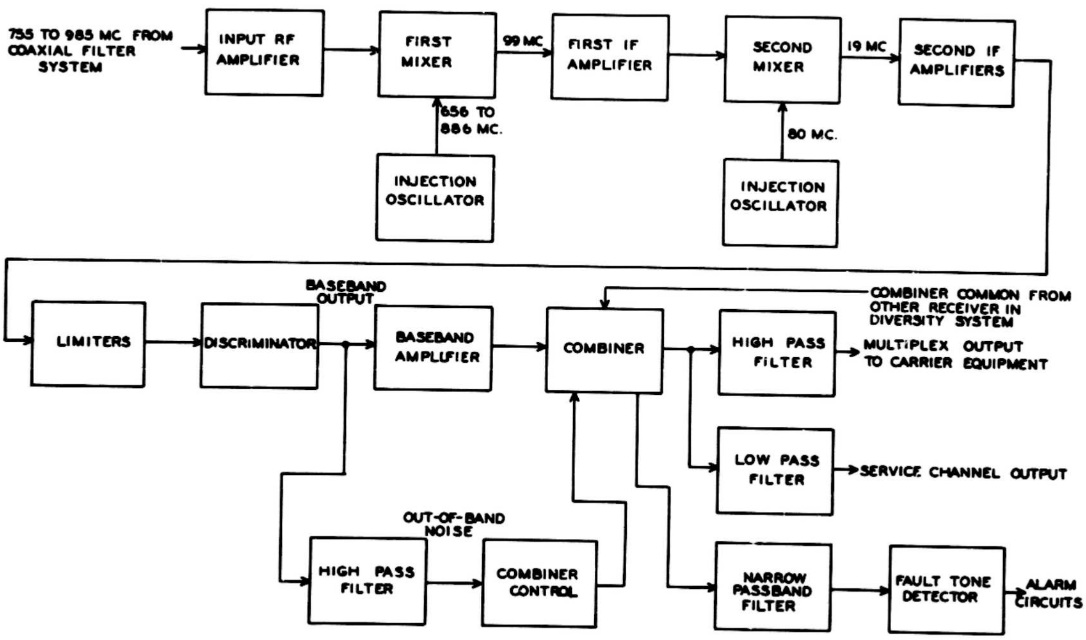


Figure 8-1. Block Diagram of Tropospheric Scatter FM Receivers

CHAPTER 8
FM RECEIVERS

Chapters 4 and 5 described the equipment used to produce an r-f signal, frequency modulated in accordance with the information-carrying baseband input. This chapter describes the equipment which receives the frequency-modulated signal at a distant terminal and detects the transmitted baseband signals. The operation of each of the major components of an FM receiver is described. This information is essential for understanding the basic adjustment and maintenance procedures which are required for an FM receiver.

8.1 DESIGN REQUIREMENTS OF AN FM RECEIVER USED FOR TROPOSPHERIC SCATTER.

The receiver used for tropospheric scatter must be capable of detecting very low-level signals with maximum reliability and low distortion. Sensitivity of the receiver is defined as the signal level required for a 10-db signal-to-noise ratio in the i-f output. The input signal level required to produce the 10-db signal-to-noise ratio is called threshold level. This definition of sensitivity has a practical basis because it recognizes that ultimately the noise level is the limiting factor and that a signal 10 db stronger than the noise level is required for reliable communications.

The expression for threshold signal is:

$$S_T = -144 + 10 \log BW + F + 10 \text{ db}$$

$$= -134 + 10 \log BW + F$$

where: S_T = threshold signal in dbm

BW = receiver bandwidth in kilocycles

F = receiver noise figure in db.

This expression for signal threshold level shows that receiver sensitivity is a function of receiver bandwidth, and the receiver noise figure which is determined by internal noise generated in the receiver. Therefore, for maximum sensitivity, the bandwidth must be kept as narrow as possible consistent with modulation requirements, and the receiver input circuits must be designed for low noise operation.

Another receiver requirement is stable oscillator operation. The frequency drift of the local oscillator used must be kept within very narrow limits. If the frequency is allowed to drift, the bandwidth must be made larger to allow for this drift. This causes a corresponding increase in noise with a reduction in receiver sensitivity.

The receiver system must also be capable of being tuned to a particular frequency in the operating range, accepting the desired band of frequencies and rejecting all other frequencies. This requires filters and tuned circuits for shaping the receiver pass band and rejecting spurious signals.

The receiver output should be an exact representation of the information transmitted from the opposite terminal. This requirement involves a system capable of detecting the baseband modulation, and amplifiers which can raise the detected baseband to the required output without significant distortion.

The reliability of reception is improved by operating two or more receivers in a diversity system. Each receiver must include some type of combining circuit so that it can operate with the other receivers in the system.

8.2 GENERAL OPERATION OF TYPICAL FM RECEIVER.

Figure 8-1 is a block diagram of a typical FM receiver used for tropospheric scatter communications. The frequency scheme is shown for an output frequency in the range of 755 to 985 mc. However, the general operation is the same for receivers operating in the other frequency ranges.

The r-f input is applied to an input amplifier circuit which is especially designed for low noise operation. For tropospheric scatter applications, the r-f amplifier must be capable of easy tuning over a wide frequency range. It must provide adequate selectivity over the entire range. The bandwidth must be at least as great as the i-f bandwidth.

The r-f input is converted to fixed intermediate frequencies for further amplification. The conversion to lower frequencies allows the use of conventional amplifier circuits. The i-f amplifiers are made to provide satisfactory uniform response over the required bandwidth. Since the i-f stages are operated at a fixed frequency, the selectivity will not vary over the receiver tuning range. A double-conversion system is used. The first i-f is made high enough for increased image rejection, and then this i-f is converted to a lower frequency for greater selectivity.

The r-f signal and an injection signal from the first injection oscillator are applied to the first mixer. The difference frequency in the mixer output is selected for amplification in the first i-f stages.

The selected mixer output contains a signal component for each component in the incoming signal. Amplitude, frequency, and phase relations between these components are preserved as the signal passes through the mixer. The signal entering the i-f amplifier therefore contains the same modulation as the r-f signal, but is centered at the i-f. Only those components lying within the pass band of the i-f amplifiers will continue through the receiver.

The oscillator injection frequency is kept sufficiently low so that the image frequency will be rejected in the tuned i-f stages For example, suppose that the r-f input frequency is 755 mc. The oscillator frequency will be set to 656 mc, and the desired i-f is the difference frequency which is 99 mc. The image frequency is the r-f frequency which is 99 mc lower than the oscillator frequency, or 557 mc. The image frequency is separated from the desired frequency by an amount equal to twice the i-f. Therefore, the higher the i-f, the greater will be the image rejection. The i-f used in the receiver represents a compromise between image rejection and other circuit requirements.

With the frequency conversion scheme shown in figure 8-1, receiver tuning is made quite simple. All that is necessary to select a particular input frequency is to tune the input r-f amplifier and the first injection oscillator. The second injection frequency is fixed. This frequency is mixed with the first i-f, and the difference frequency is selected for amplification in the second i-f stages.

Output from the second i-f amplifiers is applied through limiter stages which remove amplitude variations from the frequency modulated signal. In most cases, two limiter stages are used to provide thorough limiter action. Output from the limiters is applied to an FM discriminator stage which demodulates the r-f and produces the original transmitted baseband signal. This baseband signal is applied through amplifier stages to a combiner circuit. Output from the combiner is connected in parallel with output from the combiner in the other receiver in a diversity reception system. This produces a composite signal free of switching transients and with a signal-to-noise ratio improvement of up to 3 db over the individual receiver.

The various component signals in the baseband output are separated by filter circuits. The multiplex output is obtained from a high-pass filter which eliminates all frequencies below the low end of the multiplex band. The service-channel output is obtained from a low-pass filter which eliminates all frequencies above the upper limit of the service channel.

Receiver alarm circuits are operated by a fault tone transmitted from the opposite terminal. This tone is passed from the combiner through a narrow-band filter to a fault-tone detector which operates alarm circuit relays. An alarm is indicated when the fault tone is not present. Provisions are usually made for both local and remote alarm indications.

A squelch circuit is sometimes included to disconnect the receiver from the diversity system when the receiver noise exceeds some predetermined level. This noise level corresponds to the minimum allowable signal-to-noise ratio in the receiver output. The squelch circuits are operated by noise present in the i-f amplifier pass band above the baseband signal range.

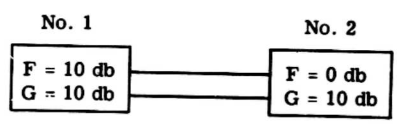
The preceding paragraphs have provided a general discussion of receiver operation. Detailed discussions of the various circuits involved in an FM receiver for tropospheric scatter are given in the following paragraphs.

8.3 R-F SECTION.

8.3.1 GENERAL NOISE FIGURE CONSIDERATIONS.

The receiver noise figure is determined almost entirely by noise introduced by circuits in the r-f section of the receiver. For this reason, it is very important that the r-f section be designed and components selected for low noise. For some receivers operating in the 1000-mc range, an r-f amplifier is included ahead of the first mixer as shown in figure 8-1. Noise introduced by amplifiers now being used increases rapidly as the signal frequency is increased above 1000 mc. Therefore, the r-f input is normally applied directly to the first mixer without amplification in receivers operating above 1000 mc.

The following analysis shows why the receiver composite noise figure is determined chiefly by the first stage, and also why it is advantageous to place a low-noise amplifier ahead of a comparatively noisy mixer. Assume that there are two circuits: one with a noise figure (F) of 0 db, and the other with a noise figure of 10 db. Both have a power gain of 10 db. First, place the noisy circuit ahead of the quiet one and check on composite noise figure.



The formula for the composite noise figure of two circuits is:

$$F_{1+2} = F_1 + \frac{F_2 - 1}{G_1}$$

where: the noise figures are given in actual power ratio instead of in db. With the above circuit arrangement, F_1 equals 10 db or a power ratio of 10, and F_2 equals 0 db or a power ratio of 1. The composite noise figure is then:

$$F_{1+2} = 10 + \frac{1-1}{10} = 10 \text{ or } 10 \text{ db}$$

Reverse the circuits and place the quiet one first and the noisy one second:

F = 0 db	F = 10 db
G = 10 db	G = 10 db

$$F_{1+2} = 1 + \frac{10 - 1}{10} = 1.9 \text{ or } 2.8 \text{ db}$$

By comparing the two arrangements, it can be seen that a noise figure of 2.8 db is much better than one of 10 db, and that it is much more advantageous to place noise-free gain ahead of noisy gain. This shows that when a noiseless r-f amplifier is placed ahead of a mixer, the composite noise figure is reduced. However, no noise reduction is obtained unless the r-f amplifier has a noise figure appreciably less than the mixer.

8.3.2 RF AMPLIFIER.

The tube selected for use in the r-f amplifier circuit must be capable of power gain across the receiver bandwidth with low noise operation. Figure 8-2 shows a typical tube selected for its low noise figure. This is a type 6280 tube which is a special purpose uhf tube with close grid-cathode spacing for lowering the transit time. The construction of this tube makes it suitable for use with coaxial cavities.

Figure 8-2 also shows the tube used in a typical coaxial cavity amplifier unit. A grounded-grid amplifier circuit is used. Coaxial cavities are used for both the input and output resonant circuits. The input grid-cathode cavity is folded back over the outside of the grid-plate cavitv to shorten the overall length and facilitate tube replacement. Separate movable shorting plungers in each cav:ty are used to tune the r-f amplifier to a particular frequency

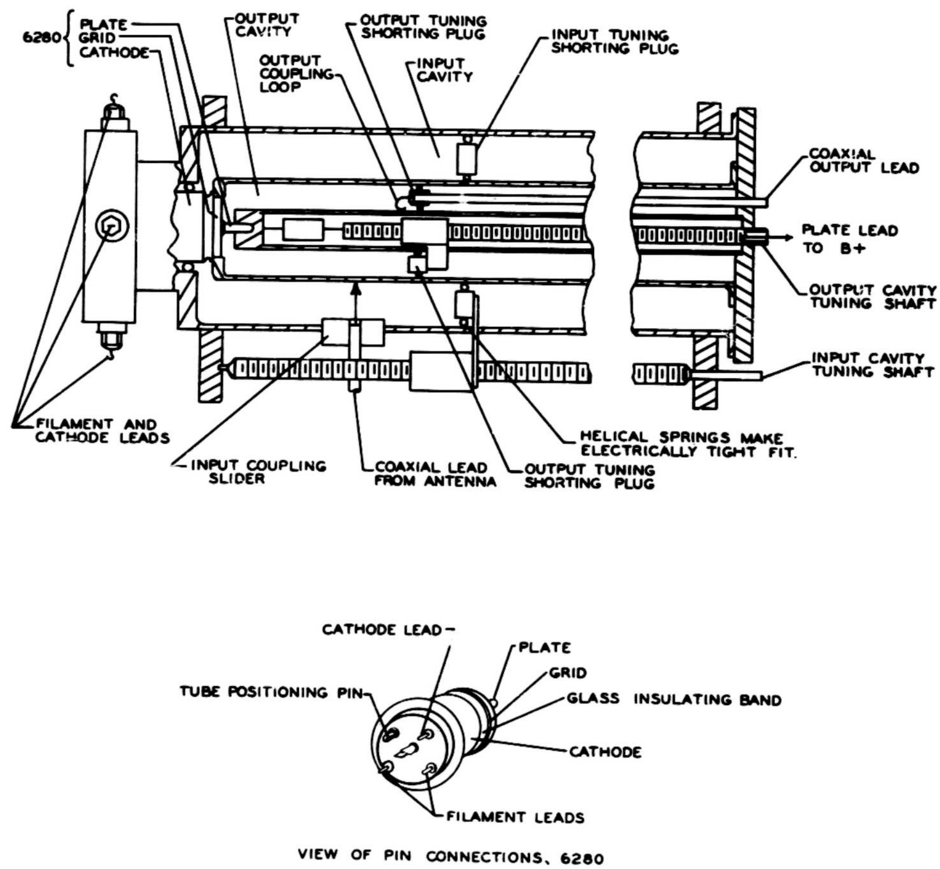


Figure 8-2. Cross Sectional View of R-F Amplifier Assembly

in the operating range. To provide maximum selectivity, the cavities are constructed to act as a quarter-wave line at the resonant frequency.

The input signal is coupled into the grid-cathode cavity by means of a sliding tap. The tap is adjusted along the cavity to provide optimum impedance match to the input coaxial line. Output is taken from the plate cavity by a one-turn link attached to the inside of the shorting plunger. This provides magnetic coupling.

Figure 8-3 is a schematic representation of the amplifier circuit. The cathode is coupled to the input cavity with a capacitor (C1) which is part of the tube construction. The plate is coupled to the output cavity with a Teflon-dielectric fabricated capacitor (C2). A filter circuit consisting of C3 and L3 keeps r-f signal off the d-c plate voltage line.

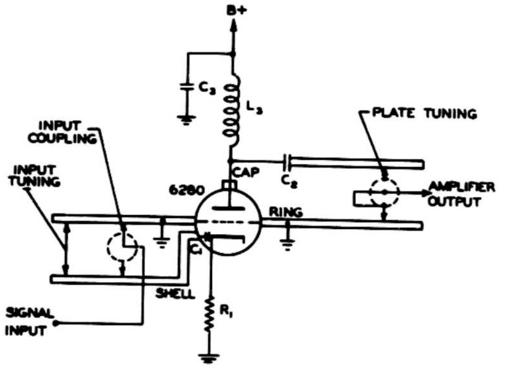


Figure 8-3. Schematic Diagram of R-F Amplifier

8.3.3 FIRST MIXER CIRCUIT.

The mixer circuit requires the use of a nonlinear electrical circuit for the generation of a new frequency. Figure 8-4 shows a typical d-c characteristic used for mixing. Notice that the current flow, due to negative voltages, is very small. Figure 8-4 also shows the combining of the small signal voltage and the local oscillator voltage. After distortion by the mixer circuit, the output contains not only these two frequencies, but frequencies equal to the sum and differences of the two input frequencies. The tuned circuits in the i-f amplifier select the difference frequency for amplification.

Any detector circuit with a nonlinear characteristic will work as a mixer. However, for uhf applications, vacuum tubes become impractical because of transit time effects and increased thermal noise. Crystals are used because they introduce less noise into the receiver and have no transit time difficulties.

The most sensitive mixer is a crystal type in which a contact on a semiconductor results in the required nonlinear characteristic. The semiconductor may be one of many different materials, such as silicon or germanium. For mechanical protection, the contact and semiconductor are mounted in a cartridge as shown in figure 8-5. The silicon crystal is soldered to a screw. A pointed piece of tungsten which makes

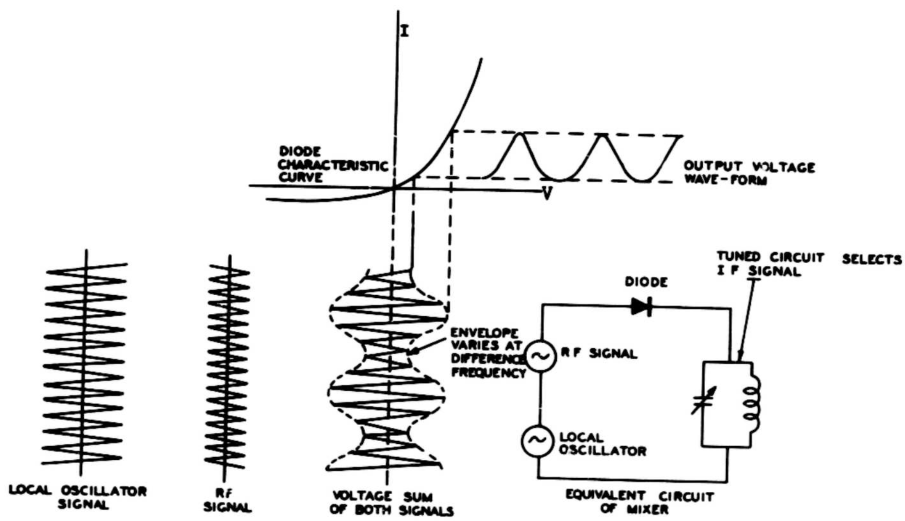


Figure 8-4. Mixer Action and Equivalent Circuit

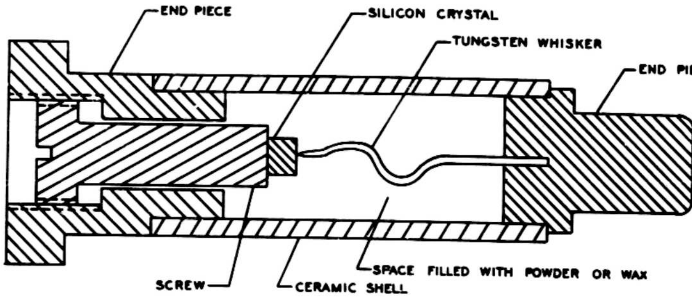


Figure 8-5. Cross Section of a Crystal Cartridge

contact with the crystal is embedded in metal. The whole assembly is held together by the brass end-pieces and the ceramic shell. The pressure on the contact is maintained by a springlike bend in the tungsten and is adjusted during manufacture by the adjusting screw. The cartridge is filled with a powder or wax to hold the pieces in place.

To make up a complete mixer, a mechanical arrangement is required that will provide r-f circuits, i-f terminals, and suitable means for plugging in the crystal cartridge. Figure 8-6 is a cross-sectional view of a typical crystal mixer assembly with the cartridge mounted in a coaxial tuned line. A schematic representation of the assembly is included in figure 8-7. The r-f input is coupled to the line by means of a sliding tap which is adjustable for correct impedance match. The line is tuned to resonate at the r-f frequency by the mixer input tuning adjustment. This is a shorting bar which is moved to determine the electrical length of the line. The injection signal is capacitively coupled to the line by means of a disc inserted into the line. The coupling is varied by adjusting the depth of penetration of the disc. The amount of injection signal coupled to the mixer circuit determines the crystal current indication on the associated meter. The mixer output appearing across C1 is coupled to the first i-f amplifier.

Two significant characteristics of a crystal mixer are the conversion loss and the noise temperature factor. These are discussed in the following paragraphs.

Conversion loss is the ratio of maximum available r-f power input to the maximum available power output at the i-f frequency. Conversion loss includes both the loss due to mismatches and the actual conversion loss occuring in the crystal. The injection signal level also affects conversion loss. The loss is infinite with no injection input, but decreases rapidly as the injection power is increased. As the injection power input is increased further, a point is reached where the conversion loss tapers off and gradually approaches a minimum value. For even larger inputs, the loss begins to increase again because of the decrease in back resistance for large applied voltages.

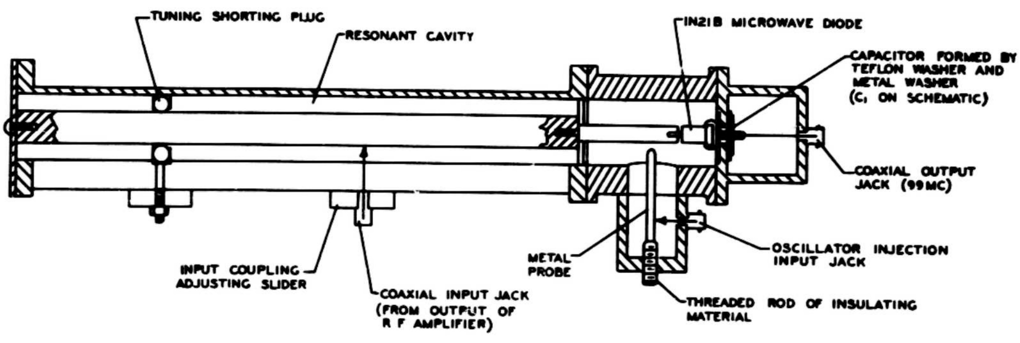


Figure 8-6. Cross Sectional View of Crystal Mixer Assembly

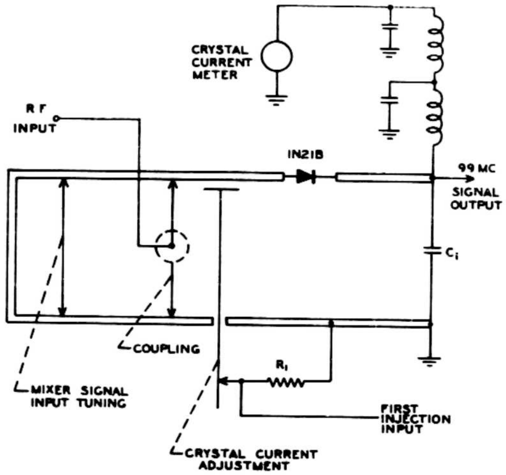


Figure 8-7. Schematic Diagram of Crystal Mixer Assembly

The reason for the relationship between conversion loss and injection level can be explained by referring again to the crystal d-c characteristic curve shown in figure 8-4. The conversion loss depends on the ratio of the slopes of the d-c characteristic at the negative peak and at the positive peak of the injection input. Up to a certain level of injection signal, the ratio will increase, and conversion loss will decrease accordingly. There is no significant increase if the injection level is increased to a point where the positive peaks go beyond the straight portion of the d-c characteristic curve.

The crystal noise temperature factor is a measure of the actual noise generated by the mixer. The level of the injection signal also affects the noise temperature. The noise increases as injection level is increased because of the increasing amount of noise applied to the crystal, and the increasing amount of noise actually generated in the crystal because of increasing back voltage.

The over-all noise figure of the mixer varies with both the conversion loss and the noise temperature factor. The increase in noise temperature factor with increase in injection level is offset by a corresponding decrease in conversion loss. This results in a fairly constant noise figure for large variations in injection signal level. The optimum injection power level for most crystals is usually about 0.5 milliwatt, which results in a continuous rectified crystal current of about 0.5 milliampere.

8.4 I-F SECTION.

The receiver i-f section consists of the first i-f amplifiers, the second mixer, the second i-f amplifiers, and the limiters. These circuits are discussed in the following paragraphs.

8.4.1 I-F AMPLIFIERS.

The intermediate-frequency amplifier stages must pass and amplify the desired signal and reject any undesired signals. The receiver bandwidth must be sufficiently wide to accept the transmitted signal without introducing distortion. The bandwidth required is determined by the system channel capacity. The smallest bandwidth consistent with system requirements should be used since the noise power increases with bandwidth. For this reason, several i-f amplifiers, each with a bandwidth to correspond to a particular channel capacity, are provided with the receiver system. Figure 8-8 shows a typical receiver bandwidth and gain relationship for the r-f and i-f stages. The r-f section accepts a wide band of frequencies which is narrowed in the first and second i-f stages. Since the second i-f pass band is the narrowest, the over-all receiver pass band is determined by these stages.

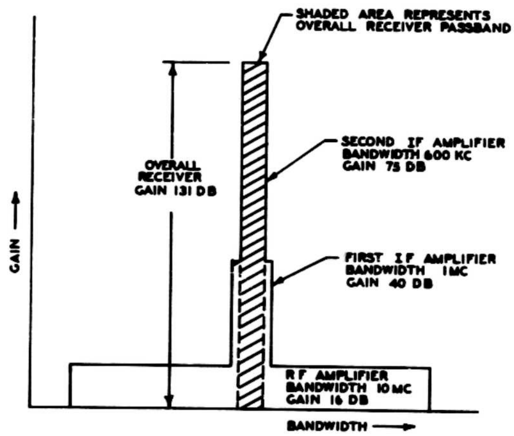


Figure 8-8. Typical Receiver Bandwidth and Gain Relationships

Figure 8-9 shows an ideal i-f response characteristic. The response should be as flat as possible within the pass band. Also, maximum rejection should be provided for signals outside the pass band. A typical response curve is also shown on figure 8-9. This is the response of a second i-f with a center frequency of 19 mc. The pass band is usually defined as the total frequency range included between the two extreme 3-db or half-power points, using the gain at center frequency as reference. This means that across the pass band, the maximum difference in gain for the various frequency components is 3 db. The rejection to frequencies outside the band is called skirt selectivity. The skirt selectivity is usually defined as the ratio of the bandwidth at the 60-db points to the bandwidth at the 6-db points. As this ratio is decreased, the skirt selectivity is increased. The ratio for an ideal response curve with steep skirts and excellent rejection for out-of-band signals is 1.

Figure 8-10 shows circuits used for obtaining a required response characteristic in the second i-f amplifiers. A system of stagger tuning is used. This system retains the simplicity of single-tuned networks, but has an increased gain-bandwidth product and greater skirt selectivity. Four stages of amplification are used. Each stage has a single response characteristic as shown in figure 8-11. To obtain the required over-all wide-band characteristic, each stage is tuned to a slightly different frequency. The resultant over-all characteristic is the product of all four stages.

The skirt selectivity is improved by parallel-resonant tuned circuits in the cathode circuits. Each circuit is tuned to a particular frequency at the skirt of the response curve. Since the circuits are parallel tuned, the cathode voltage increases at the resonant frequency. This introduces degenerative voltage to reduce the gain at the frequency to which the circuit is tuned.

One common problem in i-f amplifiers is undesired coupling which causes regeneration. One source of this coupling is through common impedances between circuits. An example of common impedance coupling is that produced in the power supply leads. This is reduced by using series decoupling networks as shown in figure 8-10. Each section in the line supplies the attenuation required to prevent feedback between

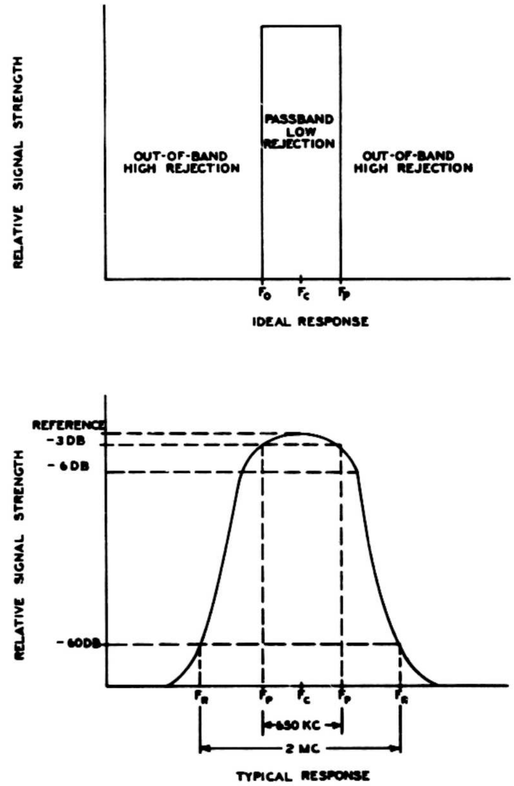


Figure 8-9. I-F Amplifier Response Characteristics

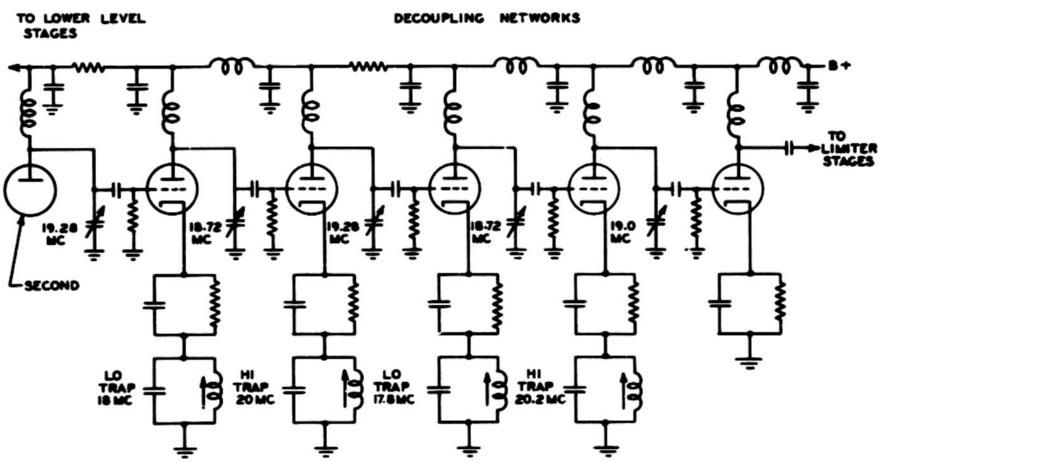


Figure 8-10. Typical I-F Circuit Showing Circuits for Formation to Band-Pass Characteristic

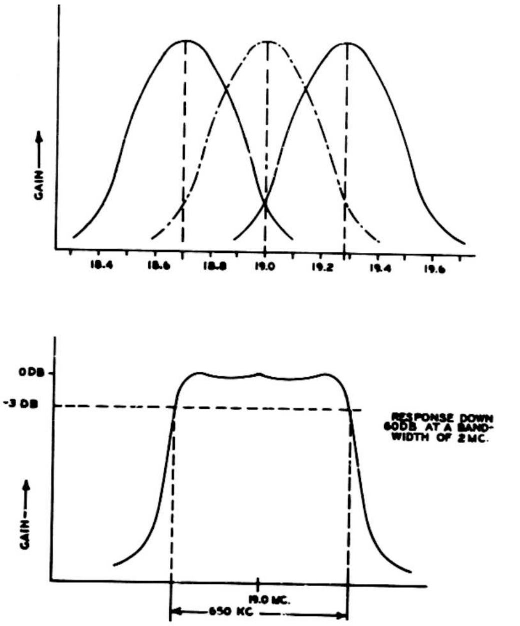


Figure 8-11. Combination of Stagger-Tuned Circuits Used to Obtain Required Over-all Response Characteristics

adjacent stages. The actual type of impedance depends on the frequency being amplified and the particular tube element supplied through the network. Another type of coupling resulting from common impedances is the r-f currents in chassis between stages of different levels. These currents are decreased by choosing grounding points so that r-f currents of a given signal level flow only in the circuit handling that level.

Another major cause of regeneration is coupling between stages through stray magnetic and electric fields. This effect increases with frequency, since radiation from any length of straight lead increases with frequency. Coupling arising from these sources is reduced by keeping leads as short as possible and using low-inductance bypassing circuits. Feed-through and button capacitors are used for reduction of coupling resulting from stray fields.

One other major cause of regeneration is feedback through interelectrode capacitance. This feedback is undesirable not only because it may cause serious distortion, but because its effect is a function of the tube transconductance. One solution is to limit stage gain; another is to use neutralizing circuits to reduce the effect of the unwanted feedback.

8.4.2 AUTOMATIC GAIN CONTROL.

An automatic gain control is sometimes used in receivers to prevent overloading of the i-f circuits. The agc voltage is obtained by rectifying a portion of the i-f output and feeding this voltage back in the proper polarity to adjust receiver gain. In doing this, it is very important that the agc line be connected so that it doesn't change loading on the tuned circuits and therefore change the pass-band characteristic. In some receivers, this effect is avoided by regulating injection level into the second mixer rather than directly changing gain of an amplifier stage. The time constant in the agc line is set so that the derived voltage will follow fades in the received signal. This allows maximum gain for low-level signals.

8.4.3 LIMITERS.

In a frequency modulation receiver, the transmitted information is recovered from frequency changes rather than from amplitude changes. To detect the frequency changes properly and to realize the full advantages of an FM system--FM improvement and noise suppression--the amplitude variations in the signal must be removed. This is accomplished in the receiver limiter stages.

To see how the amplitude variations can be removed without affecting frequency, refer to figure 8-12. Think of frequency as being proportional to the number of zero crossings of the signal per second. Amplitude variations in the signal are chopped off by limiter action. This forms a square wave having the same zero-crossing positions as the original wave. This wave is passed to a tuned circuit for reforming the sine wave. The resultant is a sine wave having the same zero crossing (frequency) as the original signal, but having a fixed amplitude.

There are various types of limiters, but they all have the same general characteristics and requirements. The circuit must overload in such a way as to make the output roughly independent of the input level once a certain input level is reached. The

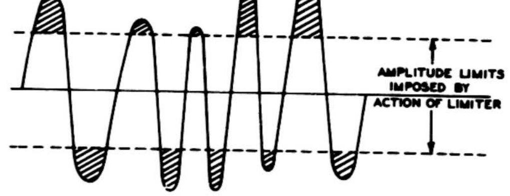


Figure 8-12. Limiter Action on Frequency-Modulated Wave

output must remain constant over a large range of input voltage. The operation time of the limiter is another important feature. The operation time must be small enough so that the limiter can resume limiting action after a sharp fade, and yet not act so fast that t follows the individual i-f cycles.

Ideal limiter action results in a reduction of noise and an improvement in receiver output signal-to-noise ratio over the input signal-to-noise ratio. Since the FM detector which follows the limiter reacts to frequency changes, any condition which produces erroneous frequency information results in noise in the detector output. Therefore, noise reduction is realized only if sufficient signal level is present at the limiter to cause the frequency information zero crossings to be controlled by the signal rather than by noise.

Limiter action with various signal levels is shown in figure 8-13. The limiter and receiver gain are set so that limiting action starts at the noise level of the receiver. Therefore, with no signal present, the noise peaks will be clipped as shown in figure 8-13A. With no signal present, the zero crossings are controlled entirely by noise, and the noise in the receiver output is at the maximum level. When a signal is applied, the noise rides on the signal wave as shown on figure 8-13B. As the signal level is increased, the number of crossovers caused by noise is reduced. This causes a reduction in noise in the receiver output. This reduction in noise with an increase in signal level is called receiver quieting.

As the signal level is increased further, a point is reached where the signal peak value equals the noise peak value as shown in figure 8-13C. This causes a significant reduction in noise because the signal overrides the noise and controls the crossover points. The rms value of signal level required for the signal peaks to equal the noise peaks is called receiver threshold.

Since the peak-to-rms ratio for the sinusoidal signal is 3 db and the peak-to-rms for noise is 13 db, the rms signal-to-noise ratio required for the improvement threshold is 10 db. The expression for the level required for threshold is:

$$S_T = -144 \text{ dbm} + 10 \log BW + F + 10 \text{ db}$$

$$S_T = -134 \text{ dbm} + 10 \log BW + F$$

where: S_T is input signal level in dbm

BW is bandwidth in kilocycles

Noise figure in db.

The receiver bandwidth comes into the expression since the input noise level increases with an increase in bandwidth. When the signal level drops below threshold, there is a rapid increase in noise level. As the signal level is increased above threshold, there is a constant reduction in receiver noise.

Figure 8-14 shows two typical limiter circuits. The top circuit uses a pentode with the operating point set so that the input signal of proper level can drive the tube beyond cutoff and saturation. The negative

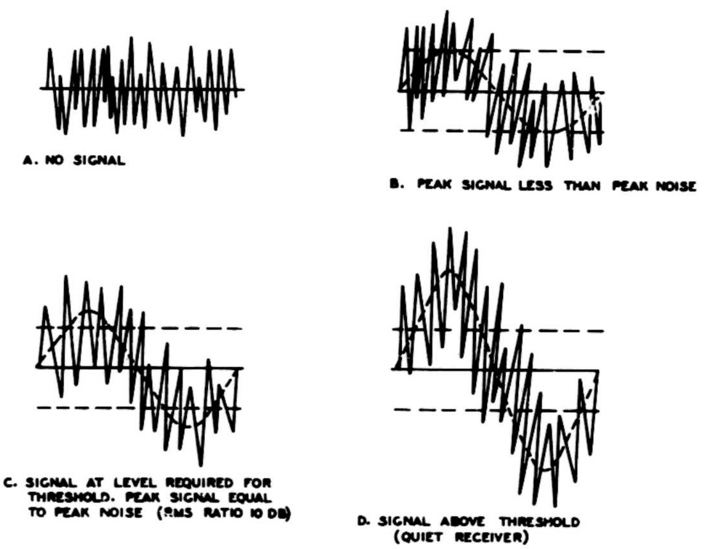


Figure 8-13. Noise Reduction by Limiter Action as Signal Level is Increased

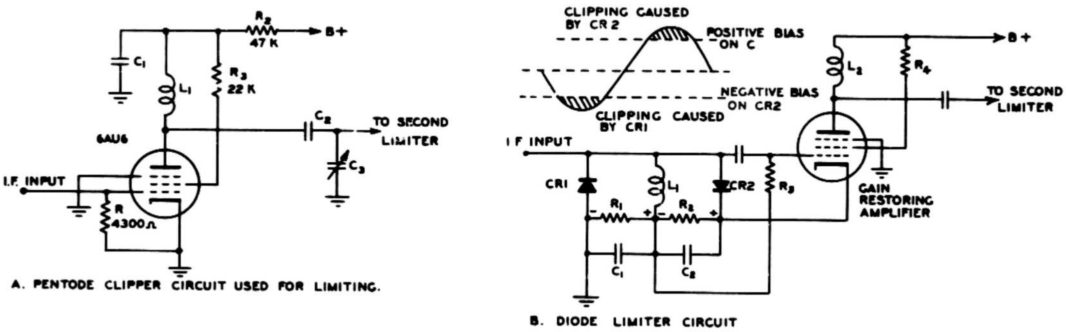


Figure 8-14. Two Typical Limiter Circuits

swing of signal voltage cuts off the tube, while the positive swing brings the plate current to the saturation value. The output circuit is tuned to the i-f frequency with variable capacitor C3 in the plate circuit. For complete limiter action, this stage is followed by another similar limiter stage.

Another limiter circuit is shown in figure 8-14. This circuit uses two diodes connected in parallel across the i-f input to provide the clipper action. Output from the diodes is connected to a pentode amplifier which provides gain to compensate for losses incurred in the clipper action. Bias for operation of the diodes is provided by plate current flowing through resistors R_2 and R_1 connected in the cathode circuit. The diodes are connected so that CR1 acts as a negative limiter and CR2 acts as a positive limiter. All portions of the input signal exceeding the bias voltage on each diode will be clipped as shown in figure 8-14. The point at which the limiting action occurs is set by the bias level on each diode which is determined by the amplifier plate current.

8.5 DISCRIMINATORS.

8.5.1 GENERAL OPERATION.

The discriminator translates frequency deviations in the received FM signal into an audio output voltage. This is done in two steps: first, the frequency changes are converted to amplitude variations, and then these variations are rectified. The resultant output voltage varies in proportion to the frequency changes from center frequency. The receiver output with change in frequency from the center frequency should follow the frequency change produced in the transmitter for a given modulating voltage input. This relationship between the transmitted signal and discriminator output is shown in figure 8-15. When the modulating voltage is zero, the frequency change is zero, and output from the discriminator is at a minimum reference point. The figure shows that the discriminator operation must be linear within the pass-band of the receiver to accurately reproduce the transmitted signal.

The following paragraphs describe the operation of two types of discriminators used in tropospheric scatter receivers.

8.5.2 FOSTER-SEELEY DISCRIMINATOR.

In the discriminator, shown in figure 8-16, the primary and secondary of L are tuned to the center frequency, which for this example is 19 mc. The primary circuit is tuned with C3; the secondary circuit is tuned with a powdered iron slug. The center of the secondary circuit is connected to the top (high potential side) of the primary by capacitor C4, which has negligible reactance at the center frequency. This arrangement develops voltages V_{12} and V_{13} which vary with frequency as shown in figure 8-12B. These voltages are applied to two matched 1N198 crystal diodes which are arranged so that the discriminator output (V_o) is the difference in the outputs of the two diodes and is therefore proportional to $V_{12} - V_{13}$. As a result, the output voltage varies with the frequency deviations above and below the center frequency as shown on figure 8-16C.

Phase relations of the various voltages developed in the discriminator are shown in figure 8-16 (D, E, and F) to illustrate discriminator action. Because the center of the secondary is connected to the primary, the following relations exist:

$$V_{12} = V_1 + V_{ac} \text{ and } V_{13} = V_1 - V_{bc}$$

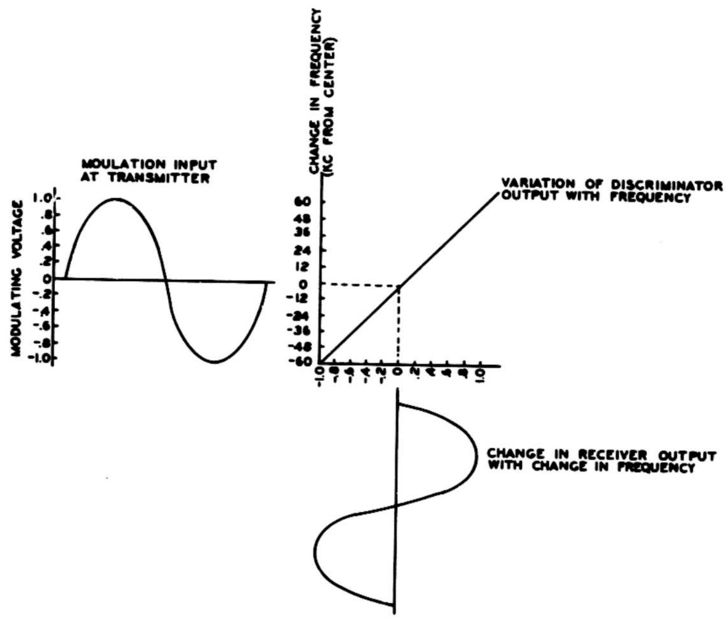


Figure 8-15. Relationship between Modulation Input at Transmitter and Discriminator Output at Receiver

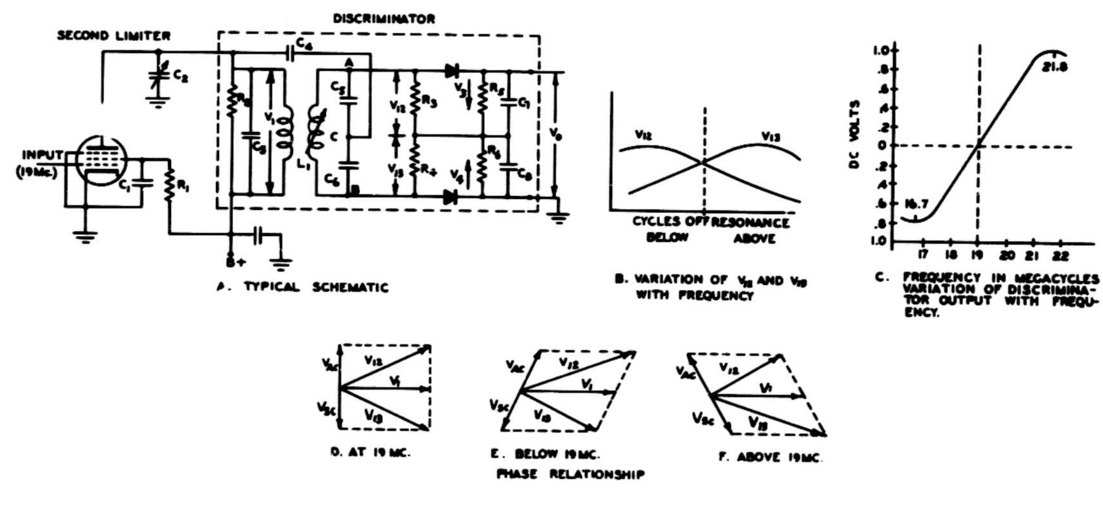


Figure 8-16. Operation of Foster-Seeley Discriminator

where: V is the voltage across the primary, and V_{oc} and V_{bc} are the voltages across each half of the secondary. These secondary voltages are 180° out of phase at all frequencies. At the resonant frequency of the tuned circuit, both the primary and secondary are resistive with no reactive component, and therefore V_{oc} and V_{bc} are 90° out of phase with the primary voltage. This makes the resultant voltages V_{12} and V_{13} applied to the diodes equal in amplitude. At frequencies off resonance, the impedance of the secondary has a reactive component. This makes the voltages across each half of the secondary either slightly less or slightly more than 90° out of phase with the primary voltage, depending upon which side of resonance the frequency deviates. As a result, the two resultant voltages (V_{12} and V_{13}) applied to the diodes are unequal in amplitude. At frequencies below resonance, V_{12} is increased in amplitude, and V_{13} is decreased; at frequencies above resonance, the reverse is true.

The output voltage V_0 is the difference between the voltages developed across the diode load resistors R_5 and R_6 and is, therefore, proportional to the difference between the input voltages V_{12} and V_{13} applied to the diode. This results in the output voltage versus frequency curve shown in figure 8-16C.

8.5.3 PULSE COUNTER DISCRIMINATOR.

Figure 8-17 is a schematic diagram of another type of discriminator. With this type of discriminator, the i-f output is applied to a phase inverter circuit. The two outputs from the phase inverter drive a pulse counter discriminator. The tube used for the discriminator is a 6BN6 gated-beam tube. In this type of tube, grid 1 and grid 2 have equal control over the plate current, and both have a very sharp cutoff characteristic. The signal from grid 1 is obtained directly from the cathode of the phase inverter and has sufficient amplitude to cut off the plate current on its negative swings. The signal from grid 2 is obtained from the phase inverter and also has sufficient amplitude to cut off the plate current on its negative swings. Since the plate and cathode signals from a phase inverter are 180° out of phase, if the signals were applied directly to the grids of the discriminator, no plate current would flow and no output would be obtained. However, the signal applied to grid 2 is delayed by a delay line in the grid circuit which provides for a constant time delay. This allows a pulse of plate current to flow.

The time relationship between the two grid inputs and the plate output is shown in figure 8-18. The duration of the pulse of plate current is equal to the delay introduced by the delay line, and the repetition rate is equal to the frequency of applied signal. Thus, as the frequency of the FM signal applied to the discriminator varies, the number of pulses-per-unit time varies, while the duration of each pulse remains constant.

The plate output is passed through a low-pass filter to produce an output with amplitude varying in accordance with the instantaneous frequency of the signal applied to the discriminator. The frequency of the output varies in accordance with the rate at which the instantaneous frequency of the input changes.

8.6 COMBINER.

A tropospheric scatter receiver terminal normally uses at least two receivers, each connected to a separate antenna, for diversity reception. A combiner output stage, which is essentially a cathode follower, is included in each receiver. These output stages are connected in parallel to produce the combined output signal. This combined signal, consisting of both the signal components and random noise components, is applied to a common load. An improvement of up to 3 db over the best of the signal-to-noise ratios of the individual receivers is possible in the composite output.

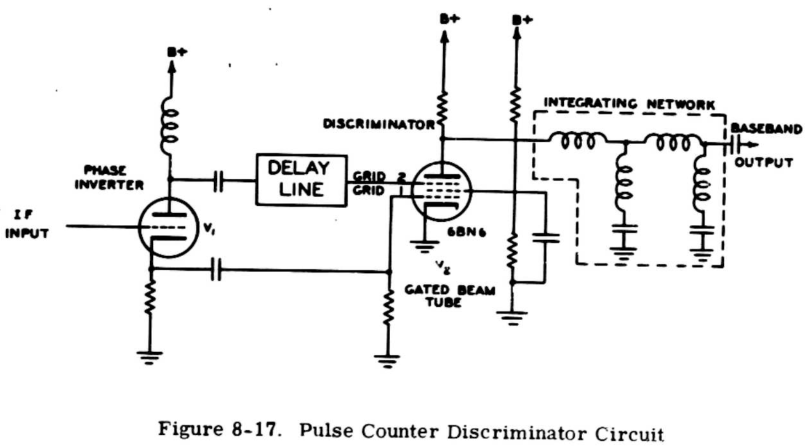


Figure 8-17. Pulse Counter Discriminator Circuit

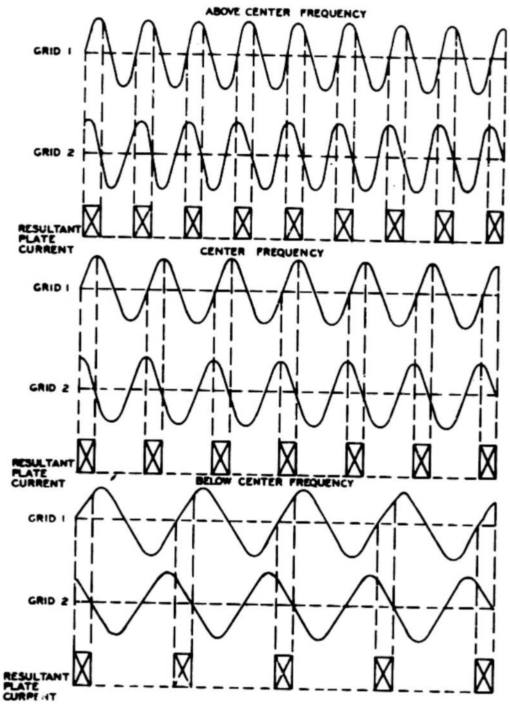


Figure 8-18. Relationship of Signal in Pulse Counter Discriminator

Figure 8-19 is an equivalent circuit of a dual-diversity combiner system which illustrates how the receiver outputs are combined. In this figure, the output impedance of each combiner is designated Z_1 and Z_2; and the output impedance of the diversity system, which is much larger than either Z_1 or Z_2, is designated Z_L.

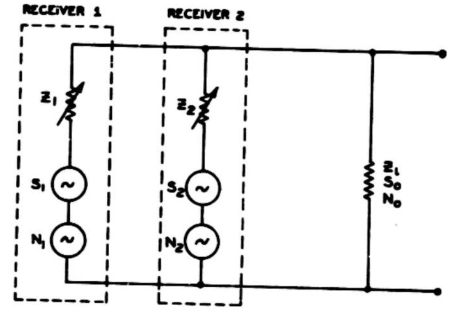
For proper combiner action, the two receiver signals (S_1 and S_2) must be in phase and be maintained at equal amplitudes. Signal output from each receiver is held constant over a wide range of receiver input signal levels by action of the limiter circuits. Since the signal components are in phase and equal in amplitude, neither combiner tube can load the other. Thus, the only load presented to the signal voltage from the combiner tubes is Z_L, and the signal output voltage (S_0) is equal to the output voltage from either of the combiner tubes alone.

Noise level in the receiver output (N_1 and N_2) depends on receiver quieting which varies with the strength of the r-f signal being received. Since the receivers are connected to diversity antennas with little correlation between the received signals, the noise output is rarely equal. By taking a sample of this noise at the output of the discriminator and applying it as d-c bias to the combiner, the output impedance of each combiner is made to vary as the square of the noise power present. With equal signals, the output impedance of each of the receivers in a diversity system is a few hundred ohms. However, if a fade is encountered by one of the receivers which increases its noise level by 10 times, its output impedance will increase 100 times, and the high noise, instead of appearing across the output load, will be dissipated in the high output impedance of the receiver so that the total noise output remains essentially unchanged.

The noise components in each receiver output have no fixed phase relationship and therefore vary randomly in phase with respect to each other. Because of this random phase relationship, a potential difference exists between N_1 and N_2, and noise current flows between the two combiners. Since Z_L is large compared to Z_1 or Z_2, the load on the first combiner tube is Z_2, and the load on the second combiner tube is Z_1. Therefore, the noise voltage across each combiner tube is given by the following expression:

$$\frac{Z_2 N_1}{Z_1 + Z_2} \quad \text{and} \quad \frac{Z_1 N_2}{Z_1 + Z_2}$$

These voltages are applied in parallel across the output impedance Z_L. Because they have a random phase relationship, they combine as the square-root



S = DESIRED SIGNAL ON FIRST COMBINER TUBE
N = NOISE SIGNAL ON FIRST COMBINER TUBE
Z = OUTPUT IMPEDANCE OF FIRST COMBINER TUBE, A FUNCTION OF THE BIAS ON THE TUBE
S = DESIRED SIGNAL ON SECOND COMBINER TUBE
N = NOISE SIGNAL ON SECOND COMBINER TUBE
Z = OUTPUT IMPEDANCE ON SECOND COMBINER TUBE, A FUNCTION OF THE BIAS ON THE TUBE
Z_L = LOAD IMPEDANCE, VERY HIGH WITH RESPECT TO Z OR Z
S_0 = OUTPUT SIGNAL VOLTAGE

Figure 8-19. Equivalent Circuit of Dual-Diversity Combiner System of the sum of the squares to give the total noise output:

$$N_O = \frac{\sqrt{(Z_2 N_1)^2 + (Z_1 N_2)^2}}{Z_1 + Z_2}$$

The signal-to-noise ratio in the output is equal to:

$$\frac{S_O}{N_O} = \frac{S_O (Z_1 + Z_2)}{\sqrt{(Z_2 N_1)^2 + (Z_1 N_2)^2}}$$

The maximum 3-db improvement in the composite output signal-to-noise ratio over that of the best receiver occurs when the signal-to-noise ratios in the two receivers are equal. This can be shown as follows. Since the individual receiver signal outputs are always equal, when the receiver signal-to-noise ratios are equal the noise outputs are also equal. When N_1 and N_2 are equal, the input impedance Z_1 and Z_2 are also equal since they are made to vary as the square of the noise voltages. Using these conditions and assuming that receiver no. 1 has the highest signal-to-noise ratio, the expression for output signal-to-noise ratio can be simplified as follows:

$$\frac{S_O}{N_O} = \frac{S_1 (Z_1 + Z_2)}{\sqrt{(Z_2 N_1)^2 + (Z_1 N_2)^2}}$$

With equal signal-to-noise ratios: $Z_1 = Z_2$ and $N_1 = N_2$. Therefore:

$$\frac{S_O}{N_O} = \frac{S_1 (2Z_1)}{\sqrt{2 (Z_1 N_1)}} = \frac{S_1 (2Z_1)}{Z_1 N_1 \sqrt{2}} = \frac{2 S_1}{\sqrt{2} N_1}$$

$$= 1.414 \left(\frac{S_1}{N_1}\right)$$

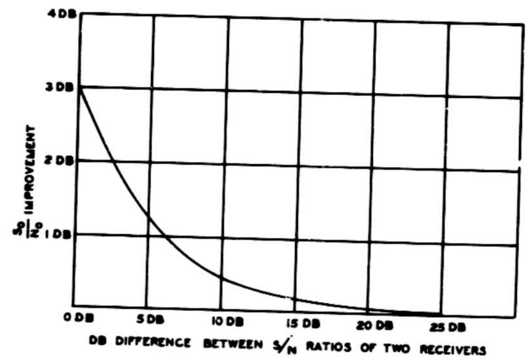


Figure 8-20. Improvement in Combined Output Over the S/N Ratio in the Best Receiver

In decibels this is $= 20 \log 1.414 + \frac{S_1}{N_1}$

$$= 20 \times .15 + \frac{S_1}{N_1}$$

$$= 3 \text{ db} + \frac{S_1}{N_1}$$

As the difference between the signal to noise ratios increases, this improvement in output signal-to-noise ratio resulting from combiner action decreases as shown in figure 8-20. As the signal level of one receiver decreases, the noise level in the output of that receiver increases. This causes a corresponding increase in the receiver output impedance. The portion of the total noise output contributed by each combiner varies inversely as the ratio of the output impedances of the combiner tubes. Therefore, when the noise of one receiver is high as compared to the other, the output signal-to-noise ratio will be that of the receiver receiving the strongest signal.

8.7 AUTOMATIC BANDWIDTH CONTROL CIRCUITS.

In the discussion of limiter action, it was shown how the output signal-to-noise ratio degrades very rapidly when the signal drops below threshold level. A recent development in FM receiver design is a bandwidth control circuit that automatically reduces the receiver bandwidth as the signal level decreases. This lowers the threshold level for low-level signals and prevents the rapid degradation of signal-to-noise ratio.

The receiver output signal-to-noise ratio as a function of received signal level for a conventional FM receiver is shown in curve A of figure 8-21. If the receiver input level exceeds the threshold value, the output signal-to-noise ratio is directly proportional to the input level. However, when the input level drops below threshold, the output signal-to-noise ratio decreases very rapidly as shown in curve A.

Suppose the bandwidth of the FM receiver is made to decrease automatically as the signal level decreases. As the signal input approaches normal threshold level, because of fading, the bandwidth of the receiver i-f is caused to decrease. This results in progressively lower threshold levels as shown in curve B of figure 8-21.

In actual receivers, the bandwidth can only be reduced a limited amount. Therefore, a threshold effect is still present with automatic bandwidth control. However, the threshold occurs at a much lower input signal level. A bandwidth reduction of 5 to 1, which reduces threshold level by 7 db, is obtained in receivers now in use.

Figure 8-22 is a block diagram of a typical automatic bandwidth control system. Output from the second mixer is applied to the bandwidth control circuit through a band-pass filter. This filter determines the maximum bandwidth of the entire receiver and thus controls the channel capacity of the receiver. The bandwidth variations within the maximum bandwidth limits set by the filter are made in a variable bandwidth stage. This stage is a cathode follower with a tuned circuit in the cathode. The bandwidth of the tuned circuit is dependent upon the cathode impedance of the cathode follower. This impedance is dependent upon the bias on the tube. Thus, if the bias upon the cathode follower is varied as a function of signal strength, the bandwidth will vary as a function of signal strength.

The gain through the bandwidth control circuit must remain constant with variations in bandwidth. A constant gain will result if the unloaded Q of the cathode tank circuit is sufficiently high to present a high impedance at the narrowest bandwidth. A Q multiplier circuit is used to meet this requirement. This circuit applies a feedback voltage in phase with the signal at the cathode of the bandwidth control circuit. This feedback voltage reinforces the signal, thereby giving an apparent increase in the Q of the tank circuit.

8.8 BASIC ADJUSTMENT AND TEST PROCEDURES FOR RECEIVERS.

The detailed step-by-step test adjustment procedures for a given receiver are included in the instruction book for that receiver. The following paragraphs describe the general procedures which apply to all receivers, and give background material for understanding the basic steps.

8.8.1 TUNING THE RECEIVER TO OPERATING FREQUENCY.

The receiver is tuned to the operating frequency by tuning the first injection system and the r-f system preceding the first mixer. The second injection frequency and the i-f stages are fixed in frequency, and therefore these stages normally require no adjustment when the station receive frequency is changed.

Before applying power to the receiver, the injection level control should always be set for minimum crystal current. This will prevent the crystal current from

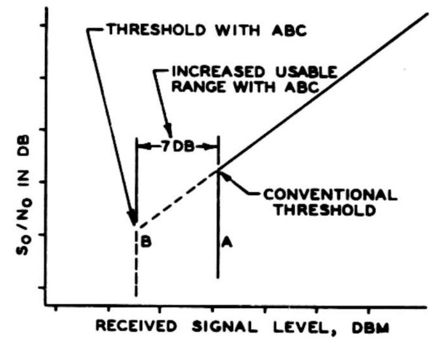


Figure 8-21. S/N Ratio as a Function of Received Signal Level for an FM Receiver with and without Automatic Bandwidth Control Circuit

operating at an excessive level and thus damage the crystal or shorten its life.

The first injection is tuned for operation at the desired receive frequency by inserting the correct crystal in the oscillator circuit and then tuning the multiplier stages which follow the oscillator. The crystal frequency is determined by the following formula:

$$\text{Crystal frequency} = \frac{\text{Receive frequency - first IF}}{\text{Multiplication factor following oscillator}}$$

For example, assume that the receive frequency is 1758 mc, the first i-f is 70 mc, and the multiplication factor is 72. The crystal frequency is then:

$$\frac{1758 - 70}{72} = 23.4444 \text{ mc.}$$

Assume another typical case for 1000-mc equipment with the receive frequency equal to 857 mc, the first i-f equal to 99 mc, and the multiplication factor is 12. The crystal frequency is then:

$$\frac{855 - 99}{12} = 63.00 \text{ mc.}$$

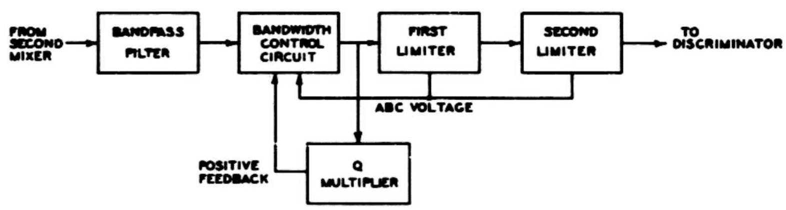


Figure 8-22. Block Diagram of Typical Automatic Bandwidth Control Circuit

After the correct crystal is inserted, the oscillator-multiplier stages are tuned for maximum output. The injection level control is then normally set for the correct level.

The r-f section can be tuned by using either the associated exciter or a stable uhf signal generator as a signal source. The associated exciter can be accurately tuned to the station receive frequency and then connected to the receiver input through an attenuator to prevent overloading the receiver. The output of the r-f section should be terminated in the correct impedance. This can be done by connecting the r-f section to the i-f input or by using a coaxial termination.

8.8.2 DETERMINATION OF RECEIVER NOISE FIGURE.

Before going into the procedures for determining noise figure, review the meaning of this receiver characteristic. An ideal receiver is defined as one that introduces no internal noise. The output noise of such a receiver will be only the gain of the receiver times the noise introduced from the antenna. In practice, however, receivers generate internal noise, primarily in the input stages. As a result, the noise output of an actual receiver is increased considerably over the amplified antenna noise developed by an ideal receiver. The receiver noise figure is the ratio of the actual noise output to the noise output of an ideal receiver with the same bandwidth. This ratio is conveniently expressed in decibels.

To determine the noise figure, the actual receiver noise is compared to a calibrated and adjustable source of noise power having a uniform frequency spectrum over the receiver bandwidth. A simplified schematic diagram of a noise generator used for this application is given in figure 8-23. Such a noise generator uses as the noise source a diode for which the noise component of the plate current is proportional to the direct current through the tube. Therefore, the noise output of the generator can then be determined by measuring the direct current. The noise output is varied by controlling the diode current produced by thermal emission from a tungsten filament. The output resistance R matches the receiver input impedance and serves as the equivalent antenna resistance and source of input noise when the noise figure

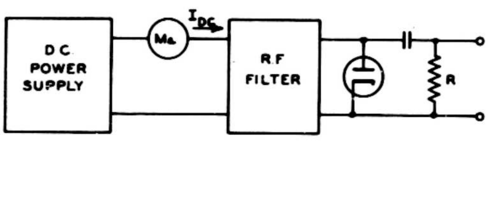


Figure 8-23. Simplified Schematic Diagram of Noise Generator

tests are being made. The meter in the noise generator is calibrated to indicate the ratio of the noise generator output to the resistor thermal noise in decibels. The calibration is made, assuming a particular value of R and that R is at room temperature.

Figure 8-24 illustrates the basic procedure for determining noise figure. The noise generator is connected to the receiver input, and the noise generated in the generator output resistance R is used as the reference. This reference noise is shown as N_i in figure 8-24. With an ideal receiver, this would be the only noise to be considered, and the output noise would be equal to GN_i, where G is the receiver gain. The internal noise generated by an actual receiver is shown as N_R in figure 8-24. The actual noise output, assuming the internal noise is generated primarily in the receiver r-f section, is $G(N_i + N_R)$. The receiver noise figure is then equal to:

$$\frac{\text{Actual noise output}}{\text{Ideal noise output}} = \frac{G(N_i + N_R)}{G(N_i)} = \frac{N_i + N_R}{N_i}$$

The noise figures for various values of N_R are given in figure 8-24. Since N_i is a known quantity determined by the value of R and its temperature, the actual value of N_R for various noise figures can also be determined.

When determining receiver noise figure, a calibrated output indicator is used as shown in figure 8-25. The noise output is first measured with the noise generator turned off. The total noise output is then $N_i + N_R$. The noise generator output is then applied and adjusted until the output indication is doubled. When this is done, the total noise output is $N_i + N_R + N_G$ and $N_G = N_i + N_R$. The noise meter indicates the ratio $\frac{N_G}{N_i}$ in db. This is also the receiver noise figure since $F = \frac{N_i + N_R}{N_i}$ and $N_i + N_R = N_G$.

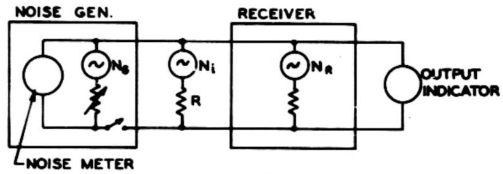
There are several factors to be considered when using this procedure for measuring noise figure. The receiver output indicator should be connected at a point which includes all significant noise sources in the receiver. Also, the indicator should be calibrated to indicate accurately when the noise power is doubled with the addition of noise generator output.

There are several additional correction factors which must be taken into account when the noise generator is used. One of these is the transit time effect in the noise diode. This effect becomes large in the uhf range when the time of travel of the electrons in passing from the cathode to the anode of the noise diode becomes an appreciable part of the period of the signal frequency. To take this effect into account, a correction factor must be subtracted from the noise figure indication to obtain the correct noise figure. This effect varies with frequency and particular type of noise generator. Another factor to be considered is mismatch error since, in general, neither the noise generator nor the receiver is perfectly matched to the line. This error will be a function of line length between the noise generator and the receiver.

A high noise figure may be an indication that the receiver input circuit is not tuned properly. Frequently, a mismatch between the input circuit and the line may produce optimum noise figure.

Figure 8-25 shows a typical equipment setup for measuring receiver noise figure. In this case, an external bias supply is used to vary the gain of the i-f amplifiers fcr calibration of the output metering circuit. A d-c vtvm is connected through a rectifier circuit to the first limiter stage. The gain control is first set for minimum gain, and then the noise generator output is adjusted for a low convenient reference indication on the d-c vtvm. The noise generator output is then doubled (increased 3 db) and the indication on the d-c vtvm is noted. This procedure calibrates the system and provides a point on the d-c vtvm which corresponds with a 3-db increase in noise level.

After the system is calibrated, the noise generator is turned off, and the gain is adjusted until the original reference point is indicated on the d-c vtvm. This provides an indication corresponding to the actual noise output of the receiver ($N_R + N_i$). The noise generator is then turned on and adjusted until the vtvm indication is increased to the 3-db reference point. When this indication is obtained, the noise generator output equals the actual receiver noise ($N_G = N_i + N_R$). The meter on the noise generator then provides a direct indication of the receiver noise figure, since the meter is calibrated to indicate the ratio of $\frac{N_G}{N_i}$ in decibels. This is derived as follows:



RECEIVER NOISE FIGURE $(F) = \frac{N_i + N_R}{N_i}$ $NR = N_i\ (F-1)$.

NOISE METER CALIBRATED TO INDICATE $\frac{N_G}{N_i}$ IN DB.

WHEN N IS SET TO EQUAL $N_R + N_i$, NOISE METER INDICATES RECEIVER NOISE FIGURE IN DB.

VALUE OF NR	NOISE FIGURE (ACTUAL RATIO)	NOISE FIGURE (DB)
$N_R = 0$	1	0
$N_R = N_i$	2	3
$N_R = 2N_i$	3	4.8
$N_R = 3N_i$	4	6
$N_R = 4N_i$	5	7
$N_R = 5N_i$	6	7.8
$N_R = 6N_i$	7	6.5
$N_R = 7N_i$	8	9
$N_R = 8N_i$	9	9.5
$N_R = 9N_i$	10	10

Figure 8-24. Basic Noise Figure Relationships

$$F = \frac{N_R + N_i}{N_i}$$

$$N_G = N_i + N_R$$

$$F = \frac{N_G}{N_i} \text{ (the noise generator indication)}$$

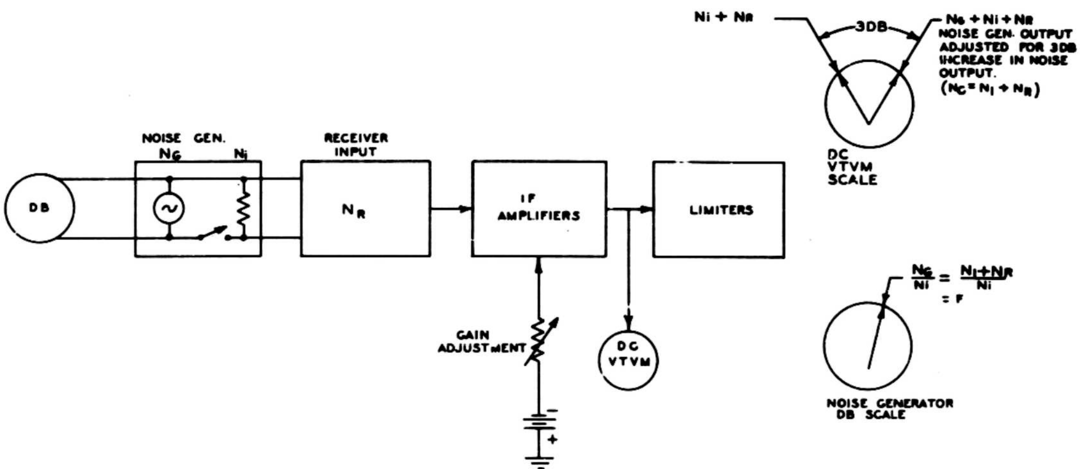


Figure 8-25. Equipment Setup for Determination of Noise Figure

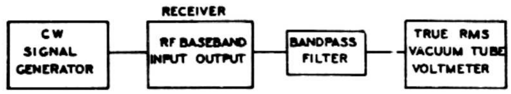


Figure 8-26. Equipment Setup for Checking Receiver Quieting

The indicated noise figure must, of course, be correct for the transit time and line length effects. The noise figure is corrected for line length effects by reducing the noise figure given in db by line attenuation in db.

8.8.3 RECEIVER QUIETING CHECK.

Due to limiting action, the noise level in the receiver output decreases with an increase in input signal level. This decrease in output noise with an increase in signal level is called receiver quieting. A receiver quieting test provides a good check for overall receiver operation, tuning, and adjustments.

Figure 8-26 shows a typical setup for performing a receiver quieting test. A signal generator, capable of operating at the receive frequency, with an adjustable output level and an accurate output meter is required for the signal input. This should be connected to the r-f input so that the quieting test checks operation of the entire receiver. A true rms vacuum-tube voltmeter is connected through a band-pass filter to the receiver baseband output. The band-pass filter is selected to pass only the noise included in the baseband range of frequencies. The true rms meter is used to give an accurate indication of noise level.

The receiver quieting is checked by first measuring the noise output with no signal input. The signal input is then increased in convenient steps, and the output noise level is recorded for each value of signal input.

Figure 8-27 shows two typical receiver quieting curves. Signal level in -dbm is shown increasing from left to right on the horizontal scale. Output noise level is given along the vertical scale in db below the noise level obtained with zero signal input. These curves show that as the signal level is increased from zero, there is very little change in

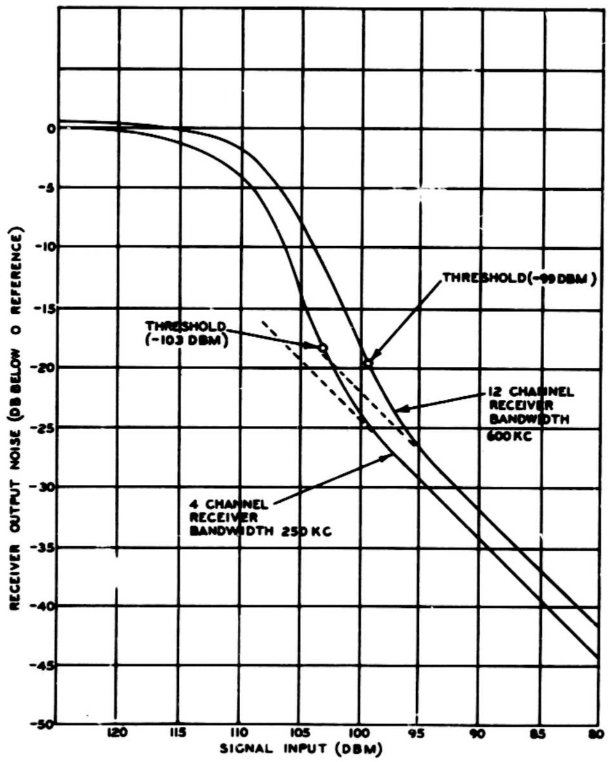


Figure 8-27. Typical Receiver Quieting Curves

output noise level until the level required for quieting is obtained. At this point, there is a sudden decrease in noise level with increase in signal level. As the signal level is increased further, a point is reached where a one-dbm change in signal level results in a one-db decrease in output noise. This change in the quieting curve is called receiver threshold. The actual threshold is defined as the point at which the output noise level is 3 db above the projected db-for-db portion of the curve.

The signal level required for threshold depends on receiver bandwidth and noise figure in accordance with the following equation:

$$S_T = -134 + 10 \log BW + F$$

where: S_T = signal required for threshold in dbm

BW = receiver bandwidth in kc

F = receiver noise figure in db.

Assuming an 8-db noise figure, the threshold for various bandwidths is as follows:

VOICE CHANNEL CAPACITY	BANDWIDTH	THRESHOLD
4	250 kc	-103 dbm
12	600 kc	-99 dbm
24	1200 kc	-95 dbm

The change in threshold with change in bandwidth for a given receiver is shown on the two quieting curves in figure 8-27. These curves also show that the basic shape of the quieting curve remains unchanged with a change in bandwidth. The only characteristic that changes is the point at which threshold occurs. The two curves of figure 8-27 also show the significance of the threshold level. Whenever the signal fades below the threshold level, a sudden increase in output noise level results.

CHAPTER 9
TRANSMISSION LINES AND ANTENNAS

9.1 INTRODUCTION.

Tropospheric scatter propagation requires the use of high-gain directive antennas and transmission lines capable of operating efficiently in the uhf and shf ranges. This chapter provides a review of the physical and electrical properties of transmission lines and discusses the types of lines especially adaptable for use in tropospheric scatter terminals. The use of transmission line sections as filters is also described.

General antenna characteristics and the requirements for tropospheric scatter communications are discussed in this chapter, along with descriptions of typical antenna systems.

9.2 PROPERTIES OF TRANSMISSION LINES.

9.2.1 BASIC REQUIREMENTS.

Transmission lines are used at tropospheric scatter terminals to conduct electrical energy from the power amplifier to the transmit antenna and from the receive antenna to the receivers. They are also u ;ed as tuned filter circuits in the power amplifier output and receiver input lines. To perform these functions, the lines must operate in the uhf and shf ranges with low losses. The transmit lines must be capable of operation with high levels of input power.

The following paragraphs describe general transmission line properties and the characteristics of transmission lines used in tropospheric scatter terminals.

9.2.2 EQUIVALENT CIRCUIT.

At high radio frequencies, a transmission line exhibits entirely different characteristics than it does at lower frequencies. As frequency increases, the effects of capacity and inductance increase. These reactive components retard the transfer of energy to such an extent that complete cycles may be generated at the source before the voltage at the start of the cycle reaches the load.

The electrical characteristics of an r-f transmission line consist of inductance series with the line, capacitance across the line, series resistance and shunt conductance per unit length of line. In ordinary circuits, these quantities are present in definite lumps. In an r-f transmission line, however, these quantitites are distributed throughout the line. However, a typical r-f transmission line can be resolved into a simple equivalent circuit as shown in figure 9-1. At high frequencies, the effect of the series resistance and shunt conductance is small compared to the effects of L and C. Therefore, for most practical purposes, they can be neglected and are not shown on figure 9-1.

9.2.3 CHARACTERISTIC IMPEDANCE.

Suppose that the line shown in figure 9-1 is of infinite length and that a voltage is applied across the input terminals AB. A certain current proportional to the voltage will flow in the line. The ratio of the voltage to the current is the line impedance ($Z = \frac{E}{I}$). The impedance presented by a line of infinite length is called the characteristic impedance and is usually designated Z_0. The characteristic impedance determines the amount of current that can flow when a given voltage is applied to an infinitely long line, in exactly the same way that resistance connected across the source limits current flow.

In actual practice, the characteristic impedance of any line is a definite constant value determined by the inductance L and capacitance C per unit length of line. Assuming negligible resistance losses, the expression for characteristic impedance is: $Z_0 = \sqrt{L/C}$ where: Z_0 is in ohms, and L and C are expressed in the same magnitude of units. To illustrate the meaning of this expression, assume that the characteristic impedance of an actual line must be determined. The procedure is as follows.

a. Using a capacity bridge, measure the capacity of a convenient length of open ended transmission line.

b. Short the open end of the same line, and measure its inductance with an inductance bridge.

c. Substitute the measured values in the formula for characteristic impedance. For example, assume

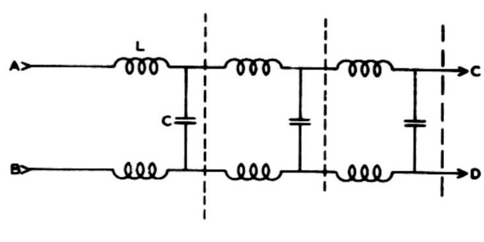


Figure 9-1. Approximate Representation of R-F Transmission Line

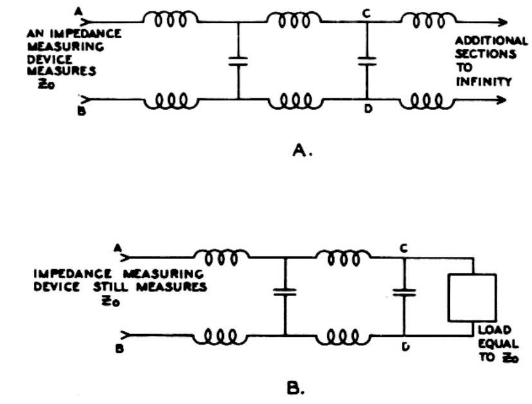


Figure 9-2. Development of Characteristic Impedance

that C = 0.00003 microfarad and that L = 0.08 microhenry.

$$Z = \sqrt{\frac{0.08}{0.00003}} = \sqrt{2707} = 52 \text{ ohms}$$

The value of inductance and capacitance per unit length of line is determined by the cross sectional geometry of the line and the nature of the dielectric medium. The characteristic impedance of typical r-f lines ranges from 50 ohms in the coaxial type to over 600 ohms in the open-wire type.

Actual transmission lines do not extend to infinity, but have a definite length and terminate in a load at the output end. If the load is a pure resistance of a value equal to the characteristic impedance, the line has the characteristics of a line infinitely long. The impedance is then constant along the line and equal to the characteristic impedance. This can be explained by referring to figure 9-2. A theoretical line is shown in figure 9-2A with an infinite number of sections extending to the right. The impedance appearing across the input terminals AB is then Z_0. Now, if the line is cut at CD, an infinite number of sections still extend to the right, since the line is endless in that direction. Therefore, the impedance appearing at CD is also Z_0. Thus, a load equal to Z_0 connected to CD makes the line appear to be of infinite length, and the impedance across the input terminals AB is still Z_0. Therefore, whether the line is infinitely long or fixed in length and terminated in the characteristic impedance, the impedance at the sending end is the same. Such a line is said to be matched. In a matched transmission line, power travels outward until it reaches the load, where it is completely absorbed.

9.2.4 WAVE MOTION ON A MATCHED LINE.

Waves exist on a transmission line as it takes a certain amount of time for energy to travel down the line. This wave motion is illustrated in figure 9-3. Assume that an r-f voltage is applied to the line simulated by the circuit shown in 9-3A. The generator voltage starts from zero and varies with time in a sinusoidal manner as shown. As soon as a small voltage change is produced, it starts its journey down the line, and the generator in the meantime continues to produce new voltages. At time T_3, the generator voltage is 100 volts. However, at this time the voltage at point 1 is 0 and just starting to increase. At time T_5, the first small voltage arrives at point 2 on the line, and finally at time T_7 this small voltage arrives at the load. Meanwhile all the changes in the sine wave pass each point in turn.

At times T_7 and T_8, the voltage at the various points on the line are as follows:

	T_7	T_8
Generator	-100 V	-70 V
Point 1	0 V	-70 V
Point 2	+100 V	+70 V
Point 3	0 V	+70 V

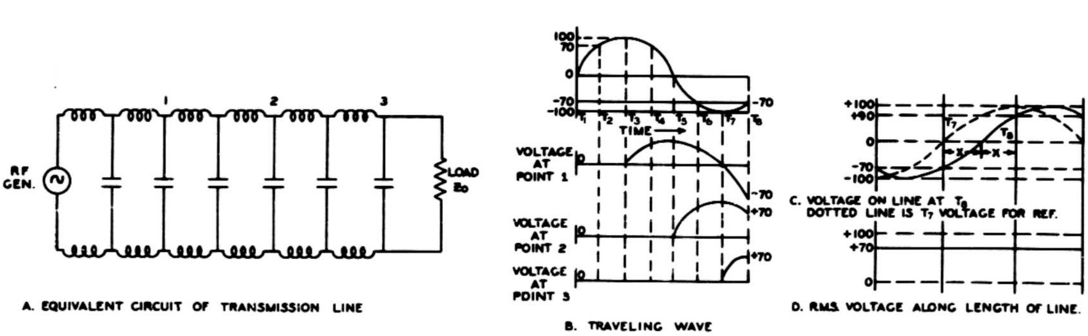


Figure 9-3. Wave Motion on Matched Transmission Line

Figure 9-3C shows these voltages for each time plotted along the length of the line. The two curves have exactly the same shape, but the one for T_8 has moved down the line by the distance X. Another plot at time T_9 would result in a new curve identical in shape to the one at T_8, but removed to the right by distance X.

On a matched line, the voltage and current are always in phase throughout its entire length. Therefore, although figure 9-3 refers only to voltages, the same analysis also applies to the current on the line.

An analysis of figure 9-3 points out the following characteristics of wave motion on a matched line.

a. All parts of the sine wave produced at the generator travel down the line in the order produced.
b. At any one point on the line, a sine wave will be obtained if the instantaneous voltages or currents passing the point are plotted at successive time intervals.

An important point to remember is that the wave motion represents instantaneous values. A voltmeter or ammeter connected to the line will average out the variations in a cycle and read constant values. With low line losses, the ammeter and the voltmeter readings on a matched line will be constant throughout the length of the line as shown in 9-3D. The ratio of the constant voltage to the constant current is the characteristic impedance of the line. When the load is equal to the characteristic impedance of the line, energy is absorbed as fast as it is delivered from the line. Therefore, no change in the phase relationship between the voltage and current results at the load.

9.2.5 WAVELENGTH AND VELOCITY FACTOR.

The distance traveled by the start of one cycle, during the time required to generate the entire cycle, is one wavelength. Therefore, the faster the wave travels, the greater will be the wavelength. The higher the frequency, the less time the start of the cycle has to travel, and therefore the shorter the wavelength. This relationship can be summarized by the equation:

$$\lambda = \frac{V}{f}$$

where: f is the frequency, λ is wavelength, and V is the velocity of the wave.

In free space, the velocity of radio energy is 186,293 miles per second, or 300,000,000 meters per second. However, r-f energy travels slower in transmission lines than in free space. The result is that the wavelength is less on a transmission line than it would be in free space at the same frequency.

The velocity factor of a transmission line is the ratio of the actual propagation velocity along the line to the propagation velocity in free space. This ratio will always be less than one. Values of the velocity figure will differ for the various types of transmission line.

Whenever a line is designated as so many wavelengths long, this means the electrical length of the line. Its actual physical length will be somewhat less, depending on the velocity factor of the transmission line. This is because the wave velocity is decreased on the line, and therefore a wave will be shortened. The physical length corresponding to an electrical wavelength is given by:

$$L \text{ (feet)} = \frac{984}{f} V$$

where: f = frequency in megacycles

V = velocity factor.

9.2.6 STANDING WAVES.

When a transmission line is terminated in a load equal to the characteristic impedance, the current and voltage are in phase and constant in level along the line, and the ratio of voltage to current is equal to the characteristic impedance. Under these conditions, energy is absorbed by the load as fast as it is delivered by the line. However, if the load is not equal to the characteristic impedance, the in-phase relationship between voltage and current is upset, and energy is not absorbed as fast as it arrives. This results in reflection of energy from the load. Under these conditions of mismatch, there are two wave trains on the line: one traveling toward the load called the incident wave, and one traveling away from the load back to the generator called the reflected wave.

The reflected wave travels on the line in the same manner and with the same velocity as the incident wave. As the two waves travel, they combine to produce sinusoidal variations of voltage and current on the line. These waves are fixed in position on the line and are therefore called standing waves.

To illustrate the formation of standing waves, take the case where the line is open at the load end. Under these conditions, the reflection is complete, causing the reflected wave to be equal in amplitude to the incident wave. An open circuit presents an infinite load impedance. Therefore, the resultant voltage at the load should be maximum, and the resultant current should be zero. This is shown in figure 9-4. The reflection takes place without change of phase of voltage so that the voltages of the reflected and incident waves add at the load to give a total voltage twice the voltage contributed by the individual waves. However, the current is reflected in opposite phase so that the resultant load current is zero.

Moving down the line away from the load, the phase relationship between the incident and reflected waves

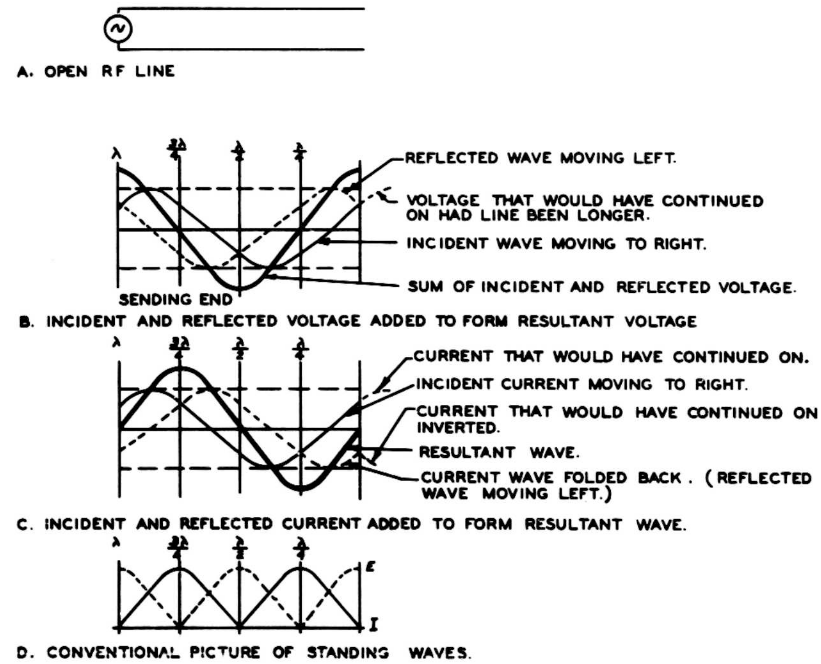


Figure 9-4. Formation of Standing Waves in Open Line

changes as shown in figure 9-4. The incident wave advances in phase, while the reflected wave drops back in phase. At a distance from the load equal to a quarter wavelength, the incident wave has advanced 90° in phase, while the reflected wave has dropped back a similar amount. This causes the voltages of the two waves, which were in phase at the load, now to be in phase opposition. The resultant line voltage at this point is the difference between the voltages of the individual waves. Under conditions of negligible line loss, the resultant voltage at the quarter-wave point will be zero, since the two voltages are equal and completely cancel. However, if there is line loss, attenuation of the line causes the two waves to be unequal in amplitude at a distance from the load, even though they are equal at the load. This causes incomplete cancellation and results in some minimum voltage level other than zero. In either case, we can see from figure 9-4 that a minimum voltage point occurs at a distance of a quarter wavelength from the load end of an open line. At this quarter-wave point, the currents which were in phase opposition at the load now have the same phase and add to produce a large resultant line current.

As the distance from the load is increased to a half wavelength, the voltages again add to produce maximum voltage, while the currents subtract to produce a minimum current. The result is a voltage and current distribution on the line as shown in figure 9-4. Notice that the voltage is maximum at points that are even multiples of a quarter wavelength and minimum at odd multiples. The corresponding current is maximum when the voltage is minimum, and vice versa.

Figure 9-5 illustrates the formation of standing waves when the load end of the line is short circuited. The reflection is complete, as it is with an open circuit, making the reflected wave equal to the incident wave. However, with a short circuit, reflection takes place with the voltage reversed in phase, and the current phase is unchanged. Therefore, the two voltages add up to a resultant of zero voltage, while the two currents add for a maximum current. Since a short circuit represents zero impedance, these conditions of minimum voltage and maximum current are as required by a short-circuited termination. As the distance from the load is increased, one wave advances in phase, while the other lags in phase, exactly as in the case of the open-circuited load. However, since it is now the currents that add at the load end and the voltages that subtract, maximum voltage and minimum current occur at a distance of a quarter wavelength from the load as shown in figure 9-5. Notice that the voltage is maximum at points that add multiples of a quarter wavelength and minimum at even multiples. The corresponding current is maximum at even multiples and minimum at odd multiples.

Another factor to keep in mind is that when an a-c meter is used to determine current or voltage on an r-f line, it will indicate only the magnitude. The polarity will not be indicated. Therefore, if all the readings obtained along the line are plotted, curves similar to those shown in 9-4D and 9-5D will be obtained. These represent the conventional method of showing current and voltage on an r-f line.

When a transmission line is terminated in a resistive load, rather than an open or a short, part of the power is absorbed in the load and the reflected wave is reduced in amplitude. In this case, neither the voltage nor current cancel completely at any point along the line. However, the velocity at which the incident and reflected waves travel is not affected by the change in amplitude, so the phase relationships are similar to those in open or short-circuited lines. The only factor changed is the relative amplitude of the minimum and maximum points as shown in figure 9-6. A smaller value of reflected voltage causes a standing wave with smaller variations.

If the resistive load equals the characteristic impedance, there is no reflection, and therefore there are no standing waves. If the resistive load impedance is greater than the characteristic impedance, the voltage is maximum at the load. If the load is less than the characteristic impedance, the current is maximum at the load.

The resistive load is seldom an actual resistor. The most common terminations are resonant circuits or resonant antenna systems, both of which have essentially resistive impedances. If the load is reactive as well as resistive, the standing waves produced resemble those shown in figure 9-6G and 9-6H. The presence of reactance in the load causes the maximum and minimum points to shift along the line. Also, the amplitude of the waves reflected from a reactive load is greater than the reflected wave produced by a purely resistive load of the same impedance. This is because a reactive load absorbs less of the incident energy.

Any irregularities resulting in sudden changes of impedance in a transmission line will cause reflections on the line. Typical causes of irregularities are bends or turns, insulating supports, extraneous objects that affect the electric or magnetic field such as probes, dielectric or metal bodies, and secondary coupled circuits.

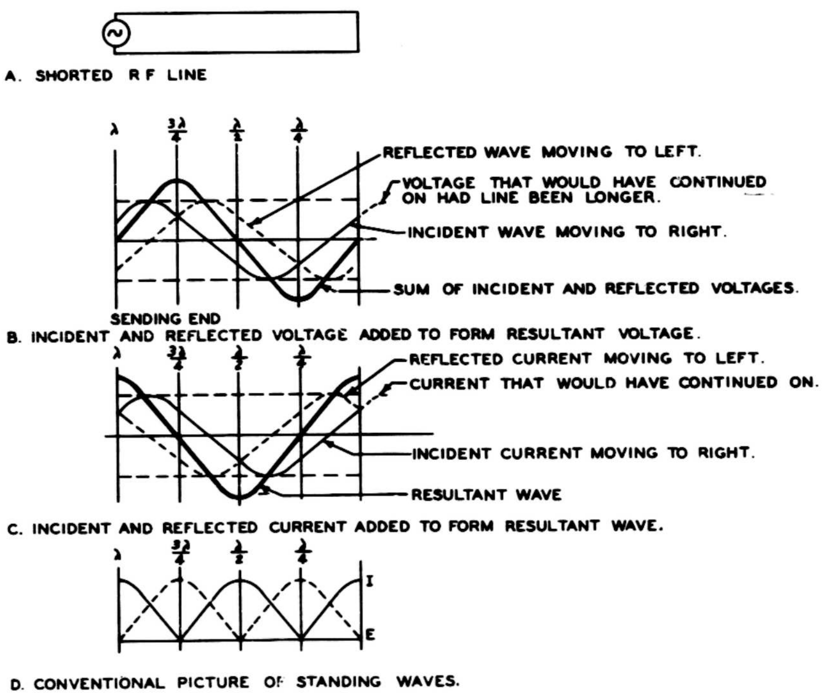


Figure 9-5. Formation of Standing Waves in Shorted Line

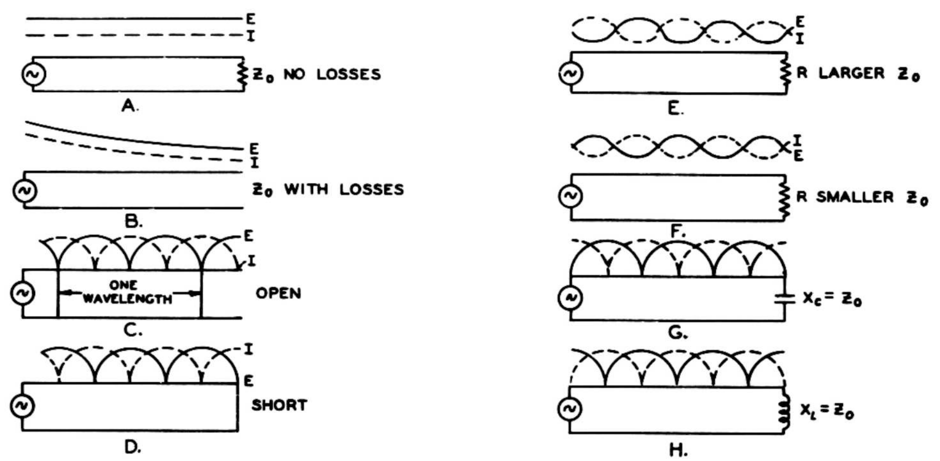


Figure 9-6. Effects of Various Terminations on Standing Waves

9.2.7 STANDING-WAVE RATIO.

The ratio of the maximum voltage (or current) to the minimum voltage (or current) on the line is determined by the degree of mismatch between the load and the characteristic impedance of the line. This ratio is called the standing-wave ratio (swr). The higher the ratio, the greater the mismatch. When the load impedance is equal to the characteristic impedance, the voltage and current are constant along the line, and the swr is 1:1.

When the load is purely resistive, the expression for standing-wave ratio is:

$$swr = \frac{Z_R}{Z_O} \text{ or } \frac{Z_O}{Z_R}$$

where: swr = standing-wave ratio

Z = impedance of load

Z_0 = characteristic impedance of the line.

The larger of the two impedances is placed in the numerator of the fraction so that the swr is never less than 1.

Swr refers to voltage or current unless power is specifically mentioned. The voltage standing-wave ratio (vswr) must be squared to convert to power relationships. The power delivered to the load is equal to the difference between the energy contained in the incident wave and the energy contained in the reflected wave, which is determined by the degree of mismatch at the load end of the line. When both incident and reflected power are known, swr can be determined from:

$$swr = \frac{\sqrt{\frac{P_i}{P_r}} + 1}{\sqrt{\frac{P_i}{P_r}} - 1}$$

where: P_i = incident power

P_r = reflected power.

The power lost in a line is least when the line is terminated in a resistance equal to its characteristic impedance, and increases with an increase in standing-wave ratio, because the effective values of both current and voltage become greater. The increase in effective current raises the ohmic losses in the conductor, and the increase in effective voltage increases the losses in the dielectric. The effect of swr on line loss is shown in figure 9-7. Excessive standing wave currents may cause overheating of the transmission line. Warm spots on the line are a positive indication of standing waves and, consequently, of a mismatched condition between the line and the antenna.

In addition, swr affects the power handling capability of a given transmission line as limited by the insulation breakdown voltage. This is of particular importance in transmission lines of the coaxial type. Most manufacturers derate their lines in direct proportion to the swr; so, in order to operate the line efficiently, it is necessary to maintain a low swr.

9.2.8 TRANSMISSION LINE IMPEDANCE.

The impedance of a transmission line at any particular point on the line is the ratio of the voltage to the current at that point. This is the impedance that a generator must work if it is connected to the line at that point. The amount of variation in impedance along a transmission line depends upon the degree of mismatch between the load and line. If the load is perfectly matched to the line, the impedance of the line is equal to the characteristic impedance at all points. However, if there is a mismatch, there are standing waves, and the line impedance may have a wide range of values. The greater the swr, the greater is the range of impedance.

Figure 9-8 shows the change in impedance along the line for open and short circuited lines. As shown on this figure, the impedance is large at points where the voltage is high, and small when the voltage is low. As a generator is moved down the line, the impedance that it sees changes; or, if the generator remains fixed in position and the frequency is changed, the input impedance will also change since the electrical length of the line changes with frequency.

Study of figure 9-8 shows that if the line is one-half wavelength long, the impedance is exactly the same at the input terminals as they are for the load. This is true for all integral multiples of a half wavelength and for all forms of loads. When a line is a quarter wavelength long or an odd multiple of a quarter wavelength, the input impedance is opposite from the load impedance in magnitude.

In addition to the variation in magnitude of impedance with line length, the presence of standing waves on the line also causes the input impedance to be capacitive or inductive even though the load may be purely resistive. The input impedance is resistive only at the points on the line where the voltage goes through maximum or minimum. At all other distances along the line, the current either leads or lags behind the voltage. The effect is exactly the same as though a capacitance or inductance were part of the input impedance.

Take the case of an open circuit line. For all lengths of open line less than a quarter wavelength, the input impedance is capacitive. When the line is exactly $\frac{\lambda}{8}$, the impedance is capacitive and equal to the characteristic impedance of the line. The reactance increases with the line length up to the quarter-wave point. Beyond the quarter-wave point, the impedance is inductive, maximum near the quarter-wave point and decreasing toward the half-wave point. Midway between the quarter-wave and half-wave points, the impedance is inductive and equal to Z_0. The impedance then alternates between the inductance and capacitance in successive quarter wavelengths.

The impedance characteristic for all lines terminated in a resistive load larger than Z_0 is similar to the

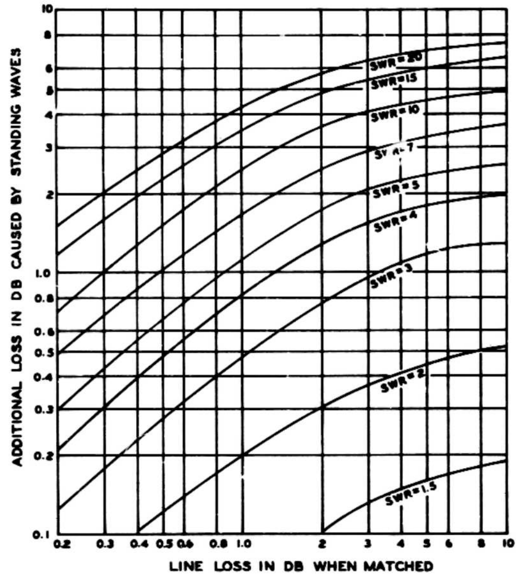


Figure 9-7. Effect of SWR on Line Losses

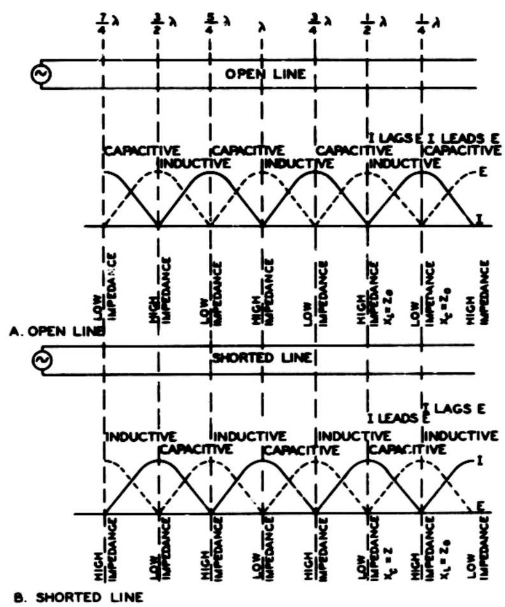


Figure 9-8. Change of Impedance Along Open and Shorted Transmission Lines

open-circuit case. However, with an actual load the reflection is incomplete, and the reflected wave is reduced. This causes the addition of a resistive component to the reactive component.

A shorted line has impedance characteristics exactly opposite to that of an open line. A shorted line less than a quarter wave in length has an inductive reactance. The inductive reactance increases with distance from the short up to the quarter-wave point. Beyond this point, the reactance alternates between inductance and capacitance in successive quarter-wave sections as shown in figure 9-8.

9.2.9 TRANSMISSION LINE SECTIONS AS CIRCUIT ELEMENTS.

The characteristics of low-loss resonant sections of transmission lines make them useful as high Q circuit elements. These applications are discussed in the following paragraphs.

9.2.9.1 LINE SECTIONS AS PARALLEL RESONANT CIRCUITS.

At the minimum current points of the standing wave, the impedance appears as a pure resistance of very high value. This resembles the impedance of a conventional high Q parallel resonant circuit consisting of lumped inductance and capacitance. As can be seen from figure 9-9, various types of lengths of line sections will produce this circuit. However, a quarter-wave shorted section is most commonly used to produce the effects of a parallel resonant circuit. The input impedance will be high, and the line section will act as a parallel resonant circuit only at the high frequency for which the line length represents a quarter wavelength. If the frequency is changed, the electrical length of the line will be changed accordingly, thus effectively detuning the circuit.

9.2.9.2 LINE SECTIONS AS SERIES RESONANT CIRCUITS.

At the minimum voltage points of the standing wave, the impedance appears as a pure resistance having a very low value. This impedance resembles the impedance of a conventional high Q series resonant circuit. Figure 9-9 illustrates that for series resonance, an open-circuited line may be any odd number of quarter-wave lengths long, while a short-circuited line may be any even number of quarter-wave lengths long. A half-wave section of shorted line is most commonly used to produce the effects of a series-resonant line.

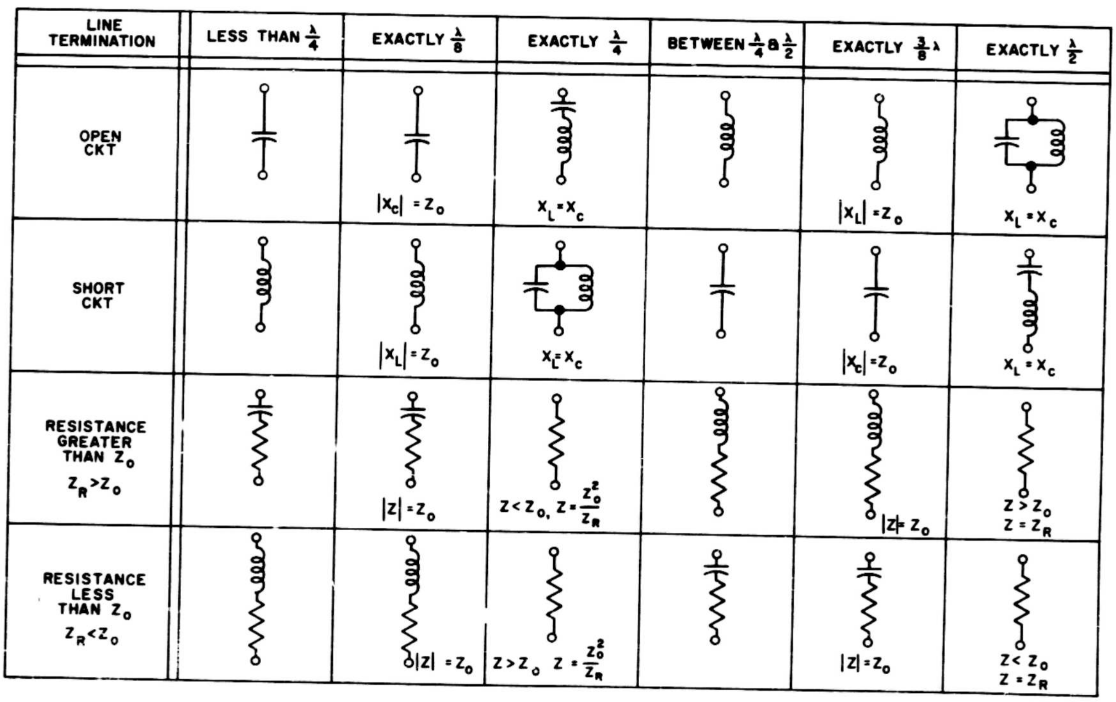


Figure 9-9. Equivalent Circuit for Various Lengths of Transmission Line

9.2.9.3 LINE SECTIONS AS LOW-LOSS REACTANCES.

The reactance produced by linear circuit elements can have almost any magnitude and can be either capacitive or inductive, depending upon the electrical length of the line section and upon the type of termination (open or short-circuited). This is illustrated in figure 9-9 for the lossless line. Losses present in actual line sections result in a small resistive component. In a well-constructed line section, however, the losses are so low that this resistive component usually can be neglected. The characteristics shown in figure 9-9 are repeated for any multiple of an electrical half-wave length that is added.

One of the most common uses of line sections as low-loss reactive elements is for impedance matching on antenna feedlines. In this application, they are ordinarily referred to as stubs and are employed to cancel or eliminate undesirable reactive components of impedance as shown in figure 9-10. Although only current standing waves are shown in the figure, standing waves of both current and voltage exist due to the mismatch at the line termination. For every wavelength on the line, there are four points where the reactive component is equal to the characteristic impedance of the line. If a stub is placed across the line at one of these points and is adjusted so that its reactance is equal in magnitude but opposite in sign to that of the reactive component existing at that point, the line will become "flat" or matched from that point back to the generator. As shown in figure 9-10, the standing waves have been removed only between the stub and the generator. Therefore, to operate a mismatched line with the highest efficiency, the stub should be placed as close to the load as practicable. In addition, this placement will allow operation of the over-all antenna system over a broader frequency range.

9.2.9.4 LINE SECTIONS AS IMPEDANCE TRANSFORMERS.

The properties of a transmission line that enable it to perform as an impedance transformer are illustrated in figure 9-11, where the resultant or apparent impedance is shown to be a function of characteristic impedance and length of the line section.

Theoretically, any value of load impedance can be transformed to any desired value of impedance by the quarter-wave line. In actual practice, the range of impedance transformation is limited by the value of Z_0, which in turn depends on the physical size of the line. Practical lines yield a range of characteristic impedance from about 50 to 600 ohms.

9.3 COAXIAL LINES.

9.3.1 GENERAL DESCRIPTION.

A coaxial line consists of two concentric conductors separated by either air or a solid dielectric material.

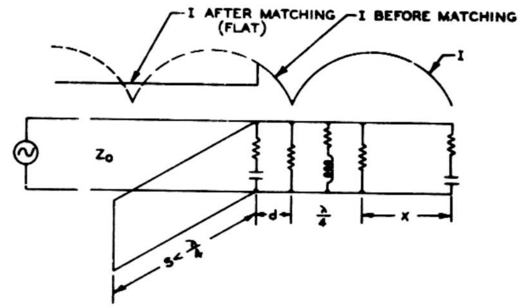


Figure 9-10. Effect of Shorted Stub on SWR Between Stub and Generator

This configuration prevents radiation from the line and pickup of energy from sources external to the line. Coaxial lines are the type used almost exclusively with tropospheric scatter equipment operating in the 1000-mc and 2000-mc ranges because they have a number of advantages over other types of transmission lines at these frequencies. Since radiation and pickup is very low, coaxial lines can be installed anywhere without causing interference to other systems or being influenced by other strong fields. With coaxial construction, it is possible to achieve a much lower characteristic impedance than is practical with other types of line. Also, coaxial lines can be constructed to handle high power with relatively low losses in the uhf range.

9.3.2 AIR DIELECTRIC COAXIAL LINES.

Typical air dielectric lines are shown in figure 9-12. Air dielectric lines are available in the rigid or flexible type and in various diameters, depending on the power handling requirements. Some common sizes for the outside diameter are: 7/8 inch, 1-5/8 inch, 3-1/8 inches, and 6-1/8 inches. The characteristic impedance of air dielectric lines is determined by the dimensions of the line as shown in figure 9-13.

The rigid type shown in figure 9-12A is constructed of two concentric metallic conductors, usually copper. The two conductors are separated by supports made

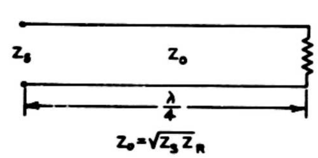


Figure 9-11. Quarter-Wave Line Section as Impedance Transformer

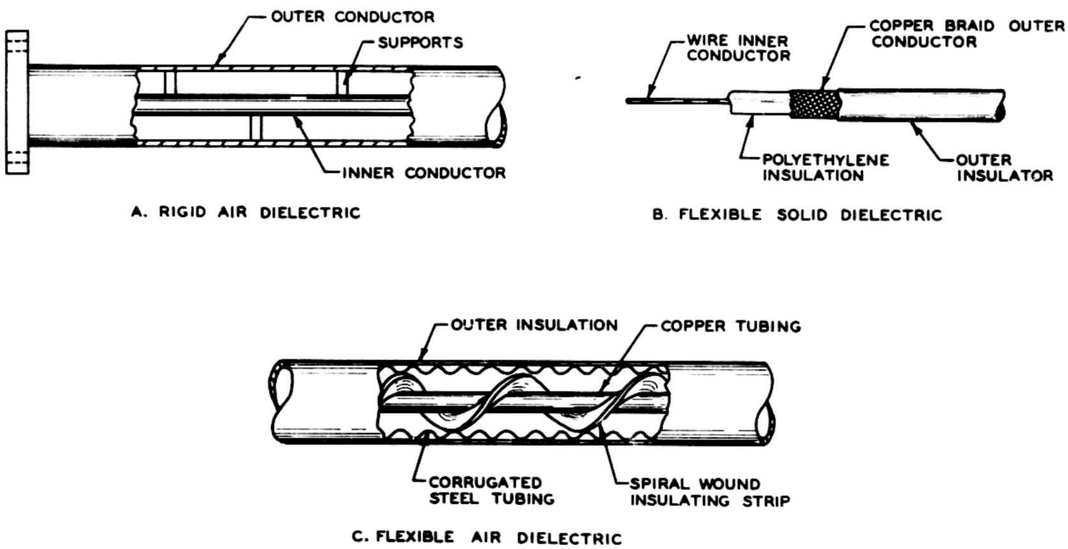
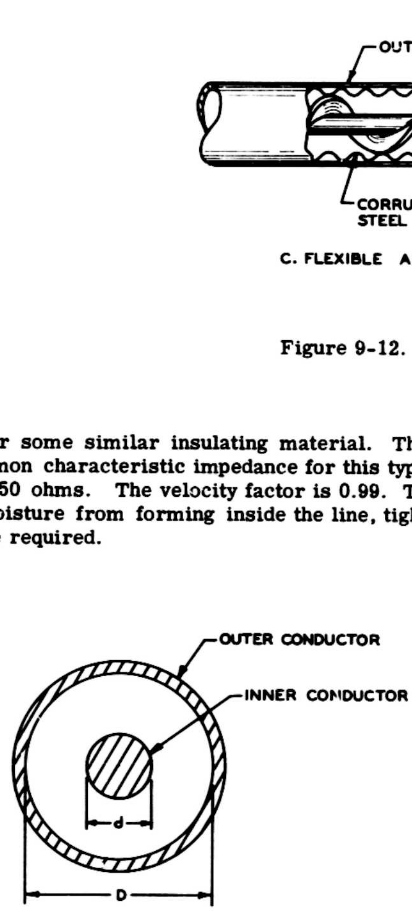


Figure 9-12. Typical Coaxial Lines

of Teflon or some similar insulating material. The most common characteristic impedance for this type of line is 50 ohms. The velocity factor is 0.99. To prevent moisture from forming inside the line, tight fittings are required.



NOTE: INNER CONDUCTOR IS SUPPORTED BY DIELECTRIC BEADS OR SPIRAL SPACERS.

$$Z_0 = 138 \log_{10} \frac{D}{d}$$

WHERE:

D = I.D. OF OUTER CONDUCTOR

d = O.D. OF INNER CONDUCTOR

Figure 9-13. Method for Determining Characteristic Impedance of Air Dielectric Lines

The flexible air dielectric, although having a higher loss than air dielectric for the same size line, has the advantage of ease of installation. A typical line of this type, called Heliax by Andrew Corporation, is shown in figure 9-12C. The flexibility is obtained by making the outer conduction out of corrugated steel tubing and using a spiral insulator which centers the inner conductor within the outer conductor. The characteristic impedance of this type of cable is 50 ohms. The velocity factor is approximately 0.92. Heliax, like all other air dielectric cables, should be maintained under dry pressure to prevent moisture from accumulating inside the cable. Cable fittings are pressure tight.

9.3.3 SOLID DIELECTRIC COAXIAL LINES.

Solid dielectric coaxial lines differ from air dielectric types in that the inner and outer conductors are separated by a dielectric material, such as polyethylene or Teflon. Such cables have the advantage of easy handling and installing; but for a given size, the attenuation is higher, and power handling capabilities are lower than that of air dielectric types. Solid dielectric coaxial lines are frequently used in short lengths for connecting antennas and transmitters to air dielectric cables. A typical cable construction is shown in figure 9-12B. The velocity factor for lines using polyethylene is 0.659, and for lines using Teflon is 0.695. The characteristic impedance of solid dielectric cable may be determined from figure 9-14. The characteristic impedance is reduced from an air dielectric cable of similar dimensions by a factor proportional to the square root of the dielectric constant. Two common types of solid dielectric cable are described in the following table.

ARMY-NAVY TYPE NO.	OUTSIDE DIAMETER (inches)	CHARACTERISTIC IMPEDANCE
RG-8/U	0.405	50
RG-17/U	0.870	52

9.3.4 ATTENUATION OF VARIOUS TYPES OF LINE.

The attenuation of the various types of line increases with an increase in frequency as shown in figure 9-15. Attenuation is given in db per 100 feet of line. The curves show the decrease in attenuation of a given type of line as the size is increased. For example, the attenuation of 3/8-inch rigid air dielectric at 1000 mc is 3.9 db per 100 feet of line. With a 3-1/8 inch line, this attenuation decreases to 0.59 db. Figure 9-15 also shows the high attenuation incurred with solid dielectric lines.

9.4 WAVEGUIDES.

9.4.1 GENERAL DESCRIPTION.

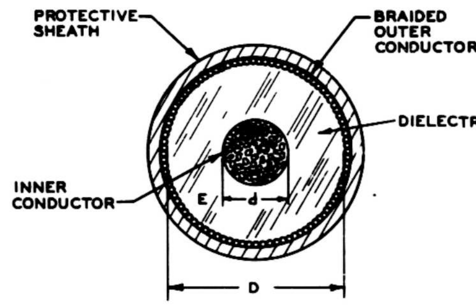
Figure 9-15 shows how losses in a coaxial line increase with frequency. For frequencies above 3000 mc, these losses become prohibitive for long lengths of line. For these frequencies, wave guides provide a practical alternative to coaxial transmission line.

A wave guide is a hollow metal tube, usually rectangular, which has the correct dimensions for conducting uhf radio energy with very little loss. Most of the characteristics of wave guides are similar to those of ordinary two-wire or coaxial transmission line. In particular, the concepts of reflection from a load impedance or an irregularity, standing waves, and impedance matching, previously discussed in connection with transmission line, can be applied directly to wave guide systems. However, wave guides also have a number of properties which do not apply in transmission lines.

One of these properties is the cutoff frequency characteristic. A wave guide does not transmit waves having a frequency less than a critical or cutoff value determined by the guide dimensions. The main requirement for conduction is that one dimension must be more than a half wavelength. This requirement makes wave guides impractical for frequencies other than uhf and shf.

9.4.2 WAVE MOTION IN A WAVE GUIDE.

When a small antenna is placed in the wave guide and excited at an r-f frequency, both positive and negative half-cycles are radiated as shown in figure 9-16A. The fields are the same as those radiated



$$Z_0 = \frac{138}{\sqrt{E}} \log_{10} \frac{D}{d}$$

WHERE:
E = DIELECTRIC CONSTANT
D = I.D. OF OUTER CONDUCTOR
d = O.D. OF INNER CONDUCTOR

Figure 9-14. Method for Determining Characteristic Impedance of Solid Dielectric Line

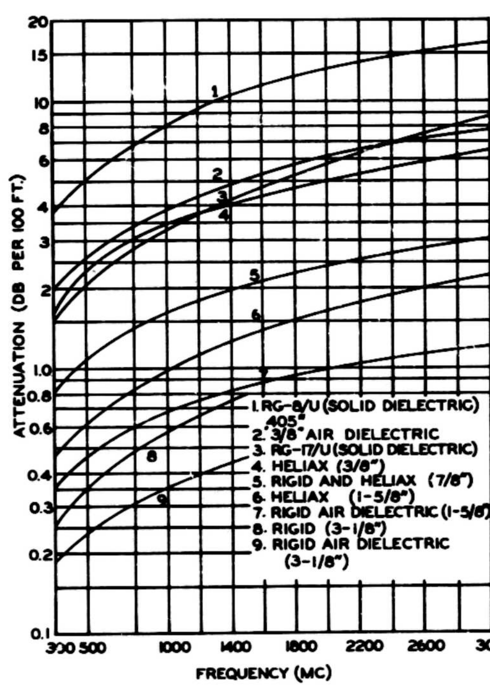


Figure 9-15. Attenuation of Various Types of Coaxial Transmission Lines

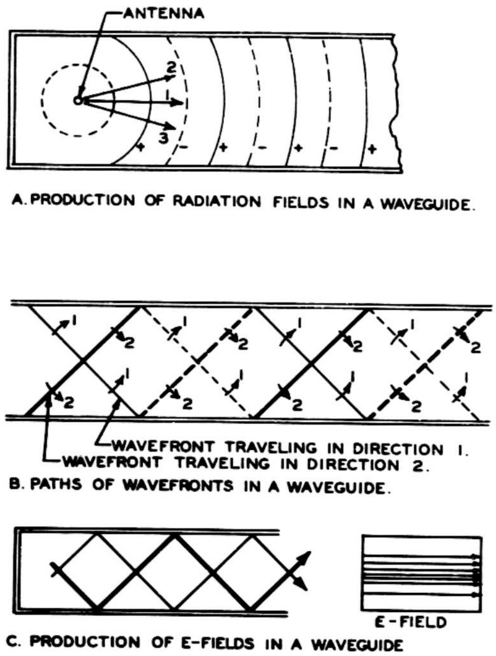


Figure 9-16. Propagation of Radio Waves in Wave Guides

into space by an antenna. The electrostatic line of force or s-lines are parallel to the antenna, and the electromagnetic line of force or h-lines are perpendicular to the antenna. They move away from the antenna at the speed of light. At each half-cycle, the polarity is reversed. Therefore, at half-wave intervals, the fields are in opposite direction (or polarity). The wave front produced is like an expanding circle. The part which travels in the direction of arrow 1 goes straight down the wave guide and is quickly attenuated. However, the part of the wave front which travels in the direction of arrow 2 is reflected from the wall. The wall is a short circuit and causes

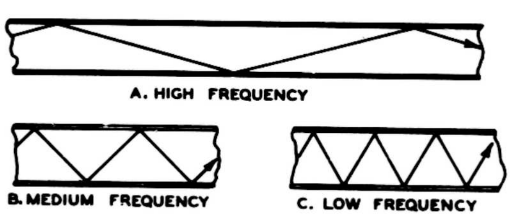


Figure 9-17. Change with Frequency of Angle at which Fields Cross Wave Guide

the wave front to be reflected in reverse phase. Meanwhile, the wave front which travels in direction 3 is reflected from the other wall and proceeds in opposite phase. Thus, the radiation fields are contained in the wave guide.

In the side view of the wave guide at B in figure 9-16, the light solid and broken lines represent the wave front going in direction 1. The heavy lines and dashes represent the wave front going in direction 2. Note that all parts of wavefront 1 are traveling at an angle across the guide. Wave front 2 is traveling at the same angle but downward.

When the wave travels in this fashion in a wave guide, propagation is possible. Note that the positive wave front (represented by solid lines) occurs simultaneously throughout the center of the guide. These fronts add and cause a maximum voltage to occur at the center. (The e-field is shown maximum at the center in diagram C.) The negative wave front adds in the same manner as the positive wave front. When the negative wave front meets the positive wave front at the walls, the two wave fronts cancel each other, making the total voltage equal to zero. This verifies the e-field condition shown at C. With the e-field zero at the edges, it is possible for the e-field to exist in the wave guide.

The angle at which a wave front crosses a wave guide is a function of the wave length and the cross-sectional dimension of the wave guide. At some intermediate frequency, the reflection is as shown at B on figure 9-17; but as the frequency increases, the angle of incidence becomes less, and the signal travels farther before it reaches the other side (see figure 9-17A). At lower frequencies, the wave front crosses the guide at more nearly right angles to the walls. At a certain frequency, the angle will be 90°. At this point, the wave travels back and forth across the guide until the energy is dissipated by the resistance of the walls of the guide. At this cutoff frequency, the distance from side to side is one-half wavelength for the wave guide. At the cutoff frequency, the attenuation is a linear function of length and is very high.

The velocity of propagation of a wave along a two-wire line is less than its velocity in air. The same is true in a wave guide. Movement of a wave along a two-wire line is slower than its movement in air because of the retarding effect of the d-c resistance, the conductors, and conductance of the insulation. In the wave guide, the lower velocity is due to the way the field travels. As shown in 9-17C, the path of a wave front at a relatively low frequency is along the zig-zag arrow at the velocity of light; but due to the long path, the wave front actually travels very slowly along the wave guide. In figure 9-17A, the frequency is higher, and the wave front or groups of waves actually travel a given distance in less time than those at lower frequencies.

The axial velocity of a wave front or a group of waves is called the "group velocity." The relationship of the group velocity to diagonal velocity causes an unusual phenomenon. The velocity of propagation appears to be greater than the speed of light. As is shown in figure 9-18, during a given time a wave front will move from point 1 to point 2, or a distance of L at the velocity of light (V_T). Due to this diagonal movement (direction of the arrow), during this time the wave front has actually moved down the guide only the distance C, which is necessarily a lower velocity. This is called group velocity (V_g); but if an instrument were used to detect the two positions at the wall, they would be the distance P apart. This is greater than the distance L or G. The movement of the contact point between the wave and the wall is at a greater velocity. Since the phase of the r-f has been changed over the distance P, this velocity is called the phase velocity (V_p). The mathematical relationship between the three velocities is stated by the equation:

$$V_L = \sqrt{V_p \; V_g}$$

where: V_L = velocity of light meters/second

V_p = phase velocity

V_g = group velocity.

This equation indicates that it is possible for the phase velocity to be greater than the velocity of light. As the frequency decreases, the angle of crossing is more of a right angle. In this condition, the phase velocity increases. For measuring standing waves in a wave guide, it is the phase velocity which determines the distance between voltage maximum and minimum. For this reason, the wavelength measured in the guide will actually be greater than the wavelength in free space. From a practical standpoint, the different velocities are related in the following manner: If the r-f frequency being propagated is sine wave modulated, the modulation envelope will move forward through the wave guide at the group velocity, while the individual cycles of r-f energy will move forward through the modulation envelope at the phase velocity. Since the standing-wave measuring equipment is affected by each r-f cycle, the wavelength will be governed by the rapid movement of the changes in r-f voltage. Since intelligence is conveyed by the modulation, the transfer of intelligence through the wave guide will be slower than the speed of light, as is the case in other types of r-f lines.

9.4.3 SYSTEM FOR NUMBERING MODES.

The normal configuration of the electromagnetic field within a wave guide is called the dominant mode of operation. For ease in identifying modes, any field configuration can be classified as either a transverse electric mode or a transverse magnetic mode. These modes are abbreviated TE or TM, respectively. In a transverse electric mode, all parts of the electric field are perpendicular to the length of the guide, and no e-line is parallel to the direction of propagation. In a transverse magnetic mode, the plane of the h-field is perpendicular to the length of the wave guide.

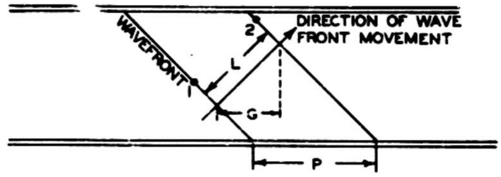


Figure 9-18. Relation of Phase, Group, and Wave Front Velocity

No H-line is parallel to direction of propagation. It is interesting to note from these definitions that the wave front in free space or in a coaxial line is TEM mode, since both fields are perpendicular to the direction of propagation. This mode cannot exist in a wave guide.

In addition to the letters TE or TM, subscript numbers are used to complete the description of the field pattern. In describing field configurations in rectangular guides, the first small number indicates the number of half-wave patterns of the transverse lines which exist along the short dimension of the guide through the center of the cross section. The second small number indicates the number of transverse half-wave patterns that exist along the long dimension of the guide through the center of the cross section. For circular wave guides, the first number indicates the number of full waves of the transverse field encountered around the circumference of the guide. The second number indicates the number of half-wave patterns that exist across the diameter.

In the rectangular wave guide illustrated in figure 9-19A, note that all the electric lines are perpendicular to the direction of movement. This makes it a TE mode. In the direction across the narrow dimension of the guide parallel to the e-line, the intensity change is zero. Across the guide along the wide dimension, the e-field varies from zero at the top, through maximum at the center, to zero on the bottom. Since this

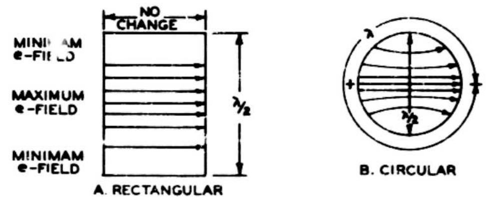


Figure 9-19. Method of Counting Wavelengths for Numbering Modes

is one-half wave, the second subscript is one. Thus, the complete description of the mode is TE_{01}.

In the circular wave guide at figure 9-19B, the e-field is transverse and the letters which describe it are TE. Moving around the circumference starting at the top, the field goes from zero, through maximum positive (tail of arrows), through zero, through maximum negative (head of arrows), to zero. This is one full wave, so the number is one. Going through the diameter, the start is from zero at the top wall, through maximum in the center, to zero at the bottom, one-half wave. The second subscript is one. The complete designation for the circular mode becomes $TE_{1,1}$.

9.4.4 INTRODUCING FIELDS INTO A WAVE GUIDE.

A wave guide, as was explained before, is a single conductor. Therefore, it does not have the two connections which ordinary r-f lines have, and it is necessary to use special devices to put energy into a wave guide at one end and to remove it from the other end.

Wave guides may be excited by three principal methods: electric fields, magnetic fields, and electromagnetic fields.

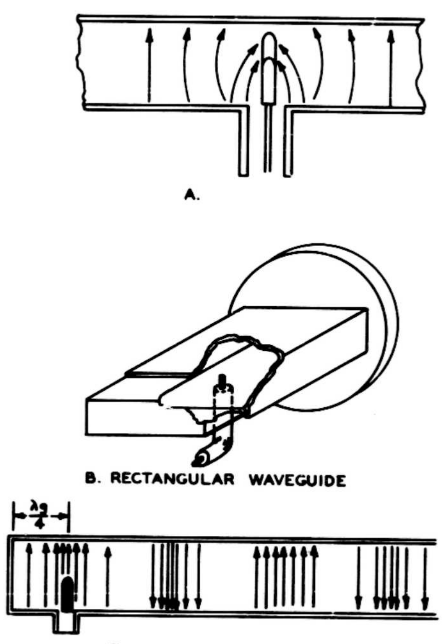


Figure 9-20. Exciting Wave Guide with Electric Field

9.4.4.1 EXCITATION WITH ELECTRIC FIELDS.

When a small probe or antenna is placed in a wave guide and fed with an r-f signal, current will flow in the probe and set up an electrostatic field such as shown in figure 9-20. This causes the e-lines to detach themselves from the probe and to form in the wave guide. When the probe is located in the right place, a field having considerable intensity will be set up. The best place to locate the probe is in the center, parallel to the narrow dimension and one-quarter wavelength away from the shorted end of the guide, as shown at C. Note here that the field is strongest at the quarter-wave point. This is the point of maximum coupling between the probe and the field. Of course, the probe will work equally well in the center of any unidirectional field. For example, a 3/4-wave distance from the shorted end will also be a good spot to place the probe.

Usually, the probe is fed with a coaxial cable. In comparison with the wave guide, this cable is extremely short. This ensures that the greatest benefit will be derived from the wave guide. The device used to connect a coaxial line to the wave guide is called a transition. Impedance matching between the coaxial cable and the wave guide is accomplished by varying the distance from the probe to the end of the wave guide (by moving the shorted end) and by varying the length of the probe. A mismatch will cause unwanted reflections in the wave guide.

The degree of excitation can be reduced by reducing the length of the probe, moving it out of the center of the e-field, or shielding it. Where it is necessary to vary the degree of excitation frequently, the probe is made retractable and the end of the wave guide fitted with a movable plunger.

9.4.4.2 EXCITATION WITH MAGNETIC FIELDS.

Another way of exciting a wave guide is by setting up a magnetic h-field in the wave guide. This can be accomplished by a small loop which carries a high current and placing the loop in the wave guide. This is what happens. A magnetic field builds up around the loop. The field expands and fits the guide. If the frequency of the current is correct, energy will be transferred from the loop to the wave guide. A loop for transferring energy into a guide is shown at A and B in figure 9-21. Notice that the loop is fed

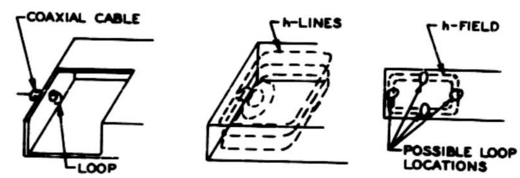


Figure 9-21. Excitation with Magnetic Fields

by a coaxial cable. The location of the loop for optimum coupling to the guide is at the place where the magnetic field which is to be set up is of greatest strength. There are series of places where this is true. Several are shown in figure 9-21C. When less coupling is desired, the loop can be rotated until it encircles a smaller number of lines of force.

When a loop is introduced in a guide in which an h-field is present, a current will be induced in the loop itself. When this condition exists, the loop will take energy out of the wave guide as well as put energy into it.

9.4.4.3 EXCITATION WITH ELECTROMAGNETIC FIELDS.

It might be assumed that a good way to either excite the wave guide or to let energy out of it is simply to leave the end open. This will not work, for when energy leaves a guide, fields form around the end of the guide and cause an impedance mismatch as shown in figure 9-22A. In other words, reflections and standing waves will result if the end is left open. Simply leaving the end open is not an efficient way of letting energy out of the wave guide.

In order for energy to move smoothly in or out of a guide, the opening of the guide may be flared like a funnel as shown in figure 9-22B. This makes the guide similar to a V-type antenna. The "funnel," in effect, eliminates reflection by matching the impedance of free space to the impedance of the wave guide. When the mouth of the funnel is exposed to electromagnetic fields, they enter the funnel where they are gradually shaped to fit the wave guide. The funnel is directional in characteristic. It sends or receives the greatest amount of energy from in front of the opening.

Another method for either putting energy into or removing it from wave guides is through slots or openings. This method is sometimes used when very loose coupling is desired. In this method, energy enters the guide through a small aperture, as you can see in figure 9-22. Any device which will generate an e-field may be placed near the aperture, and the e-field will expand into the wave guide. A single wire is shown in figure 9-22D. On it e-lines are set up parallel to the wire due to the voltage difference between different parts of the wire. The e-lines, in expanding, will exist first across the aperture, then across the interior of the wave guide. If the frequency is correct and the size of the aperture properly proportioned, energy will be transferred to the wave guide with a minimum of reflection.

In order for energy to move from one end of a waveguide to the other without reflections, the size, shape, and dielectric material of the wave guide must be constant through its entire length. Any abrupt change in its size or shape results in reflections. Therefore, if no reflections are desired, any change in the direction or size of the wave guide must be gradual. When it is necessary that the change in direction or size be abrupt, then special devices, such as bends, twists, joints or terminations, must be used.

9.4.5 BENDS.

Wave guides may be bent in several ways to avoid reflections as shown in figure 9-23. One is to make the bend gradual. It must have a radius of bend greater than two wavelengths in order to minimize any reflections. A bend can be made in either the narrow or wide dimension of a guide without changing the mode of operation. In a sharp 90° bend, normally reflections will occur. To avoid this, the guide is bent twice at 45° one-quarter wave apart. The combination of the direct reflection at one bend and the inverted reflection from the other bend will cancel and leave the fields as though no reflection had occurred.

To permit using any special bend which an installation might require, sections of a wave guide are often made

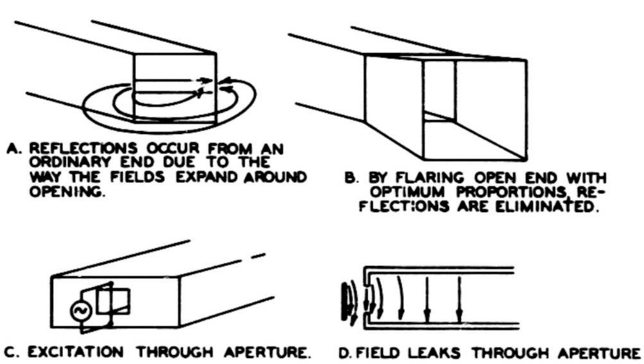


Figure 9-22. Excitation with Electromagnetic Fields

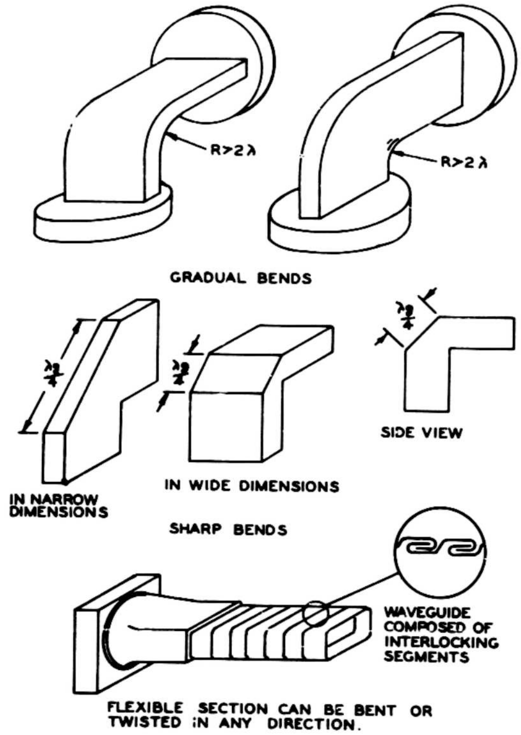


Figure 9-23. Types of Wave Guide Bends

flexible. These sections can be bent or twisted in any desired direction. They consist of a spiral wound ribbon of brass. In cross section, the winding is exactly the same size as a wave guide. The entire assembly is like a spiral spring in that it can be bent or twisted into any desired shape. As skin effect keeps the current at the inner surface of the wave guide, the inside surfaces of the flexible section are chromium plated. This provides for maximum current conductivity. The outside of the section is covered with rubber. This gives the section flexibility, and at the same time makes it both air and watertight.

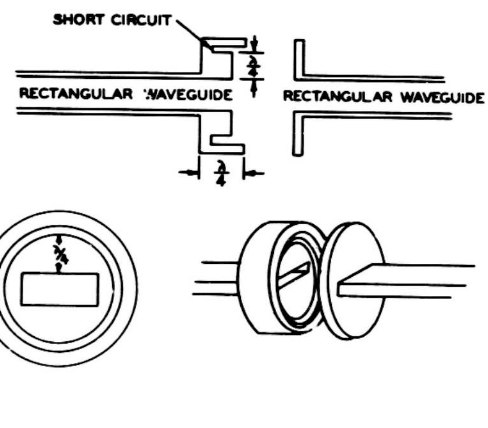


Figure 9-24. Wave Guide Choke Joints

Sometimes it is desired to rotate the electromagnetic fields so that they are in the proper direction for matching. This may be accomplished by twisting the wave guide. The twist should be gradual and extend over two wavelengths or more to prevent excessive reflections. Flexible sections also are used to rotate fields.

9.4.6 JOINTS.

Since it is impossible to mold an entire wave guide system into one piece, it is necessary to construct it in sections and then to connect the sections together by joints. There are three main types of joints: these are the permanent, the semipermanent, and the rotating joints.

On the surface, it would appear that joining two waveguide sections together would only require that the sections be the same size and fit tightly at the joint. However, irregularities at the joints set up standing waves and allow energy to escape. One kind of joint, which affords a good connection between the parts of a wave guide and which has very little effect on the fields, is the permanent type. This joint is made at the factory. When it is used, the wave guide sections are machined within a few thousandths of an inch and then welded together. The result is a hermetically sealed and mirror-smooth joint.

Where it is necessary that sections be taken apart for normal maintenance and repair, it is impractical to use a permanent joint. To permit portions of the wave guide to be taken apart, they are commonly connected together with semipermanent joints. The most common type of semipermanent joint is the choke joint shown in figure 9-24. It consists of two flanges which are connected to the wave guide at the center. The right-hand flange is flat, and the one at the left is slotted a quarter wave deep at a distance a quarter wave from the point where the walls of the guide are joined. The quarter-wave slot is shorted at the end. The two quarter waves together become a half wave and reflect a short circuit at the place where the walls are joined together. Electrically, this creates a short circuit at the junction of the two wave guides. The two sections actually can be separated as much as a tenth of a wavelength without excessive loss of energy at the joint. This separation allows room to seal the interior of the wave guide with a rubber gasket for pressurization. The quarter-wave distance from the walls to the slot is modified slightly to compensate for the slight reactance introduced by the short space and the open circuit from the slot to the periphery of the flange.

9.4.7 T-JUNCTIONS.

Sometimes it is desirable to connect a section of wave guide into the side of another wave guide. This type of connection forms a T-junction. It may be connected either in the narrow side as shown in figure 9-25A, or in the wide side of the wave guide as shown in diagram 9-25B. When the T-junction is in the plane of an h-field, it is called an h-type junction; and when the junction is in the plane of the e-lines, it is called and E-type junction.

The h-type junction effectively is a parallel connection with the main line. For example, when the end of the T-joint shown at 9-25A is short circuited at a distance of a half wave from the center of the waveguide, the result is the equivalent parallel circuit shown in the figure. Note that this figure shows a half-wave section which is connected to a wire line. This section will reflect a short circuit at the line and will not allow any energy to pass. Similarly, in the wave guide itself, a short circuit is reflected to the center where the e-lines are supposed to be. Since an e-line cannot exist at a short circuit, no energy will pass that point. If the shorted end of the T-section were only a quarter wave from the center of the waveguide, an open section would be reflected there, and passage of energy would be unaffected.

The E-type joint effectively is a series connection with one side of the main line, as can be seen in figure 9-25B. In this case, if the section added to the wave guide is a half wave long, it will act as a short at the junction, but will allow the energy to pass. The length of the added section will have to be a quarter wave in order to open the circuit at the junction and stop the ordinary flow of energy past it.

In practice, the actual length of the sections in T-junctions is not an exact quarter- or half-wave length because the e-fields and h-fields are not perfect around the T-junction. A distortion around the fringe called "fringe effect" requires some variations from exact wavelengths.

9.4.8 MATCHING DEVICES.

Effects of undesired inductances and capacitances can be tuned out with small fins or plates in a wave guide. Figure 9-26 shows a number of reactive plates which are deliberately used to introduce capacity or inductance in a wave guide. When these plates are employed as shown in figure 9-26A, they set up oscillations in the higher modes. Since a wave guide is too small for higher modes at the same frequency, these frequencies are not propagated, but remain in the vicinity of the plates. If the edges of the plates are vertical with respect to the plane of the h-field, the modes produced are the TM type. The effect of this on power flow is that of inductance across the two-wire line. This causes reflections and a shift in the standing-wave pattern. The wider the space between the plates, the greater the inductive reactance.

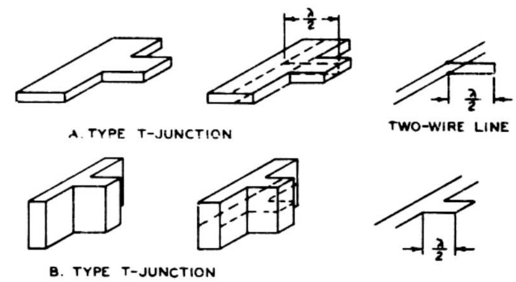


Figure 9-25. T-Junctions

When the partitions are arranged perpendicularly to the e-field as in figure 9-26B, a local e-field and the higher modes of oscillation are set up between the edges of the plates. These oscillations cannot be propagated, but do change the dominant mode to a TE mode and introduce capacitive reactance. As with the TM mode, the wider the opening, the greater the reactance.

From these facts, it would seem that by combining both types of plates and leaving a small opening in a large guide, as in figure 9-26C, would produce a resonant circuit. This is approximately true, provided the dimensions are correct. At resonance, a resonant circuit acts like high resistance. In this condition, a small opening would introduce a high shunt resistance, and the guide would, in effect, have connected across it a resonant circuit, since at resonance a resonant circuit acts like a high resistance.

9.4.9 TERMINATING A WAVE GUIDE.

Since a wave guide is a single conductor, it is not as easy to define its characteristic impedance (Z_0) as it is for a coaxial line. Nevertheless, you can think of the characteristic impedance of a wave guide as being approximately equal to the ratio of the strength

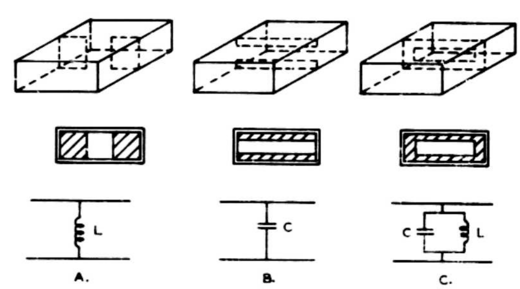


Figure 9-26. Reactive Plates in Wave Guide

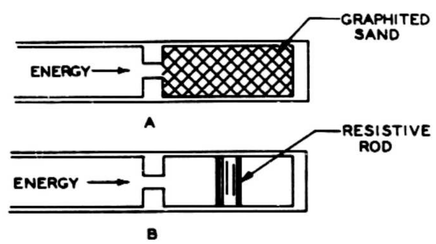


Figure 9-27. Terminations for Minimum Reflection

of the electric field to the strength of the magnetic field for energy traveling in one direction. This ratio is equivalent to the voltage-to-current ratio in coaxial lines on which there are no standing waves.

The lowest characteristic impedance of a circular wave guide is about 350 ohms. In a rectangular wave guide, it may be any value, depending on the dimensions of the wave guide and the frequency of the electrical energy. In this guide, it is directly proportional to the narrow dimension when the other dimension and the frequency are fixed and may vary from approximately zero to 465 ohms.

On a wave guide, there is no place to connect a fixed resistor to terminate it in its characteristic impedance, as there is on a coaxial cable. However, there are a number of special arrangements which accomplish the same thing. One consists of filling the end of the wave guide with graphited sand as shown in figure 9-27A. As the fields enter the sand, currents flow in it. These currents create heat, which is instrumental in dissipating energy. None of the energy thus dissipated as heat is reflected back into the guide. Another arrangement, shown in figure 9-27B, uses a high resistance rod which is placed at the center of the e-field. The e-field (voltage) causes current to flow through the rod. The high resistance of the rod dissipates the energy as an I^2R loss.

9.5 TRANSMISSION LINE FILTERS.

9.5.1 GENERAL DESCRIPTION OF FILTER SYSTEM.

Two antennas are used at a tropospheric scatter terminal to provide dual-diversity operation. When the transmitter and receiver are tuned to different frequencies, a system of filters in the transmission lines permits one antenna to be used for simultaneous transmission and reception. This is called diplex operation. The transmit and receive frequencies must be separated by a given amount, depending on filter design characteristics.

Figure 9-28 is a block diagram of a typical filter system for a terminal. A diplexing filter connected in one antenna line isolates the transmitter and receiver when the antenna is used for simultaneous transmission and reception. A receive band pass filter connected in each receiver input line acts as a preselector for the receiver to prevent interference from adjacent transmitter or antenna systems. Harmonics in the power amplifier outputs are attenuated by low-pass filters in the power amplifier output lines.

The diplexing filter may be replaced by a bipolarized feed system described in paragraph 9.6.3.2. This system used with the receiver input filters provides the required isolation between the transmitter and receiver.

9.5.2 DIPLEXING FILTER.

The diplexing filter is a two-section filter shown in figure 9-28. Each section can be tuned to pass one band of frequencies and reject another.

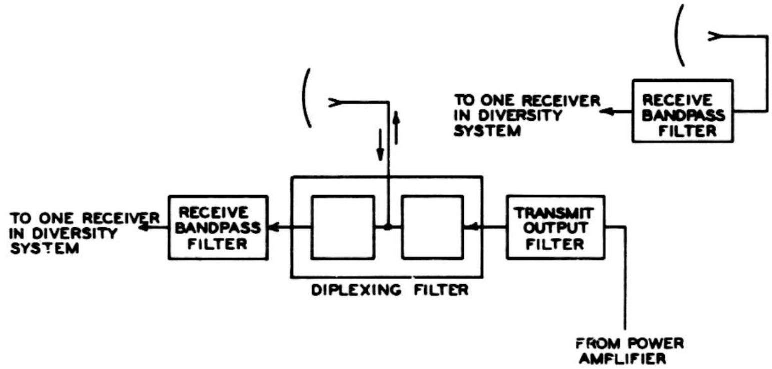


Figure 9-28. Typical Filter System for a Tropospheric Scatter Circuit

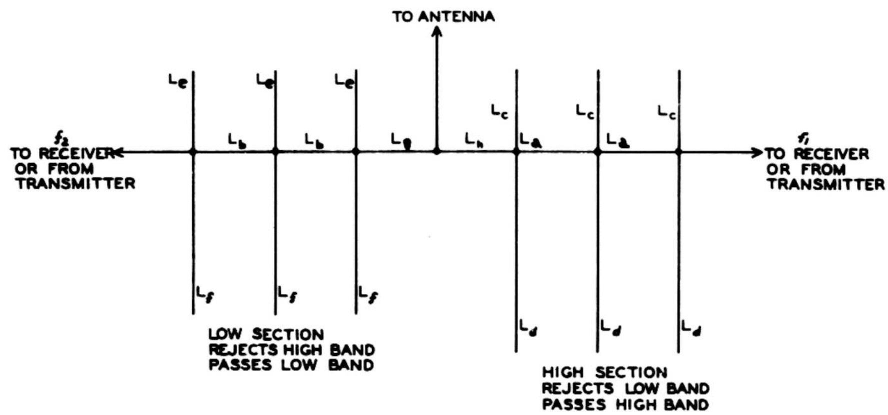


Figure 9-29. Schematic Diagram of Coaxial Diplexing Filter

Figure 9-29 is a schematic diagram of a diplexing filter constructed of coaxial line sections. The solid lines on this figure represent the center conductors of the coaxial lines. The filter sections are tuned with shorting discs on both ends of the cross lines. The long stubs of the cross lines (Lf and Ld) are tuned for maximum rejection at the rejection frequencies for the filter section. The short stubs of the cross lines (Le and Lc) are tuned for the best match at the frequencies to be passed by the filter section. The three longer stubs of the cross lines of each filter section must be an integral number of the half wavelengths at the rejection frequency for the filter section. The total length of each cross line must be an integral number of half wavelengths at the frequency to be passed by the filter section. The number of half wavelengths is dependent upon the separation between the transmit and receive frequencies.

In the pass band, the voltage and current distribution along the cross lines is a sine function with maximum current and minimum voltage at the shorted ends. The ratio of the main line current to the maximum current in the line is proportional to the ratio of the length of the line Le or Lc to the wavelength at the pass frequency for the filter section. The power rating of the cross lines is limited by the heat generated due to ohmic losses. The power dissipated as heat is proportional to the current squared. For minimum heat dissipation in the cross lines, the short stubs should be near a quarter wavelength at the pass frequency. For a given frequency separation, heat dissipation in the cross lines is controlled by setting the ratio of the short stubs to the pass frequency wavelength.

The filter is fabricated from copper or copper alloys. The movable shorts and spring fingers are silver plated, while the surfaces on which the plungers slide are gold plated to maintain good contact in retuning even after long periods of time. The gold finish on the exterior prevents corrosion.

Fittings are standard 3-1/8 inch OD flanges with locating pins. The center conductor connector is a standard bullet for 50-ohm, 3-1/8 inch OD coaxial line. The receiving end may be equipped with adapters for 1-5/8 inch rigid coaxial line or RG-8/U.

9.5.3 RECEIVE BAND PASS FILTER.

Additional attenuation of transmitter interference in diplexer systems is provided by the receive bandpass filter which acts as a preselector for the receiver input circuit on any preset frequency.

One filter used for this purpose consists of four coaxial cavities, each approximately a quarter wavelength long, which are coupled by 5/8-inch rigid coaxial line. The cavities are spaced approximately three-quarters wavelength apart. Tuning is accomplished by threaded capacitive tuning slugs at the top of each cavity. The cavities are coupled by tapping the center conductor in each cavity at a precise position from the shorted end.

A flat band pass response is produced by the four filter cavities which act as high Q resonant circuits. A high impedance appears across the transmission line connecting the cavities when the cavities are tuned to resonance. With frequencies off resonance, the low impedance of the tapped section nears zero, and high reflections occur on the line, preventing the signal from passing through the filter. If the reflected power from each cavity is 97 per cent of the incident power, attenuation would total about 60 db, which is the approximate response of the filter at 50 mc from center frequency.

Power is dissipated in the cavities as well as in the lines because currents are higher in the cavities than in the line. This loss amounts to approximately 0.3 db over most of the band.

The cavities are constructed from copper sheet, and the connecting lines and tuning slugs, from brass. All are gold-plated to prevent oxidation. The filter assembly is clamped on a rigid base plate.

The two end sections of the filter are fitted with tapered connectors to fit with 50-ohm, 1-5/8 inch rigid coaxial line. An adapter is used when connection to 3-1/8 inch rigid coaxial line is required for the branching filter.

9.6 ANTENNAS FOR TROPOSPHERIC SCATTER PROPAGATION.

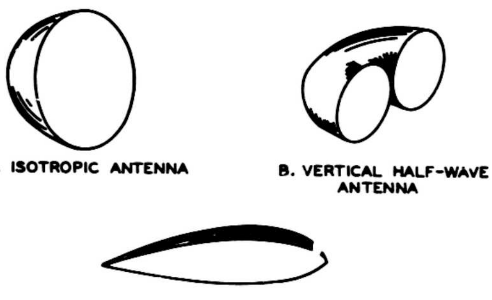
9.6.1 GENERAL ANTENNA CHARACTERISTICS.

9.6.1.1 FUNCTION OF ANTENNA.

The receiving and transmitting antennas, together with the intervening medium, perform the function of the transmission line in a wire communication system. The function of the transmitting antenna is to radiate the radio-frequency energy that is generated in the transmitter and guided to the antenna by the transmission line. In this capacity, the antenna acts as an impedance matching device to match the impedance of the transmission line to that of free space. In addition, the transmitting antenna should direct the most energy in the desired direction and suppress the radiation in other directions. The receive antenna intercepts the electromagnetic fields from the required direction.

9.6.1.2 ANTENNA RECIPROCITY.

Antenna characteristics are essentially the same regardless of whether an antenna is transmitting or receiving electromagnetic energy. This property of identical characteristics is called reciprocity. Because of antenna reciprocity, the same antenna can be used for both transmitting and receiving. This is done at tropospheric scatter terminals, where one of the antennas used for diversity reception is also used for transmitting.



A. ISOTROPIC ANTENNA

B. VERTICAL HALF-WAVE ANTENNA

C. DIRECTIONAL ANTENNA

Figure 9-30. Solid Radiation Patterns

9.6.1.3 POLARIZATION.

The energy leaving an antenna is in the form of electric (E) and magnetic fields (H). These fields are perpendicular to each other, and both are perpendicular to the direction of propagation. The polarization of the radio wave is defined as the orientation of the electric field with respect to the earth's surface as follows:

Horizontal polarization: electric field lies in a plane parallel to the earth's surface.

Vertical polarization: electric field lies in a plane perpendicular to the earth's surface.

The E plane of an antenna is the plane in which the electric field lies. The H plane is the plane in which the magnetic field lies. The H plane is perpendicular to the E plane.

9.6.1.4 ANTENNA DIRECTIVITY.

If the radiation is measured from a single point in space, it will be found that this point radiates equally well in all directions. The strength of the radiated field varies inversely with distance. If all the points where the energy is of the same strength are plotted, the points will form a sphere, with the radiating point at the center. This hypothetical point which produces equal radiation in all directions is called an "isotropic antenna." The solid radiation pattern of such an antenna is shown in figure 9-30A.

An actual antenna does not radiate uniformly in all directions. A plot of all points of equal strength around an actual antenna will produce a three-dimensional figure which will differ from a sphere, depending on the directive properties of the antenna. For example, a plot of all points of equal strength around a half-wave antenna in space would produce a doughnut-shaped, three-dimensional figure as shown in figure 9-30B. A highly directive antenna would produce a radiation pattern similar to that shown in figure 9-30C. The radiation pattern of an antenna will be affected by the terrain characteristics and immediate surroundings of the antenna.

The variation of signal strength around an antenna system can be shown graphically by polar diagrams. These are circular charts which resemble the face of a compass. The antenna position is marked as the center of the circle, and the circumference of the circle is marked in degrees. A polar diagram is constructed by measuring the field strength at the same distance in all directions from the antenna and plotting the field strength for the various angular positions.

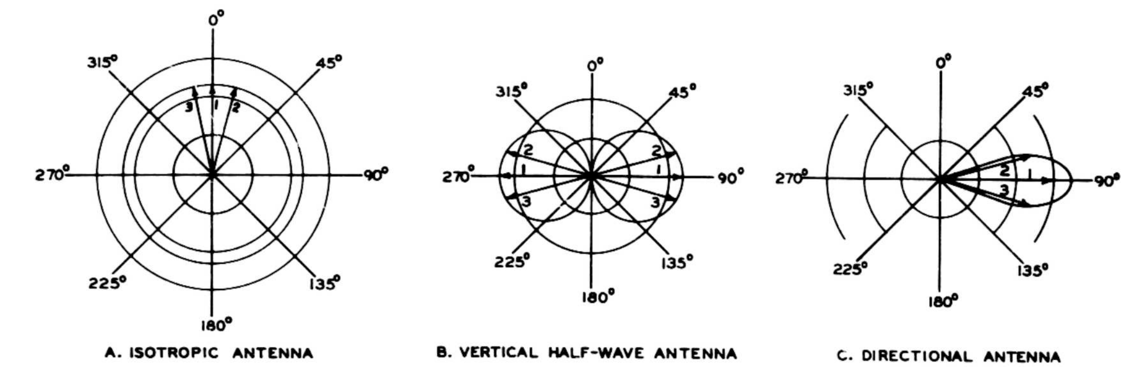


Figure 9-31. Polar Diagrams

Polar diagrams corresponding to the solid radiation patterns shown in figure 9-30 are included in figure 9-31. Note that each polar diagram is a cross-sectional slice of the corresponding solid radiation pattern in a plane including the antenna. The relative intensity of radiation in any given direction is proportional to the length of the line drawn from the antenna to the parametric of the diagram In the case of the isotropic antenna, the field strength is the same for all directions from the antenna. In the other two diagrams, lines 2 and 3 indicate points of lower intensity than line 1.

The more that antenna radiation is concentrated in a certain direction, the greater will be the field strength produced in that direction for a given amount of total energy radiated. For example, refer to the polar diagrams in figure 9-31. Suppose that the isotropic antenna and the directive antenna are radiating the same total power. The radiation along line 1 will be much higher for the directive antenna than for the isotropic antenna since the energy is concentrated in that direction for the directive antenna.

The gain of an antenna is a measure of how well the antenna concentrates its radiated power in the desired direction. It is the ratio of the power radiated in a given direction to the power radiated in the same direction by a standard reference antenna. The gain can be expressed with the reference to any antenna if the reference is specified. The reference antenna is commonly the hypothetical isotropic antenna which is assumed to radiate equally well in all directions. Another reference antenna is the simple half-wave dipole, which has a gain of 1.64 in power over the isotropic radiator. This means that, in the direction of maximum radiation, the dipole will produce the same field strength as an isotropic radiator which is radiating 1.64 times as much power. For convenience, the power gain of an antenna system is frequently expressed in decibels:

$$G = 10_{1} \log \frac{P_1}{P_2}$$

where: $\frac{P_1}{P_2}$ = power gain

9.6.2 BASIC REQUIREMENTS FOR TROPOSPHERIC SCATTER.

Tropospheric scatter propagation is a high-loss process. Therefore, to obtain reliable communications between two points, the radiated energy must be concentrated in a narrow beam. This requires highly directive antennas. At uhf frequencies, the required directivity is most easily obtained with parabolic antenna systems. Besides being highly directive, these systems are comparatively simple, with no critical tuning adjustments, and are broadband devices. The following paragraphs describe operation of parabolic antenna systems.

9.6.3 PARABOLIC ANTENNA SYSTEMS.

9.6.3.1 GENERAL DESCRIPTION OF OPERATION.

A parabolic antenna system consists of a primary feed system and a parabolic reflector. The primary feed system is located at the focus of the reflector. The reflector concentrates radiation from the focus to form a beam in the same way that a searchlight reflector produces a light beam. The focusing action of a parabolic reflector is illustrated in figure 9-32. If a point source of radiation is placed at the focus, spherical wave fronts are produced. Parts of the

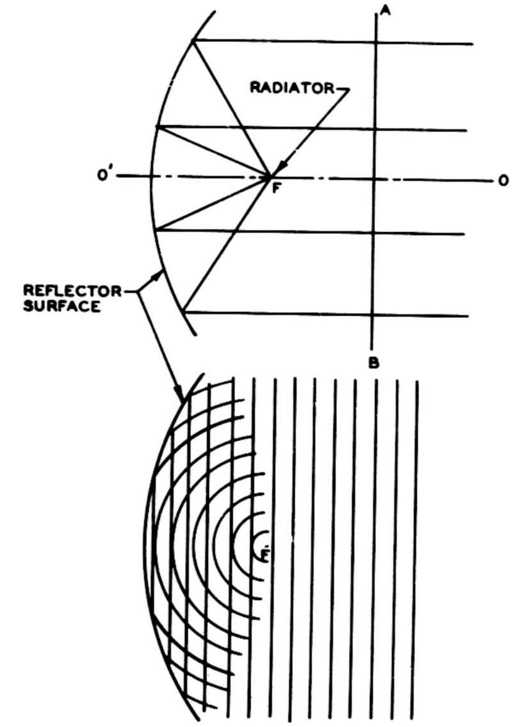


Figure 9-32. Focusing Action of Parabolic Antenna

wave front reaching the reflector are reflected at an angle of reflection equal to the angle of incidence. All parts of the field will arrive at line AB at the same time after reflection because all paths from F to the reflector and then to the line AB are equal in length. This will be true for any line drawn perpendicular to the line 00'. For any line not perpendicular to the line 00', the wave fronts will arrive out of phase and cancel. The cancellation increases as the angle increases. The result is that energy is radiated in a narrow beam in a direction parallel to line 00'.

High directivity is not obtained until the diameter of the parabolic reflector is made many wavelengths long. This prohibits the use of parabolic reflectors at low frequencies.

Another factor which determines directivity of a parabolic system is the radiation pattern of the feed system. All of the energy from the feed system should be directed against the reflector. Energy that is not directed against the reflector does not contribute to the directive beam. This reduces the gain of the antenna system. Loss of energy is prevented by shaping the parabolic reflector so that the focus lies outside the mouth and then using a directional antenna for the feed.

9.6.3.2 FEED SYSTEM.

The feed system must be properly located and have the correct radiation pattern for the most efficient illumination of the reflector. If the radiation pattern of the feed system is too sharp, very little power is lost outside the reflector; but the edges of the reflector do not contribute to the gain, and therefore the effective area of the antenna is reduced. If the radiation pattern of the feed system is too broad, power is radiated beyond the edges of the reflector without being utilized. Therefore, the radiation pattern of the feed system is a compromise between uniform illumination of the reflector and minimum power loss beyond the edges. The directive feed systems used produce illumination which is maximum at the center of the reflector and gradually decreases to the edges. The ratio of the illumination at the center to that at the edge of the reflector is called illumination taper. This is normally expressed in db. The usual compromise choice for taper is 10 db.

Besides providing the required illumination, the feed system must match the transmission line to space through the required operating frequency range.

The type of feed antenna most often used is the wave guide horn. (Typical horn radiators are illustrated in figure 9-33). A wave guide horn consists of a wave guide flared so that the fields in the wave guide gradually expand to produce a field distribution across the mouth of the horn.

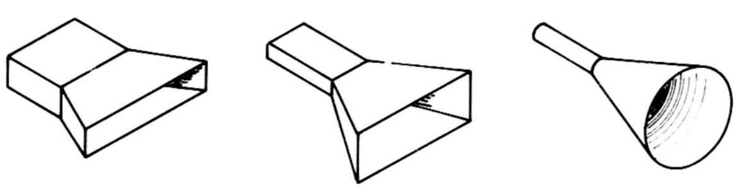


Figure 9-33. Different Types of Horn Radiators

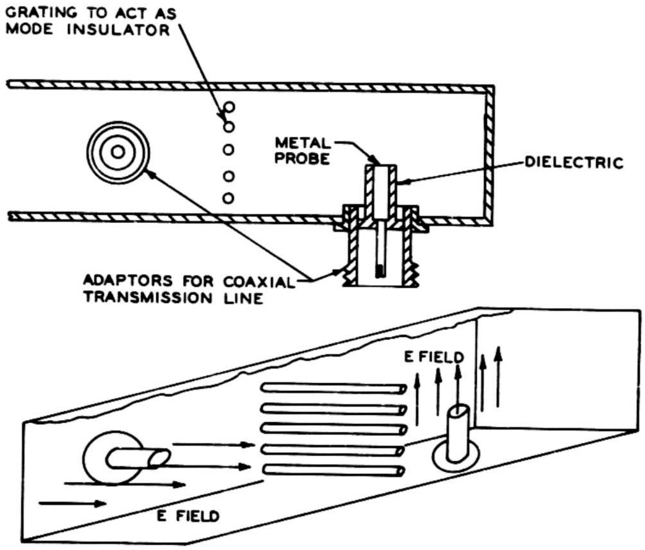


Figure 9-34. Bipolarized Feed System

The radiation pattern of the horn is determined by the flare angle and the size of the aperture. The beamwidth is decreased by an increase in size of the aperture or a decrease in the flare angle of the horn. In general, the larger the opening of the horn, the more directive is the resulting radiation pattern. Since horns do not involve resonant elements, they have the advantage of being usable over a wide frequency range.

The horn may be fed with either wave guide or coaxial line. With wave guide, the input dimensions of the horn are made the same as the wave guide dimensions. The horn then provides an impedance transformation between the wave guide and free space. When coaxial line is used, a probe transition similar to that shown in figure 9-20 is used to provide input from the coaxial line to the horn. These probes are frequency sensitive, and therefore a coaxial line input system is more difficult to use over a wide band of frequencies.

Bipolarized feed systems may be used to provide isolation between a transmitter and receiver connected to the same antenna. Figure 9-34 is a cross section of a bipolarized feedhorn. The horn must be square and flared in both planes. Energy is fed to the wave guide from coaxial cable inputs through two probes inserted into the wave guide. These probes are perpendicular to each other, so each excites one mode in the guide. The two probes are separated by a grating which passes only one mode and shorts out the other.

9.6.3.3 RADIATION PATTERNS.

The major portion of energy radiated from a parabolic antenna system is confined to a small cone of nearly circular cross section. The width of the beam is normally given as the number of degrees between the points on either side of the maximum radiation direction where the power drops to one-half maximum.

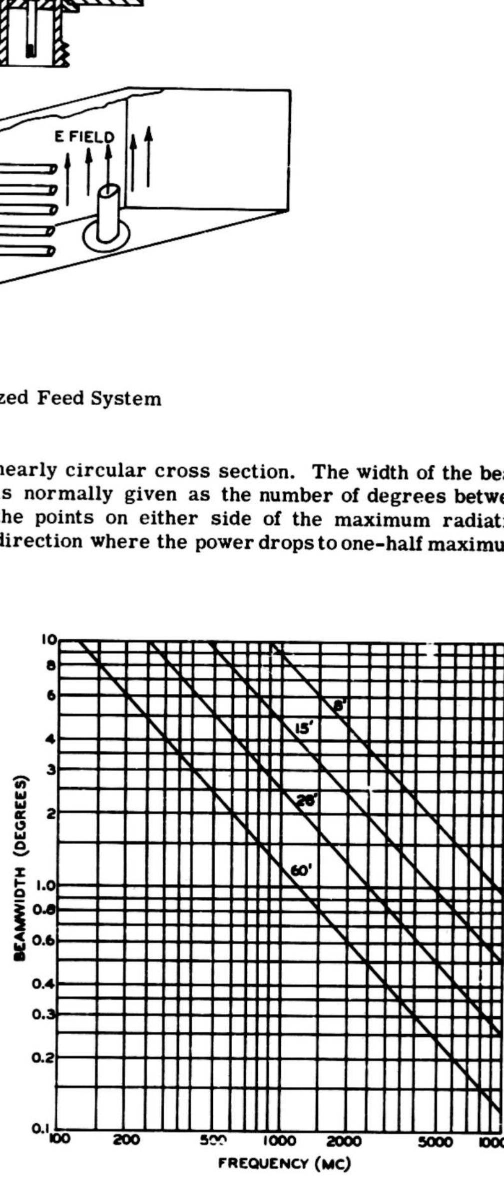


Figure 9-35. Beam Width versus Frequency for Various Reflector Diameters

These are called the 3-db or half-power points on the radiation pattern.

If it is assumed that the intensity of illumination at the edge of the reflector is 10 db down from the center, the width between the half-power points of the main radiation beam is approximately $\theta = \frac{70\lambda}{D}$, where θ is the beamwidth in degrees, is wavelength of radiated energy, and D is the diameter of the reflector. This expression shows that for a given operating frequency, the beamwidth will be decreased as the reflector diameter is increased. Also, the directivity of a given size reflector will increase with an increase in operating frequency. Figure 9-35 shows the relationship between beamwidth and frequency for various size reflectors.

Typical patterns for various combinations of frequency and reflector size are given in figure 9-36. Side lobes are centered around the main beam as shown. These are caused by imperfect aperture illumination, irregularities in the reflector surface, and phase error introduced by the feed. This phase error results in incomplete cancellation in the direction of the side lobes. Ideally, a point source at the focal point radiating perfect spherical wave fronts is required for complete elimination of side lobes. A practical feed will give certain phase error, causing side lobes. Figure 9-36 shows that the position of the side lobes changes as the ratio of $\frac{\lambda}{D}$ is changed.

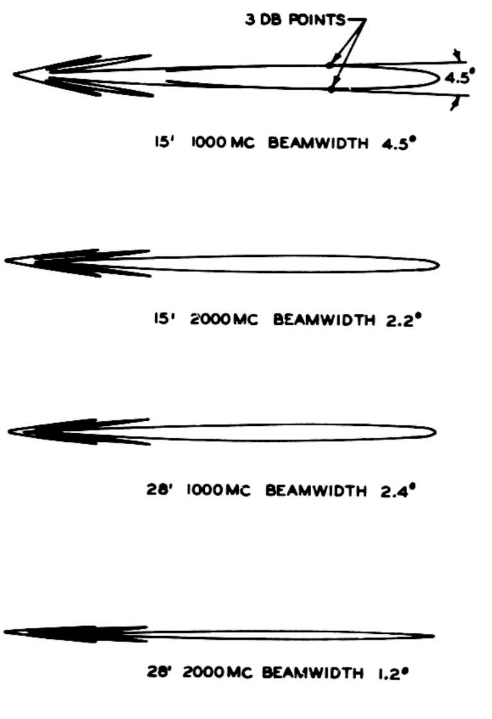


Figure 9-36. Typical Radiation Patterns for Parabolic Antenna Systems

Wide angle lobes caused by spill-over energy may also occur 80 to 120 degrees away from the main beam. Back lobes caused by edge diffraction of the reflector may occur opposite the main beam. All of these lobes may be reduced by narrowing the primary feed pattern at the expense of antenna size.

9.6.3.4 GAIN OF PARABOLIC ANTENNA SYSTEMS.

As the beamwidth is decreased, more of the radiated energy is concentrated in a beam aimed in the desired direction. This increases the antenna gain. If H is the beamwidth in the horizontal plane and V the beamwidth in the vertical plane, the total number of square degrees in the beam is H x V. The total number of square degrees in a sphere is approximately 41,000. Therefore, the gain relative to an isotropic radiator which radiates equally in all directions is:

$$G = \frac{41000}{H \times V}$$

This expression is not exact because it fails to take into account the exact shape of the beam and minor lobes. However, it does show that gain increases as the beamwidth decreases. Since beamwidth is determined by the ratio of $\frac{D}{\lambda}$, the gain is also determined by these factors. The gain of a parabolic antenna may be expressed as $G = E\,\frac{\text{reflector area}}{\lambda^2}$, where G is antenna gain normally given in db relative to an isotropic radiator, λ is the wavelength corresponding to the operating frequency, and E is the gain factor or efficiency of the antenna. The gain factor depends on the illumination taper and reflector tolerances. The gain factor of most parabolic antennas is about 55%. The curves given in figure 9-37 show the theoretical gain of various size reflectors based on an efficiency of 55 percent.

9.6.4 TYPICAL ANTENNA SYSTEMS USED FOR TROPOSPHERIC SCATTER.

The type of antenna systems used for a particular tropospheric scatter installation depends on the frequency, gain requirements, environmental conditions, and mobility requirements. The types range from 120-foot antennas for fixed installations to air-inflatable types for mobile terminals.

Figure 9-38 is a photograph of a high-gain antenna system used for tropospheric scatter systems operating in the frequency range of 700-1000 mc. The reflector is a four-sided steel structure which is shaped into a parabolic reflector at its face by steel

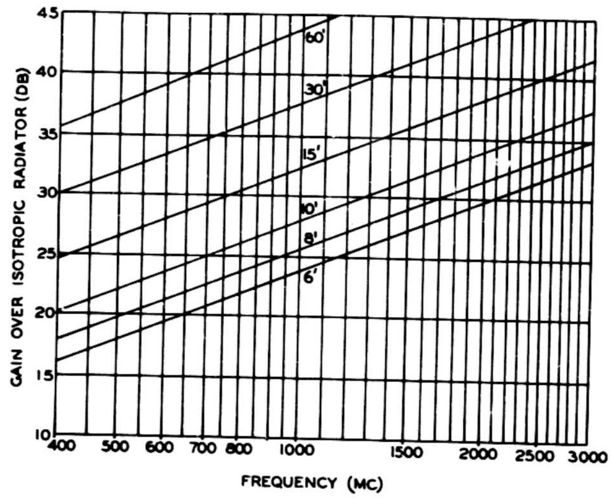


Figure 9-37. Parabolic Antenna Gain Curves

plates. The structure is 60 feet wide and has an over-all height of 65 feet. At 1000 mc, this 60-foot reflector will provide a gain of approximately 42 db. The feed system, consisting of wave guide, wave guide horn and supporting tower, is located approximately 30 feet directly in front of the reflector. The wave guide feedhorn is mounted at the top of the tower and is located at the focal point of the reflector. When required by environmental conditions, the reflector and feedhorn are provided with deicing equipment.

Figure 9-39 shows a different type of feed system. In this case, the feed system extends through a hole in the center of the reflector instead of being supported in front of the reflector. The energy is directed out to the feed system and then reflected back to the reflector. The parabolic reflector is 30 feet in diameter. The focal length is 9 feet. Gain provided by this system at 1000 mc is approximately 36 db.

Figure 9-40 shows a 28-foot parabolic reflector and feed system. The wave guide feed system is supported in front of the reflector. The weight and wind resistance is reduced considerably by using a grating surface for the reflector instead of a solid reflector. The openings can be regarded as short wave guides designed to be far beyond cutoff for the frequency band over which the antenna is to be used.

An antenna system designed for use with transportable scatter terminals is shown in figure 9-41. The 15-foot parabolic reflector of this type of antenna consists of a circular envelope, formed of two Fiberglas fabric sections. The inner surface of the rear fabric section is coated with aluminum to serve as the reflector. The reflector assumes the shape of a true parabolic form when the envelope is inflated. Focal length of the reflector is six feet. The inflation equipment consists of a motor-driven centrifugal blower, relief valve and check valve, and hose connections. The blower and accessories are mounted on the supporting tower.

The feed system consists of a wave guide horn supported in front of the reflector and fed with a coaxial line. The horn is designed to illumine the reflector with a 10-db taper. The flared open end of the horn is sealed with a plastic which is electrically transparent.

9.7 TESTS AND ADJUSTMENTS.

9.7.1 SWR MEASUREMENTS.

The standing-wave ratio on the transmission line can be determined by measuring the incident power and reflected power on the line. The relation between these quantities and swr is:

$$\text{swr} = \frac{\sqrt{\frac{P_i}{P_r}} + 1}{\sqrt{\frac{P_i}{P_r}} - 1}$$

where: P_i = incident or forward power

P_r = reflected power.

For convenience, this relationship is plotted in figure 9-42 for typical values of power and swr.

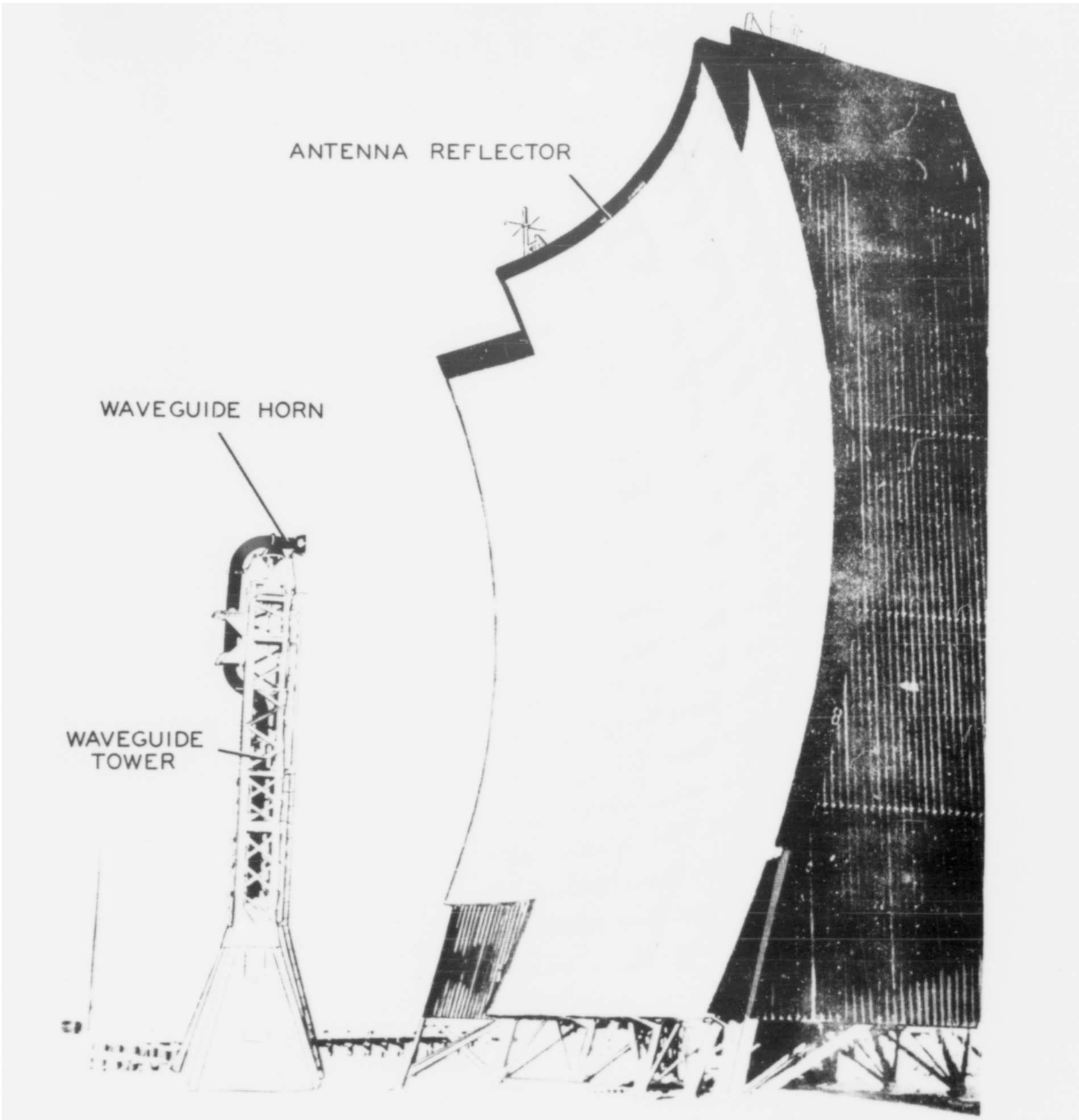


Figure 9-38. Sixty-Foot Parabolic Antenna System

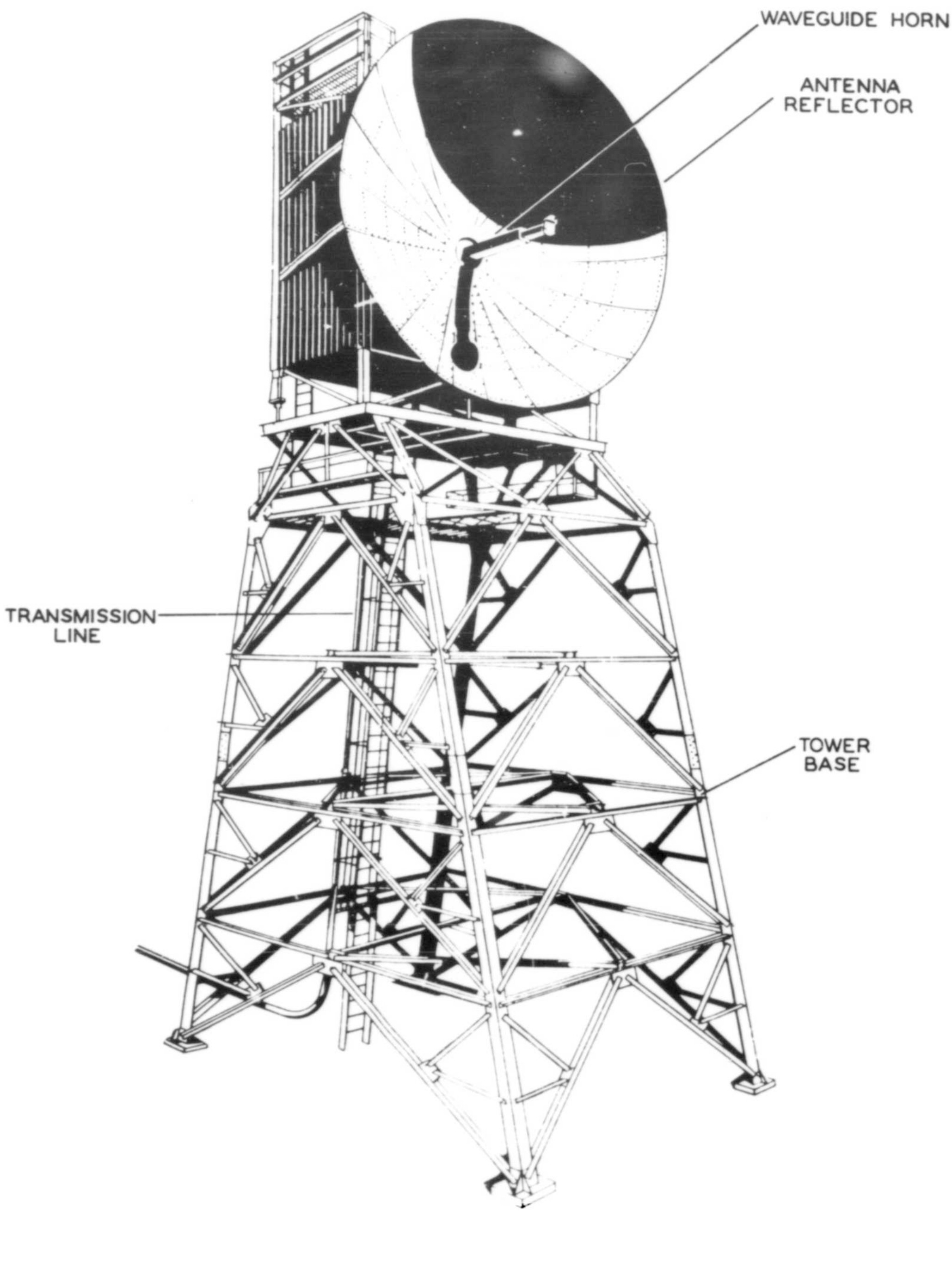


Figure 9-39. Thirty-Foot Parabolic Antenna System

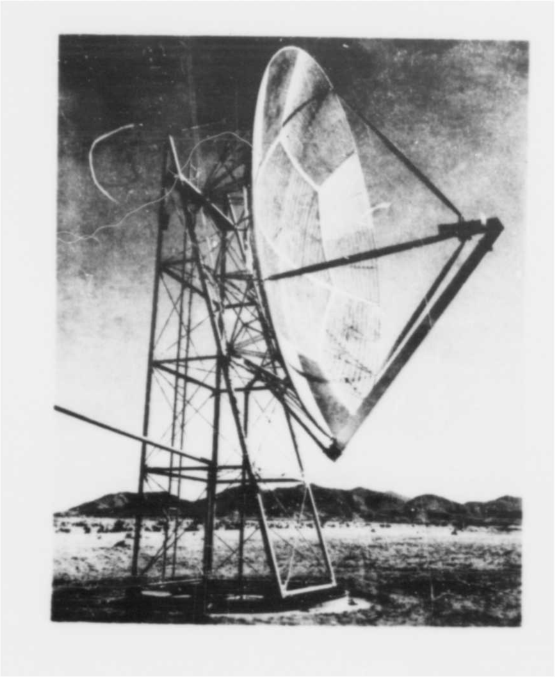


Figure 9-40. Twenty-Eight-Foot Parabolic Antenna System

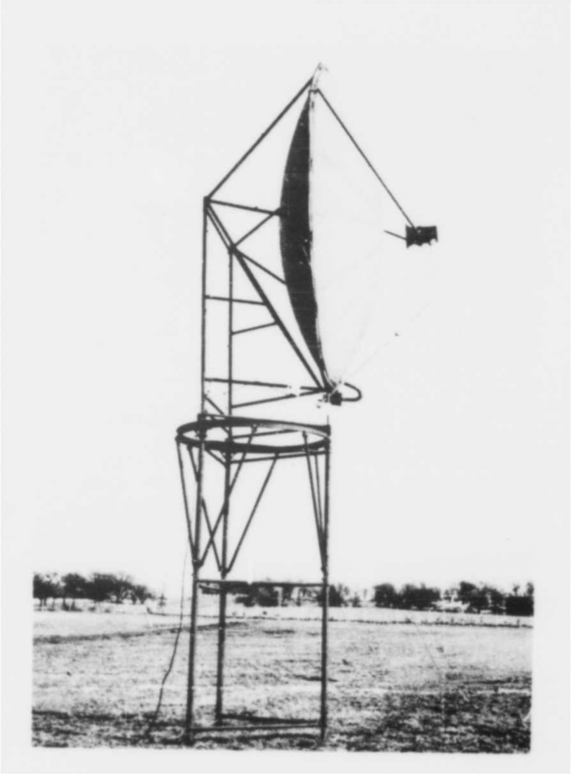


Figure 9-41. Air-Inflatable Antenna System

An important factor to consider when making swr measurements is the effect of line attenuation on the actual swr. On a lossy line, the measured swr is less than the actual swr.

The incident and reflected power indications are normally provided by a power meter on the associated power amplifier. This meter is connected to directional couplers in the transmission line which respond to power in only one direction.

9.7.2 ANTENNA DIRECTIVITY.

In order that the directional characteristics of a given antenna be completely known, the relative field must be measured in both azimuth and elevation. This requires that field strength measurements be taken in both planes around paths of constant range from the antenna.

In most cases, the azimuth pattern can be taken relatively easily. If the antenna under test can be rotated, a pickup antenna and field strength meter are located at a fixed remote point. The test antenna is then rotated and the field strength is recorded for each angular position. The pattern may be affected by reflections from ground obstructions.

An airplane is required to determine accurately the elevation pattern. However, a check on the elevation pattern can be made if the antenna is adjustable in elevation. The elevation of the antenna can be changed and the field change noted for each position. Maximum field strength should be recorded when the antenna is aimed directly at the recording position.

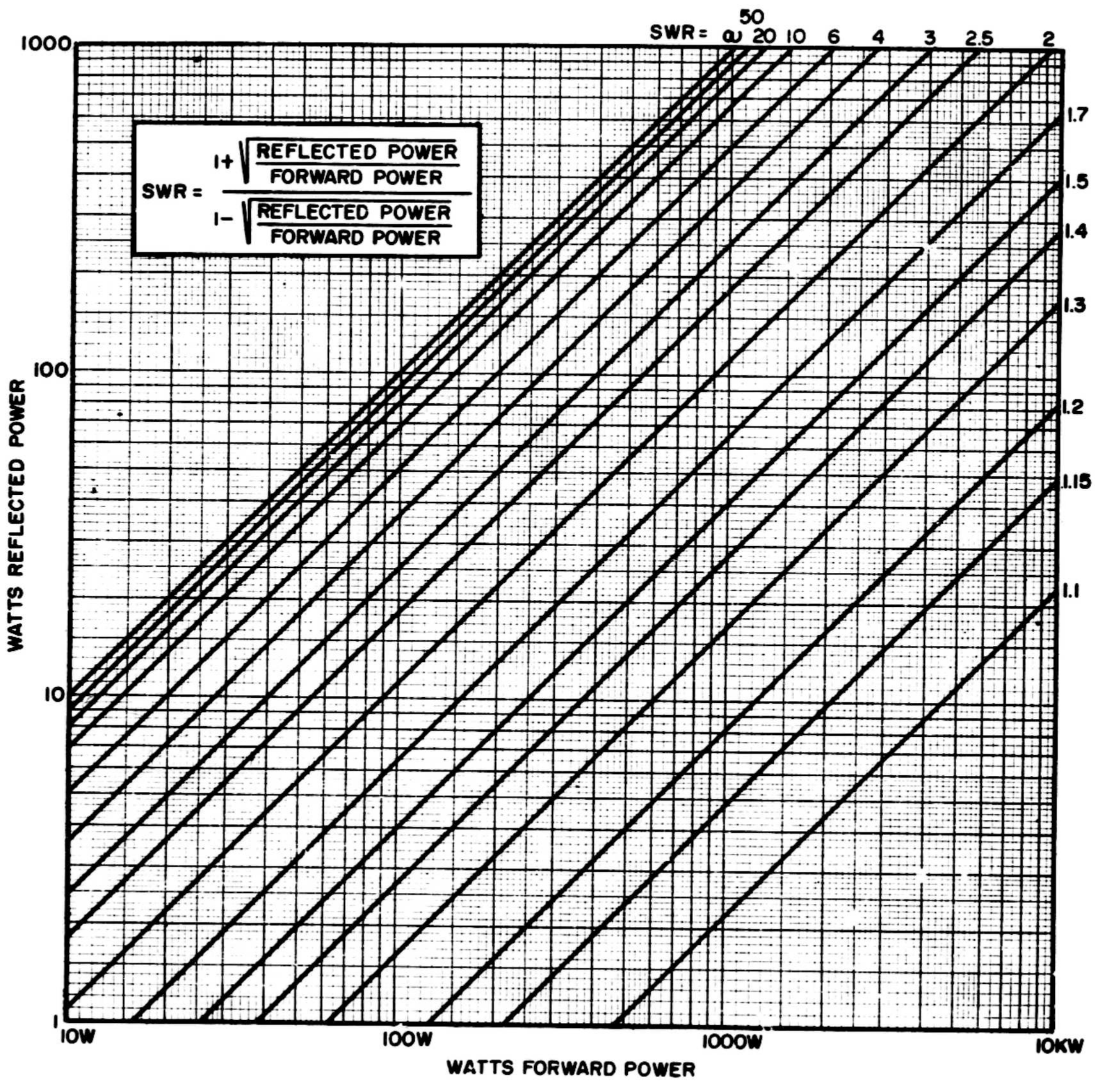
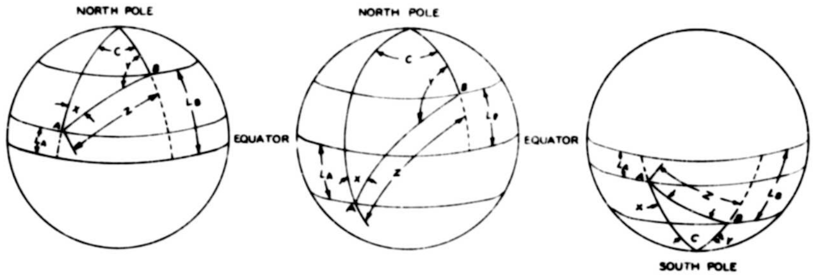


Figure 9-42. Relationship of SWR to Incident and Reflected Power



B = PLACE OF GREATER LATITUDE, i.e., NEARER THE POLE. L_A = LATITUDE OF A. L_B = LATITUDE OF B, AND C = DIFFERENCE OF LONGITUDE BETWEEN A AND B

$$\tan\frac{Y-X}{2} = \cot\frac{C}{2}\,\frac{\sin\frac{L_B-L_A}{2}}{\cos\frac{L_B+L_A}{2}} \quad \text{AND} \quad \tan\frac{Y+X}{2} = \cot\frac{C}{2}\,\frac{\cos\frac{L_B-L_A}{2}}{\sin\frac{L_B+L_A}{2}}$$

$$\frac{Y+X}{2} + \frac{Y-X}{2} = Y \qquad \frac{Y+X}{2} - \frac{Y-X}{2} = X$$

$$\tan\frac{Z}{2} = \tan\frac{L_B-L_A}{2}\left(\sin\frac{Y+X}{2}\right)\Big/\left(\sin\frac{Y-X}{2}\right)$$

Figure 10-1. Calculations for Determining Great-Circle Path Length and Bearing

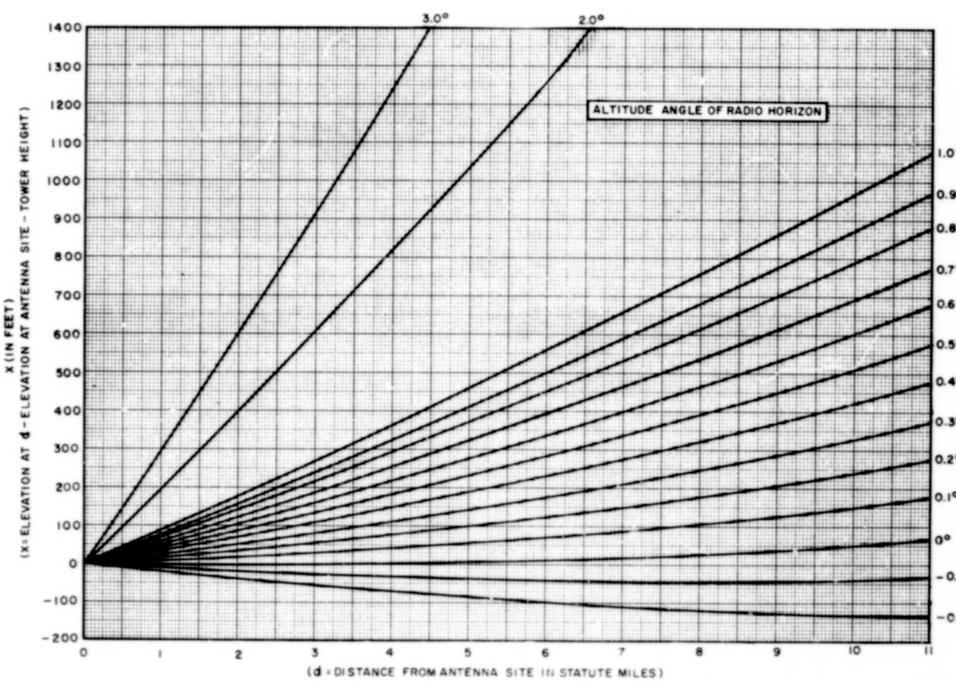


Figure 10-2. Horizon Angle Determination Curves for Short Distances

CHAPTER 10
PLANNING PREDICTION, AND INSTALLATION PROCEDURES FOR TROPOSPHERIC SCATTER CIRCUITS

10.1 INTRODUCTION.

This chapter includes simplified step-by-step procedures for establishing a tropospheric scatter circuit to satisfy certain ccmmunication requirements between two points. The information included in this chapter is based on a survey of existing material and results of circuit tests performed by Collins Radio Company under Signal Corps contract DA 36-039-sc-67491.

Procedures for establishing a tropospheric scatter communications circuit can be organized into four major steps. These are: (1) determination of the basic circuit requirements, (2) analysis of the proposed circuit to determine bearing, path length, horizon angle, and optimum location, (3) prediction of circuit performance based on type of equipment used, path length, horizon angle, and required channel capacity, and (4) actual installation procedures. Each of these major procedure steps is described in the following paragraphs.

10.2 DETERMINATION OF BASIC CIRCUIT REQUIREMENTS.

The first step in planning a tropospheric scatter circuit between two given site locations is clarification of the circuit requirements. The following questions should be answered.

a. Will the circuit be used for voice, teletype, or high-speed data transmission?

b. What is the required voice channel capacity?

c. Will the sites be established on a permanent basis or should transportable scatter equipment be used?

10.3 ANALYSIS OF PROPOSED CIRCUIT.

After the circuit requirements are established, the path length and horizon angle at each of the sites should be determined. These physical parameters will determine if a tropospheric scatter circuit, satisfying the basic requirements, can be operated with the required reliability between the proposed sites.

10.3.1 DETERMINATION OF PATH LENGTH AND BEARING.

Path length for purposes of predicting circuit performance and the bearing between the two sites, can generally be determined with sufficient accuracy from maps of the area. However, if maps are not available or if greater accuracy is required, great-circle calculations can be used. The mathematical method of great-circle calculations is outlined in figure 10-1.

If the latitude and longitude of the sites A and B are known, the angles X and Y and the distance Z between the sites along the great-circle path can be calculated using the formulas given in figure 10-1.

In the formulas, north latitudes are taken as positive, and south latitudes as negative. For example, if B is latitude 50° N and A is latitude 20° S,

$$\frac{L_B + L_A}{2} = \frac{50 + (-20)}{2} = \frac{50 - 20}{2} = \frac{30}{2} = 15°$$

$$\frac{L_B - L_A}{2} = \frac{50 - (-20)}{2} = \frac{50 + 20}{2} = \frac{70}{2} = 35°$$

Figure 10-1 includes an equation for finding the angular distance Z in degrees between the two sites. This may be converted to linear distance as follows:

Z (in degrees) x 111.195 = kilometers

Z (in degrees) x 69.093 = statute miles

Z (in degrees) x 60.000 = nautical miles

In multiplying, the minutes and seconds of arc must be expressed in decimals of a degree.

10.3.2 DETERMINATION OF HORIZON ANGLES.

The horizon angles can be determined by using a transit at each site and sighting along the circuit path. Another method, which eliminates the necessity of going to the proposed sites, involves the use of topographic maps and the horizon-angle curves given in figures 10-2 and 10-3. To use this method, proceed as follows.

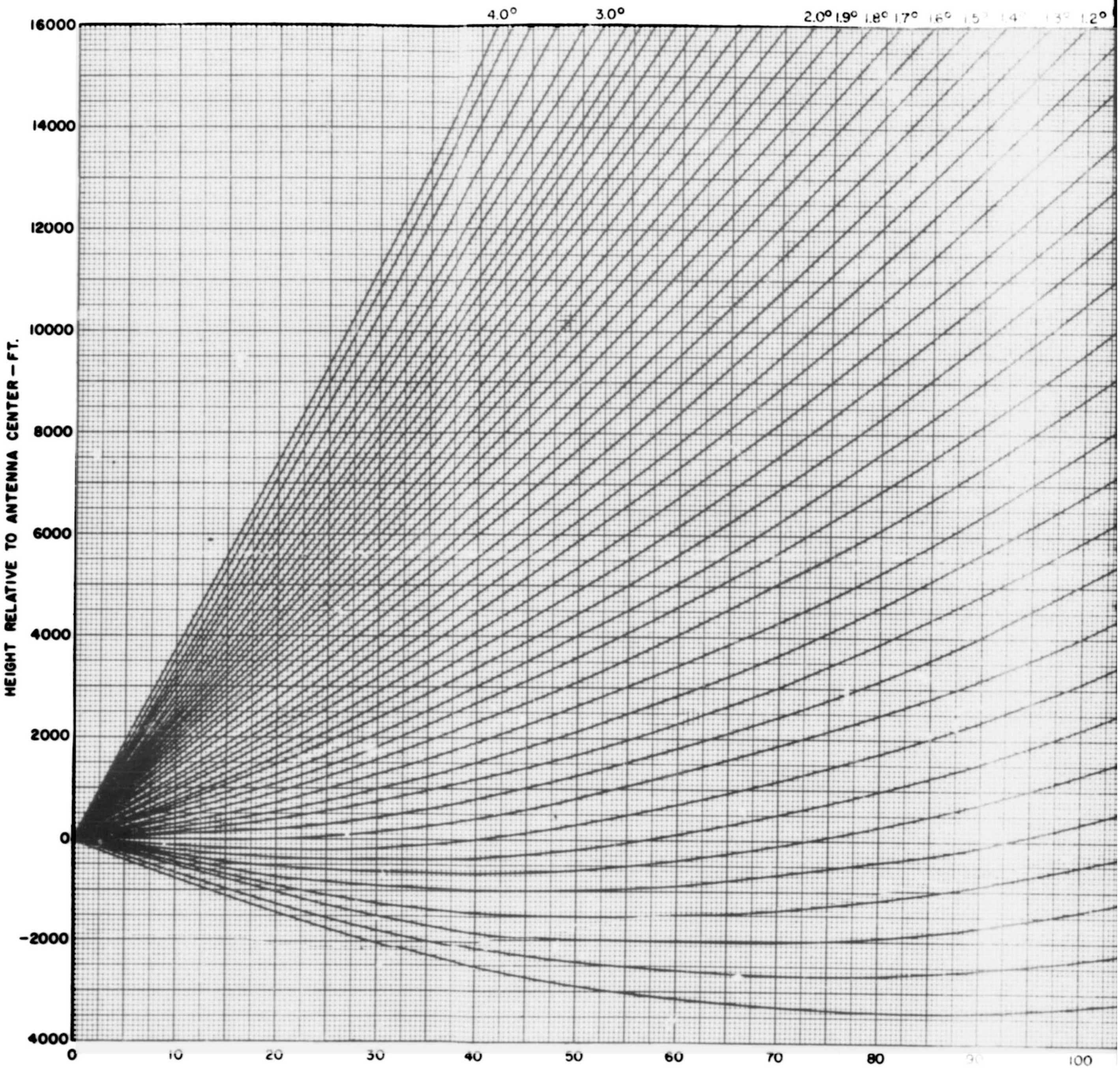
a. Draw a line between the two proposed sites on a topographic map.

b. Starting at one site, find the highest elevation points along the line toward the other site.

c. Find the difference between the elevation at one of these points (d) and the elevation at one site and the distance from the site to the elevation point. Depending on the distance, use either figure 10-2 or 10-3 to determine horizon angle.

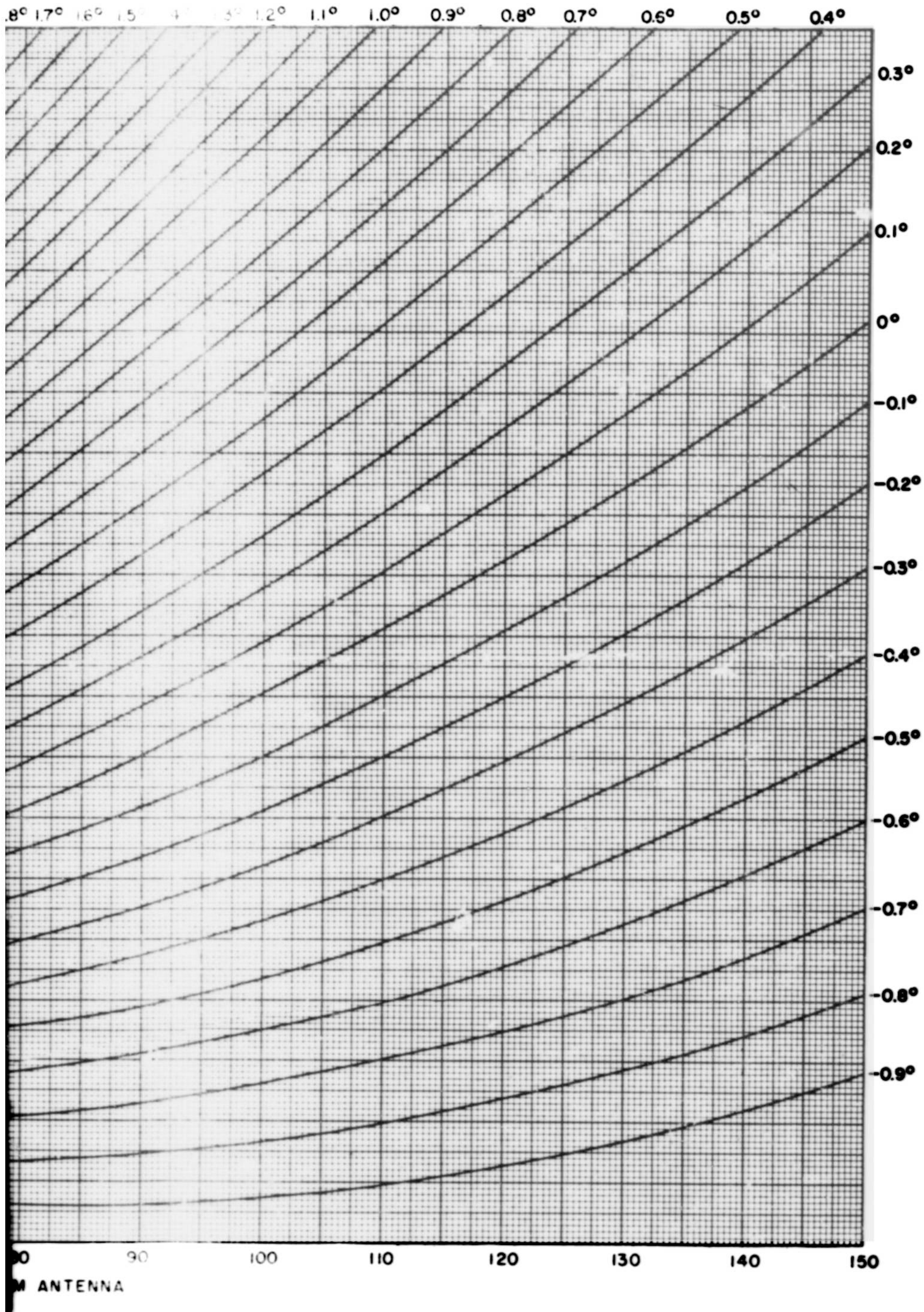
d. Repeat step c for the other high elevation points until the maximum horizon angle for the site is determined.

e. Repeat steps c and d for the other site, and add the horizon angles to find the sum of horizon angles for the circuit.



1

Figure 10-3. Horizon Angle Determination Curves for Long Distances



.8° 1.7° 1.6° 1.5° 1.4° 1.3° 1.2° 1.1° 1.0° 0.9° 0.8° 0.7° 0.6° 0.5° 0.4°
0.3°
0.2°
0.1°
0°
-0.1°
-0.2°
-0.3°
-0.4°
-0.5°
-0.6°
-0.7°
-0.8°
-0.9°
90
100
110
120
130
140
150
M ANTENNA

2

If 4/3 earth profile paper (Form DA 11-47 or DA 11-48) is available, a path profile can be drawn for the proposed circuit. Distance and elevation values necessary for using the horizon angle curves can then be found from the profile.

10.4 PREDICTION OF CIRCUIT PERFORMANCE.

10.4.1 FACTORS WHICH DETERMINE CIRCUIT RELIABILITY.

After the horizon angles and path length have been determined, the next step is to predict the circuit reliability which will be obtained with the equipment selected to meet the basic circuit requirements. Circuit reliability is determined by the following factors.

a. Horizon angle at each site.
b. Path length.
c. Transmitter power.
d. Antenna reflector size.
e. Receiver threshold, which is determined by the bandwidth required to match the system channel capacity, and receiver noise figure.

10.4.2 SIMPLIFIED PROCEDURE TO DETERMINE FEASIBILITY OF CIRCUIT.

The predicted reliability of a circuit is based on a prediction of the percentage of time that the received signal level will be above threshold. This "above-threshold" time is determined by both the long-term and short-term fading characteristics of the circuit. Experimental results have shown that the following short-term reliabilities are required for reliable voice communications and data error rates of 1 bit in 10 for high speed data and 2 bits in 10 for teletype.

Voice	98%
Teletype	99.99%
High-Speed Data	99.99%

Figures 10-4 and 10-5 show maximum path length and horizon angle combinations which will result in the required reliability 95% of the total circuit operating time for various channel capacities. These curves are drawn for equipment operating in the 2000-mc range, with 1000-watts maximum transmitter output, and 15-foot reflectors. To use the curves, proceed as follows:

a. Select the correct set of curves for voice, teletype, or high-speed data transmission.
b. Find the intersection point of lines drawn perpendicular from points on the horizon angle and path length scales corresponding to values determined for the circuit.
c. The circuit will provide the required reliability for the channel capacities which are above the intersection point.

10.4.3 DETAILED PROCEDURES FOR PREDICTING RELIABILITY.

Use of figures 10-4 and 10-5 will provide a preliminary determination of the feasibility of a given circuit using certain equipment. The following paragraphs outline a procedure for using curves to predict the circuit reliability for any given set of physical conditions and equipment.

Reliability of a given tropospheric scatter circuit is predicted by completing the following basic calculations:

a. Determine effective transmitted power.
b. Determine propagation loss.
c. Obtain expected median level by subtracting propagation loss from effective transmitted power.
d. Determine signal level required for threshold.
e. Determine fade margin by subtracting threshold level from expected median level.
f. Predict circuit reliability using calculated fade margin level versus reliability curves.

Each of these steps is explained in more detail in the following paragraphs. Calculations and characteristics for a typical circuit are summarized in table 1 below. This table summarizes the calculations for a circuit with a total horizon angle of 0.5° and with 15-foot antennas.

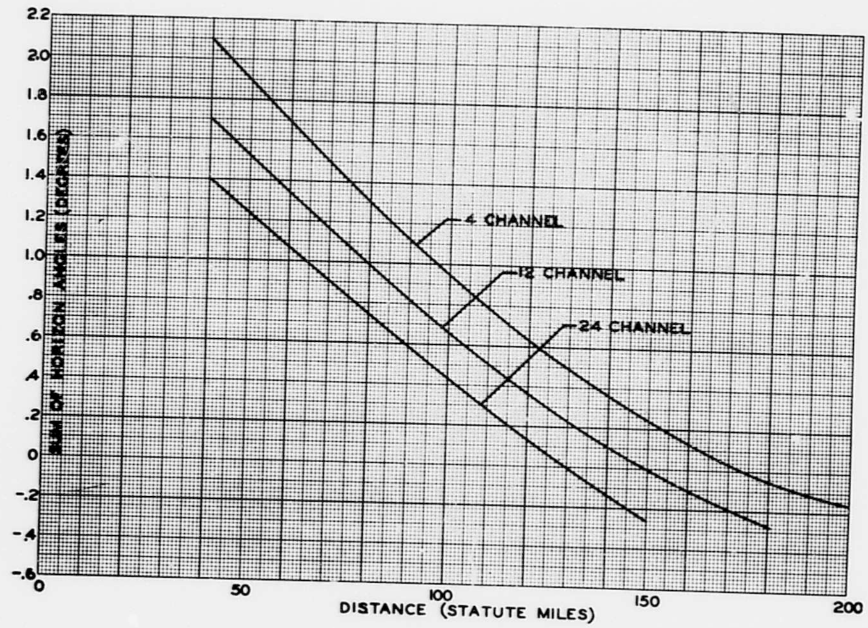


Figure 10-4. Maximum Path Length, Horizon Angle and Channel Capacity Combination for Voice Transmission

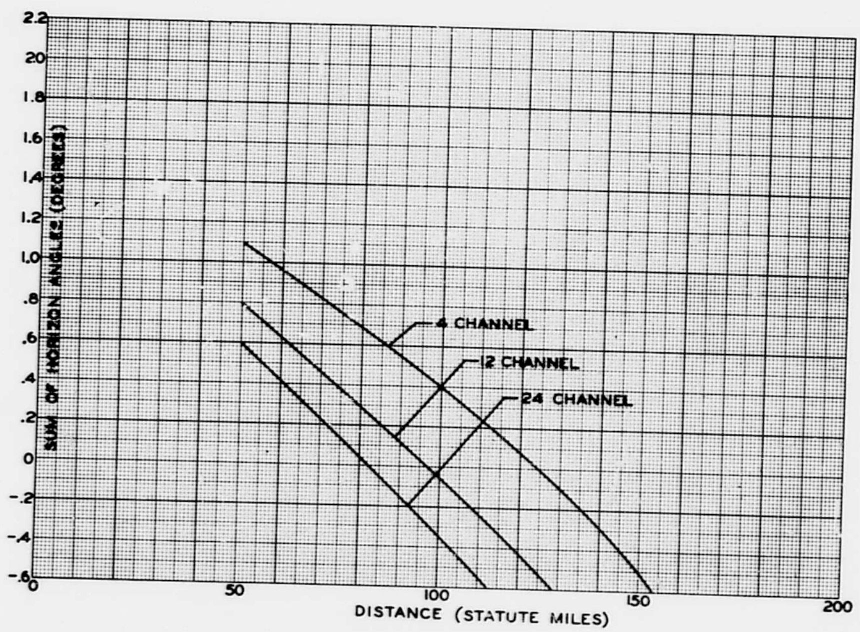


Figure 10-5. Maximum Path Length, Horizon Angle, and Channel Capacity Combinations for Teletype and High-Speed Data Transmission

TABLE 10-1. CALCULATIONS FOR A 100-MILE TRANSHORIZON CIRCUIT

ASSUMED PATH CHARACTERISTICS	
Circuit length	100 miles
R-f operating frequency	2000 mc
Antenna size	15-foot parabola
Transmitter power	1 kw
Receiver bandwidth	600 kc (12 channels)
Sum of horizon angles	0.5°
EFFECTIVE TRANSMITTED POWER	
Transmitter power (1 kw)	60 dbm
Antenna gain (from figure 10-6) for 15-foot antennas	74 db
Line loss	-1 db
Total effective transmitted power	133 dbm
PROPAGATION LOSS	
Basic propagation loss (from figures 10-7 and 10-8)	203 db
Horizon angle loss (from figure 10-9)	8 db
Aperture-to-medium coupling loss (from figure 10-10)	1 db
Total propagation loss	212 db
PREDICTED MEDIAN SIGNAL LEVEL	
Difference between propagation loss and effective transmitted power	-79 dbm
SIGNAL LEVEL REQUIRED FOR THRESHOLD	
Receiver noise figure	8 db
Noise in 600-kc i-f pass band (from figure 10-11)	-116 dbm
I-f S/N for threshold	10 db
Required signal level for threshold	-98 dbm
FADE MARGIN	
Difference between level required for threshold and predicted median level	19.0 db
CIRCUIT RELIABILITY	
For dual-diversity system (from figure 10-12)	98.6%

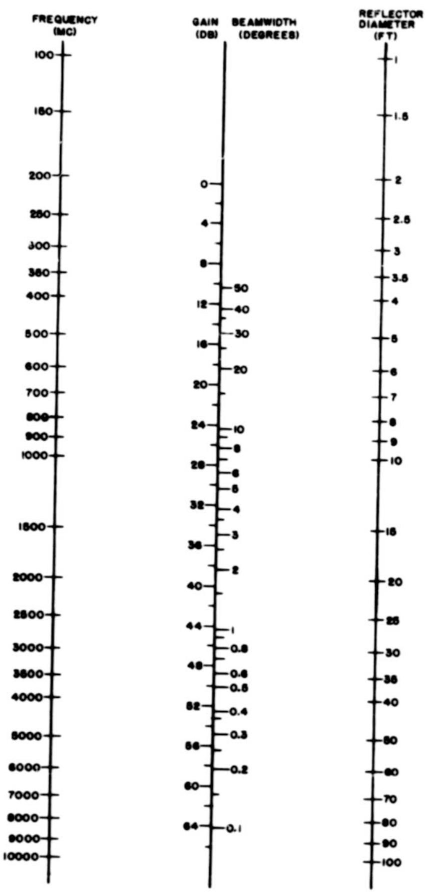


Figure 10-6. Nomogram for Determining Antenna Gain and Beam Width

10.4.3.1 EFFECTIVE TRANSMITTED POWER.

a. In the sample circuit, transmitter output is one kilowatt. Convert transmitter output to dbm using the following equation:

Output power (dbm) = 30 + 10 log output power (watts)

In the example, this becomes 60 dbm.

b. Using figure 10-6, estimate antenna gain. At 2000 mc, the gain for a 15-foot antenna is 37 db. The composite antenna gain is 74 db.

c. To calculate effective transmitted power, add transmitter output to antenna gain (60 + 74 = 134 dbm), and subtract the composite line loss from the sum. Line loss is a factor dependent on particular station installations and should be calculated for a given circuit. The composite line loss should include filter loss and both transmit and receive line losses. In this example, assume a composite line loss of 1 db. This gives an effective transmitted power of 133 dbm.

10.4.3.2 PROPAGATION LOSS.

a. From figures 10-7 and 10-8, find the basic propagation loss. Figure 10-7 shows propagation loss for 1000 mc. Figure 10-8 is a 9-db-per-octave curve to include frequency correction for both free-space loss and scatter loss. For example, for the circuit operating at 2000 mc, the correction factor is 9 db. Add this to the basic propagation loss for 1000 mc at 100 miles (194 db) obtained from figure 10-7. This gives a basic propagation loss of 203 db.

b. From figure 10-9, find the loss due to elevated horizon angles. In the sample circuit with a sum of horizon angles of 0.5°, the loss is 8 db.

c. From figure 10-10, find the aperture-to-medium coupling loss. This nomogram shows loss as a function of antenna beam width and scatter angle. The scatter angle is equal to the angle corresponding to equivalent smooth-earth statute miles plus the sum of the horizon angles.

To find the aperture-to-medium coupling loss using figure 10-10, first find the angle opposite the statute miles corresponding to the path length. This gives the scatter angle, assuming zero horizon angle at each site. For the sample circuit with a circuit length of 100 miles, the smooth-earth angle is 1.08°. To this angle, add the sum of the horizon angles to find the scatter angle. For the sample circuit, this is 1.08 + 0.5 or 1.58°.

On the nomogram, place a straight edge between the scatter angle and the antenna beam width. Aperture-to-medium coupling loss is found where the straight edge intersects the loss scale.

The antenna beam width for 15-foot antennas at 2000 mc, from figure 10-6, is approximately 2.2°. A straight edge placed between 2.2° on the antenna beam width scale and 1.58° on the scatter angle scale intersects the loss scale at approximately 1 db.

10.4.3.3 PREDICTED MEDIAN LEVEL.

To find the predicted median level, add the effective transmitted power and total propagation loss. This will be the expected long-term median over a typical tropospheric scatter circuit. For the sample circuit, the predicted median level is 133 dbm-212 db=-79 dbm.

10.4.3.4 SIGNAL LEVEL REQUIRED FOR THRESHOLD.

To find the signal level required for threshold, first find the noise power in the i-f pass band using figure 10-11. This curve is a plot of the equation:

Noise Power (in dbm) = -144 dbm + 10 log B

where B is the i-f pass band in kilocycles.

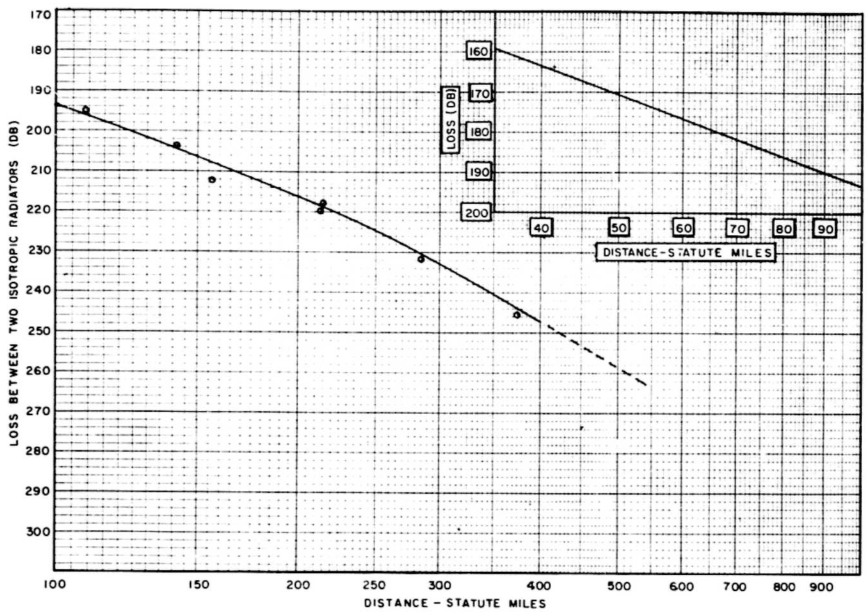


Figure 10-7. Basic Propagation Loss Curve for 1000 Megacycles

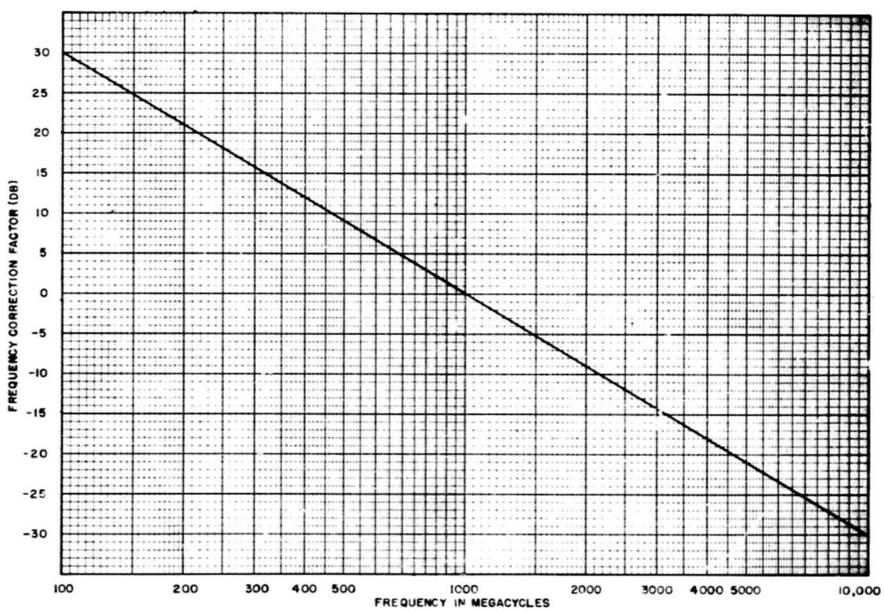


Figure 10-8. Frequency Correction Curve for Basic Propagation Loss

To the noise power found from figure 10-11, add the receiver noise figure in db to find the total receiver noise. The threshold signal-to-noise ratio for an FM system is 10 db. Therefore, the required signal level at threshold is 10 db above the noise.

In the sample circuit, with 600 kc i-f bandwidth and a noise figure of 8 db, the noise power is -108 dbm. The signal required for threshold is 10 db above the noise or -98 dbm.

10.4.3.5 FADE MARGIN.

The fade margin is equal to the difference between the predicted median level and the signal level required for threshold. In the sample circuit, this is 19 db.

10.4.3.6 RELIABILITY.

Circuit reliability versus fade margin for dual-diversity system, is given in figure 10-12. For the sample, circuit reliability is 98.6%. This is actually a measure of the time that threshold level is exceeded.

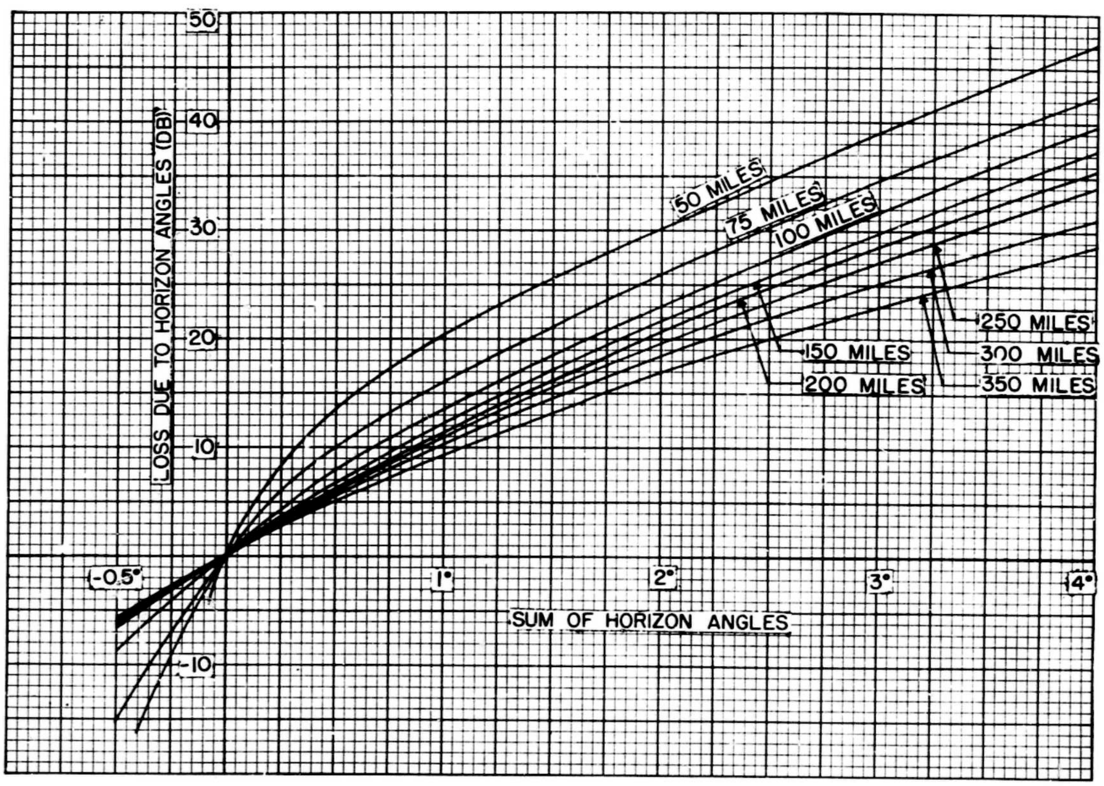


Figure 10-9. Loss Due to Elevated Horizon Angles

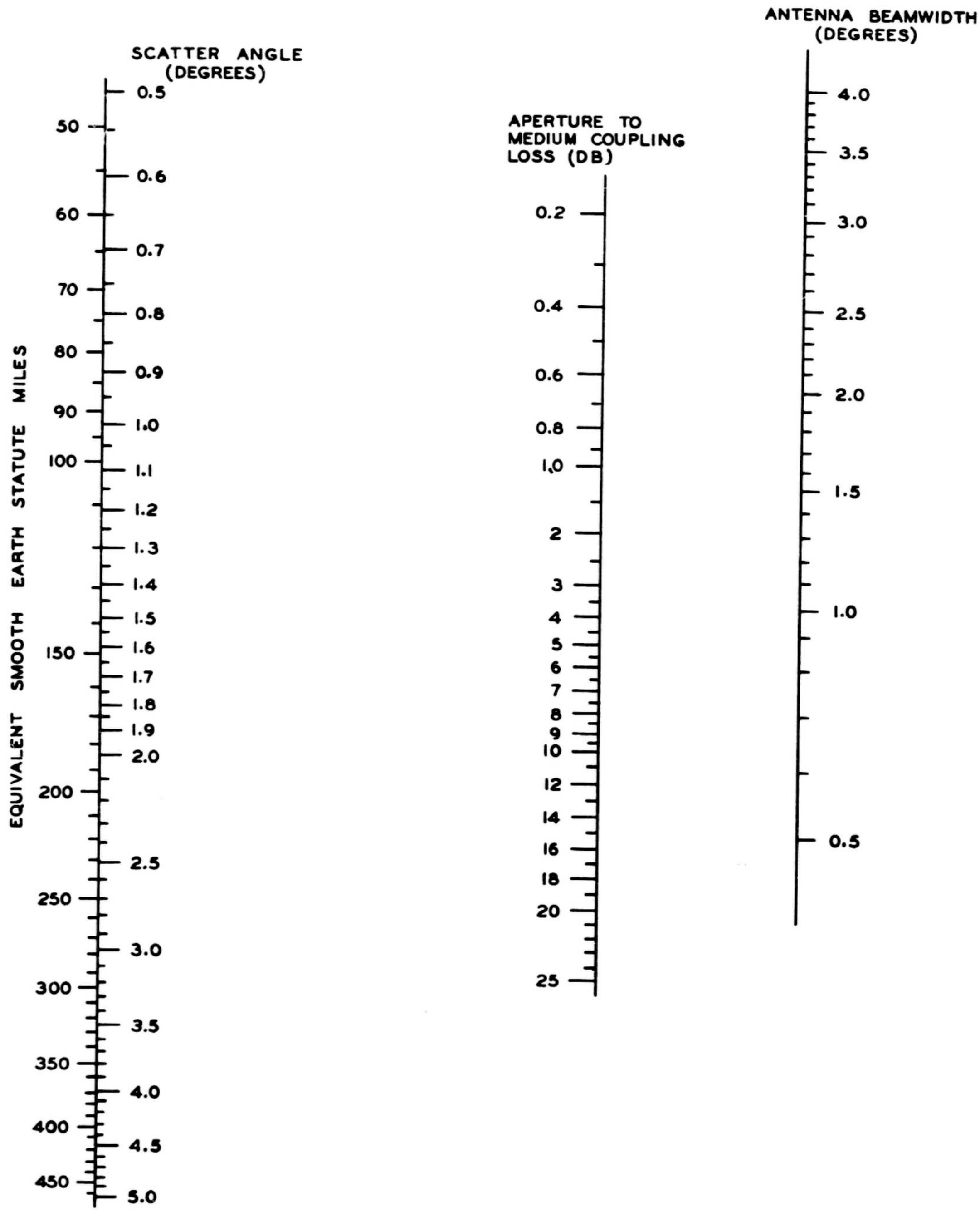


Figure 10-10. Nomogram for Determining Aperture-to-Medium Coupling Loss

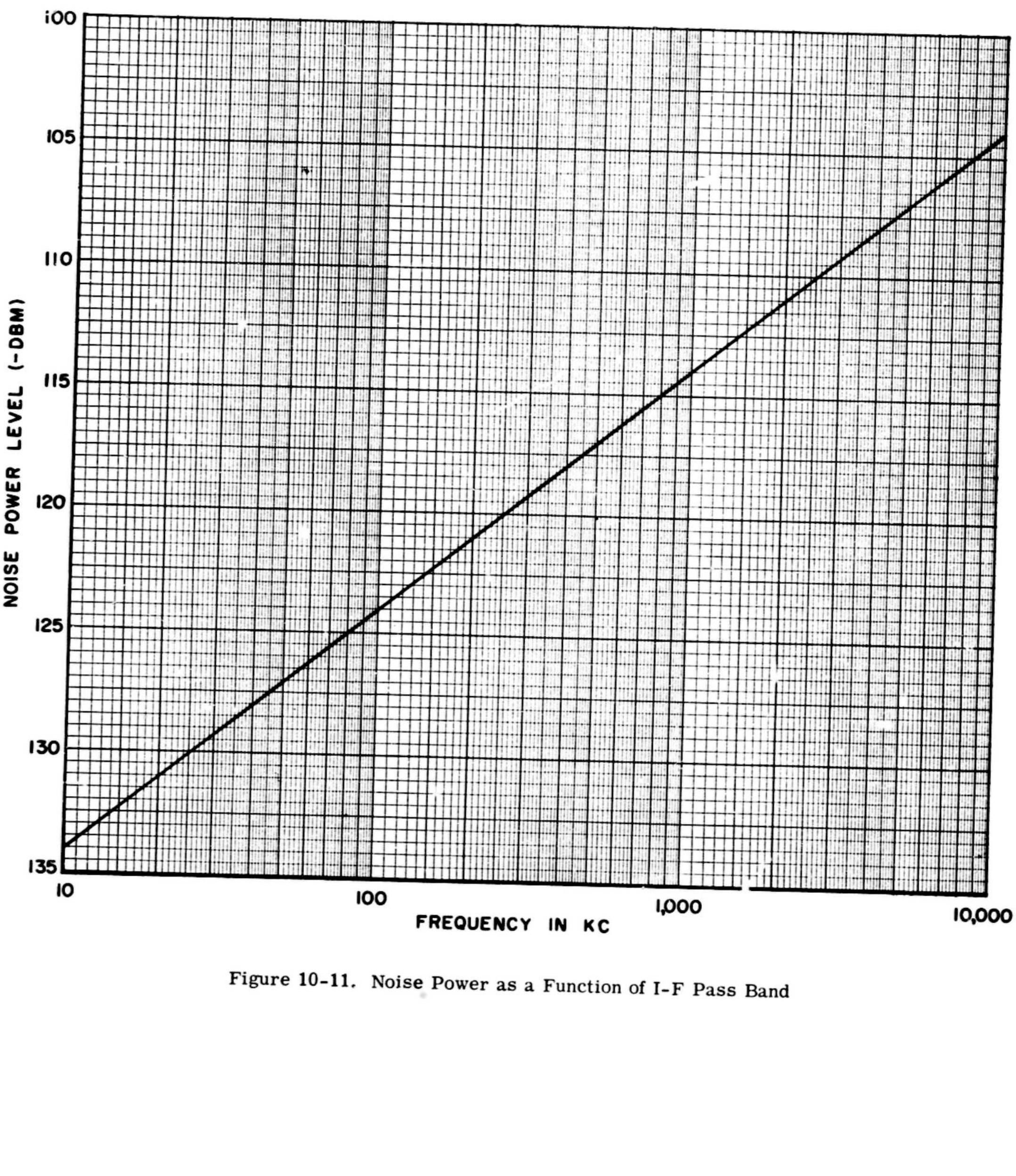


Figure 10-11. Noise Power as a Function of I-F Pass Band

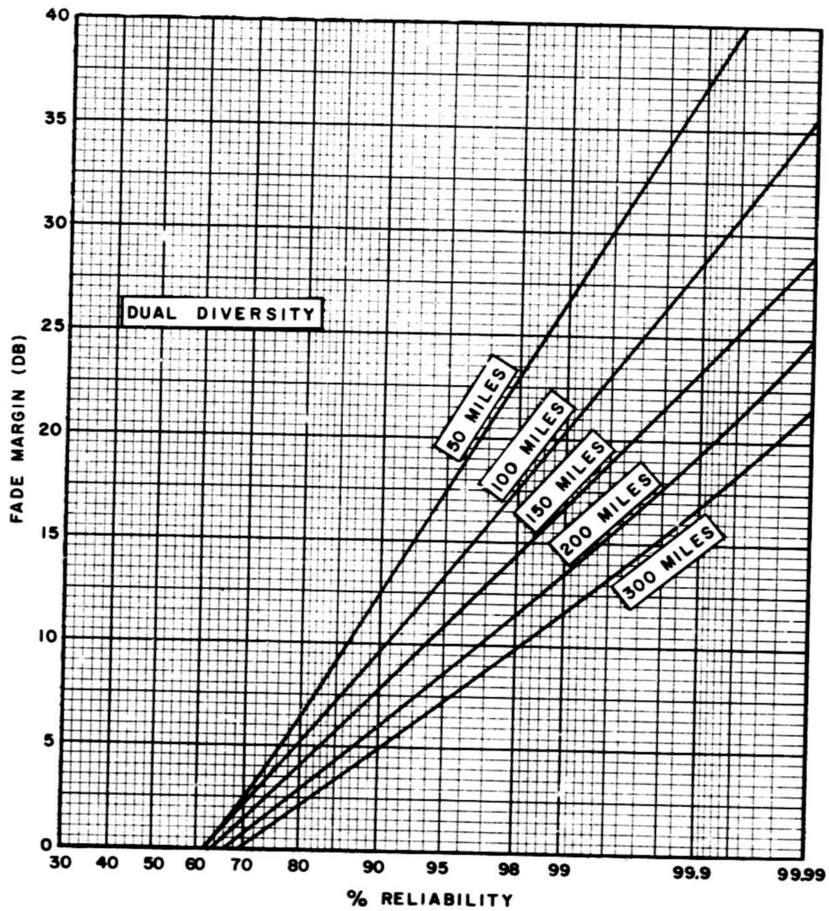


Figure 10-12. Reliability as a Function of Fade Margin for a Dual-Diversity System

10.4.4 ALTERNATIVES IF PREDICTED RELIABILITY IS TOO LOW.

If the circuit will not provide the required reliability for the specified circuit requirements, several alternatives are possible. These are discussed in the following paragraphs.

10.4.4.1 RELOCATION OF TERMINALS.

It may be possible to relocate the terminals within the same general localities and reduce the horizon angles.

10.4.4.2 USE OF OTHER TYPES OF PROPAGATION.

If horizons are extremely elevated, the mountainous terrain normally associated with these conditions offers a possibility of using obstacle gain or knife-edge diffraction circuits. A path profile made from the topographical maps should indicate if a common horizon does exist. If such a condition is available, the resultant loss will normally be much less than an equivalent length tropospheric scatter circuit.

Another possibility is the use of a mountainside as a passive reflector. Visual checks or investigation of topographical maps will indicate if reflection is possible.

10.4.4.3 REDUCTION IN CHANNEL CAPACITY.

If relocation of the terminals or use of other types of propagation is not possible, the circuit can still be established with the same equipment by reducing the channel capacity requirement. Reference to figures 10-2 or 10-3 will show the maximum channel capacity which can be used and still satisfy the reliability requirement. Of course, one alternative to reduction in channel capacity is operation at full capacity with the resultant reduction in reliability.

10.4.4.4 USE OF RELAY STATION.

In some cases, it may be necessary to use a relay station between the two terminals to satisfy the circuit requirements. This relay station will require two complete tr. smit-receive terminals.

10.4.4.5 USE OF EQUIPMENT WITH INCREASED TRANSMITTER POWER OR ANTENNA SIZE.

If the circuit is to be permanent, and transportable equipment does not have to be used, equipment with increased transmitter power and larger antennas can be used. This will provide an increase in reliability.

A tropospheric scatter calculator can be used to determine reliability provided by various equipment configurations over a given circuit. Procedures for using this calculator are given in the following paragraphs.

10.5 USE OF CALCULATOR FOR PREDICTING PERFORMANCE OF TROPOSPHERIC SCATTER CIRCUIT.

10.5.1 DESCRIPTION.

A calculator which incorporates all of the required curves for circuit prediction has been developed. One side of the calculator is for use with circuits operating in the 700- to 1000-mc range; the other side is for use with circuits operating in the 1700- to 2350-mc range.

The inner disc of the calculator includes a received median level scale in -dbm arranged opposite circuit distance scales for various size reflectors. These scales incorporate experimentally verified values for propagation losses and antenna gains for reliable predictions of received medians. The received median scale is calculated for a transmitter power of one kilowatt and incorporates a composite line loss, including filters and diplexers, of 2 db. The scale is arranged linearly in db. Therefore, median levels for transmitted power other than one kilowatt can be read from the scale if the power ratios are expressed in db. For example: for transmitted power of 10 kilowatts, add 10 db to the median level scale. Values are given for distances of 50- to 350-statute miles.

If zero horizon angles are assumed for a circuit, the median level can be obtained by using only the center disc and cursor. However, this median level must be corrected for horizon angles other than zero. The scale for adding the effect of elevated or depressed horizon angles is included on the outer disc of the calculator. This scale has mileage lines for distances of 50 to 350 miles. The mileage lines are intersected by curves for sums of horizon angles from -.5° through 4.0°. When the cursor hairline is set for the correct intersection point of mileage and horizon angle curves, the distance setting on the inner disc is corrected to correspond to the effect of elevated horizon angles. The received median level, read on the opposite end of the cursor, is then corrected to the change in distance setting.

Two reliability scales, one for dual-diversity and one for nondiversity, are also included on the outer disc. These scales are arranged to incorporate the two factors which affect reliability, namely, circuit distance and fade margin. The circuit distance is given in mileage lines on the scales. These mileage lines are intersected by reliability percentage curves. The fade margin, which is the difference between median level and FM threshold, is found by rotating the center disc for a corr . setting of FM threshold for the system channel capacity. An FM threshold index is located on the left of each of the reliability scales. The FM threshold level is found by rotating the center disc until the NONDIVERSITY or DUAL-DIVERSITY pointer is set to the correct channel scale (4 ch, 12 ch, or 24 ch). This moves the median level along the reliability scale to set the calculator for correct fade margin. The FM threshold level is indicated on the RECEIVED MEDIAN LEVEL scale opposite the FM THRESHOLD index. The reliability percentage is found at the intersection of the hairline on the cursor set on the median level, the mileage line, and the reliability curve.

The threshold levels were calculated assuming an 8-db receiver noise figure and i-f bandwidths of 200 kc, 600 kc, and 1.2 mc for the 4, 12, and 24 channel systems, respectively. The FM threshold can be determined for other channel capacities and noise figures if the ratios are expressed in db. For example, if the index is set at 24 channels, add 3 db to the threshold level for 48 channels. If the receiver noise figure is 7 db, subtract 1 db from the threshold level.

10.5.2 DETERMINING MEDIAN LEVEL FOR ZERO HORIZON ANGLES.

a. Select the proper frequency range scale. Note one side covers 700-1000 mc; the other, 1700-2350 mc.

b. Set the hairline on the cursor to path length on mileage scale corresponding to the reflector size used.

c. The median received level in -dbm appears under hairline at the opposite end of the cursor.

10.5.3 DETERMINING MEDIAN LEVEL FOR HORIZON ANGLE OTHER THAN ZERO.

a. Perform steps a and b in paragraph 10.5.2.

b. Rotate cursor and center disc together, making certain that mileage setting is not disturbed, until hairline on mileage end of cursor is set to $\psi = 0$.

c. Holding center disc stationary, move cursor hairline to intersection of mileage line and sum of horizon angles (ψ in degrees).

d. The median received level, corrected for horizon angles, appears under hairline at opposite end of cursor.

10.5.4 DETERMINING RELIABILITY, FM THRESHOLD, AND FADE MARGIN.

a. Set calculator for median receive level. Note mileage under hairline on correct reflector scale after the setting for elevated horizon angles has been made. This new mileage reading indicates the increase in effective path length due to elevated horizons.

b. Set DUAL-DIVERSITY or NONDIVERSITY pointer to system channel capacity. Make certain that the cursor and center disc are moved together without disturbing the median level setting.

c. The reliability percentage appears under hairline at intersection of hairline, mileage line, and reliability percentage line. Use the mileage line which corresponds to the mileage under the hairline after the setting for elevated horizon angle has been made. It may be necessary to interpolate for intersection of distance and reliability curves.

d. The FM threshold level for the channel capacity setting appears on the RECEIVED MEDIAN LEVEL scale opposite the FM THRESHOLD index.

e. The fade margin is the difference between the FM threshold and the median level read under the hairline of the cursor.

10.5.5 EXAMPLE OF TYPICAL CALCULATION.

Assume a circuit with the following characteristics.

Circuit length:	100 miles
Size of reflectors:	15 ft
Channel capacity:	12
Order of diversity:	dual
Sum of horizon angles:	+0.3°
Frequency:	800 mc

a. Use 700- to 1000-mc side of calculator.

b. Set cursor to 100 on statute mile scale for 15-ft REFLECTOR (see figure 10-13A).

c. Rotate cursor and center disc together until hairline is set to $\psi = 0$ horizon angle scale (see figure 10-13A).

d. Hold center disc stationary and move cursor until hairline is set to intersection of 100-mile line and curve for $\psi = 0.3°$. (See figure 10-13B.) Note new mileage reading on 15-foot reflector scale (120).

e. Read median level (-78.5 dbm) under hairline on opposite end of cursor.

f. Rotate cursor and center disc together until DUAL-DIVERSITY pointer is set to 12 CH.

g. Read threshold level (-99 dbm) opposite FM THRESHOLD index. See figure 10-13C.)

h. Read reliability (99.6 percent) at intersection of reliability curve, hairline, and curve interpolated for 120 miles. (See figure 10-13C.)

10.5.6 DETERMINATION OF MINIMUM ANTENNA SIZE REQUIRED.

Suppose that all the necessary calculations for a given circuit assuming transmitter output of 1000 watts and 15-foot antennas have been made, and a reference to figures 10-4 or 10-5 shows that the required reliability will not be provided. The minimum antenna size that can be used to meet the reliability requirements can be determined by using the calculator and by following the procedure in the following paragraphs.

a. Select the proper frequency scale.

b. Set DUAL-DIVERSITY or NONDIVERSITY pointer to required channel capacity.

c. Holding center disc stationary, move cursor hairline to intersection of mileage line and reliability percentage line for the reliability required for the circuit (98% - voice, 99.9% - teletype, 99.9% - high-speed data transmission).

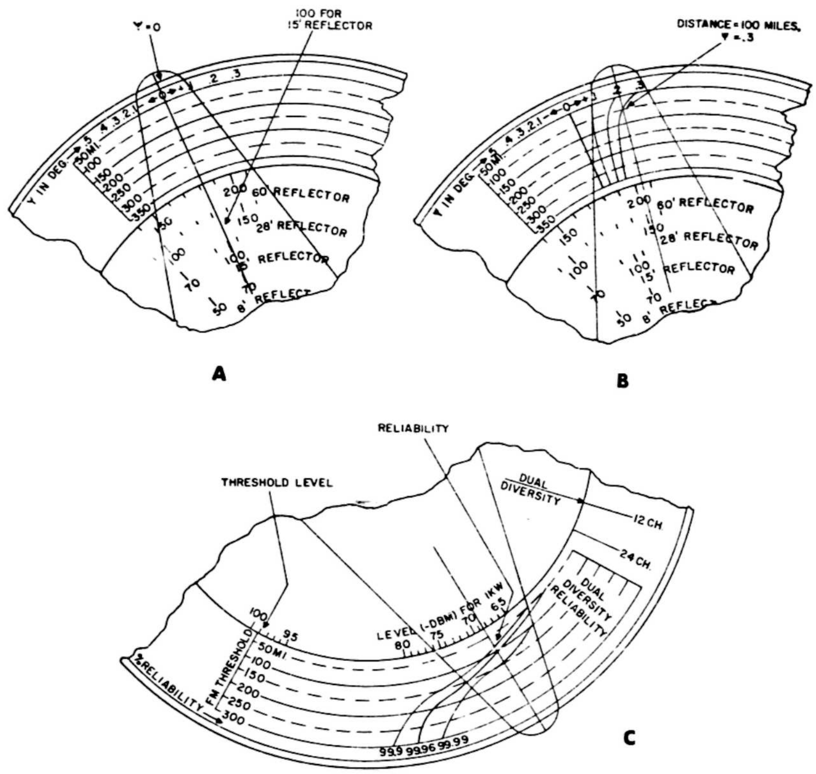


Figure 10-13. Settings of Circuit Calculator

d. The median level required for this reliability appears under the cursor hairline on the RECEIVED MEDIAN LEVEL scale.

e. Move the cursor hairline to intersection of mileage line and sum of horizon angles (in degrees).

f. Holding the cursor stationary, move the center disc until the required median level noted in step d appears under the hairline on the opposite end of the cursor.

g. The minimum reflector size can now be found by noting the relative position of the mileage scales for each reflector size and the cursor hairline. The minimum reflector size is smallest reflector which has the mileage corresponding to the path length to the right of the cursor hairline.

For example, assume a circuit with the following characteristics.

Path length:	150 miles
Channel capacity:	12
Order of diversity:	dual
Sum of horizon angle:	1.0°
Frequency:	1830 mc
Reliability required:	99%

(1) Use 1700- to 2350-mc side of calculator.

(2) Set DUAL-DIVERSITY pointer to 12 CH position.

(3) Holding center disc stationary, move cursor hairline to intersection of 150-mile line and 99%-reliability percentage line.

(4) The median level required for this reliability appearing under the cursor hairline is -83 dbm.

(5) Move the cursor hairline to intersection of 1.0° horizon angle curve and 150 mileage line.

(6) Hold the cursor in this position and move the center disc until -83 dbm appears under the hairline on the opposite end of the cursor.

(7) Note the position of the mileage scale for each reflector in relation to the cursor hairline. Note that the maximum distance possible with an 8-foot reflector is 100 miles; 15-foot reflector, 140 miles; and 28-foot reflector, 175 miles. Therefore, the smallest reflector size which will provide the required reliability for 150 miles is the 28-foot reflector.

10.6 PROCEDURES FOR ESTABLISHING CIRCUIT.

After it has been determined that the circuit is feasible and the required equipment has been selected, the next step is to install the equipment and establish communications. The following procedures are given, assuming that transportable equipment is to be used. All procedures should be carefully planned to reduce setup time at the sites.

10.6.1 FREQUENCY AND POLARIZATION ASSIGNMENT.

Before the equipment is moved to the sites, the operating frequencies and polarizations should be assigned to each terminal. If possible, each terminal should be preset for operation with the assigned frequencies and polarization to reduce setup time at the site. Procedures for tuning and setting polarization are given in the instruction book for the equipment.

10.6.2 SITE LAYOUT AND EQUIPMENT INSTALLATION.

The installation crew should have with them the best available topographic map with the desired site and alternate sites marked. A pocket transit is also required for antenna alignment and site layout.

After confirming that the area is accessible to the equipment, an exact spot should be selected for the shelters and antenna towers. The site should be reasonably level, about 75 feet long and 100 feet wide with the greatest dimension perpendicular to the great-circle path. Allowance should be made for diversity spacing of the antennas. A general requirement is that it should be approximately 100 λ for the operating frequency. For 1000-mc equipment, this is approximately 100 feet; and for 2000-mc equipment, approximately 50 feet.

The detailed instructions for installation will depend on the equipment to be used. However, the basic steps for all types of equipment are as follows.

a. Move equipment into site and unload at preassigned and marked locations.

b. Connect power equipment and external carrier equipment.

c. Energize equipment and tune to operating frequencies if this has not been completed prior to location at site. Steps c and d can be completed simultaneously for reduction of setup time.

d. Erect antennas at assigned locations.

e. Roughly align antennas using hand transit.

f. Complete alignment of antennas by adjusting for maximum signal over circuit.

10.6.3 ANTENNA ALIGNMENT PROCEDURES.

The antennas are set initially for predetermined azimuth and elevation; and then after the stations are in operation, they are adjusted for maximum received signal. The initial alignment has to be only of sufficient accuracy to provide measurable received signals at each site. This can be done with a pocket transit.

To align the antennas initially in azimuth and elevation, it is necessary to know the approximate bearing on one site with respect to the other, and also the approximate horizon angle. The azimuth bearings can be obtained directly from maps. The elevation angle can be calculated from topographic maps or by optical methods. The antenna at one of the sites is adjusted for maximum received signal, while the antenna at the other site remains at the preset position. The antenna at the second site is then rotated for maximum signal. This process is repeated until no further improvement is obtained in the received levels. The entire process of final alignment can be completed in a few minutes.

HEADQUARTERS
U. S. ARMY ELECTRONIC PROVING GROUND
Fort Huachuca, Arizona

SIGPG-OTW

TO: See List

USAEPG-SIG 960-67, "Instruction Manual Tropospheric Scatter Principles and Applications", March 1960 is revised as follows:

Page 6-1	Paragraph 6.1.4, first equation, change "F" to "ΔF".
Page 6-19	Second paragraph, last sentence, change the word "returned" to "retuned".
	Third paragraph, second sentence, change "uhf" to "MOD".
Page 6-21	Change equation in right-hand column to: "20 (log 150 - log X) = 16 where X equals deviation per channel".
Page 6-22	Table 6-1, last item in bandwidth column, change "1.5 kc" to "1.5 mc".
	Right-hand column, first sentence, change to "In the typical 12-channel system, deviation is . . ."
	Right-hand column, second sentence, change "FM" in equation to "F_m".
Page 10-3	Paragraph 10.4.2, last part of sentence above table in left-hand column, change to read ". . . . error rates of 1 bit in 10^5 for high speed data and 2 bits in 10^4 for teletype".

FOR THE COMMANDER:

OFFICIAL
MARSHALL CONNELL
Captain AGC
Asst Adjutant General

F. N. MILLER
Colonel GS
Chief of Staff

1 Incl
Distribution List

UNCLASSIFIED