antenna
manual

antenna manual

by

WOODROW SMITH

formerly

Editor, *Radio* Magazine

Editor, *"Radio" Handbook*

Radio and radar engineer, U.S. Navy

Radar officer, U. S. Marine Corps Reserve

PUBLISHED AND DISTRIBUTED BY

Editors and Engineers
LIMITED

1300 KENWOOD ROAD, SANTA BARBARA,
CALIFORNIA, U.S.A.

This book will be revised periodically as needed to keep it abreast of new developments, new discoveries, and advancements in the art as regards techniques and standard practices.

CONTENTS

PREFACE

The purpose of the "Antenna Manual" is to present in a single volume a comprehensive compilation of antenna, transmission line, and propagation data of vital interest to those associated with the more practical aspects of radio broadcasting and communication. The term "radio" in this case embraces television and certain navigational aids which employ the same basic techniques.

The book is more concerned with such things as good engineering practice and how to arrive at or choose an antenna type, site, or dimensions for a particular practical application than it is with such relatively abstruse and academic things as the method of calculation of the radiation resistance of an antenna of unusual configuration. The book is primarily intended for the man who is more interested in the answer than in the method of arriving at it, but who nevertheless realizes that in many cases a certain familiarity with or appreciation of the method greatly facilitates evaluation, exploration and solution of his own problem.

The reader will find that the emphasis has been placed on engineering approximations and physical reasoning, rather than rigorous mathematical treatment. While more is desirable, no background other than elementary electric circuit theory and a familiarity with the simpler technical aspects of radio communication is essential. This was done in order to make the book useful to the greatest possible number of readers. It is felt that the qualitative discussion method of presentation does not make the data contained therein any less useful to the advanced engineer, should he find need for any of the information. On the other hand if the presentation relied upon mathematical developments to the extent of requiring, say, calculus as a prerequisite, many technicians who came up the hard way and (successfully) rely upon rule of thumb, cause and effect, and trial and error would not be able to assimilate readily all the data contained therein. The same would apply to many of the younger students.

The book does not treat wave guides, electromagnetic horns, and other techniques employed almost exclusively above approximately 1000 megacycles, as the author considers the techniques and applications of microwaves to be sufficiently different as to justify—if not actually require—a different and somewhat more advanced approach than is employed in this book, with more attention paid to fields and less to ordinary electric circuit theory.

Much of the content is simply a codified compendium of information available in technical and scientific periodicals and journals (particularly the *Proceedings of the Institute of Radio Engineers*), various general electronics texts, numerous publications of the Central Radio Propagation Laboratory of the National Bureau of Standards, and publications of the Federal Communications Commission. Except in a very few instances, no claim towards originality can be made other than choice, correlation, and presentation of material. Because of the wide scope of the source of material, embracing almost all currently available literature bearing on the subject (both American and British), it is felt that an attempt to give specific credits throughout would unduly clutter up the book.

WOODROW SMITH

Santa Maria, California

June, 1948

CHAPTER ONE

Radiation and Propagation of Radio Waves

INDUCTION vs RADIATION

In ordinary alternating current theory when the "action at a distance" effect of a *field* is considered, the effect is considered as taking place *instantaneously*, and as far as such applications go this is to all intent and purposes true. However, it does take a finite time for the effect of the field to be felt, the speed with which the effect of a disturbance of the field travels being the same in empty space as the speed of light.

This is not of pertinent interest as regards normal applications of the induction field to transformers, motors, etc., but it *is* important as regards propagation of radio waves, because if the action of a field were everywhere instantaneous, there could not be electromagnetic *radiation*, and without such radiation, a prohibitive amount of power would be required to induce, by induction alone, a signal in a receiving apparatus more than a few miles away. It is true that the induction field is employed in certain signaling apparatus, particularly remote control devices, but they seldom are designed or intended for use over distances of more than a few hundred feet.

It is obvious that signaling by means of an oscillating *induction* field cannot be accomplished over a distance any farther than it is possible to transmit a detectable stationary magnetic or electric field. Either an electric or magnetic field of force in the static form represents energy which is stored in space. When the current or charge is removed, the field collapses, and the energy is returned to the circuit which produced the field.

The strength of a stationary field falls off very rapidly (inversely as the square of the distance, to be specific). Therefore, an enormous amount of current would be required in the generating coil, or "primary", to produce by induction alone a detectable amount of current in the receiving coil, or "secondary", at distances of a few thousand yards. And as previously mentioned, the resistance loss in the wire carrying such a heavy current would represent a prohibitive amount of power.

If alternating current were fed to a large, single turn coil of wire well removed from surrounding objects, and the inductance of the wire were at all times effectively neutralized by means of a power factor correcting capacitor so that the coil acted as a purely resistive (non-reactive) load on the alternator, and if the power dissipated in the wire as resistance loss could be measured, it would be found that at very low frequencies the power delivered by the generator would apparently equal the power dissipated in the coil as heat.

However, if the frequency were raised sufficiently, a point would be found where not all of the power delivered by the alternator or oscillator showed up as heat; there would be a measurable difference. If the frequency were raised still more, it would be found that less and less of the output of the alternator shows up as heat, until eventually (assuming a well insulated coil of high conductivity) only an insignificant percentage of the alternator output shows up as heat loss in the conductor.

Where is the power from the generator going? It has to go someplace, and it does.

It is *radiated,* in the same manner as heat or light. It leaves the coil never to return, and when traveling through empty space the amplitude of the wave is attenuated only by simple dispersion. This means that in empty space the amount of radiation arriving at a distant point (strength of the radiation field) varies directly as the reciprocal of the distance. This attenuation is much less than the previously mentioned attenuation for an induction field, and causes the radiation field to predominate at distances beyond a small fraction of a wave length.

The attenuation of the radiated wave through normal atmosphere will be slightly higher than through a vacuum, but until the frequency of the alternations begins to approach that of radiant heat, the loss through ordinary air is negligible and the attenuation may be considered essentially that which would occur in free space.

If the wire in the single turn loop is stretched out into a straight line and fed either in the center or at one end, the percentage of the alternator or oscillator output which shows up as heat loss in the wire is still less. This indicates that the linear conductor is a better *radiator*, or *antenna,* than the coil, which is explained by the fact that a coil—even a single loop—tends to confine the field.

As it is difficult by ordinary circuit analysis to conceive of electrical power, fed to a wire, vanishing into empty space, a fictitious term, *radiation resistance*, is used to represent a load of equivalent resistance and is a measure of the radiating effectiveness of an antenna. To minimize the amount of power dissipated as heat in an antenna, the radiation resistance should be high and the loss resistance low, the ratio determining the percentage of power which is wasted as heat. Radiation resistance will be treated in more detail later in the book.

WHY AN ANTENNA RADIATES

What causes some of the energy fed to an antenna to be radiated, to be projected indefinitely into space and continue on its way even after the antenna may be taken down? A rigorous and comprehensive discussion of the phenomenon of electromagnetic radiation is highly complex, and not within the scope of this book. However, by simplified physical analogies an attempt will be made to give the reader some idea of what transpires when a wave train of energy leaves a radiator on a one-way ticket for points known and unknown.

First, it will be necessary to assume that the reader is acquainted with the elementary principles of magnetic and electrostatic induction; if he is not, it is suggested that he refer to the subject in any standard text on basic electricity or radio before attempting to proceed further.

One law of magnetic induction states, in effect, that an e.m.f. is induced in a conductor which is *moved* in a magnetic field in such a way as to cut "lines of force". It is interesting to note that if a higher and higher resistance conductor is substituted, there still will be an e.m.f. induced in it under these conditions of relative motion between the magnetic field and conductor. The only difference is that relatively little current can be drawn from the conductor having high resistance. In fact, the e.m.f. even will be induced in an *insulator,* though virtually no current could be made to flow by closing the circuit. The insulator may even be free space, a perfect insulator, and still the e.m.f. is induced, in the form of a *field*.

The same statements apply to a *changing* magnetic field. A changing magnetic field induces an electric force not only in a conductor, but in an insulator, and empty space as well.

The reverse also is known to be true: a changing electric field, such as exists between the two plates of a vacuum dielectric capacitor across which the potential difference is made to vary, creates a *displacement current,* which in turn creates a magnetic field. Therefore we can say that not only does a disturbed magnetic field produce an electric field, but that conversely a disturbed electric field produces a magnetic field. But keep in mind that these two fields do *not* produce each other in space *instantaneously*. A changing magnetic field creates a changing electric field, which then in turn creates a changing magnetic field, and so on; but there is a *finite time lag*. The action is said to be *retarded*. Next consider that a field which is generated by another does not exist in exactly the same place but an infinitesimal distance *beyond*.

Thus we see that when an antenna is fed from an oscillating source of power, the oscillating fields which are set up near the antenna in accordance with ordinary electric circuit theory do not return all their energy to the antenna, but spend some of it in inducing other fields farther out in space, which in turn induce other fields still farther out, and thus are detached or "shaken loose" from the antenna, never to return their energy to the parent circuit.

In this way does the energy permanently detached from the antenna propagate itself through space, the resultant effect being that of *wave motion*. Thus we speak of *propagation* of radio *waves*.

If the reader has difficulty in visualizing the action of retarded fields, in view of the fact that they have no physical actuality, it may be some comfort to know that even the best mathematicians and physicists have never met one in person, and usually consider them simply as a mathematical convenience in explaining observed effects and predicting results. As hard as they are to visualize, the author finds them easier to accept than the premise of an omnipresent conducting medium, or "ether," having no mass. Rather than worry about the matter (or just matter) too long, it is suggested that the reader accept the existence of retarded fields on faith the same as he accepts that of the Creator whom he also cannot see. It is simpler that way, because it is then possible to proceed without further ado to the practical aspects of radiation, where we are concerned more with the effects of a field than we are with the fact that you can neither see it nor weigh it and that it takes finite time to go places.

RECEPTION

The radiation of power into space would be of academic interest if something were not done to utilize the power. The strength of the field at a distance of even a few wave lengths removed from the radiator is too attenuated to be useful for anything but signalling purposes, and it is for such purposes that much money is spent in radiating intelligence-carrying radio-frequency power into space.

It is reasonable to assume that any device which is an efficient radiator of radio-frequency power should also be highly effective in extracting or intercepting power from a passing wave of the same wavelength, and such is the case. This *recipro-*

TABLE I

STANDARDIZED CLASSIFICATION OF RADIO WAVES BY FREQUENCY

Frequency Range	Designation	Abbreviation
Under 30 kc.	Very low frequencies	V.L.F.
30 kc. to 300 kc.	Low frequencies	L.F.
300 kc. to 3000 kc.	Medium frequencies	M.F.
3000 kc. to 30 Mc.	High frequencies	H.F.
30 Mc. to 300 Mc.	Very high frequencies	V.H.F.
300 Mc. to 3000 Mc.	Ultra high frequencies	U.H.F.
3000 Mc. to 30,000 Mc.	Super high frequencies	S.H.F.

city characteristic and special considerations pertaining to receiving antennas will be dealt with in detail later on. For the time being let us concern ourselves with the problem of radiation and propagation of the wave to the point where it is desired to intercept or "receive" it, without worrying too much about how it is intercepted and utilized.

WAVELENGTH AND FREQUENCY

In many respects a train of radio waves in space can be compared to a wave train in water. The radio wave has what correspond in water to peaks and troughs. *Frequency* describes the number of wave cycles which pass a given point in one second. *Wavelength* is a measure of the distance between two successive peaks or troughs, which also is the distance the wave travels during one cycle of oscillation of the parent circuit. Normally it is expressed in *meters* except when a small fraction of a meter is involved, in which case the *centimeter* is more common.

It is apparent that the more waves there are passing a fixed point in a second, the closer together are the peaks and troughs. Also, the distance between succesive peaks or troughs for a given frequency will depend upon the speed of propagation, because the greater the velocity, the farther apart the peaks and troughs. For propagation in empty space (or for all practical purposes in normal air) the velocity is very close to 300,000,000 meters (186,000 miles) per second. The relationship between wavelength and frequency for such a propagation medium may be expressed as follows:

$$F_{kc} = \frac{300{,}000}{\lambda}$$

where F_{kc} is the frequency in kilocycles

and

λ is the wavelength in meters.

The relationship between frequency in kilocycles and free space wavelength *in feet* is shown in simple chart form in figure 3-2 (Chapter 3). The curves are for one quarter and one half wavelength, as these are more useful. For a full wavelength, simply multiply the length given for a half wavelength by two.

POLARIZATION

Radio waves, unlike sound waves, are *transverse*. This means that the stress exerted by the traveling wave is in a direction at *right angles* to the direction of travel. Also, the electric stress is at right angles to the magnetic stress, so that the electric field, the magnetic field, and direction of travel are mutually perpendicular.

Like light waves, which also are electromagnetic waves and therefore transverse, radio waves may be definitely *polarized*. Light waves as emitted normally are random polarized, and must be passed through a polarizing medium before assuming a definite or *linear* polarization. With radio waves, on the other hand, the reverse is true: the radiated wave tends to have linear polarization as it leaves the antenna, unless special means are employed to produce *eliptical* polarization. The latter types of complex polarization always can be broken down into two component waves having definite but different polarization and *not exactly in phase in space*. If the two components were in phase, then the resultant would have linear polarization, instead of being eliptical.

Arbitrarily, the polarization of a radio wave has been defined as the orientation of the electrostatic component. This is convenient, because the electrostatic component is in the same plane as a linear radiator (the most common form of radiator). Thus it is seen that a simple linear radiator oriented horizontally with respect to the earth will emit horizontally polarized waves. It is for this reason that antennas themselves often are referred to as being "horizontally polarized" or "vertically polarized".

WAVE FRONTS

If a "point" on one cycle of a wave train is followed from the time it leaves a radiator, it will be seen to occupy a rapidly expanding sphere. Because of the directional

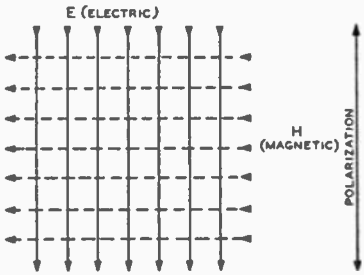


Figure 1-1.

GRAPHICAL REPRESENTATION OF AN ELECTROMAGNETIC WAVE FRONT.

The wave is travelling AWAY from the reader. A half cycle of time earlier or later the direction of both fields of force would be reversed. If the diagram is viewed from the opposite side of the paper by holding it to the light, the effect of an APPROACHING wave front is obtained. Note that ONE field is then reversed.

properties of most radiators, there probably will be two or more directions in which the amplitude of the wave is zero, but nevertheless any finite amplitude of radiation, regardless of how small, will at any instant be contained in an imaginary perfect sphere so long as the transmission medium is homogenious.

The imaginary expanding sphere may be considered as a *wave front.* For this reason radiated radio waves are correctly termed *spherical waves,* even though at considerable distance a section of an imaginary sphere will resemble a plane surface. For practical purposes such as designing a directional antenna array for reception, an advancing wave front may be considered as lying in a plane surface if it is a large number of wavelengths removed from the radiator, just as a very small section of the earth's surface may be considered flat.

For the same reason a comparatively small area of wave front is considered to be not only of equal phase but of equal amplitude, though strictly speaking this would be true only in the case of an isotropic radiator (one radiating equally well in all directions).

Thus we see that (1) an ultimately extended wave front in a uniform medium forms a spherical surface, on which all points are of the same phase but *not necessarily* of the same amplitude, and (2) in actual practice a relatively small section of the wave front is considered as being an equiphase, flat surface or *plane wave* of uniform amplitude.

If we could for a moment look through a magic telescope attached to a radiator of electromagnetic waves and somehow manage to see such a "flat" section of wave front as it travelled away from the radiator through empty space it would look at times like figure 1-1. If, as long as we are assuming supernatural powers, we could instantly transport ourselves to the moon and watch the wave front approach instead of recede, the direction of one set of the lines would appear reversed, just as they do if you hold figure 1-1 to the light and view it from the back side of the paper. Reversing *both* sets of lines simply represents an earlier or later time, or a 180 degree rotational change in the orientation of the viewer, and *not* a reversal of direction of travel.

REFLECTION

From a previous statement that an electromagnetic wave consists at one instant of an electric field and at the next instant of a magnetic field as it propagates itself through space, it might be assumed and correctly so that half the energy resides in the electric field and half in the magnetic field.

Thus when a plane wave strikes a flat surface of infinite conductivity (zero resistance) at normal incidence, the electric field cannot exist in the plane of the conductor, because otherwise infinite current would flow, an obvious impossibility with a finite amount of power. If an electric field cannot exist on the surface of a perfect conductor and there can be no resistance loss because of the perfect conductivity, one might wonder what happens to the energy contained in the electric field when the wave strikes the conducting sheet. It is *converted into a magnetic field* by virtue of the conduction current induced in the reflector by the incident electric field, thus causing the strength of the magnetic field at the surface of the conductor to be doubled.

This oscillating conduction current acts as a radiator of electromagnetic waves and radiates in the only direction it can: away from the conducting sheet. Thus we have an electromagnetic wave travelling in the *opposite direction,* away from the sheet, and with the same amplitude as that arriving at the sheet. No energy has been dissipated; it is all accounted for. In effect the electric component of the incident wave is *reflected* with a reversal in phase and the magnetic component is *reflected* with no change in phase, the result being an electromagnetic wave travelling in the opposite direction in accordance with the explanation of figure 1-1. As will be described later, the reflected waves combine with the incident waves to produce a most interesting resultant or interference pattern called a *standing wave.*

If the intercepting sheet is not a perfect conductor, or if it is a dielectric or leaky dielectric, not all of the energy will be reflected. A sheet consisting of or plated with silver or copper or aluminum is a nearly perfect conductor, and reflection will be almost complete. For most other materials, however, some of the energy will be reflected and some transmitted, the ratio being determined by the permeability, resistivity, and dielectric constant of the material as compared to that of the first medium. The transmitted portion will suffer attenuation as a result of *absorption* by the new medium unless the new medium is a perfect dielectric having infinite resistance.

If, instead of an abrupt change or sharp boundary in the properties of the propagation medium, there is a gradual change, then the amount of reflection is reduced; in fact the reflection will be quite small if the transition is so gradual as to occur over a distance of many wavelengths.

The foregoing discussion assumes that the wave strikes the reflecting surface or plane at normal incidence. If the wave strikes the plane at oblique incidence, reflection still occurs, and as in optics the angle of reflection is equal to the angle of incidence.

REFRACTION

The transmitted portion of a wave striking at *oblique incidence* (the portion that penetrates the new medium) is bent or *refracted* in the same manner as light waves. In fact, the *index of refraction* as applied to optics is applicable also to radio waves, which is not so surprising when it is remembered that they both are electromagnetic waves.

While a sharp boundary condition is necessary for strong reflection, such is not the case for refraction. A change in the characteristics of the propagation medium too gradual to produce appreciable reflection will still do an excellent job of refraction if the total change is sufficiently great. It should be noted that the *path* or *direction* taken by the wave front, sometimes called the *wave normal,* will in such cases *curve* upon entering the zone of gradient, rather than make the sharp angular change in direction observed in the case of a sudden discontinuity in the medium.

Variations in the temperature, density, and water vapor content of the earth's atmosphere cause variations in the dielectric constant which, though of small magnitude, are sufficient to refract radio waves when they strike the gradient at an angle sufficiently near grazing, and thusly cause them to follow a slightly curved or bent path. This will be treated later in detail.

DIFFRACTION

Radio waves, like light waves, tend to follow a more or less straight path when travelling through a uniform medium. However, when either light waves or radio waves impinge upon an opaque object, the shadow is not complete and sharply defined unless the object is large in terms of wavelength. Also, for an object of given transverse cross-sectional area, the shadow will be more sharply defined if the surfaces forming the silhouette outline are comparatively smooth and "deep" in terms of wavelength. For instance, if a point source of light is used, the cutting edge of a sharp knife will not cast as sharp a shadow as the back edge, because the "depth" of the sharp edge is more comparable to the wavelength of light.

This is an important consideration in the study of the propagation of radio waves, because we are vitally concerned with "shadows" or areas of no reception caused by mountains or other obstructing objects. This will be treated in detail later on.

STANDING WAVES

Under "Reflection" it was shown how a wave reflected from a large, perfectly conducting sheet at normal incidence causes the strength of the *resultant* electric component to be zero at the surface of the conductor and the strength of the *resultant* magnetic component to be doubled at the surface of the conductor, because in one case the phase is such as to cancel and in the other to add. If the primary source of radiation is sufficiently far away that attenuation due to the difference in path length between the direct and reflected waves can be ignored, then the same condition will be found to exist at exact integral numbers of half wavelengths from the perfect reflector. Thus, if we should walk away from the reflecting sheet towards the source of radiation with a device for indicating the strength of the oscillating electric field, we would find the amplitude of the resultant to be zero at distances an integral multiple of a half wavelength from the reflecting sheet, and maximum half-way between these points, because at the halfway points the two fields are *in* phase. The points of zero field strength are called *nodes;* the points of maximum strength, *loops*.

At the electric field loops, which are seen to be odd multiples of a quarter wavelength from the reflecting sheet, the *magnetic* components of the two waves are in phase *opposition,* causing *nodes* in the resultant *magnetic* component. Likewise the resultant magnetic component will be maximum at the electric field nodes.

If the distance from the primary source of radiation to the reflector is considered finite, then, except at the surface of the reflecting sheet, the amplitude of the reflected wave will be weaker than that of the direct wave, even though we assume a perfect reflector. If the reflector is not perfect, which is of course the case in actual practice, the amplitude of the reflected wave will be still further reduced. The result is that the resultant field will not drop to zero at a "node", but rather to a *minimum,* because there cannot be *complete* cancellation when the direct and reflected components do not have the same amplitude.

The periodic variation in electric and magnetic field strength with distance as a result of reflection and recombination is known as a *standing wave.* When they result from oblique reflections from imperfect conductors or dielectrics, of irregular shape, the standing wave pattern becomes quite complex. However, we need not concern ourselves with this complexity except to appreciate how the standing waves are formed and why the problem becomes complex when such factors are involved. Complex standing wave patterns are more correctly termed "interference patterns", because in many such cases they do not conform to the simple idealized example previously cited. But in every case a standing wave is fixed in space even though oscillating in magnitude (assuming a fixed radiator and fixed reflecting objects or media).

Standing waves are of pertinent interest in connection with radio frequency transmission lines, and will be treated in more detail in a chapter devoted to them.

INTENSITY INDEX FOR RADIATION FIELD

The yardstick commonly used to measure or specify the amount of energy in a radio wave is the intensity of the electrostatic component expressed in "*microvolts per meter*" or "*millivolts per meter*", the term referring to the dielectric stress existing between two points in the wave front 1 meter apart and lying on a line parallel to the electric lines of force. R.M.S. values are assumed unless otherwise specified.

Because the electric and magnetic field intensities are always equal in free space, the intensity of an undisturbed radiation field is proportional to either the electric field or the magnetic field component. This of course assumes that the distance to the radiator is great enough that the induction fields can be ignored. It also assumes *no reflected wave components,* because the resulting standing wave or interference pattern may show large variations in the electric field intensity between two points only a fraction of a wavelength apart, while in fact there is no such variation in the effective field strength of the wave or *radiation* field when both electric and magnetic components are considered. The same naturally applies to the magnetic field intensity under the same conditions.

An ideal electromagnetic field strength meter for use in areas of standing waves would respond equally well to the magnetic and electric components, and be calibrated in terms of an equivalent electric component in free, undisturbed space. Unfortunately the instruments usually employed do not have equal response to the two component fields, and therefore are strictly accurate only where reflections are insignificant.

RADIO WAVE PATHS

Radio waves of a very low frequency do not behave the same as radio waves of a very high frequency. The waves are basically the same, but the *effect* of physical objects and variations in the propagation medium upon the wave depends upon their relative size, *size* in the case of the wave referring to the *wavelength* and not to the amplitude.

In empty space the waves would act alike, but in radio communication we have to deal with the effect of a spherical earth which doesn't even have a smooth surface—the latter characteristic being aggravated by man's construction efforts—as well as refracting gradients and reflecting media in the earth's atmosphere. For these reasons certain portions of the electromagnetic spectrum are better suited than others for certain applications, and frequencies or "channels" are allocated by the regulating authority in accordance with a master plan which takes these things into consideration in the public interest.

Even the simplest practical radiator will exhibit some inherent *directivity*, which means that it will not radiate equally well in all directions even in empty space. In fact, radiators usually are deliberately designed so as to increase the directivity, either in the vertical plane or horizontal plane or both, in order to avoid the waste of power and the interference represented by radiation in useless or undesired directions.

Because the *basic* directivity of the antenna system proper is a function of the antenna design rather than the presence of surrounding objects, it therefore will be treated in detail under *antenna systems* rather than under *propagation*. By divorcing the two and studying them separately, a better appreciation can be obtained of the directivity due to surrounding objects and variations in the propagation medium as contrasted to the basic or "free space" directivity due to the configuration of the radiating system.

For the present, while studying the effect of the medium and surrounding objects, we shall assume a hypothetical isotropic radiator (one which in free space radiates equally well in all directions), but it is well to keep in mind that the directivity obtained in actual practice is modified by the inherent or basic directivity of the radiating system.

Radio wave paths useful for communication are grouped into three* main categories:

(1) Ground waves

(2) Sky waves, often called "ionospheric" waves for reasons which will become apparent

(3) Low frequency or "guided" waves, applying to frequencies below about 300 kilocycles

It should be kept in mind that in the following discussion of waves (1) and (2), frequencies well above 300 kilocycles (wavelengths below 1000 meters) are assumed in order to facilitate the explanations. Obviously there is no exact frequency where propagation by means (1) and/or (2) suddenly ceases and propagation by means (3) begins, but rather a gradual transition in the mode of propagation in the range of frequencies between 100 and 600 kc.

THE GROUND WAVE

A "ground wave" is a radio wave which

*A *tropospheric* wave, when it exists, may possibly be considered as a fourth category, or it may be considered as a spurious component of the ground wave.

travels along or over the earth and is substantially affected by the presence of the ground. It consists of a *surface wave,* a *direct wave,* and a *ground reflected wave,* the latter two producing a resultant known as a *space wave.*

The predominant component of the ground wave depends upon the wavelength of the radiations and also the height of the radiating and receiving antennas above the earth in wavelengths. Soil or water of given resistivity acts like a better conductor at long wavelengths than it does at short wavelengths. This is understandable when it is considered that the *effective* conducting or reflecting surface of the earth has finite thickness, and that this thickness is a function of wavelength, rather than being a fixed dimension. If the wavelength is ten times as long, the effective layer is several times as thick, and it is apparent that the thicker layer of soil or water has the lower resistance. At sufficiently short wavelengths even salt water acts more like a dielectric than it does a conductor.

THE SURFACE WAVE

At frequencies between approximately 300 and 3000 kc., where the antenna usually is but a fraction of a wavelength above effective earth, the chief component of the ground wave is a component that is guided along by the sudden discontinuity at the juncture of earth and atmosphere. Because it exists only close to the surface of the earth, it is called a *surface wave.*

A typical surface wave has predominantly vertical polarization, but the wave front is given a slight forward tilt as a result of the slower propagation constant of the earth. The amount of tilt depends upon the wavelength and upon the resistivity, permeability, and dielectric constant of the earth over which the wave is passing. In the AM broadcast band, where propagation of the surface wave is of prime importance, the tilt usually runs between approximately 6 and 12 degrees. At very high frequencies, over soil of poor conductivity the wave can have a tilt of more than 20 degrees; while at very low frequencies, over salt water the tilt will be hardly perceptible.

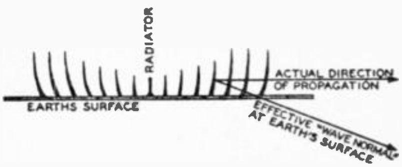


Figure 1-2.

SIMPLIFIED PHYSICAL REPRESENTATION OF A SURFACE WAVE

The wave front assumes practically full tilt by the time it has travelled one or two wavelengths, then maintains this tilt so long as the earth constants remain unchanged. The intensity of a surface wave decreases rapidly with height, becoming insignificant a few wave lengths above ground.

The effect of the tilt is both good and bad: the tilt assists the surface wave to follow the curvature of the earth, but on the other hand it causes a part of the energy contained in the surface wave to be continually deflected downward into the earth and absorbed by virtue of the earth's imperfect conductivity. As this energy must continually be replenished from the upper portion of the surface wave, the absorption represents an undesirable loss of energy.

It should be noted that only a very small amount of tilt, corresponding to a fraction of a degree, is all that is required to facilitate the travel of a surface wave around the earth's curvature. A wave will get down close behind a high range of hills better, however, if the wave has considerable tilt. This means that under certain conditions of wavelength and topography the strength of the surface wave at a shadowed location might actually be increased by poor soil conductivity. For this condition to obtain, the increased bending would have to more than offset the effects of the increased absorption.

Because of the tilt of a surface wave travelling over the earth, the electric vector (and therefore the polarization) has a small component in the direction of propagation and horizontal to the earth's surface. These two components, vertical and longitudinal, are not exactly in phase, with the result that the longitudinal component causes the wave to be *eliptically* polarized rather than linearly polarized. The phase difference of the two components is small under typical conditions, and usually is not of any great consequence from a practical standpoint.

Surface wave propagation (attenuation) formulas and charts have been worked out which combine theory with experimentally observed results, and are surprisingly accurate in determining what field strength may be expected from a transmitter of given power and frequency over various types of terrain. Such data, available from the Federal Communications Commission and the National Bureau of Standards, is widely used to predict the coverage of AM broadcast stations.

EARTH CHARACTERISTICS

It was noted under the discussion of absorption of the surface wave by the earth that the dielectric constant and conductivity of the soil or water determined how a wave of certain frequency will be affected by the presence of the earth. Even at frequencies low enough that the earth acts like a nearly perfect conductor, earth currents of substantial magnitude will flow well beneath the surface of typical soil, or more specifically at depths on the order of 0.05 to 0.5 wavelength. The 0.05 figure applies to the longer wavelengths.

Thus the character of the topsoil is no more important than the nature of the subsoil, 10, 25, or even 50 feet down, depending upon the wavelength and characteristics of the soil. Currents will not flow as deeply in a soil of high conductivity as in a soil of low conductivity; for this reason it is better to have a layer of high conductivity soil over a layer of low conductivity soil than the reverse condition. Another way of stating it would be to say that the importance of the character of the soil 25 feet down becomes less as the conductivity of the soil above this level goes up.

An idea of the order of dielectric constants and relative conductivities of various types of terrain may be obtained from table II. Soil with a conductivity greater than 10×10^{-14} e.m.u. is considered good. Soil with a conductivity of less than 5×10^{-14} e.m.u. is considered poor. The exact value of dielectric constant is not important at frequencies for which surface wave propagation predominates, because it is small com-

TABLE II

Effective ground constants (approximate) in the AM broadcast band for typical types of terrain.

Terrain	Conductivity ($\times 10^{-14}$ e.m.u.)	Dielectric Constant
Sea Water, normal salt concentration	4600	81
Flat farm land or low rolling hills, rich soil	10 to 20	17
Average soil, flat	5 to 10	15
Fresh water ponds and lakes	6	80
Rocky New England hills	2	14
Flat, dry, sandy country	2	10
Towns, city residential area	2	5
Large City (industrial area)	1 or less (widely variable)	4

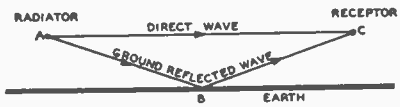


Figure 1-3.

PATHS TAKEN BY DIRECT AND GROUND-REFLECTED COMPONENTS OF A SPACE WAVE

The resultant amplitude of the space wave at C depends upon the amplitudes and relative phase of the two components, which in turn are affected by antenna elevation and ground constants.

pared to the conductivity in any case.

The foregoing discussion is concerned primarily with the effect of earth constants upon the propagation of the surface wave, particularly the attenuation. However, as will be seen later, these constants also have a material influence upon the vertical-plane *directivity* of an antenna system, particularly when vertical polarization is used, and therefore are of importance even under conditions where propagation by means other than the surface wave is employed.

THE SPACE WAVE

For ground wave communication at frequencies above approximately 30 Mc., the radiating and receiving antennas usually are placed several wavelengths above earth. Under these conditions, the amplitude of the surface wave is quite small, partly due to the height of the antennas and partly due to the high absorption of the surface wave by the earth at these frequencies.

The predominant wave under such conditions is the *space wave*, which comprises a direct wave and a ground-reflected wave as illustrated in figure 1-3. The same type of wave is the predominant wave at much lower frequencies (say 3 Mc.) when one end of the circuit happens to be a nearby airplane, because under such conditions of elevation there either is but little surface wave or else the receiving antenna is not within the zone of the surface wave. The term "nearby" is used to qualify the statement because, as will be taken up later, communication over long distances in the wavelength range 3 to 30 Mc. is by neither the surface wave nor the space wave, but by a sky-reflected space wave. But that is getting ahead of our story. For the present discussion of the space wave component of the ground wave we shall assume frequencies in the range 30 to 1000 Mc.

QUASI-OPTICAL CONCEPTS

Waves in the ultra high frequency (u-h-f) range (300 to 3000 Mc.) and in the upper portion of the very high frequency (v-h-f) range (30 to 300 Mc.) have a length small in comparison to commonly encountered objects such as buildings, trees, etc., and therefore such everyday objects cast noticeable shadows. The kinship of such waves to visible light is apparent, and therefore they often are referred to as "quasi-optical" in their behavior, and ascribed rectilinear or "line of sight" characteristics. This is only an approximation, because even visible light waves do not stick strictly to rectilinear paths when dimensions comparable to the wavelength are considered.

Just as the use of geometrical ray construction can be applied to most problems in optics without introducing appreciable error, so it can be applied to most v-h-f and u-h-f wave propagation problems so long as certain limitations are observed and the reader appreciates that use of the method consists of making simplifying assumptions in order to avoid almost insurmountable calculations. As an example, it was assumed in the discussion of figure 1-3 that the ground-reflected component reaching point C was reflected entirely from point B. Strictly speaking the reflection from an area (called a "zone plate") extending for some distance on either side of point B is instrumental in producing a reflected wave at point C. This is because a point in space is "illuminated" or excited at any instant not by a remote radiator of light or radio waves, but by the wave front or field immediately adjacent at the moment, and so on all the way back to the radiator (Huygens' principle).

Fortunately, calculations based on rectilinear or "ray" propagation will in the case of most problems give an answer in close agreement with the exact (and more complicated) treatment in which interference effects are considered in detail. Also, ray propagation is more easily visualized, the mind being somewhat at a loss when it comes to conjuring up a mental picture of an almost infinite number of wave fronts simultaneously acting upon each other to give interference effects resulting in reflection and other phenomena associated with electromagnetic wave propagation.

There is one catch to leaning entirely upon geometrical optics for explanations and solutions: we cannot ignore diffraction, which illuminates areas which the geometrical ray construction will show to be dark. To keep matters reasonably simple, let us keep the rays but assume that even in a homogenious medium they can be coaxed away from straight lines by a small amount by the mere presence of objects of certain size and configuration.

SPACE WAVE PATTERNS

Referring to figure 1-3, it is apparent that the phase difference between the direct component and ground-reflected component arriving at C will cause either cancellation* or reinforcement, and also that the phase difference is dependent upon the difference in path length and also upon the phase shift caused by reflection. For a given distance between A and C, the difference in path length will vary with the height of either A or C or both. The phase shift encountered at B, the reflection point, depends upon the ground constants, the wavelength, the polarization, and the angle of incidence.

For zero angle of elevation, which is the case when two vertical antennas are at the surface of the earth and are short in terms of wavelength, the phase shift always is 180 degrees. Because the difference in path length in this case is effectively zero, it may be seen why the surface wave is the predominant wave in AM broadcasting; the two components of the space wave virtually cancel out, being effectively of equal amplitude and opposite phase.

Actually there is a reduction in amplitude at the reflection point, except in the hypothetical case of a perfectly conducting, perfectly smooth earth. However, the reduction in amplitude for *small angles* of incidence, say a degree or less, is very small even under the worst conditions of conductivity and dielectric constant encountered in actual practice, and for either vertical or horizontal polarization, provided that the terrain is reasonably smooth in terms of wavelength. This means that when points A and C are separated by a distance which is very large compared to the antenna heights, the direct and ground-reflected waves are effectively of the same amplitude, and the resultant is entirely a function of the phase angle and can approach zero for a phase difference of 180 degrees.

For angles not approaching zero, the magnitude of the reflection coefficient varies with polarization, with the angle of incidence, with the conductivity and dielectric constant of the reflecting surface, and with frequency.

If one or both of the antennas should be raised sufficiently it is obvious that the difference in path length eventually would reach a point where the phase shift due to this difference plus that occurring at the reflection point would equal 360 degrees, giving maximum signal strength. If the magnitude of the reflection coefficient approaches 1, then the field strength is substantially twice that which would occur if there were no reflection from the ground. If the antenna or antennas should be raised still further, the resultant would then get weaker, until eventually a null* point would be reached which corresponded to a total of 180 degrees phase angle.

While it is true that with either vertical or horizontal polarization a wave which strikes the earth *at a grazing angle* is reversed in phase without a reduction in amplitude, regardless of the ground characteristics (so long as the ground is perfectly smooth), *the similarity ends there.*

For *horizontal* polarization the magnitude

**Cancellation* does not necessarily mean complete cancellation; it may be either partial or complete.

*The term "null" as applied to radiation patterns, standing waves, etc., is commonly employed to indicate a *minimum* which may either be zero or have a finite value. Maxima are sometimes called *loops.*

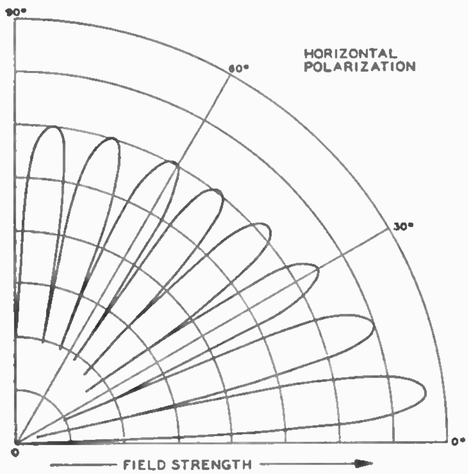


Figure 1-4.

POLAR DIAGRAM SHOWING VERTICAL DIRECTIVITY OF A HYPOTHETICAL ISOTROPIC RADIATOR, HORIZONTALLY POLARIZED, ELEVATED SEVERAL WAVELENGTHS ABOVE TYPICAL SOIL WHICH IS LEVEL IN TERMS OF WAVE-LENGTH

Any vertical directivity contributed by the radiator itself would modify this basic pattern. Only one quadrant of the hemisphere is shown, the other being a mirror image. In comparing with figure 1-5, note that the nulls are more pronounced than is the case with vertical polarization, particularly at low angles. Except over water, the regularity and depth of the nulls are much less pronounced in actual practice because of terrain irregularities. As explained in the text, the number of lobes is a function of the height of the radiator above ground.

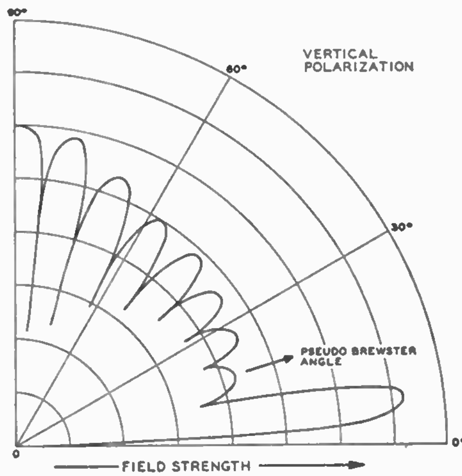


Figure 1-5.

POLAR DIAGRAM SHOWING VERTICAL DIRECTIVITY OF THE HYPOTHETICAL ISOTROPIC RADIATOR OF FIGURE 1-4 WHEN VERTICALLY POLARIZED

Vertical directivity contributed by the radiator itself will modify the basic pattern. In comparing with figure 1-4, note particularly the severe irregularity or distortion of the lobes at what corresponds to the Brewster angle in optics. Also observe that the angle of the lowest lobe corresponds closely with that of the lowest lobe of figure 1-4, but that the highest lobes coincide approximately with the high angle nulls of figure 1-4. This is explained by the fact that with vertical polarization the phase angle of the reflection coefficient changes from 180 degrees at grazing angles to 0 degrees at the normal, while for horizontal polarization the phase angle is substantially constant

of the reflection coefficient falls off gradually and smoothly as the angle is increased, until it reaches a value at vertical incidence which is determined by the frequency, and by the dielectric constant and conductivity of the ground. This value usually will be between 0.5 and about 0.98 for typical types of earth (including sea water) at v-h-f and u-h-f frequencies. The phase angle gradually decreases, but for all types of terrain it is so small that for practical purposes it may be ignored, and the wave may be considered to be reflected with no change in phase for all vertical angles.

The result, for a well-elevated, horizontally-polarized antenna over all types of smooth earth encountered in actual practice, is a series of nulls in the vertical directivity pattern, nulls which are very pronounced and follow a simple formula as related to the elevation of the radiator in wavelengths: The number of nulls between 0 and 90 degrees is equal to the number of half wavelengths of radiator elevation plus 1. Thus, with an ultra high frequency radiator 50 half wavelengths above earth (not uncommon at very high frequencies), there will be 51 nulls between zero and 90 degrees angular elevation. This assumes level terrain.

Actually the "effective elevation" is not easily predicted under conditions of unlevel terrain, where for instance the height of the reflection point varies with the angle CAB of figure 1-3. Also, for irregularities of certain configuration in the terrain there will be scattered reflections which prevent substantial cancellation and thus obscure or reduce the magnitude and regularity of the nulls.

An example of the pattern of lobes and nulls obtained with a horizantally polarized isotropic radiator located many wavelengths above terrain that is level for considerable distance in every direction is shown in figure 1-4*. The representation is not strictly accurate, because it ignores the curvature of the earth, refraction (bending) due to atmospheric gradients, and also ignores the fact that the angle made by the two rays converging at the receiving point varies with distance. These factors, which will be discussed in more detail under v-h-f and u-h-f propagation and practice, tend to make the angle of elevation at which a null occurs vary slightly with distance from the radiator. This condition cannot be depicted in a simple polar diagram of the type illustrated, but the curvature is sufficiently small beyond the immediate vicinity of the radiator that it may for the purpose of this discussion of nulls be ignored.

When a *vertically* polarized radiator is employed, conditions are somewhat more complex. The phase shift and the magnitude of the reflection coefficient are much more dependent upon the angle of incidence and the characteristics of the earth at the reflection point. Instead of falling away from unity gradually and only moderately as the angle is increased from 0 to 90 degrees elevation, as is the case for horizontal polarization regardless of terrain, the magnitude of the reflection coefficient drops rather rapidly to a comparatively low value at an angle corresponding to the Brewster angle in optics, then increases again to the same value as for horizontal polarization at 90 degrees elevation.

The phase shift at the reflection point drops from 180 degrees at zero elevation to approximately 90 degrees at the pseudo Brewster angle previously mentioned, then more slowly to 0 degrees at vertical elevation. For a given wavelength the elevation of the pseudo Brewster angle varies widely with the type of terrain, being much lower over sea water than over poor soil. However, the variation in the angle with different ground characteristics tends to become less as the frequency is raised.

An example of the pattern of lobes and nulls obtained with a vertically polarized isotropic radiator located many wavelengths above terrain that is level for considerable distance in every direction is shown in figure 1-5. It will be noted that the null depth varies with elevation angle, becoming less pronounced as the magnitude of the reflection coefficient becomes lower. As with the illustration for horizontal polarization, figure 1-4, figure 1-5 is not a true representation because of the impractability of showing on such a diagram how the vertical angle at which a null occurs changes slightly with distance between points A and C of figure 1-3.

TROPOSPHERIC REFRACTION AND REFLECTION

In a normal, well-mixed or "standard" atmosphere the refractive index decreases very gradually with height at a substantially constant rate. This causes space wave components radiated at low angles of elevation (below approximately one degree) to be bent downward in a slightly curved path, the curvature being constant regardless of height under the aforementioned conditions. The effect is to cause v-h-f and u-h-f waves to be propagated beyond the optical horizon without benefit of diffraction.

Under these "standard" conditions the curvature is such as to extend the *effective* horizon to where it would be if the atmosphere were homogenious (no curvature of the wave path) and the earth's radius were

*The vertical or horizontal plane directivity of an antenna ordinarily is illustrated by means of a *polar diagram.* Such diagrams show field strength vs. direction or angle for a constant distance, as in figures 1-4 and 1-5, or else show range vs. direction or angle for a constant field strength. *In free space* the curves are the same, because the field strength in free space is inversely proportional to distance. For ground distance at constant antenna heights, the curves will not be the same.

increased by a factor of 1.33. This means that the refractive index of a "standard" atmosphere has a gradient such as to cause the wave path to follow a curve having a radius four times that of the earth. The extended horizon is called the "radio path horizon".

Actually the factor of "standard" refraction varies slightly with latitude, but the variation is so small that the figure of 1.33 is generally used without regard to latitude.

It should be kept in mind that "standard" refraction is an average or mean and that usually the variation in refractive index with height will be greater or less than that of "standard" atmosphere, being dependent upon meteorological conditions in the troposphere. The distance to the radio path horizon then will be greater or less than that obtained under conditions of standard refraction. However, so long as the curvature of the path is substantially independent of height, the propagation is known as ground wave propagation.

Under some conditions, which occur rarely in some geographical areas and often in others, the variation in dielectric constant with height is not substantially constant, but tends to be *stratified*. In some cases the stratification is very pronounced, being caused by relatively sharp discontinuities which represent the boundaries of dissimilar air masses. In the case of a sufficiently sharp discontinuity the mechanism whereby the wave is returned downward resembles reflection more than it does refraction, there being no sharp line of distinction between the two.

A wave refracted or reflected back to earth by a stratified layer in the troposphere (lower atmosphere) is called a *tropospheric wave*. The curvature of the path of such a wave is not substantially constant at all heights and throughout its length, but rather tends to be concentrated at one or more points.

A tropospheric wave may exist in combination with the regular ground wave (within the radio path horizon or on into the diffraction region), or may exist alone at distances so great that the regular ground wave is too weak to be of significance. When both a tropospheric wave and the regular ground wave exist in combination, the two will add or subtract (depending upon the relative phase, which is constantly changing). When the two components are of the same order of magnitude, interference fading will result. The same condition can exist beyond the regular ground wave range when there are two or more tropospheric paths.

Under certain conditions of tropospheric propagation variously known as "guided" propagation, "duct" transmission, or "trapping", a u-h-f tropospheric wave will travel far beyond the distance at which the regular ground wave no longer can be received. This will be treated in further detail under a discussion of v-h-f and u-h-f practice, along with other aspects of tropospheric propagation at v-h-f and u-h-f.

Even though tropospheric waves may travel far beyond the regular ground wave range, the field strength varies with air turbulence and weather conditions, resulting in a fading or fluttering signal even when there is but one tropospheric wave present and no significant ground wave exists. Thus, it may be seen that while tropospheric waves may permit reception well beyond the distance at which the regular ground wave ceases to be useful, they usually produce an unsteady, unreliable signal and therefore are of no great benefit or usefulness from the standpoint of commercial broadcast or communication applications. On the contrary, the fading effects produced by them within the range of the regular ground wave makes their existence undesirable in most instances. Generally speaking, the effects tend to increase in magnitude slowly with increased frequency.

Tropospheric waves sometimes are considered components of the regular ground wave, and at other times classed separately as a form of sky wave that is not an ionospheric wave. If by arbitrary definition they are not part of the ground wave, they at least should be considered as closely associated with the ground wave, because they seldom reach a height of more than 8000 feet and usually do not exceed 4000 feet.

It is true that at times it is difficult to draw the line or make a distinction between an "extended" ground wave and a tropospheric wave, the designation being determined by arbitrary and unstandardized limits with regard to variations in the curvature over the wave path. However, to avoid confusion, waves which unquestionably are refracted or reflected from a stratified "layer" of atmosphere or the interface between two dissimilar air masses will be considered as falling in a separate category, neither strictly ground wave nor sky wave, and will be referred to specifically as tropospheric waves.

SKY OR "IONOSPHERIC" WAVES

Radio waves which leave the earth at too great an angle to be reflected or refracted back to earth by gradients or discontinuities in the troposphere are not always destined for points in outer space. There is an ionized region called the *ionosphere* starting about 30 miles above the earth and extending out to the limits of the earth's atmosphere at about 300 miles, a region which reflects or refracts radio waves, oftentimes sufficiently well to return a usable signal to earth at great distances from the transmitter.

The manner in which the waves are sent back towards earth by the ionosphere is not a matter of simple refraction or reflection, but the end result is the same, and therefore for all practical purposes the wave is considered to be reflected as though from a conducting mirror. However, to understand and appreciate some of the peculiar effects which occur, it is desirable that the reader at least be familiar with the rather complex manner in which radio waves are bent by the ionosphere. Such bending is responsible for medium and long range transmission in the approximate range 2 to 30 Mc. While sky waves are not as reliable or predictable as ground waves, they are much more so than tropospheric waves, and are widely utilized for long range communication.

THE IONOSPHERE

The ionosphere consists of stratified "layers" of ionized atoms and molecules (particularly nitrogen and oxygen) in the earth's upper atmosphere. The primary source of this ionization is solar radiation, particularly ultra violet radiation, though there are other factors contributing to the ionization.

When an atom of nitrogen or oxygen is struck by electromagnetic energy from the sun in the form of ultra violet radiation of a "sympathetic" frequency, some of the electrons in the atom or molecule are set into such violent oscillation that an electron is shaken loose, resulting in a "free" electron and an ion. This free electron will eventually collide or *recombine* with a positive ion, the "mean free path" or distance it travels and the time it takes to recombine being much greater at high elevations where the atmosphere is less dense and the atoms, molecules, electrons, and ions are farther apart.

Until it does recombine with a positive or "heavy" ion, or temporarily attach itself to an atom, the electron is capable of being set into oscillation by radio waves, some of the energy from the passing radio wave being expended in setting the electron into oscillation. The electron then *reradiates* the energy, with little loss, provided it has time to do so before recombining with a heavy ion or hitching a ride with a passing atom. If not, the energy imparted to the electron by the radio wave is largly dissipated and randomly scattered in the collision, thus reducing the reradiated energy. The manner in which a radio wave representing the energy reradiated by oscillating free electrons gets deviated from the course of the exciting wave will be taken up later; first let us examine the nature of and reasons for the ionization at different altitudes.

At a height of about 300 miles where it begins (or ends) the earth's atmosphere consists primarily of molecular and possibly atomic nitrogen. As the altitude is reduced, this gas becomes more dense, and at about 100 miles some atomic oxygen begins to occur. The gases continue to become more dense with decreasing altitude, and at about 60 miles the oxygen begins to take molecular rather than atomic form. At about 40 miles a layer of ozone is encountered.

The ionizing radiations (and possibly particles) from the sun have a *selective* effect, so that the particular type of gas existing at a certain altitude is susceptible to certain radiations (and possibly particles) but not to others. Thus when the energy from the sun enters the earth's atmosphere, *certain portions* of it are spent in ionizing the gases present at that altitude. As what is left proceeds towards the earth, other portions are successively spent in ionizing the particular gases through which the radiations must pass.

Because of this selective effect and the variation in temperature, composition and density of the atmosphere with altitude, the rate at which free electrons *are formed* does not change uniformly and gradually with altitude, but rather tends to be *stratified*.

The rate at which free electrons recombine with heavy ions is just as important in determining the number of free electrons *existing* at a certain altitude as is the rate at which they are formed. The variation in the rate of recombination with altitude tends to accentuate the stratification of the density of ionization, and also to make the height of maximum *density* of free electrons in each "layer" slightly *higher* than the height of maximum *production* of free electrons.

To get a better idea of how the ionized strata are formed, let us examine the formation of one of the principle "layers". When the sun's radiations first strike atoms of nitrogen about 300 miles above the earth, rays within a certain band of wavelengths ionize these particular atoms. As the rays travel towards the earth, the density of the nitrogen (atoms per unit volume) increases and therefore the rate of ionization (free electrons formed per unit volume in a given time) also increases *at first*. However, because the "resonant" rays expend energy in ionizing the nitrogen atoms, the ionizing energy becomes weaker as the density of the gas becomes greater. Consequently, below a certain altitude the formation or production of free electrons decreases rapidly to zero even though the density of the gas is still increasing.

The result is a distribution of free electron formation or production similar to the curve of figure 1-6. This *Chapman curve*, as it is called, is for but one narrow band of ionizing frequencies and one type of gas, and is a representation of the *rate* at which the free electrons are formed. It should be kept in mind that the number of free electrons *existing* depends not only upon the rate at which they are formed, but also upon the rate at which they recombine.

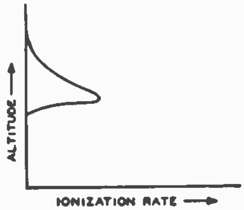


Figure 1-6.

TYPICAL CHAPMAN CURVE

Graphical representation of the variation with altitude of the production of free electrons in the earth's atmosphere by solar radiation. This curve applies to but one type of atom and one kind of radiation, and does not take recombination into account.

It also should be appreciated that there is a similar curve for each type of atom and ionizing wavelength, and that a composite curve for *all* radiations and existing gases which takes the *recombination rate* into consideration is necessary for a true picture of the stratification of ionization in the atmosphere. Such a composite curve is shown in figure 1-7, which represents ionization density on a typical summer day. Four distinct maxima or "humps" will be observed. These are the principal layers of the ionosphere in the *daytime* during *summer*. It is apparent that the layer whose formation we just studied is the F_2 layer, also that there is no zone of zero ionization or well defined dividing line between the ionized regions; they *overlap*.

It might seem that, if the principal ionizing agent is solar radiation, then the ionization density of all layers might drop to zero or nearly so when the sun went down. If there were rapid recombination at all altitudes and if the sun were the only ionizing agent, this would be true, and there would be no ionosphere after sunset. However, at high altitudes, where the gas is rare and atoms are far apart, and the mean free path is therefore long, recombination is so slow that a fairly high electron density still is

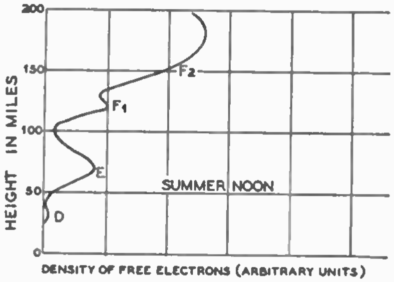


Figure 1-7.

TYPICAL DAYTIME DISTRIBUTION OF FREE ELECTRONS IN THE ATMOSPHERE

The E and F_2 layers are the only permanent layers in the ionosphere. The D and F_1 layers appear only in daytime, and during part of the solar cycle only in summer.

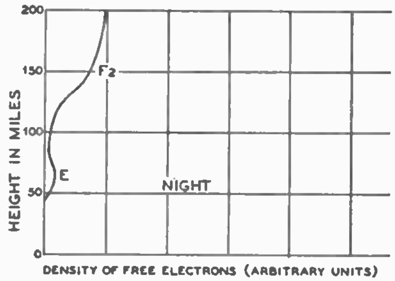


Figure 1-8.

TYPICAL NIGHTTIME DISTRIBUTION OF FREE ELECTRONS IN THE ATMOSPHERE

The D and F_1 layers (fig. 1-7) disappear at sundown and the density of the E layer drops to a low value, where it remains all night long. Under normal conditions, the F_2 layer slowly decreases in density during the night.

present just before sunrise. So the F_2 layer exists all night long; but because the free electrons forming the F_1 layer have recombined, the F_1 layer disappears or merges into the F_2 layer,* as is illustrated in figure 1-8. The same condition sometimes prevails during the daytime in midwinter; the F_1 layer does not always appear because the free electrons may recombine as fast as they are formed, leaving only the F_2 layer.

At the level of the E layer, recombination is very rapid, and this layer would disappear shortly after sunset except for other ionizing factors which maintain a low but effective magnitude of ionization all night long. The nighttime E layer is "splotchy" or "porous", and therefore rather ineffective at angles of incidence approaching the normal. The causes of this nighttime ionization at the level of the E layer are not definitely known, though there is evidence to indicate that meteorites are one contributing factor. For equatorial regions the nighttime E has a steeper gradient, or sharper bounary, on its underside and therefore will return higher frequencies to earth than in higher latitudes.

It will be noted from figure 1-8 that the D layer completely disappears at night. Recombination at this comparatively low altitude is very rapid, and the sun seems to be the sole ionizing agent.

To summarize, the two permanent layers or ionized regions are the F_2 layer, and the E layer. Under normal conditions the former is strong at all times, but the latter is strong only in the daytime. There are two layers that are present only under certain conditions: the F_1 layer, which is just under the F_2 layer and weaker, and the D layer, which is just under the E layer and weaker.

The height of the F_2 layer usually is between 125 and 250 miles, and varies appreciably from year to year, season to season, day to day, and even hour to hour. The height of the E layer is very constant at approximately 70-75 miles. The height of the D layer varies between 20 and 55 miles when it is present, and usually is between 40 and 50 miles.

It should be kept in mind that the foregoing discussion is for normal or "undisturbed" conditions, and that, as will be discussed later, there are occasional times when the foregoing picture does not apply at all. Also bear in mind that the magnitude as well as the height of maximum density of free electrons can vary with time.

*The nighttime F_2 layer was formerly designated the F layer.

Before going deeper into the character and behavior of the stratified regions composing the ionosphere, let us first examine the manner in which a "layer" in the ionosphere effectively acts like a reflecting mirror to radio waves.

"BENDING" PROPERTIES OF AN IONIZED MEDIUM

There are two ways to explain how a radio wave gets bent back to earth by the ionosphere. The first explanation, while not strictly accurate, is simple and may satisfy the reader who is not too curious and is more interested in effects than causes.

It is a well known fact that ions are a good conductor of electricity. Under "reflection" it was shown how a radio wave is reflected from a conducting sheet by setting up conduction currents in the sheet. A layer of the ionosphere, containing many ions, acts as a conducting sheet to the radio waves and reflects them back to earth in accordance with the laws of reflection. It is such a simple explanation that it is a shame that it isn't strictly accurate except possibly for very long wavelengths, where the "thickness" of an ionized layer is negligible in terms of wavelengths.

What actually transpires is much more complex. Upon striking the fringe of an ionized region the velocity of the electromagnetic energy is slowed down, because the propagation constant of the ionized medium is slower than that of atmosphere which is not ionized. It might seem that under these conditions the tranmitted portion of a wave striking at oblique incidence would be refracted towards the normal, or upwards. Fortunately this does not occur.

Instead the electric component of the wave sets the free electrons into oscillation, *oscillation of the same frequency but not in phase with the exciting wave.* The result is that while the travelling electromagnetic disturbance or bundle of energy actually has been slowed down, the *phase* at a given point in the medium has been shifted until it is the same as it would be if the propagation had been speeded up. The result is that the wave reradiated by the free electrons acts as though it *had* been speeded up, and the net effect is the same as refraction *away* from the normal. In one sense the wave *has* been speeded up, even though the velocity of a "bundle" of energy (such as a short spurt or pulse of radiation) has been slowed down. By definition a wave front is an equi-phase surface, and by shifting the phase of a portion of a wave front with respect to the remainder, we can *effectively* advance or retard that portion of the wave front with respect to the remainder. This brings up the distinction between phase velocity and group velocity.

PHASE VELOCITY AND GROUP VELOCITY

It is impossible for electromagnetic energy to travel faster than the characteristic velocity of space (the speed of light in a vacuum). The velocity of an infinitesimally short pulse, or unit function, can never exceed V_c, or about 186,000 miles per second.

If the pulse is lengthened to include several cycles, the wave train or group of wave fronts taken *as a whole* will travel at the same speed as the unit function or infinitesimally short pulse. However, it is possible under some conditions for the individual wave fronts comprising the pulse to "move up through the pulse" as though the pulse were equipped with caterpillar tractor treads and the wave fronts were the cleats on the top half of the tracks. The cleats on the top portion of the tracks may be travelling faster than the caterpillar, but they don't get any place ahead of the tractor. The tractor (including cleats as a group) may be said to be travelling at *group* velocity, while an individual cleat may be said to be travelling at *phase* velocity.

In free space, the group velocity and phase velocity of a wave train are the same. In a complex medium such as the ionosphere, they differ and the phase velocity may even exceed V_c. Just as the lateral direction of travel of a tractor is determined by the relative speed of the left and right hand tracks rather than the speed of the tractor as a whole, so the direction of travel of a bundle of electromagnetic disturbance or energy is determined by the relative phase velocity of different portions of the disturbance. When a wave strikes the ionosphere at oblique incidence it is effectively deflected downward as it passes into progressively denser ionization and greater phase velocity. If the density of free electrons reaches a sufficiently high magnitude, the wave will be "turned over" and sent back towards the earth.

The density of free electrons required to bend the wave back to earth is a *function of the wave frequency and the angle of incidence.* The higher the frequency or the steeper the angle, the greater the required density of free electrons. This is explained by the fact that for a given electromagnetic field strength, a free electron will vibrate with less amplitude at high frequencies, because of the mass or inertia of the electron. Also, less bending is required to return a grazing wave than is required to return one of less oblique incidence.

To summarize: as the frequency is raised, a point finally is reached where the free electrons are not excited sufficiently to cause the wave to be refracted back to earth. The frequency above which this occurs is determined by the density of free electrons and the angle of incidence of the exciting wave. Conversely, the elevation angle above which this occurs (or the angle of incidence below which this occurs) is determined by the wave frequency and the ionization density; or, the magnitude of ionization density below which this occurs is determined by the wave frequency and the angle of incidence.

TERRESTRIAL MAGNETIC EFFECTS

The electric component of a plane-polarized electromagnetic wave will cause a free electron to oscillate in a line parallel to the electric field. Once the electron is moving, it is subject to influence or deflection by a magnetic field, in accordance with the basic laws of electricity. For this reason the earth's magnetic field, when it is acting in a transverse direction on the electron, will cause it to deviate from a straight line. The pulling effect of the magnetic field is proportional to the velocity of the free electron. The velocity of the free electron is greater at low frequencies, because a cycle lasts longer and the electron has a longer time in which to accelerate before the direction of the exciting electric field reverses. Hence, the effect of the earth's magnetic field on an oscillating free electron becomes greater as the frequency is lowered.

At very high frequencies the transverse pull of the magnetic field on the moving electron is small compared to the force exerted by the exciting wave. The resultant path of a free electron subjected to the two forces is a rather flat elipse, which becomes more circular as the frequency is lowered.

At low frequencies the pull is so great that the electron makes several tight loops during one back-and-forth journey or cycle. These loops are made only when the transverse pull due to the magnetic field is strong enough, just as a reed which vibrates in a strong wind will only bend over slightly into a steady position when the wind is not sufficiently strong.

At some intermediate frequency the electron doesn't seem able to make up its mind whether to travel in an elipse (one loop) or to make two loops per excitation cycle. The result is that it makes a *spiral*, gathering momentum until it finally collides with an atom or ion. Because of the high speed it reaches under these conditions, it is very effective in absorbing power from the exciting wave, but because of the spiral nature of the path the reradiation tends to be random and therefore comparatively ineffective; also, because of its continual accelleration, it will collide with an ion and dissipate its energy much sooner than a free electron of slightly different frequency which does not spiral and therefore is limited in its velocity.

The frequency at which a free electron is disposed to take a spiral path is termed the *gyro frequency*, and the frequency at which this resonant effect occurs varies with geographic location, being dependent upon the earth's magnetic field. The frequency is that at which free electrons tend to rotate around the lines of magnetic force of their own accord when accellerated with an infinitesimally short "kick". For Washington, D.C. it is approximately 1400 kc., but elsewhere may be anything between 700 and 1600 kc., depending upon the strength of the magnetic field at the particular location.

The earth's magnetic field, by the action just described, affects not only the absorption and polarization of radio waves travelling in the ionosphere, but by a complex process causes an ionosphere layer to exhibit two refractive indexes. This means that a wave entering an ionosphere layer will be split into two components, one travelling more deeply into the layer than the other.

The one travelling more deeply is called the ordinary ray, and at vertical incidence (but not oblique incidence) acts as though the magnetic field were absent . The component which is more readily bent is called the extraordinary ray and suffers the greater absorption, particularly near the gyro frequency. The two merge into a single signal having random polarization upon leaving the ionosphere, provided that the density is sufficient to refract the ordinary ray at the particular frequency and angle of incidence involved. The splitting is more pronounced and of more import in the F_1 and F_2 layers, particularly the latter, than in the D and E layers.

ABSORPTION

It is obvious that a very important factor in the propagation of electromagnetic energy through the ionosphere is absorption, because the bending of a radio wave back to earth is of academic interest if the absorption is so great that the ionosphere-reflected signal has been attenuated to the extent that it cannot be utilized by a sensitive receiver.

When a free electron which has been excited by a radio wave collides with a particle, some of the kinetic energy which the electron received from the wave is reradiated in random phase and the rest is dissipated as heat. The energy which is reradiated from free electrons in random phase contributes nothing to the orderly reradiated wave induced by those free electrons which are oscillating rythmically and have not yet collided or recombined. The random radiation produced by collision therefore represents wasted energy.

It was previously stated that when a free electron is excited by a passing wave, its oscillation is not in phase with that of the exciting wave. If the free path of the electron (distance it travels before collision or recombination) is such that the electron has had time to absorb considerable energy from the wave but has not had time to reradiate it, *absorption will be maximum.* Thus, for a high frequency wave in a given medium, the absorption per unit of time will vary with frequency, becoming less as the frequency is raised because it takes less time for a free electron to pick up and reradiate some energy, and its chances of reradiating appreciable energy before colliding with an atom, molecule or ion are improved.

The regular variation in absorption with frequency is disturbed, however, by the effect of the earth's magnetic field. We have already noted how the gyro or "precession" effect causes almost complete absorption of a wave at the gyro frequency. At very low frequencies, below the gyro frequency, the geo-magnetic field tends to reduce absorption by slowing down the mean velocity of the free electrons as a result of the looping action previously described under the discussion of the effects produced by the earth's magnetic field. Therefore, below some frequency near the gyro frequency, the absorption *decreases* with frequency.

Because of the comparatively high density of nonionized particles in the lower portion of the E region and on down through the D region, absorption is greatest at those levels, particularly in the daytime. Thus it may be said that the major portion of the absorption suffered by an ionospheric wave occurs at the lower fringe of what is commonly called the ionosphere. It is readily apparent that for effective high frequency communication via the ionosphere it is necessary to use a frequency sufficiently high that the wave can reach and leave the main refracting portion of the ionosphere without excessive attenuation. Under normal conditions there is little absorption below the D region because while the particle density (gas pressure) is comparatively high, there are no free electrons present to extract energy from the passing wave and dissipate it by premature collision.

VIRTUAL HEIGHT

We have referred at various times to the height of an ionosphere layer. Yet we have indicated, as is apparent from figures 1-7 and 1-8, that the "layers" of ionization do not have sharply defined boundaries. Also,

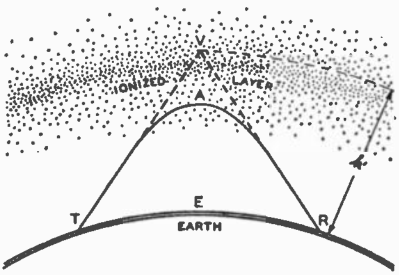


Figure 1-9.

VIRTUAL VS. ACTUAL HEIGHT OF IONOSPHERE LAYER

The wave takes the actual path TAR, but the angles of arrival and departure, and the time taken for the energy to reach point R, are the same as though the ionized layer were not present and the wave were reflected from a mirror surface at point V. Distance AE is the maximum height actually reached by the wave; distance VE is the "virtual height" of the layer, sometimes called equivalent height or effective height, and abbreviated h'. The layer illustrated is the E layer

it was pointed out that as the frequency of a wave is increased, it penetrates more deeply into a layer before being refracted back to earth.

Fortunately we are more concerned with effects than with causes; so it is expedient to ascribe to a layer a *virtual height,* which is the equivalent or effective height if the wave is considered as being *reflected* from a *sharp boundary* rather than refracted by a gradient. Figure 1-9 illustrates the relationship between virtual and actual heights. In travelling from transmitter T to receiver R the wave takes the *actual* path TAR, but the angles of departure and arrival are the same as though the wave had taken the *virtual* path TVR by virtue of mirror reflection at point V. By a coincidence or plan of nature, the time taken for electromagnetic energy to travel the path TAR is for all practical purposes the same as it would take to travel the path TVR if the latter path were all through a nonionized medium. This is explained by the fact that the reduction in group velocity of the wave while it is travelling through the ionized layer is just sufficient to make up for the shorter path length. The distance VE is the virtual height of the layer for the frequency involved; AE is the maximum height reached by the wave.

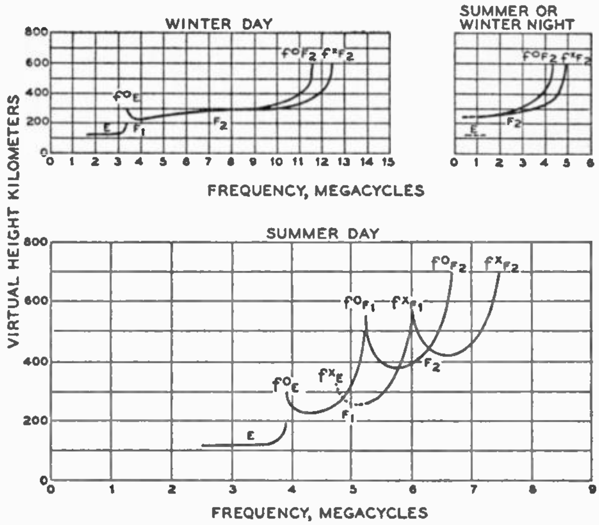
The height a wave penetrates into an ionosphere layer before being refracted downward depends upon the frequency of the wave and the angle of incidence. The less the angle of incidence (the closer to vertical or normal incidence) and the higher the frequency, the deeper the maximum penetration. For reasons of standardization and common reference, ionosphere virtual heights are based upon the wave striking at vertical incidence. The height either is given graphically as a function of frequency or else the lowest observed critical height is assumed.

At vertical incidence, the wave connot be visualized as being bent back to earth by refraction in accordance with the physical picture of refraction presented in figure 1-9. The wave *is* returned to earth at vertical incidence if the density of ionization is high and the frequency is sufficiently low; it doesn't at some angle closely approaching vertical incidence suddenly have difficulty deciding which way to bend and therefore keep right on going. The return at vertical incidence can be explained by an optical analogy by simply saying that the wave penetrates the layer until it reaches an ionization density having an index of refraction equal to zero. The following simplified physical picture may assist in understanding what transpires when a wave is returned vertically.

When a wave penetrates an ionized layer at vertical incidence, the wave reradiated by the oscillating free electrons travels in the same direction (vertical) as the incident exciting wave and combines with it. But after the wave penetrates to a certain ionization density, the phase shifted wave which is reradiated by the free electrons reaches an amplitude sufficiently great that, when combined with the exciting wave, the resultant electromagnetic wave has its electric component shifted in phase with respect to its magnetic component to the extent that the direction of wave travel is reversed. (Refer back to the discussion of figure 1-1.) In some respects this may be a specious explanation, but the process is not one which is easily explained in everyday terms or by simple physical analogies.

Figure 1-10.

TYPICAL h'-f CURVES SHOWING VARIATIONS IN VIRTUAL HEIGHTS WITH FREQUENCY



CRITICAL FREQUENCY

If the frequency of a wave striking an ionosphere layer at vertical incidence is raised, a frequency finally will be reached where the wave penetrates to the maximum ionization density of the layer before its direction is reversed. This frequency is called the *critical frequency*. A higher frequency will pass the region of maximum density without being reversed in direction, and therefore on out the upper side of the layer.

At a frequency about 80 or 85 per cent of the critical frequency, the *virtual* height is equal to the height of maximum ionization density, assuming a normal distribution of ionization density in the layer and negligible retardation of the wave while passing through lower layers. Below this frequency the virtual height falls off very slowly with frequency, which means that the virtual height is never greatly different from the actual height until the critical frequency is approached.

The earth's magnetic field will cause a layer to have two critical frequencies, one for the ordinary ray and one for the extraordinary ray. The effect is particularly noticeable in the case of the F_2 layer. For critical frequencies well above the gyro frequency, the difference between the critical frequency for the ordinary ray and the critical frequency for the extraordinary ray will be approximately half the gyro frequency.

Thus it is seen that the critical frequency is primarily a function of the ionization density, but also is affected by the strength of the geomagnetic field. The critical frequency of the ordinary ray, which is not affected by the geomagnetic field at vertical incidence, will be found to vary as the square root of the number of free electrons.

It will be noted in figures 1-7 and 1-8 that the maximum density of ionization of any layer is greater than the maximum density of all lower layers. This always is the case under normal or "undisturbed" conditions, and it means two things: first, that the critical frequencies of the layers increase in the same order as their heights, and secondly (as a result of the foregoing) a wave having a frequency just barely too high to be returned by one layer will in all probability be returned by a higher layer, provided there is a higher layer.

Obviously a wave which is sent straight up and comes directly down is of little use in long distance communication. However, vertical incidence data are more readily obtained than oblique incidence data, because the transmitting and receiving apparatus composing the ionospheric measuring or "sounding" equipment are more readily handled when located together. And the data obtained at vertical incidence can be applied to oblique incidence propagation by means of factors which in practice have withstood the test of time, as will be discussed under the practical aspects of high frequency propagation. Suffice it to say for the moment, the vertical incidence data permit an accurate estimation of the lowest "perforating" frequency for any layer and angle of departure

In the vicinity of the critical frequency, the *virtual* height of a layer increases rapidly to a comparatively high value before the wave finally penetrates the layer completely, or "perforates" it. It might first appear that near the critical frequency a slight increase in frequency greatly increases the depth of penetration into a layer. This is not the case; the depth of actual penetration is increased only slightly, but the *retardation* of the wave (reduction in group velocity) becomes great at frequencies which permit the wave to penetrate to the region of maximum electron density before being bent back to earth. Because of the increase in the length of time the wave spends in the ionized region, there is a sharp increase in absorption at or near the critical frequency.

IONOSPHERE MEASUREMENTS

The importance and utility of comprehensive data on the ionosphere at different points in the world is obvious. When sufficient data are available, it is possible to predict the performance of long distance communication circuits with a high degree of accuracy.

Because of its many advantages, one method of "sounding" the ionosphere is employed almost exclusively. It is the "pulse-echo" system, developed by Messrs. Breit and Tuve many years before the advent of radar but utilizing the same principles.

A pulse of radio frequency power, so short that it is measured in microseconds, is sent skyward. The time taken for it to return to earth is observed visually on a calibrated cathode ray oscillograph, or else recorded on a recording oscillograph.

For most of its ionospheric data the National Bureau of Standards employs fully automatic, multi-frequency recorders which plot the virtual heights of the layers against frequency. The transmitter and receiver move automatically in small increments over the band of frequencies to be observed, so that a substantially continuous plot of height vs. frequency is obtained every few minutes. From these data, charts can be made up showing the variation in the nominal virtual height or the critical frequencies of the various layers with time, the time scale being hours, days, years, or whatever is desired. If years are used, for example, the values of course are averages, or daytime averages, or summer daytime averages, or whatever will best illustrate the information it is desired to convey. By means of such plots it has been possible to establish the existence of definite correlations between the behavior of ionospheric propagated waves and certain regular cycles of nature, as well as certain sporadic phenomonon.

Typical frequency sweep curves (called h′-f curves) showing virtual height plotted against frequency are shown in figure 1-10. Critical frequencies are indicated by f, with the superscript indicating the ordinary (o) or the extraordinary (x) ray, and the following character(s) the layer. Thus f^xF_2 indicates the critical frequency of the extraordinary ray of the F_2 layer.

Typical curves showing the variation in virtual heights and critical frequencies with time are shown in figure 1-11. These particular graphs illustrate the diurnal (daily) variations encountered in the USA during summer and during winter for both the peak and bottom of the sunspot cycle. These variations, as well as others, will now be discussed in detail.

NORMAL VARIATIONS IN THE IONOSPHERE

Because the characteristics of the ionosphere are so dependent upon the amount of solar radiation falling upon it, it is to be expected that there will be cyclic variations in the heights and critical frequencies of the various layers which correspond to diurnal (daily) and seasonal variations in the amount of radiation striking any particular portion of the ionosphere. There also are recurrent variations which correspond to the

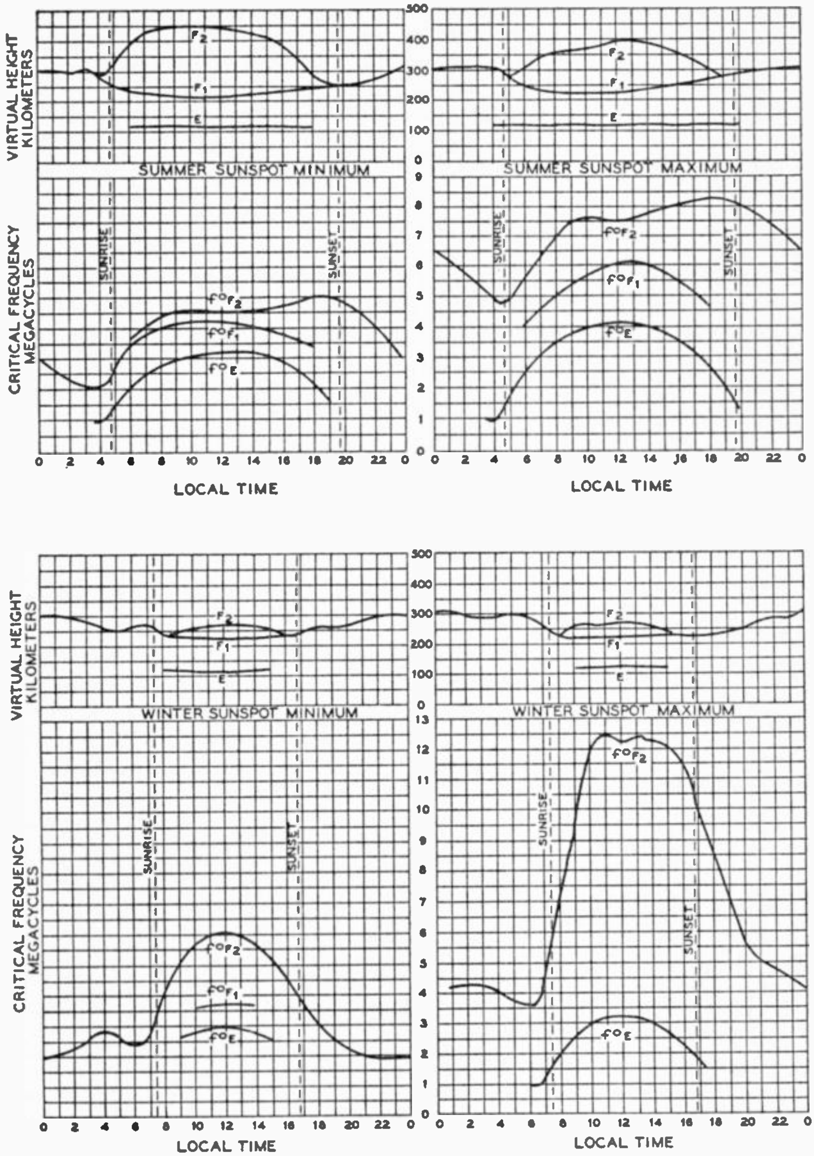


Figure 1-11.

TYPICAL CURVES SHOWING DIURNAL VARIATION OF VIRTUAL HEIGHT AND CRITICAL FREQUENCY IN THE U.S.A. AT DIFFERENT SEASONS AND FOR MAXIMUM AND MINIMUM PERIODS OF THE SUNSPOT CYCLE.

27-28 day solar rotational cycle and to the 11.1 year sunspot cycle. It is also to be expected that the character of the ionosphere at the equator will differ from that in the polar regions.

Diurnal Variations

Diurnal variations are most pronounced in the D, E, and F_1 layers. The density and therefore the critical frequency of the E layer and of the F_1 layer follows the altitude of the sun quite closely because of the rapid recombination in these regions, but the E layer does not die out completely at night because of other ionizing factors. The critical frequency of the E layer is maximum at high noon, as would be expected. The critical frequency of the D layer is not considered because the frequency is so low, and because the D layer is primarily an absorbing layer.

Because of the slow recombination of the F_2 layer, the diurnal variations in critical frequency may lag behind the altitude of the sun by as much as several hours. In the winter the maximum is reached shortly after noon, but in summer it is not reached until late in the afternoon.

Diurnal variations in the virtual heights of the layers are not as pronounced as the variations in critical frequency. The virtual height of the E layer remains practically constant over a 24 hour period, while the F_1 and F_2 virtual heights vary only slightly over a like period, the greatest diurnal variation in the Northern Hemisphere being that of the daytime increase of the F_2 in summer.

It should be kept in mind that the foregoing applies only to normal or "undisturbed" conditions.

Solar Rotational Cycle

All portions of the sun's surface are not equally effective in emitting ionizing radiations or particles. "Quiescent" (stable) sun spots have an appreciable effect upon the ionosphere and tend to occur in patches, the distribution of the patches varying over a long period of time but being fairly constant over a few 27 day solar rotational cycles. This means that superimposed on all other variations there is a tendency for ionosphere conditions to repeat every 27 or 28 days.

It should be kept in mind that the foregoing applies only to normal or undisturbed conditions.

Seasonal Variations

There is considerable variation in the critical frequency of the F_2 layer with season and appreciable variation in the critical frequency of the E layer with season. The F_1 layer may even appear and disappear with change of season. There is considerable variation in the virtual height of the F_2 layer with season, but the virtual height of the E layer remains substantially constant.

Seasonal variations in the F_2 layer do not follow the same pattern in both hemispheres, though the variations in the lower layers do. A possible contributing factor in the case of the F_2 layer is the fact that due to the eliptical character of the earth's orbit around the sun, the earth is some 3 million miles closer to the sun on January 1 than it is on July 1. This causes the rays in the northern hemisphere to be more slanting when the earth is closest to the sun, while in the Southern Hemisphere they are more slanting when the earth is farthest from the sun, thus possibly affecting either directly or indirectly the outer regions of the ionosphere.

The critical frequencies of the E and F_1 layers tend to be higher in the summer than in the preceding or succeeding winter. This applies to both hemispheres.

The critical frequency of the *nighttime* F_2 layer tends to be lower in the winter in both hemispheres, except for a pre-dawn rise or "hump" which sometimes reaches considerable proportions and is not fully understood.

The critical frequency of the *daytime* F_2 layer tends to be *higher* in *winter* in the *Northern* hemisphere, and in the *spring* and *fall* in the *Southern* hemisphere. This is another one of the anomilies of the ionosphere which has not been satisfactorily explained, but must be accepted because of its regularity.

Unlike the weather, ionosphere conditions tend to follow the sun's season position with no lag. Another interesting fact is that in the Northern hemisphere, conditions tend to be definitely "summer" or definitely "winter" most of the year. During the relatively short transitional periods at the time of the equinoxes, conditions tend to be somewhat erratic.

It should be kept in mind that the foregoing discussion of seasonal variations applies only to normal or undisturbed conditions.

The Sun Spot Cycle

Ionizing radiation is increased by (or at least occurs at the same time as) maximum sunspot activity. The intensity of ionization varies with the intensity, number, and solar latitude of sunspots, and the aggregate effect shows a pronounced 11.1 year periodicity. High frequency propagation also may be affected *indirectly* by the sunspot cycle; ionospheric waves are affected somewhat by the earth's magnetic field, and the latter is affected either by sunspots or by the same force that causes periodic sunspot variations. In any event there is a definite correlation between magnetic activity and the sunspot cycle.

The intensity of solar radiations does not reach the same magnitude at the peak or trough of each cycle, and sometimes a cycle is slightly accellerated or decellerated, but nevertheless the periodicity is very pronounced. The minima tend to lag behind the maxima by slightly more than 6 years, rather than half the 11.1 year period. This means that the sunspot activity dies out somewhat more slowly than it builds up.

The most recent "bottom" occurred in the spring of 1944, and the next peak probably will occur in the spring of 1948 and be more intense, with the next bottom in 1954 or 1955. Because of minor fluctuations in the periodicity, predictions are necessarily approximations or estimates, particularly when made many years in advance. Short-time predictions (three months or less) are comparatively accurate.

The critical frequencies of all of the ionosphere layers tend to follow the sunspot cycle, being highest at the peak of the cycle. The variations due to the sunspot cycle can be isolated from those due to other factors by taking the yearly mean critical frequencies for normal (undisturbed) conditions. When this is done, it is seen that the sunspot cycle has been responsible for a variation of about 1.25 to 1 in the critical frequency of the E layer, 1.5 to 1 for the summer daytime F_2, and about 2 to 1 for the F_2 at other times. The variations during the next sunspot cycle may be greater or less than these values, but they do give some idea of the relationship and magnitude to be expected.

Effect of Solar Eclipse

A complete solar eclipse produces an effect on the E layer which is much the same as the diurnal effect except that it is more rapid. This is a logical expectation, and therefore the effect is considered "normal", in spite of the rather rare occurence of a solar eclipse at any particular locality. Because of the short duration of the eclipse, little effect is produced upon the F_2 layer.

Geographical Variations

When the zenith angle of the sun is low, the ionizing effect of solar radiations is reduced. This causes a *latitude effect,* which results in higher critical frequencies in equatorial regions than in polar regions. Also, as would be expected, there is less seasonal variation near the equator.

There is also a *longitude effect;* ionosphere characteristics are not the same at all points on a degree of latitude for the same local time. The effect is related to geomagnetic latitude, which varies considerably from regular latitude because of the appreciable geographical distance between the magnetic poles and the true poles. Thus the close relationship between terrestrial magnetism and solar radiation as regards ionospheric propagation again is brought to our attention.

The variation with longitude is not appreciable in the case of the regular E layer. It is the critical frequency of the F_2 layer which is considerably affected by geomagnetic latitude, and therefore subject to a "longitude effect".

ABNORMAL IONOSPHERE VARIATIONS

The variations discussed up to this point are the normal variations which occur with regularity as a result of nature's cyclic tendencies, and can be predicted well in advance as to approximate time, magnitude, etc.

But in addition to her cyclic tendencies, nature gives us some very vexing sporadic phenomona to cope with. While some of these effects can be predicted as more likely to occur at certain times than others, they nevertheless occur with little or no warning, and because the magnitude oftentimes is sufficient to cause a complete cessation of sky wave communication, it is important that one know a little about them even if nothing can be done about them.

Days during which abnormal ionosphere variations occur are called "disturbed", and when none occur, "undisturbed".

The Dellinger Effect

Sporadic solar eruptions of light gasses, called hydrogen flares, result in solar radiation containing frequencies which pass through the upper layers of the ionosphere with little absorption but have a particularly strong ionizing effect on the earth's atmosphere in the D region near the top of the ozone layer, at about 38 miles. This ionization causes very high absorption of high frequency waves, but reduced absorption in the very low frequency range where the waves are reflected rather than refracted. The effect is more pronounced the more perpendicular the sun's rays, and does not occur in the hemisphere which is not illuminated by the sun at the moment. If the entire wave path is in total darkness, little if any effect will be noted.

The onset is very sudden, which is responsible for the effect being clasified as a "sudden ionosphere disturbence", commonly referred to as a "drop out" or "fade out". High-frequency signals drop to a very low intensity or out altogether within a fraction of a minute. The eruption itself seldom lasts more than a few minutes, but the effects are prolonged for a total time of from 15 minutes to an hour (or occasionally longer) because of the time taken for recombination of the absorbing ions.

Above the gyro frequency range, the absorption by the flare-activated absorbing screen varies inversely with frequency, with the result that the higher frequencies are the first to penetrate the screen when it starts to subside. Thus, after the initial eruption of the flare, communication sometimes can be restored immediately by shifting to a higher frequency.

Prolonged Excessive Absorption

The same type of ionizing radiation sometimes is emitted from solar disturbances which build up and die out gradually, unlike the sudden burst of a hydrogen flare. The deleterious effect upon the D layer and high frequency propagation is identical to that of a "drop out", but the onset and decay are more gradual and the maximum absorption usually does not reach as high a peak value. Quite often frequencies near the perforation frequency of the F_2 layer will not be affected appreciably, and communication can be maintained during one of these prolonged periods of low-layer absorption by utilizing a frequency sufficiently high that absorption is not excessive.

Ionosphere Storms

Certain types of solar disturbances apparently cause the emission of "clouds" of electrified particles or corpuscles which travel at very high velocity and sometimes reach the ionosphere. These clouds have been detected by pulse measurements, approaching at very high speed well beyond the maximum height of the F_2 layer.

Upon reaching the ionosphere, they produce erratic effects upon the layers of the ionosphere, and at the same time appear to create fluctuations in the earth's magnetic field, which abnormalities are called a "magnetic storm".

The particles enter the ionosphere some 30 hours after emission by the sun, travel through the F_2 and F_1 layers, under some conditions creating violent turbulence in these layers, and finally spend themselves in the E layer, which also becomes disturbed and erratic under some conditions. The necessary conditions in each case are a sufficiently dense particle cloud and sufficiently high geomagnetic latitude.

An ionosphere storm may be light, moderate, or severe, and the effects may last from a few hours to several days, depending upon the severity of the storm and the frequency of the wave.

While the Dellinger effect (drop out) is more pronounced in the equatorial regions, where the sun's rays are more perpendicular, the effects of an ionosphere storm are much more pronounced in the *regions of the magnetic poles,* because of the relationship between ionosphere storms and terrestrial disturbances. When a signal path lies entirely in the equatorial region, the signal will not be appreciably affected during ionosphere storms.

While a severe ionosphere storm is in progress, the regular ionosphere layers tend to become diffused, shift in height and density erratically, and demonstrate a general turbulence which increases with geomagnetic latitude. The F_2 layer is most strongly affected, and it will show wide and rapid variations in virtual height and critical frequency during the storm, sometimes even disappearing entirely for short periods at the peak of the storm in the higher latitudes. At other times the F_2 layer appears to have not a single virtual height, but a *band* of virtual heights. These anamalous reflections are called "spread echoes" and are responsible for severe distortion of a signal, as well as an ambiguous critical frequency.

Usually there is increased absorption of high frequency waves in the lower ionosphere regions during an ionosphere storm. The absorption generally will increase with latitude until in regions near the magnetic poles there oftentimes is complete absorption, just as in a "drop out". These polar regions in which complete absorption is likely to occur are the *auroral zones,* or regions in which aurorae occur. The aurorae seem to be activated by the same solar particles which have such a disturbing effect upon the ionosphere. These regions are called "magnetic blankets" during an ionosphere storm.

The maximum intensity of an ionosphere storm will depend generally upon the character of the solar particle cloud and upon geomagnetic latitude. The maximum effect may be the extreme turbulence previously mentioned, or the only observed effects may simply be a slight reduction in critical frequency and increase in virtual height of the F_2 layer. During the climax of a mild storm, conditions will be similar to those existing towards the tail end of a severe storm, the ionosphere disturbances going through more or less the same phases before subsiding.

Ionosphere storms are more frequent during periods of high solar activity, which means they are more numerous (and more severe) during the peak of the sun spot cycle. Also, they tend to recur in accordance with the 27 day solar rotational cycle for a few cycles, after which a new pattern evolves.

The general effect of ionosphere storms upon high frequency propagation is to increase fading and absorption, increase virtual heights, and lower the critical frequencies.

At *very low* frequencies, below approximately 100 kc., the effect of the increased ionization in the D layer during an ionosphere storm tends to *reduce* the absorption of the wave, because waves of such low frequency do not ordinarily penetrate the D layer anyhow.

The effect on AM broadcast band signals is not great within the ground wave range, but causes erratic variations in signal strength at distances large enough that the sky wave ordinarily predominates. At times the sky wave signal may actually be increased, but ordinarily it is decreased by a rapidly varying amount.

The ultimate intensity of an ionosphere storm is not necessarily related to the rate at which it builds up. A severe storm may take from several minutes to several hours to reach its maximum intensity. If the storm is slow in developing, its approach is first heralded by irregularities in the F_2 layer in the auroral region quite some time before the effects are apparent elsewhere.

Sporadic E Region Ionization

Just slightly above the maximum height to which waves ordinarily penetrate the E layer before being bent earthward, there occur at odd times and odd intervals very dense patches of free electrons. These shallow clouds of free electrons comprise a relatively thin "layer" but have very sharply defined lower boundaries, and tend to produce *reflection* of waves in the lower high frequency range, and refract waves at grazing incidence well up into the v-h-f range (50 to 70 Mc.) when the ionization reaches peak values.

These patches may be from a few hundred yards to several hundred miles across. At times they appear to travel with considerable velocity; at other times they are almost stationary. They may exist for only a few minutes, or may last for hours. The season and time of day of maximum occurance vary with geographical location and other factors, not all of which are known or understood.

Whereas the virtual height of a normal layer increases rapidly when the critical frequency is approached (due to greatly increased retardation of the wave), no such effect occurs in the case of sporadic E reflections; the waves act as though they were returned by partial or complete reflection from a mirror surface. When the reflection is partial, not all the wave is reflected. This is more likely to occur the more normal the incidence. Thus it is possible to have return waves at normal incidence from both the sporadic or abnormal E layer (E_s) and the regular F_2 layer simultaneously at the same frequency. At oblique incidence it is possible to have waves reflected simultaneously from the E_s and from some other layer for another reason: there may be a small patch of E_s ionization at the reflection point for the E layer, but no patches to interfere with the higher angle F_1 or F_2 refracted wave path to the same receiving point.

There seem to be two distinct types of sporadic E reflections. They both occur at the same height and both return waves to earth having a frequency much higher than the critical frequency for the normal E layer, yet they differ somewhat in certain other respects.

The first type, sometimes called the *blanketing* type, appears to be impervious to waves of all frequencies, producing complete reflection. It is most prevalent in the auroral zones, during periods of high sunspot activity, yet also occurs in temperate latitudes during the peak of severe ionosphere storms. In all localities it occurs almost exclusively at night. During the daylight hours, it seems that sporadic E ionization of this type takes on the nature of an absorbing screen rather than a mirror, and reflections are seldom noted in the daytime.

The second type, sometimes called the *reflecting* type, produces anything from complete or almost complete reflection to partial reflection in which only a very small proportion of the wave is reflected, depending upon the frequency. As the frequency is lowered, the reflection is more complete. This type is the more prevalent in the United States and localities of equivalent latitude. It is more likely to occur during periods of *low* sunspot activity, and unlike the blanketing type, often occurs during the day as well as at night.

Much remains to be learned about sporadic E ionization and its propagation characteristics, but rapid progress is being made as more and more data become available for correlation.

PERTINENT IONOSPHERIC WAVE PHENOMENA

There occur in connection with propagation of waves via the ionosphere several phenomena which can be considered neither strictly "normal" nor "abnormal", and in some cases are so related that it is unfortunate they must be taken up singly.

Polarization

A wave which has definite or linear polarization will exhibit eliptical polarization (having both vertical and horizontal components) after being reflected or refracted by the ionosphere. This is caused by the twisting effect exerted by the earth's magnetic field upon a wave traveling oblique to it when the wave strikes an ionized medium. Ordinarily the relative magnitude of the vertical and horizontal components, as well as the phase difference between them, will vary randomly with time because of the unstable character of the ionosphere and the earth's magnetic field.

Skip Distance and MUF

There is a fairly exact relationship between the critical frequency for a given

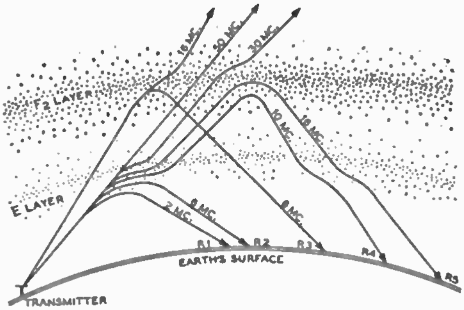


Figure 1-12.

SIMPLIFIED PRESENTATION OF REFRACTION IN THE IONOSPHERE, ILLUSTRATING SKIP DISTANCE AND THE MANNER IN WHICH THE PATH OF A WAVE IS DETERMINED BY THE FREQUENCY

This ray representation indicates how a wave of given frequency is bent back to earth only if the angle of departure and the frequency are low enough. The distance between the transmitter and receiver R5 is the skip distance for a 16-Mc. wave, as it is not possible to refract the wave from the F_2 layer at a steeper angle and thereby return a 16-Mc. wave to earth at shorter distances. The 16-Mc. wave arriving at R5 has penetrated to the maximum electron density of the F_2 layer. The ionosphere conditions depicted are what might occur during a winter day.

layer and the minimum "perforation" frequency for any oblique angle. This will be discussed in detail later under *High-Frequency Communication Practice.*

Any wave is more readily bent as the angle of incidence is increased, and a wave which perforates an ionosphere layer at vertical incidence may be readily refracted back to earth at some smaller angle of departure. Or, for a given angle of departure, it will be found that waves below a certain frequency (here called the "lowest perforating frequency") are refracted back to earth by a given layer while higher frequencies are not. Obviously the lowest perforating frequency at normal incidence is the critical frequency, the latter expression being reserved exclusively for normal incidence in order to abide by convention and avoid possible confusion.

It is apparent that a wave of a frequency just a little too high to be returned from any layer at normal incidence will in all probability be returned from the F_2 layer (the most dense) if the angle of departure is lowered sufficiently. The angle at which this occurs, together with the virtual height of the refracting layer, determine the minimum distance at which a direct sky wave of this frequency can be returned to earth under these ionosphere conditions. This distance is called the *skip distance* for that frequency, and the area encompassed by the skip distance in all directions but beyond the ground wave range is called the *skip zone.* Actually, as far as the sky wave is concerned, the skip zone embraces everything within the skip distance; but by convention "skip zone" excludes the area covered adequately by the ground wave, unless otherwise stated.

Skip distance is represented pictorially in figure 1-12. The distance between the transmitter and point R5 is the skip distance for a 16 Mc. wave. For reasons to be discussed, the signal is not necessarily absent in the skip zone, but it usually is from 50 to 100 db lower in intensity than the field strength just beyond the skip distance, and is subject to severe distortion.

Just as the distance T-R5 is the skip distance for 16 Mc., so 16 Mc. is the *maximum usable frequency* for direct sky wave transmission for the distance T-R5. A higher frequency will not be refracted sufficiently to be returned at point R-5. Thus it may be said that the maximum usable frequency or *MUF* for a certain distance is the frequency at which the specified distance becomes the skip distance.

Scattered Reflections

Scattered reflections are random, diffused, substantially isotropic reflections which are partly responsible for reception within the skip zone, and for reception of signals from directions off the great circle path.

In a heavy fog or mist, it is difficult to see the road at night because of the bright glare caused by scattered reflection of the headlight beam by the minute droplets. In fact, the road directly to the side of the car will be weakly illuminated under these conditions, whereas it would not on a clear night (assuming flat, open country). This is a good example of propagation of waves by scattered reflections into a zone which otherwise would not be illuminated.

Scattering occurs in the ionosphere at all times, because of irregularities in the medium (which result in "patches" corresponding to the water droplets) and because of random-phase radiation due to the collision or recombination of free electrons. However, the nature of the scattering varies widely with time, in a random fashion. Scattering is particularly prevalent in the E region, but scattered reflections may occur at any height, even well out beyond the virtual height of the F_2 layer.

There is no "critical frequency" or "lowest perforating frequency" involved in the scattering mechanism, though the intensity of the scattered reflections due to typical scattering in the E region of the ionosphere decreases with frequency.

When the received signal is due primarily to scattered reflections, as is the case in the skip zone or where the great circle path does not provide a direct sky wave (due to low critical or perforation frequency, or to an ionosphere storm), very bad distortion will be evident, particularly a "flutter fade" and a characteristic "hollow" or echo effect.

Deviations from a great circle path are especially noticeable in the case of great circle paths which cross or pass near the auroral zones, because in such cases there often is complete or nearly complete absorption of the direct sky wave, leaving off-path scattered reflections the only mechanism of propagation. Under such conditions the predominant wave will appear to arrive from a direction closer to the equator, and the signal will be noticeably if not considerably weaker than a direct sky wave which is received under favorable conditions.

Irregular reflection of radio waves from "scattering patches" is divided into two categories: "short scatter" and "long scatter".

Short scatter is the scattering that occurs when a radio wave first reaches the scattering patches or media. Ordinarily it is of no particular benefit, as in most cases it only serves to fill in the inner portion of the skip zone with a weak, distorted signal.

Long scatter occurs when a wave has been refracted from the F_2 layer and strikes scattering patches or media on the way down. When the skip distance exceeds several hundred miles, long scatter is primarily responsible for reception within the skip zone, particularly the outer portion of the skip zone. Distortion is much less severe than in the case of short scatter, and while the signal is likewise weak, it sometimes can be utilized for satisfactory communication.

During a severe ionosphere disturbance in the north auroral zone, it sometimes is possible to maintain communication between the Eastern United States and Northern Europe by the following mechanism: That portion of the energy which is radiated in the direction of the great circle path is completely absorbed upon reaching the auroral zone. However, the portion of the wave leaving the United States in a *southeasterly* direction is refracted downward from the F_2 layer and encounters scattering patches or media on its downward trip at a distance of approximately 2000 miles from the transmitter. There it is reflected by "long scatter" in all directions, this scattering region acting like an isotropic radiator fed with a very small fraction of the original transmitter power. The great circle path from this southerly point to northern Europe does not encounter unfavorable ionosphere conditions, and the wave is propagated the rest of the trip as though it had been radiated from the scattering region.

Another type of scatter is produced when a sky wave strikes certain areas of the earth. Upon striking a comparatively smooth surface such as the sea, there is little scattering, the wave being shot up again by what could be considered specular or mirror reflection. But upon striking a mountain range, for instance, the reradiation or reflected energy is scattered, some of it being directed back towards the transmitter, thus providing another mechanism for producing a signal within the skip zone.

Meteors and "Bursts"

When a meteor strikes the earth's atmosphere, a cylindrical region of free electrons is formed at approximately the height of the E layer. This slender ionized column is quite long, and when first formed is sufficiently dense to reflect radio waves back to earth most readily, *including v-h-f waves which are not ordinarily returned by the F_2 layer.*

The effect of a single meteor, of normal size, shows up as a sudden "burst" of signal of short duration at points not ordinarily reached by the transmitter. After a period of from 10 to 40 seconds, recombination and diffusion have progressed to the point where the effect of a *single* fairly large meteor is not perceptible. However, there are many *small* meteors impinging upon earth's atmosphere every minute, and the *aggregate* effect of their transient ionized trails, including the small amount of residual ionization that exists for several minutes after the original flash but is too weak and dispersed to prolong a "burst", is believed to contribute to the existence of the "nighttime E" layer, and perhaps also to sporadic E patches.

While there are many of these very small meteors striking the earth's atmosphere every minute, meteors of normal size (sufficiently large to produce individual "bursts") do not strike nearly so frequently except during some of the comparatively rare meteor "showers". During one of these displays a "quivering" ionized layer is produced which is intense enough to return signals in the lower v-h-f range with good strength, but with a type of "flutter" distortion which is characteristic of this type of propagation.

FADING

To a greater or lesser degree there always are random fluctuations in the intensity of a signal received via sky wave propagation. Sometimes the variations will have components which, for a short time, appear more or less periodic; at other times no rhythm will be apparent. When the fluctuation in field intensity is rapid enough and great enough to be readily noticeable, it is called *fading*. Long time fluctuations (with a period exceeding approximately one minute) are more often called "variations in field strength", rather than fading. There is no sharp dividing line, though any variation rapid enough that the average eye or ear detects at once that the intensity is changing usually is called a fade (assuming no automatic gain control at the receiver).

There are many types of fading when classified subjectively as to aural characteristics, and several types when classified objectively as to the condition, effect, or phenomenon responsible for the fading. Fading is variously *described* as "slow", "fast", "rolling", "fluttering", "scintillating", and so on. Fading usually is *classed* as *absorption* fading, *skip* fading, *polarization* fading, and *interference* fading.

Quite often in the case of a sky wave signal as many as three of the four classes of fading listed are present to a significant degree simultaneously, though the resultant effect upon the ear usually can be aptly described by one or two adjectives such as those listed in the first category. Thus we may have a slow, rolling fade due primarily to interference effects; a scintillating fade due to a combination of absorption, polarization, and interference effects; and so on.

Some of the basic causes of fading are the same in the case of sky waves, tropospheric waves, guided (very low frequency) waves, and combinations of the above. The following discussion of fading is concerned primarily with sky-wave propagation of high frequency waves, but where noted also applies to other frequencies and modes of propagation.

Absorption Fading

When there is a rapid variation in the absorption coefficient of the transmission medium (or part of it), the signal intensity will fluctuate accordingly. Usually these fluctuations do not occur rapidly enough to be classed as fading, but when they do they constitute probobly the simplest example of the phenomenon of fading.

Skip Fading

At frequencies near the MUF (or at distances near the skip distance), a very slight change in the virtual height or critical frequency of the F_2 layer can throw the receiving point in or out of the skip zone. As the critical frequency and virtual height often do fluctuate slightly, due to turbulence in the ionosphere, this condition can prevail at certain times for certain distances from the transmitter. The resulting severe fluctuation in signal intensity is called *skip fading.*

Polarization Fading

Under the previous discussion of polarization it was explained that a wave refracted from the ionosphere is eliptically polarized, and that the phase difference between and amplitudes of the vertical and horizontal vectors all vary randomly with time. This causes *polarization fading* at the receiver. It is particularly bad when a receiving antenna having linear polarization is employed. However, even with a circularly polarized receiving antenna polarization fading still can be present, because the direction of rotation of the eliptically polarized wave reverses at random under some conditions, while the sense of the circularly polarized receiving antenna is fixed.

Interference Fading

So far the types of fading discussed can be considered a type of amplitude modulation wherein the modulation is accomplished by variation in the attenuation suffered by the wave during its propagation, even though in the case of skip and polarization fading the "attenuation" is an indirect result.

The most common and the most deteriorating (and therefore the most serious) type of fading is caused by a somewhat more complex mechanism. It is called *interference fading* because it is caused by the varying interference patterns resulting from unstable multiple wave paths. The energy at a *distant* point seldom arrives via a single path. In the case of sky-wave transmission, the constantly changing ionosphere irregularities previously discussed, together with random fluctuations in the earth's magnetic field, ordinarily result in a received signal comprising several components, with the relative phases constantly changing due to a fluctuating *difference* in phase velocity, group velocity, or path length. The same applies, except for the magnetic effects, to the troposphere and to tropospheric waves.

When energy arrives at the receiving point via multiple paths, it is called *multipath transmission.* It exists at distances where both ground wave and sky wave, or ground wave and tropospheric wave, have significant intensity. The sky-ground combination readily occurs at frequencies in the vicinity of 500 kc. to 3000 kc., because the skip zone is absent or short, and the ground wave (assuming vertical polarization) is strong. The troposperic-ground combination readily occurs in the v-h-f and u-h-f range at distances near the optical horizon, especially when the distance to the horizon is great.

Multi-path transmission also exists when refraction from more than one ionosphere layer is contributing to the received signal. For instance, a "one hop" F_2 layer signal and a "one hop" E layer signal may exist simultaneously at the same receiving point. Due to the greater density and height of the F_2 layer, together with the greater angle of departure required for the F_2 layer path, there is a narrow range of frequencies over which the F_2 signal perforates the E layer while the E signal does not.

Multi-path transmission also can exist when energy can reach the receiving point by making several "hops" between the earth and the F_2 layer. For instance, under certain conditions of frequency, virtual height, and critical frequency, energy may arrive via three hops and four hops simultaneously. This has been proved experimentally by isolating the different components by means of a receiving antenna which is both highly directional and readily adjustable in the vertical plane.

Another form of multipath transmission occurs in the case of the ordinary and extraordinary rays, the two rays taking different paths to the receiver as a result of the effective difference in refractive index.

Scattering is a perfect example of an extreme case of multi-path transmission.

There are other, less important causes of multi-path ionosphere transmission, but the examples cited will serve to illustrate how highly improbable it is that the transmitted energy will arrive at a distant point via only one path.

Interference fading is more pronounced, or "deeper," when there are just two components, of approximately the same amplitude. If the amplitudes are exactly equal, the resultant intensity can vary from zero to twice the intensity of one component as the phase difference fluctuates at random between 0 and 180 degrees. It is interesting to note that the increases due to reinforcement are not so important as the decreases due to cancellation, because in the case of reinforcement the ratio of the resultant to either component cannot exceed 2, while in the case of cancellation the ratio can be infinite.

When one of the two components is appreciably stronger than the other, cancellation never is so complete and the fading therefore is less severe.

When there are more than two paths, or components, the effects of integration and laws of probability act to minimize the severity of the fading, even though theoretically the fluctuation could be more severe. Fortunately there is an averaging effect when several paths exist, and the more paths there are the less likelihood there is of all the components being exactly in phase or alternately out of phase at the same instant. While fluctuations in the resultant intensity may be more rapid when several paths exist, the excursions are not so great.

It should be noted that only *one* of two or more paths need be disturbed for interference fading to exist. Thus a ground wave, which ordinarily is quite steady, may combine with a disturbed sky wave or tropospheric wave to produce interference fading which is quite severe when the two components are of comparable magnitude.

Selective Fading

To this point the discussion of fading has assumed the radiation of a continuous wave of constant frequency and amplitude. It is apparent that for a slightly different frequency, the phase relationships between the various components of a multi-path signal will be different, and that if the difference in path length is sufficiently great in terms of wavelength, a small change in frequency can change the phase difference by as much as 180 degrees, or from complete cancellation to maximum reinforcement. Thus it is seen that in the case of multi-path transmission the intensity of the received signal is highly sensitive to changes in the frequency of the wave.

This means that in the case of a modulated signal, consisting of sidebands, or carrier and sidebands, the various components of the signal will not all add or cancel in exactly the same manner when multi-path transmission and interference effects exist. The peculiar type of distortion produced when a modulated signal is subjected to interference fading is aptly referred to as *selective fading* because at any particular moment certain frequency components will cancel, while others will add.

In the case of single sideband, suppressed carrier, amplitude modulation transmission, the predominant effect will be to emphasize certain frequencies and suppress others, producing frequency distortion but no harmonic distortion. In the case of the more common method of amplitude modulation wherein carrier and both sidebands are transmitted, both frequency distortion and amplitude (harmonic) distortion will be produced. This is explained by the fact that the carrier can be one of the components which is selectively suppressed or attenuated by interference effects, and unless the percentage modulation of the original signal is very low, there will be times when the carrier wave intensity at the receiving point is not sufficient to "support" the sidebands without overmodulation and consequent amplitude distortion.

Phase dissymmetry of the upper and lower sidebands with respect to the carrier as a result of interference fading also will produce amplitude distortion which contributes to the characteristic aural quality of "selective fading."

GUIDED (LOW FREQUENCY) WAVES

At frequencies below approximately 300 kc. the propagation picture presented for medium and high-frequency ground and sky waves becomes less and less precise. The *effective* conductivity of even the poorest soil becomes quite high, because a much thicker layer of soil is involved, and as the frequency approaches the upper audio range the earth begins to act like a nearly perfect conductor as far as absorption and reflection are concerned. Thus the attenuation of the ground wave due to earth losses is quite small, and attenuation of a multiple-hop sky wave due to earth losses likewise is quite small.

As the frequency is lowered the effective conductivity or coefficient of refraction of the ionosphere is increased to the point where the lower edge of the ionosphere (D or lower E region, depending upon frequency) acts like a reflecting mirror at a height of from 35 to 60 miles, and for this reason the wave is considered as being *reflected* rather than refracted. The absorption is very low, particularly at frequencies below 100 kc., because the wave is reflected before it penetrates very far into the region that ordinarily is responsible for most sky wave absorption.

Another characteristic of low frequency waves is that there is no skip distance for the sky wave, because it always is returned at all angles of incidence, including the normal.

So we have a sky wave that travels great distances with no skip and little absorption, the spreading of the sky wave being confined to one plane after it once fills the space between the two concentric conducting spheres which confine and guide it.

This sky wave or "guided wave" suffers less absorption and therefore is stronger during the night than during the day, during the winter than during the summer, and during the less active portion of the sunspot or terrestrial magnetic cycle. These normal variations, especially the diurnal variation, become less pronounced with frequency, until at frequencies in the neighborhood of 15 to 25 kc. the signal intensity is of the same order of magnitude at all times, even during severe ionospheric and magnetic storms.

There is a pronounced dip in the intensity of the sky wave at or slightly after sunset, and sometimes another, less pronounced, around sunrise. Like other variations, the magnitude of the dip decreases with frequency, until at 20 kc. or so it is insignificant. The reasons for the dip in intensity are not definitely known or fully understood, though various theories have been advanced.

Like high-frequency sky-wave signals, low-frequency sky waves are propagated differently in different directions and at different geographical locations, though in most cases the effects are not so pronounced, and tend to become still less as the lowest practical radio frequency (about 15 kc.) is reached.

At distances close to a low-frequency transmitter the surface wave predominates, because the ionosphere reflected wave must travel so much farther to reach the receiving point. At extreme distances, the sky wave predominates, because the ground wave travels in more continuous and intimate contact with the ground, and while the absorption of the ground wave is low, it nevertheless is greater than that of the sky wave.

At intermediate distances of from 200 to 1000 miles (depending upon the particular frequency and the ground constants over the transmission path), the two components will be of the same order of magnitude. However, instead of producing severe inference-type fading as occurs at high frequencies under the same conditions, changes in signal strength occur gradually. This is explained by the fact that at low frequencies the sky wave path is relatively stable in terms of wavelength. At a distance where the intensity is low because of cancellation effects, the intensity an hour or a few hours later may be maximum because of a gradual shift in the effective height of the reflecting layer due to sunrise or sunset. However, this cannot be classed as fading.

The comparatively stable interference pattern which is produced will resemble a standing wave, and there will be locations where the resultant intensity is only a small fraction of the intensity a few miles closer to or farther from the transmitter. At a different time the phase difference may be such that the intensity at the same location is greater than that existing a few miles in either direction along the signal path.

It should be noted that at the extreme lower range of radio frequencies ordinarily used, or about 15 kc., the height of the reflecting medium is only about four wavelengths. Under these conditions it is difficult to make a distinction between the "ground wave" and the "sky wave," as they begin to lose their identity as such, and combine to produce a resultant mode of propagation which is more aptly described is "guided" propagation. However, interference patterns and effects still exist.

NOISE

With an ideal receiver, one generating no noise of itself, and with a hypothetical receiving antenna which generated no fluctuation or random noise as a result of its radiation resistance, the *required field intensity* for satisfactory communication would be determined only by the background noise picked up by the receiving antenna from cosmic, atmospheric, terrestrial and man-made sources, because if this noise were not present we could receive the signal on our ideal receiver no matter how weak the signal.

Of course our idealized receiver is, unfortunately, a physical impossibility. However, below 50 or 100 Mc. it is possible to build receivers with an inherent noise level low enough that, with even a modest receiving antenna, the noise picked up by the antenna exceeds that generated within the receiver. In such cases, the required signal field intensity for satisfactory communication is determined by the noise field intensity just as would be the case for our ideal receiver.

If the noise field intensity is going to determine the distance over which we can communicate with a given power, or determine the power required to communicate over a certain distance, it is a very important factor and deserving of attention.

The main types of noise may be grouped into the following categories: atmospheric static; cosmic noise; precipitation static; and man-made noise.

STATIC (ATMOSPHERICS)

The random electrical discharges which occur naturally in the earth's atmosphere cause noise which is termed *static*, or sometimes, *atmospherics*. When the discharges are of sufficient intensity, they qualify as "lightning." Because of the countless number involved they also contribute to the noise level when not of sufficient intensity to cause the visible display known as a "lightning flash," provided they are close by.

Nearby thunderstorms will produce large crashes of static, having very high intensity but short duration, coinciding with lightning flashes. The large number of such crashes occurring each second when a large area is involved cause *distant* thunderstroms to produce a constant "*rattle*" of lower amplitude. The numerous nearby random discharges of low intensity also contribute to this steady, crackling background noise which serves as an accompaniment to the intermittent crashes.

While lightning occurs at one time or another at practically every point on the globe, it is much more frequent in certain more or less well-defined tropical areas in which thunderstorms and hurricanes are common. For this reason the "steady" static observed at frequencies which are favorable to long distance transmission tends to arrive from certain more or less fixed directions.

A single pulse of static, while short in terms of audio frequency, is long in terms of radio frequency. For this reason the energy distribution is such that the intensity of a wave produced by an atmospheric discharge is greater at low frequencies, following an inverse law somewhere between the first and 3/2 power. Thus, the intensity of the noise field from a local thunderstorm varies (approximately) inversely with frequency, or directly with wavelength, becoming quite low at ultra high frequencies and quite high at very low frequencies. At frequencies in the vicinity of 15 to 25 kc., a wavelength is quite long, and during a local thunderstorm many lightning discharges may occur within a radius of a small fraction of a wavelength. Under these conditions the intensity of the disturbance at the receiver (as compared to higher frequencies) will be much greater than indicated by the inverse first or 3/2 power law, because of inductive effects. The first or 3/2 power relationship holds only for disturbances occuring at a distance great enough that the radiation field predominates.

The distribution of intensity vs. frequency in the case of *distant* static discharges is altered by the effect of frequency upon propagation characteristics. Thus, below 300 kc. or so the intensity of static caused by distant disturbances increases much more rapidly than the inverse first or 3/2 power, because the lower the frequency the less the absorption and the more efficient the propagation. The more rapid decrease in intensity with increased frequency is most noticeable during the daytime, because absorption of the sky wave increases with frequency much more rapidly in the daytime.

In the range 500 kc. to 2000 or 3000 kc. which includes the AM broadcast band, static from distant sources will have very low intensity during the daytime because of the high degree to which the skywave is absorbed. At night, when the sky wave is propagated efficiently, the effect of distant static will be appreciable.

In the high frequency range the intensity of static arriving from distant disturbances will depend primarily upon the frequency and the conditions of the ionosphere along the transmission path.

At frequencies above a certain frequency which will be determined by the conditions of the ionosphere at the moment, no sky wave transmission from the distant static source will be possible, and therefore static in the u-h-f and upper v-h-f range always is of local origin.

Most of the heavy static originating in the equitorial zone and propagated into temperate and polar latitudes by the same mechanism as regular radio waves originates in regions in or near equitorial Africa, the East Indies, and Central America. This often results in lack of reciprocity in the case of communication over a path having appreciable north-south component when identical directional antennas are employed, because the receiving antenna of the station nearest the equator will point away from these noise centers, while the other station may be beamed right at or only slightly off one of these principal noise centers.

The intensity, location, and time of maximum incidence (both diurnal and seasonal) of tropical static are closely correlated with weather, and therefore can be predicted in terms of generalities, just as it is possible to predict that one location will receive more or less rainfall than another and that it is more likely to fall during certain periods of the day and at certain times of the year.

COSMIC NOISE

At frequencies too high to be refracted or reflected by the ionosphere, we are blessed with an absense of static arriving from distant points by means of sky-wave propagation. However, the transparency of the ionosphere to waves having a frequency above the F_2 critical frequency permits a type of noise that is absent on lower frequencies: solar noise and cosmic noise arriving from outer space. This noise, which is of the nature of a steady hiss, is of maximum intensity near the upper limit of the h-f band or in the lower part of the v-h-f band and decreases approximately as the third power of the wavelength. Because of the rapid decrease in intensity as the frequency is raised, cosmic noise is not a significant factor in the u-h-f range.

In view of the fact that cosmic noise emanates from certain fixed locations in the universe, the intensity will vary appreciably from hour to hour, day to day, and season to season, the variation being more pronounced at the receiver as the directivity of the receiving antenna is increased. Likewise at any given time the intensity at the receiver input circuit will vary considerably as the orientation of a directional antenna is changed.

AIRCRAFT CORONA STATIC

An aircraft in flight is subject under some conditions to corona discharge which, unless controlled, produces a severe form of continuous static that is especially bad on the lower frequencies. The corona discharge is produced in two ways.

In the first, which is the more prevalent and more serious, the aircraft becomes electrified by frictional generation of electricity or by impact with charged particles when it flies through a precipitation in the atmosphere. Dry snow or ice crystals cause an especially bad corona condition, but rain, dust, and smoke also will cause an aircraft to acquire a charge sufficient to produce corona. The charging is called "autogenous" and "unipolar" because the charge is generated by the moving plane (self-generated) and the entire plane carries a charge of the same sign (unipolar). The static resulting from this type of discharge is called *precipitation static.*

The second type of corona discharge is called "exogenous" and "bipolar" because the free charge already exists in the medium in a form such as a thundercloud (externally generated) and the charge is transferred both to and away from the aircraft so that the charge on one part of the plane has one sign and that on another portion of the plane is of the opposite sign (bipolar). In effect the aircraft acts as a "short-circuit" for adjacent thunder clouds which are highly charged and of opposite polarity. The static produced resembles precipitation static (though it is more variable and usually of shorter duration) and the remedy is the same. Therefore, it oftentimes is considered under the general heading of "precipitation static," though strictly speaking it deserves a separate classification.

There is a particular voltage at which each of the sharp projections or edges of the main protuberances of an aircraft will ionize the contiguous air and produce corona. The charge acquired by the aircraft thus leaks off in a series of impulsive discharges; the more rapidly the plane acquires charge, the more rapid the impulses and the greater the number of discharge points, thus tending to keep the aircraft in equlibrium.

Aircraft corona static is most severe when there is corona discharge from any portion of the receiving antenna system proper. The next most serious condition is where the corona exists adjacent to any part of the antenna system. If the corona occurs some distance from the antenna, the interference is only moderate, and can be effectively suppressed by controlling the character of the corona discharge.

To prevent corona discharge from the antenna proper, the wire is covered with a low loss dielectric of high dielectric strength and suitable physical properties. Polyethylene is one such dielectric. Loop antennas often are housed in streamlined metallic or dielectric shields. All hardware or appurtenances touching or adjacent to the antenna wire are sufficently rounded off or else taped or otherwise covered with dielectric to eliminate any possibility of corona in the immediate area of the antenna. Then, to minimize the remaining intereference, special "quiet" dischargers are placed at various extremities of the plane such as the wing tips.

One such discharger which is very effective comprises a fabric wick which has been impregnated with colloidal metal. Each of the large number of fibers in the wick offers an effective "pin point" discharge point and the composite result is a smooth, quiet discharge at moderately low potential, preventing the aircraft from acquiring a charge high enough to produce corona at undesired spots.

Another form of discharger, which incidentally was more or less standard for many years, employs a short length of very fine wire which is trailed behind and isolated from the aircraft by a series resistor of sufficiently high value to "soften" the discharge and confine the generated noise to the discharge wire.

A comparatively mild form of precipitation static sometimes is observed with certain types of antennas at ground stations during a severe snow, dust, or thunder storm, particularly when a high wind is blowing. If the antenna is grounded, however, any static due to corona discharge from the antenna is more likely to be of the exogenous type, because the autogenous charge is conducted to ground as rapidly as it is generated.

MAN-MADE NOISES

Electrical apparatus which utilizes commutation (such as a d-c motor or generator), vibrating contact points (such as an electrical shaver), spark discharge (such as an ignition spark plug) or any such mechanism whereby an electrical spark is produced (however minute) will radiate damped radio frequency waves which cover a wide spectrum *unless preventative measures are taken.* Under certain conditions the noise field intensity from such sources can be high enough to prevent satisfactory reception of anything but extremely strong signals.

When practicable, the usual procedure is to confine the r-f currents developed by the spark discharge to that particular circuit, or at least to the apparatus, by shielding and by use of suitable isolating resistors or chokes in conjunction with bypass capacitors. When this cannot be done or when it does not offer complete freedom from interference, the usual procedure is to confine all pickup to the antenna proper and locate and orient it so that the noise picked up from the offending source is minimized. Pickup is confined to the antenna proper by shielding or balancing the lead-in or feed line from antenna to receiver, and by filtering the supply leads to the receiver to prevent noise currents being conducted into the receiver via the power wires. This will be treated in more detail in the chapter devoted to receiving antennas.

REDUCTION OF NOISE—GENERAL

Because the amount of noise energy appearing in the output of a receiving system increases with the band width, it is desirable in order to minimize the intensity of all types of noise to limit the band width of the receiver to the minimum which will pass the modulation with satisfactory fidelity.

Loud crashes of atmospheric static of local origin, as well as ignition interference or any interference which consists of a not too rapid succession of peaks of high amplitude but *short duration* can be greatly reduced with little or negligible distortion of the signal by means of *peak clipping* or *chopping*. This type of "limiting" is inherent in frequency modulation systems, and is highly effective except when the desired signal is quite weak. In amplitude modulation systems, the limiting must be accomplished by special auxiliary circuits, and the limiting generally is not as effective as that occurring in FM reception.

Peak chopping is effective as a means of noise reduction only on impulse type noise, in which the ratio of peak to average energy is high; it is not effective on the steady hiss type of noise termed "fluctuation" or "random" noise.

When satisfactory reception is not obtained in spite of full exploitation of directional antennas well separated from local noise sources, noise suppression on nearby electrical devices, narrowest practicable receiver bandwidth, power line filtering, and peak noise clipping or limiting, the only recourse is to overcome the residual noise by increasing the radiated power of the transmitter. This assumes, of course, that the receiver proper is sufficently well designed that the residual noise is predominantly noise which is fed into the receiver from the antenna system, and not inherent receiver background noise.

COMMUNICATION PRACTICE IN DIFFERENT FREQUENCY RANGES

Because of the different propagation characteristics of different frequency ranges, each range has its own scope of usefulness. Likewise, different techniques and practices have developed as a result of each frequency range possessing characteristics unique to that part of the spectrum, not only with regard to propagation, but to antenna equipment and method of installation as well. While this chapter is concerned primarily with the propagation aspects, they are in many ways so closely related to the location, orientation, and design of the antenna equipment that it is not practicable to ignore one while discussing the other. However, until antennas themselves are treated in specific detail, they will be referred to in connection with propagation considerations only in generalities.

BASIC ANTENNA CHARACTERISTICS AND CONSIDERATIONS

In order for the reader to appreciate the following discourse on communication practice with regard to the propagation aspects, it will be necessary for him to accept or recognize for the time being certain basic facts in connection with antenna systems, which are postulated as follows:

(1) An antenna must have sufficient physical size in terms of wavelength to possess appreciable radiation resistance; otherwise it will be an inefficient radiator and wasteful of power.

(2) All practicable antennas exhibit some directivity. They transmit or receive better in certain directions than others. This effect may take place in the vertical plane, the

horizontal plane, or both, depending upon the configuration and orientation of the antenna system. The vertical directivity is affected by the presence of the ground.

(3) The simplest practicable radiator is a linear conductor which, in order to minimize resistance losses and reactance effects, usually is made a resonant length at frequencies where this is economically feasible. Fundamental resonance of an ungrounded wire occurs at one half wavelength; of a grounded wire, one quarter wavelength. Such a radiator shows no *inherent* directivity in a plane normal to the radiator, but possesses an inherent directivity pattern resembling a figure 8 in all planes containing the radiator. The orientation of the figure 8 is such that zero radiation occurs off the ends of the radiator. The inherent directivity of this simple radiator, like any other, is modified by the presence of the ground because of reflections and interference effects.

(4) By the use of a multiplicity of radiating elements so disposed as to provide systematic exploitation of interference effects, it is possible to concentrate most of the radiation from an antenna system into a restricted angle in either the vertical plane, horizontal plane, or both. In general, the greater the cross sectional dimension of the radiating system in either plane, the greater the directivity (narrower the angle of the "beam") in that plane. The greater the *depth* of the array, or cross section of the antenna system in the direction of propagation, the greater the directivity in *both* planes. A directional antenna system is called a *directional array,* or sometimes a *beam antenna.*

(5) Assuming negligible resistance losses in any case, concentrating the radiation into a restricted solid angle which includes the desired transmission path produces the same result as raising the transmitter power. The smaller the solid angle into which the radiation is concentrated, the more intense the radiation for a given transmitter power. The increase in field strength realized due to the directive properties of an array usually is defined in terms of *equivalent power gain* or decibels improvement over a *reference antenna* which is fed the same amount of power. Unless otherwise stated the reference antenna ordinarily is a half wavelength linear radiator having the same average elevation, and oriented to have the same polarization, as the array.

(6) Generally speaking, the principle of reciprocity applies to transmission and reception with a given antenna system. The directive properties (directional pattern) in nearly all practical cases will be the same whether an array is used for transmission or reception. And generally speaking, a good transmitting array will make a good receiving array.

(7) Summarizing on directional arrays: They are commonly used to minimize interference and provide an equivalent power increase. They are used on transmitters to minimize radiation in undesired directions, thus reducing interference, and to provide an equivalent power increase by concentrating it in the desired direction(s). They are used on receivers in order to minimize reception of interfering signals from undesired directions, and to provide an equivalent increase in transmitter power or receiver sensitivity. The most efficient and protected point-to-point circuit utilizes directive arrays for both transmission and reception. In many applications the same antenna is used for both transmission and reception by means of relays.

(8) It is common practice, especially on the higher frequencies, to locate the antenna proper some distance away from the transmitter or receiver, and to employ a *feed line* or *transmission line* to convey the energy from transmitter to antenna or from antenna to receiver. The transmission line ordinarily is *balanced* or *shielded* so as to confine virtually all radiation and pickup to the antenna proper. The antenna and transmission line together commonly are referred to as an antenna *system.*

(9) When propagation is primarily by the surface wave component of the ground wave, as in the AM broadcast band, the chief consideration is the conductivity of the ground. Except as it affects radiation resistance and efficiency, the height of the antenna is not of great importance. When propagation is via the space wave component of the ground wave, as at v.h.f. and

u.h.f., the primary consideration is elevation of the antenna, the conductivity of the ground being of minor importance. When propagation is via the sky wave, the chief consideration is the condition of the ionosphere, the elevation of the antenna above the ground being relatively unimportant except as it affects the vertical directivity pattern of the antenna.

V-L-F AND L-F PRACTICE

The lower end of the radio frequency spectrum is employed for long distance circuits (several thousand miles) or general coverage over a wide area when the maximum possible reliability is absolutely necessary. While one octave in the high frequency range, say, from 6 to 12 megacycles, contains several hundred channels, the octave between 15 and 30 kc. contains comparatively few. Therefore the full potentialties of channels in this range are exploited by use of very high power transmitters and elaborate and extensive antenna installations. It is the only portion of the radio frequency spectrum where a signal can be put into almost all portions of the world simultaneously, regardless of magnetic or ionosphere conditions and under nearly all atmospheric conditions, by sheer "brute force."

At the lower end of the spectrum, the radiation resistance of any practical antenna is very low, even when the antenna system covers many acres or even several square miles and utilizes towers several hundred feet high, because a wavelength is so long at such frequencies. If the inductors which are used to tune out the reactance presented by the nonresonant radiator are of low loss construction in order to minimize losses, then the low radiation resistance results in a very high Q, which, because of the low frequency, limits the bandwidth of the signal which can be radiated efficiently to a matter of cycles rather than kilocycles. This limits the use of these frequencies to comparatively slow speed telegraphy.

The l-f band, from 30 kc. to 300 kc., is used for various applications requiring reliable ground wave coverage over distances up to several hundred miles, particularly in high latitudes where auroral effects make high frequency sky wave propagation unreliable and the distance from the equatorial "static factories" results in a low static level at low frequencies.

When "long wave" stations are used for trans-oceanic circuits, they are located as close to the ocean as practicable, because it has been found that the wave suffers appreciable attenuation if the station is set back from the coast even a few miles. When possible the antenna is located over a marshy country or soil where salt water exists a few feet below the surface, in order to obtain a low resistance path for the high amplitude earth currents which flow in the vicinity of such an antenna.

MEDIUM FREQUENCIES

The spectrum between 300 kc. and the lower edge of the AM broadcast band is employed for certain marine, aircraft, military, and other applications requiring reliable ground wave coverage over a maximum distance of from 100 to 300 miles over land and up to 500 miles or so over salt water.

The most important portion of the medium frequency band from a commercial standpoint is the AM broadcast band from 550 to approximately 1600 kilocycles. The rest of the m-f spectrum, from the upper edge of the AM broadcast band to 3 Mc., is mostly used for applications which could be better served by use of lower frequencies but for which such channels cannot be made available. Coverage by means of ground wave propagation at 3 Mc. is relatively inefficient over land, and sky-wave communication ordinarily is better accomplished on frequencies somewhat above 3 Mc. However, the range 2-3 Mc. is as satisfactory as lower frequencies for ground-wave transmission from or between small boats, where the dimensions of the antenna are strictly limited and the earth conductivity is high.

THE AM BROADCAST BAND

In the AM broadcast band an antenna system is called upon to deliver a field strength adequate for good program quality

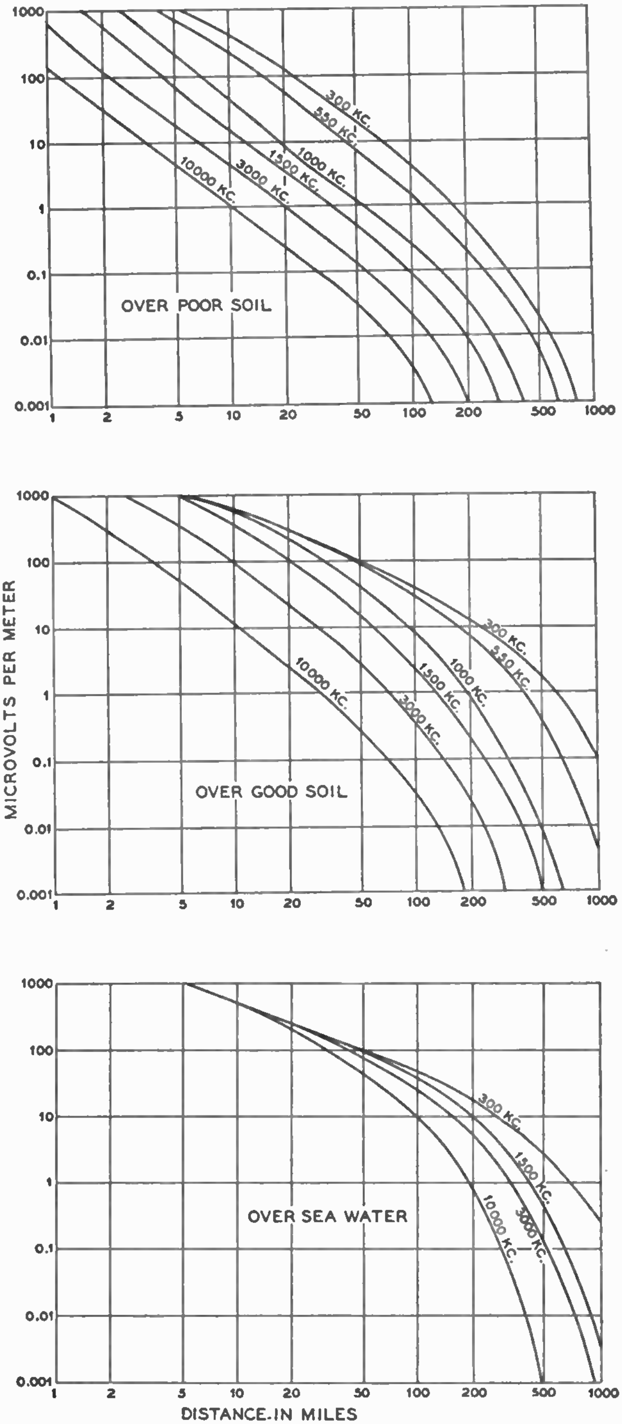


Figure 1-13.

APPROXIMATE GROUND WAVE FIELD INTENSITY AS A FUNCTION OF DISTANCE FOR A SHORT VERTICAL ANTENNA RADIATING ONE WATT OF POWER

It is assumed that the radiator is located at the surface of the earth, is not longer than one quarter wavelength, and that the terrain is fairly level (no high mountains) along the propagation path. The field strength in microvolts per meter for other values of radiated power will be proportional to the square root of the power in watts.

The radiated power for an antenna longer than 1/8 wavelength will approach the value of power fed to the antenna, provided very good coupling and grounding systems are employed. A short "whip" antenna which is but a very small fraction of a wavelength long will radiate as little as 1/100th of the power delivered by the transmitter.

A vertical radiator acquires additional vertical directivity as it is lengthened. Thus, the field strength from a half-wave vertical radiator will be about 20 per cent greater than the values shown in the charts (for a given total radiated power) because more of the energy is concentrated along the ground.

over a specified or desired area, sometimes a slightly degraded but still usable signal within a larger area, and as little signal as possible beyond the point where the signal is no longer enjoyable (in order to minimize interference to other station(s) assigned the same frequency). Furthermore, the design of the antenna system should be such that this is accomplished as economically as possible when the cost of the entire station (including the antenna system) is considered. In fact, when the F.C.C. specifies a certain power when licensing a station, it is based on the power required to cover the desired area on the specified frequency when an economical proportion of the total station cost is represented by the antenna system, assuming that the antenna system is of good engineering design regardless of cost and represents the best performance that can be obtained for the particular amount of money expended on the antenna system.

PRIMARY AND SECONDARY COVERAGE

The requirements for broadcast reception (either AM or FM) differ from other services in that it is not sufficient just to deliver a highly intelligible signal to the receiver; in commercial broadcasting the received signal (within the "primary" service area) must be not just intelligible but virtually distortionless and free from interference and noise even when received on an inexpensive receiver, using a mediocre antenna, and located in an area of high ambient noise level due to man-made electrical interference.

The standards of good engineering practice and other requirements established by the F.C.C. are such that, employing or applying known techniques, the only way to comply is to use a vertical radiator or radiators and a very excellent ground system. To ensure 100 percent coverage of the primary service area at all times the station is operating except possibly during local thunderstorms, it is necessary to rely upon ground wave (surface wave) propagation.

Use of a vertical radiator or radiators provides the strongest ground wave for a given power, and also confines the major sky-wave radiation to low vertical angles so that the amplitude of the received nighttime sky wave does not approach that of the received ground wave except at points beyond or near the outer edge of the daytime primary service area. If the amplitude of the sky wave exceeds a small fraction of the ground wave intensity, wave intereference effects and the objectionable "selective fading" type of distortion occur, degrading the quality of the signal. If the two waves should be of approximately the same magnitude at some distance, the distortion will be so severe that at times dialogue will become completely unintelligible.

The field intensity required for satisfactory primary coverage will depend upon the ambient electrical noise level. Ignoring local thunderstroms, the required signal field intensity will vary from 50 millivolts per meter in particularly noisy business or industrial districts in large cities, to about 0.1 millivolt per meter in the more quiet rural areas.

Some idea of the primary coverage to be obtained for a given frequency, power, and terrain may be had from figure 1-13, which illustrates the variation in field strength of the ground wave with distance, frequency, and terrain. The attenuation of a ground wave is much greater when passing over a skyscraper district or over dense tropical jungle than is shown for poor soil.

"Secondary coverage" defines an area in which the field strength is sufficient to provide a signal which, except under unfavorable conditions of reception, will satisfy the less critical listener most of the time and a critical listener part of the time.

In the daytime the secondary coverage area due to the ground wave is greater than at night, because the absence of sky-wave propagation in the daytime reduces atmospheric noise and heterodyne or "shared channel" interference from distant stations. If the station is powerful, there may be, in effect, two nighttime secondary coverage zones. In the outer zone the sky wave (which is completely absorbed in the daytime)* predominates and is strong enough to provide a usable signal. In the inner zone the ground wave will be a little too weak to qualify as "primary coverage," yet still be enough stronger than the sky wave that distortion is absent or tolerable. In the intermediate zone both signals will be too weak or else they will be of the same order of magnitude, causing bad selective fading and severe distortion.

At the lower end of the broadcast band, propagation of the surface wave is good and transmitter power always is substantial. There seldom will be an intermediate zone in which both the ground wave and nighttime sky wave are too weak. The ground wave is propagated so efficiently that it is strong out to and usually past the point where the sky wave reaches appreciable amplitude. In this case the primary and secondary coverage areas are separated only by an area of excessive distortion. Under some conditions this nighttime distortion area will start within the day time primary coverage area, causing the primary coverage to be less at night.

When low or moderate power is used at the higher end of the broadcast band, the ground wave intensity will drop below usable values before distances are reached where the sky wave might provide a sufficiently strong signal to qualify as secondary coverage. This is especially true when the gyro frequency is above 1200 kc. In these cases there will be no zone of severe distortion from ground vs. sky-wave interference, but rather a zone of inadequate signal strength corresponding to a "skip zone." This "skip zone" in the secondary coverage exists only when the sky wave intensity is strong enough to provide secondary coverage. The *primary* coverage in these instances of no distortion zone will be approximately the same day or night for a given transmitter power, because it is assumed that in the primary service area the signal field intensity is strong enough to override moderately severe man-made noise, which is so much stronger than nighttime atmospheric static and sky wave heterodyne interference that the presence or absence of the latter would make little difference in noise level.

*The sky wave usually is completely absorbed in so far as *practical effects* are concerned. Actually, while the sky wave ordinarily is so highly attentuated in the daytime that it is no longer of significance, it nevertheless exists, and may at times be strong enough to cause interference to distant stations on the same channel.

SKY-WAVE INTENSITY

When considering the distance at which the ground wave and nighttime sky wave are of the same order of magnitude, it is well to examine the reasons why the sky wave is comparatively weak at distances close to the transmitter. In the first place a conventional vertical radiator has a vertical radiation pattern in which the intensity is approximately proportional to the cosine of the angle of elevation. This means that the radiation at high angles is of comparatively low magnitude. Also, the sky wave travels considerably farther than the surface wave to reach a nearby point on the earth's surface, the path difference being at least 140 miles or so for zero geographical distance (assuming E layer reflection). And of course there is appreciable absorption of the sky wave while it is in the ionosphere.

These factors combine to reduce the amplitude of the sky wave to a value well below that of the ground wave at distances out to at least 30 miles over average earth at the high end of the broadcast band, and to roughly 100 miles at the lower end where the ground wave attenuation is considerably less.

DIRECTIONAL SYSTEMS

Broadcast transmitters commonly employ two or more vertical radiators as a directional array, especially when the channel is shared with another station. In the latter case the usual procedure is to locate the transmitter on the side of the area to be served toward the sharing station. The array is then designed and adjusted so as to have a rather sharp "null" in the direction of the distant sharing station. In some cases an array is employed not only to "protect" another station but to give a radiation pattern which more economically covers the desired area than would an antenna radiating equally well in all directions. A typical example is where there is a populated area spread along the seacoast with a sparsely settled interior. Obviously any energy radiated toward the sea or inland is wasted (from a commercial standpoint) and it is better to concentrate it in a direction along the seacoast and thusly provide a better service in the thickly populated area.

AM BROADCAST STATION SITES

Broadcasting transmitters usually are sited only after careful propagation surveys are made to determine the relative merits of different potential locations. Ordinarily the station will be placed in the suburbs or outskirts in a region having high soil conductivity, but in the larger cities there are a few, low power "metropolitan" stations with transmitters located right in the downtown business districts and intended to serve only the immediate area. Even though an efficient antenna system is impossible under such conditions, such a station often will override the very high electrical noise level existing in the business section of a large city even with comparatively low transmitter power, just because of the mere proximity of the transmitter. In some such cases the radius of primary coverage will not exceed a couple of miles, but a large population may be included within the area.

High powered "regional" stations with main studios in large coastal cities such as New York, Boston, Los Angeles, San Francisco, etc. often have their transmitter located as much as 25 miles out of town, usually in a salt marsh or else quite close to the beach, if not actually on a small natural or artificial island.

AM broadcasting techniques are highly developed and specialized. It is recommended that those who have a special interest in AM antenna systems, propagation, practices, etc. obtain the various publications and charts (free or nominal cost) issued by the Federal Communications Commission and the National Bureau of Standards, some of which are available direct and some only from the Government Printing Office (Superintendent of Documents). Because new publications are constantly being issued and old ones superseded, no attempt will be made to list them here.

HIGH FREQUENCIES

Except for a few isolated applications (particularly marine and military) at the low-frequency end of the range, the high-frequency band embracing 3 to 30 Mc. is used principally for medium to long distance sky-wave communication, and long distance (international) broadcasting.

When the wave does not have to travel close to the auroral zone, when the ionosphere is comparatively "undisturbed," when effective antenna systems are used, and when there are no local thundershowers, voice communication over distances of several thousand miles is possible with only a few watts of radiated power *provided the most favorable frequency is used.* C-w telegraphy can be effected with even less power, or about 1-50th the power required for voice communication of comparable intelligibility.

MUF AND OWF

The most favorable high frequency from the standpoint of signal strength is just below the maximum frequency which is returned by the ionosphere to the desired point at that particular time. It is called the *maximum usable frequency,* abbreviated *MUF,* and previously described in conjunction with skip distance. However, because the MUF is constantly changing, it would not be practicable to employ the MUF as the operating frequency. Skip fading would be very severe, and the frequency would have to be changed constantly.

To avoid this difficulty, a frequency about 15 per cent lower than the MUF, called the *optimum working frequency,* or *OWF,* ordinarily is employed rather than the MUF. The signal will be almost as strong as when the MUF is used, particularly at night, and fewer frequency changes will be necessitated as the ionosphere conditions change.

While communication is possible with very low power under favorable conditions, commercial and government circuits ordinarily employ powers of from 1 to 25 kilowatts into the antenna, or sometimes much more in the case of international broadcasting stations, and then utilize elaborate directional arrays to increase the effective radiated power still further. The use of high power in conjunction with rather elaborate antennas permits communication under less favorable conditions, and improves the reliability of the circuit. At long distances, interference fading due to multi-path transmission occurs, and under certain conditions the signal can fade to zero. The probability of this happening during any one transmission is very small. However, constant excursions of 20 db or more in signal strength are quite common. The use of plenty of transmitter power ensures adequate signal strength at the bottom of a fade.

Short wave "international" broadcast stations do not provide a service which compares with the *primary* coverage of a standard AM broadcast station as regards quality of signal. The field intensity at times will be very high, but interference fading will degrade the fidelity or quality of the signal even though the receiver a.v.c. maintains the average signal level nearly constant. However, the distortion ordinarily will not be as severe as is experienced in the case of selective fading in the standard AM broadcast band, because at the higher frequencies the carrier and various sidebands tend to fade more nearly together and overmodulation of the carrier is not likely to be as severe from this cause.

The drop in field intensity as the frequency is lowered from the MUF is more rapid in the daytime, when absorption at the lower frequencies is high. Therefore, in the daytime it is imperative that the operating frequency not be too far below the MUF. The additional absorption induced by using a frequency 3 or 4 megacycles lower than the OWF may be as high as 10 or 20 db during the daytime. The chief advantage in using the OWF at night, rather than a much lower frequency, is a matter of noise level rather than absorption. The signal field intensity may be nearly as high at the distant point at night when the lower frequency is used, but the noise level will be much higher. Thus it is desirable to use a frequency as near the OWF as possible, day and night, though it is most important in the daytime. Generally speaking, the OWF will be considerably lower at night than during the daytime for a given path.

LOWEST USEFUL HIGH FREQUENCY

For a given time and path, there will be a lower limiting frequency, determined by transmitter power and atmospheric noise, below which satisfactory communication is impossible because of excessive absorption. This frequency is called the *lowest useful high frequency*, or *LUHF*. Under some conditions, particularly in the daytime when low power is used, the LUHF may be only slightly lower than the OWF. Note that, unlike the MUF and OWF, the LUHF is not a fixed value for a given path and time; it depends upon transmitter power and other factors such as transmitter and receiver antenna gain, receiver sensitivity, and required circuit performance.

DETERMINATION OF MUF

When the critical frequencies of the various layers are known (even approximately) at certain "reflection points" or "control points" along the transmission path, it is possible to make an accurate determination of the MUF (and therefore of the OWF, as the latter is a constant percentage of the MUF). This is made possible by the fact that there is a definite relationship between the critical frequency of a given layer and the "lowest perforation frequency" for a particular oblique angle of incidence.

The MUF for a given layer is roughly proportional to the secant of the angle of incidence. Because of the curvature of the layers, the angle of incidence never can be 90 degrees, even when the wave leaves the earth at zero elevation (tangent to the earth). Because of its substantially greater height, the F_2 layer is more oblique to a wave just grazing the earth than is the E layer. This means that for the same critical frequency, the maximum MUF of the E layer is somewhat higher than the maximum MUF of the F_2 layer.

The maximum possible "one hop" distance for E layer reflection is considerably less than the maximum possible "one hop" distance for F_2 layer reflection, because of the much lower height of the E layer. For usual heights of the E and F_2 layers, the maximum one hop distances will be about 1250 miles for the E layer and about 2500 miles for the F_2 layer. This is for a wave elevation angle of about 3 degrees, as radiation below this angle suffers excessive absorption by the earth. The F_1 layer has been omitted from the discussion because seldom is it the effective or "controlling" ionosphere layer in sky wave transmission. The majority of the time the predominant wave arrives via either the E layer or the F_2 layer.

It is obvious that for a given distance there will be an MUF for E layer reflection and an MUF for F_2 layer reflection. Whichever layer has the higher MUF is the *controlling layer*, and the unqualified term "MUF" applies to the controlling layer. Also, in the case of the F_2 layer, MUF applies to the *extraordinary wave* unless stated otherwise.

The relationship between extraordinary wave critical frequency and the MUF for the E and F_2 layers for one hop transmission over various distances is shown in table III. The MUF factors given for the F_2 layer are the lowest that can be expected at any time (for instance, summer midnight). Actually the F_2 factors vary with the condition of the earth's magnetic field, time of day, and season, the variation being greater at the greater distances. For instance, the MUF factor for the lower angles of departure (larger angles of incidence) will be slightly higher during the day than at night, and slightly higher during the winter than during the summer. The MUF factors for the E layer are fairly constant at all times, diurnal and seasonal variations being insignificant. The reference factor 1 in the case of the MUF factors applies to the critical frequency (extraordinary ray in the case of the F_2 layer) at a point halfway along the transmission path for single hop propagation.

The *distance factor* is proportional to the reciprocal of the MUF factor, and has an arbitrary reference value of 1 for a 2500 mile F_2 hop or a 1250 mile E hop. Thus, if the MUF for a distance of 1250 miles or more in the case of the E layer is known, or if the MUF for a distance of 2500 miles or more in the case of the F_2 layer is known, it is possible to determine the MUF for shorter distances simply by multiplying the maximum distance MUF by the appropriate distance factor. Note that distances of 1250 miles *or longer*, and 2500 miles *or longer* were specified. This requires some explanation.

MULTIPLE OR EXTENDED HOP TRANSMISSION

When the distance is too great for a wave to travel in one simple "hop" the wave *acts* as though it were refracted from two *control points*, one at each end of the circuit and separated from it by a distance equal to one half the maximum one hop distance. This means that the "control points" in the case of F_2 layer transmission exceeding 2500 miles may be considered as being located 1250 miles from each end of the circuit. Actually, this is a specious explanation which does not account for propagation between the two control points, but the important thing is that, so far as MUF calculations are concerned, predictions based on this simplified mechanism give results which agree closely with experiment. Thus, the MUF for a multiple or extended hop circuit may be taken as the MUF of whichever control point has the *lower* MUF.

It doesn't seem quite fair to say that, "The MUF for all distances over 2500 miles is the same," and then change the subject. But after all, we are really interested in the results, and suffice it to say, there is considerable disagreement and speculation as to exactly what happens to account for the accurate results obtained from the supposition of two "control points." At one time it was assumed there was a secondary skip zone, just beyond the maximum one hop distance, for frequencies slightly too high to be refracted in two geometrical hops. This seems logical, but in practice the secondary skip zone just beyond 2500 miles does not ordinarily exist for propagation via the regular layers.

CRPL PROPAGATION PREDICTIONS

It is recommended that the reader who has use for or further interest in the determination and prediction of ionosphere conditions and their application to high frequency propagation problems take advantage of the services offered by the Central Radio Propagation Laboratory of the

U. S. National Bureau of Standards. Basic Radio Propagation Predictions for a particular month are issued three months in advance through the Superintendent of Documents, U. S. Government Printing Office, Washington, D. C. These are available either singly or by subscription at a nominal charge. Because of possible changes in the cost, and because of the possibility of additional data or services being provided at a date subsequent to the publication of this book, it is suggested that inquiry be made of the Superintendent of Documents as to cost, and to the availability of other publications bearing upon the subject.

The Basic Radio Propagation Predictions

TABLE III

MUF AND DISTANCE FACTORS (APPROXIMATE) FOR PROPAGATION VIA THE E AND F_2 LAYERS

Distance (statute miles)	MUF FACTOR		DISTANCE FACTOR	
	E LAYER	F_2 LAYER	E LAYER	F_2 LAYER
0	1.00	1.00	0.21	0.35
100	1.13	1.01	0.24	0.35
200	1.5	1.03	0.31	0.36
300	1.9	1.07	0.40	0.37
400	2.4	1.13	0.50	0.39
600	3.3	1.3	0.69	0.45
800	4.0	1.5	0.83	0.54
1000	4.5	1.8	0.94	0.64
1250	4.8	2.1	1.00	0.74
1500		2.4		0.83
2000		2.7		0.95
2500		2.9		1.00

To determine the MUF when the critical frequencies of the E layer and F_2 layer (extraordinary ray) at the mid point of the transmission path are known, multiply the critical frequency or "zero distance MUF" by the MUF factor given for the distance in question. Do this for both the E layer and the F_2 layer, and take the higher frequency. This gives the MUF for that distance.

When the MUF is known for 1250 mile (or greater) E layer transmission or for 2500 mile (or greater) F_2 layer transmission, the MUF for the distance in question can be obtained by multiplying by the distance factor.

comprise various simplified nomographic charts and other data which permit rapid and accurate determination of the MUF and other propagation information for any high frequency sky-wave path in the world at any time of day. Predictions are based on *average, "undisturbed"* conditions for the month of prediction, but are supplemented by forecasts of ionosphere disturbances.

H-F ANTENNAS AND SITES

Antennas employed for sky-wave communication within the range 3 to 30 Mc. vary widely with frequency and application. However, they are almost invariably horizontally polarized, because ground absorption losses are minimized and because man-made static is lower when horizontal polarization is employed. In the case of general coverage antennas it is necessary to exercise care in ascribing to the antenna a specific polarization, because a simple horizontal antenna, for instance, radiates horizontally polarized sky waves in a broadside direction and vertically polarized sky waves off the ends.

The vertical directivity or space wave pattern of an antenna or array is greatly affected by the presence of the earth. For a given mean elevation of the antenna or array, there will be certain vertical angles where cancellation of the direct and ground-reflected rays causes the resultant wave to be comparatively weak, and other angles where reinforcement will increase the strength to approximately twice the "free space" value. To avoid excessive cancellation at the low vertical angles of radiation generally favorable to extreme long distance work, a horizontally polarized antenna or array must be elevated at least one and preferably two wavelengths above the effective height of the surrounding terrain, even when the conductivity is high as in the case of the ocean.

For short and medium distance general-coverage work at the lower end of the high frequency range (3 to 6 or 8 Mc.) a comparatively high angle of radiation is favorable. In this case the antenna usually is elevated just enough to clear the earth and surrounding objects sufficiently to keep losses due to induced currents and low radiation resistance from being excessive.

Unless the area surrounding the "reflection point" would appear comparatively smooth in terms of wavelength as viewed from the center of the radiating system, scattering and complex reflections will result in an interference pattern of considerable complexity and make it virtually impossible to ascribe an effective height to the antenna or to predict just what can be expected.

The lateral extent of the effective "zone plate," or area of ground appreciably affecting the vertical pattern of the antenna, is determined primarily by the elevation of the radiator and the vertical angle of radiation involved. To get a rough idea of whether a certain area of terrain is appreciably affecting the radiation at a certain vertical angle in a certain direction, take an imaginary sight from the effective "center" of the radiating system in that compass direction at an angle of *depression* equal to the angle of elevation involved. This will give the geometrical "reflection point." The height of the antenna with regard to this point is the "effective height" in so far as determining the degree of cancellation or reinforcement of the direct and ground-reflected components of the space wave.

Actually the earth for some distance in back of and beyond the reflection point is of importance, but *in the case of fairly smooth terrain* the simplified assumption of a "ray" of radiation striking a "reflection point" will give a usefully accurate idea of what can be expected. For instance, an antenna located 30 feet above the top of a gradual rise might have an effective height of only slightly more than 30 feet for radiation at angles of elevation approaching the vertical, but an effective height of several hundred feet for a vertical angle of around 5 degrees. It is assumed in every case that the ground-reflected "ray" does not again strike the earth, such as a high range of mountains.

Except for high angles of radiation at the lower end of the h-f range, ground reflection may be considered as occuring at the surface of the terrain without introducing a serious error.

Frequencies in the range 8 to 20 or 25 Mc. are widely employed for long distance point-to-point work. For this application elaborate, highly directive arrays often are utilized, though there is a limit to the amount of directivity which may be employed unless the antenna is made electrically "steerable" over several degrees of arc, because signals sometimes deviate from a great circle path as previously discussed.

Commercial and government high frequency stations, or at least the more elaborate ones, usually are located in a reasonably level area some distance from any high mountain range which might interfere with low angle radiation, and free from automobile and other man-made electrical interference. The ground conductivity is of little importance. A location near the ocean is advantageous only in so far as the sea acts as a very flat zone plate or reflecting mirror and minimizes divergence and scattering of the earth-reflected component in the case of waves propagated seaward. Also, there is some evidence that the presence of the sea may facilitate reception via "long scatter" when a vertically polarized receiving antenna is employed. Under some conditions it appears that a sky wave which cannot, because of ionsphere conditions, reach the receiver directly may excite a surface wave some distance seaward which is propagated to the receiver as a ground wave with but little attenuation.

SHORT-DISTANCE SKY-WAVE TRANSMISSION

Sky-wave transmission is not employed solely for medium and long distance communication. In some cases a much stronger (though not so steady) signal can be delivered over a distance from 10 to 100 miles via sky-wave transmission (using optimum frequency). This is particularly true when the power is low, the permissible or practicable height of the antenna or supporting structure is limited, and the intervening terrain is of mountainous character and low conductivity, or is of tropical jungle, or other conditions unfavorable to ground-wave propagation exist. When sky-wave transmission is employed for such distances and low power is used, it is important that the operating frequency be near the MUF, especially in the daytime.

LONG AND SHORT PATH — ROUND THE WORLD ECHOES

Under certain conditions a stronger signal may be propagated via the "long" great circle path than via the "short" path. In fact, when the transmitter power is high and the great circle transmission paths do not pass through a zone of winter night, it is possible at times to detect echoes, or signals which have made one or more "extra laps." The echo pattern will be more complex when omnidirectional or 180 degree bidirectional antennas are employed, because then there will be signals (both direct and echo signals) arriving from both great circle directions. When a undirectional array with good front-to-back discrimination is employed on either the transmitter or receiver or both, the signal and echoes usually will arrive from only one direction.

SPORADIC E PROPAGATION

The foregoing discourse on high frequency propagation practice applies only to propagation via the "regular" layers. While sporadic E "patches" may on some days act as the controlling layer as much as 80 or 85 per cent of the time, the behavior of the E_s layer is erratic and difficult to predict accurately. Therefore, while the existence of sporadic E oftentimes greatly increases the MUF over what it would be for the regular layers, standard commercial and government practice is to employ an OWF which disregards the possible presence of sporadic E ionization, so that should the sporadic E layer suddenly disappear, communication would not be interrupted. This is practicable because of the high efficiency with which sporadic E ionization reflects waves even in the daytime; frequencies well below the sporadic E MUF show little absorption even when E_s is the controlling layer. This is not true of the other layers except at night.

REDUCTION OF FADING

The more elaborate of the fixed commercial and government communications circuits in the high frequency range often employ "diversity" reception in order to minimize interference fading effects. In one form this entails the use of two or more receivers (usually two) with their a-v-c circuits tied together and their audio outputs combined, so that the receiver receiving the weaker signal is effectively inoperative and does not contribute much to the background noise. The receivers usually are fed from different antennas, the antennas either being spaced considerably in terms of wavelength (space diversity) or else employing different polarization (polarization diversity). Interference effects will not be the same for the different antennas at the same instant, and seldom will the signal be unsatisfactory on both antennas at the same instant.

Somewhat the same effect can be realized with a single receiver in the case of c-w telegraphy or other "on-off" keying by applying a small amount of phase or frequency modulation to the carrier. A form of diversity reception utilizing a single receiver is also realized in the case of frequency shift keying, where "mark" is sent on one frequency and "space" on a slightly different frequency.

Distortion due to interference fading ("selective" fading) of amplitude modulated voice signals can be greatly reduced by use of "exalted carrier" reception, whereby the carrier is selectively amplified so as to be much stronger than the sidebands under normal conditions. The received carrier then can fade to the point where its amplitude is quite low in proportion to the sidebands, without overmodulation distortion occuring during detection, because the "exaltation" of the received carrier will permit it to "support" the sidebands.

VERY-HIGH AND ULTRA-HIGH FREQUENCIES

Because of the infinitely greater number of channels, the avoidance of sky-wave interference between widely separated stations, and the fact that a comparatively low antenna mast is suitable when an elevated site is available, the v-h-f and u-h-f portions of the radio spectrum are becoming more and more thickly occupied, in many cases taking over or sharing applications previously handled exclusively by means of surface-wave propagation on low and medium frequencies.

For ground-to-ground applications, which represent a very common usage of v-h-f and u-h-f communications, the antenna elevation is by far the most important factor except under certain unusual circumstances which will be discussed later on. While transmitter power and antenna directivity and receiver sensitivity all naturally have considerable effect upon the maximum reliable communication range and upon "primary" broadcast coverage, antenna elevation (both transmitting antenna elevation and receiving antenna elevation) ordinarily has a much greater effect.

The term *elevation* is used rather than *height,* because, unlike the case of an antenna in the 550-1600 Kc. range, the antenna (or an element of an array) usually is not more than a few feet in length, and ordinarily is elevated at least several wavelengths above earth. In the case of an AM broadcast antenna, one end of which always is at ground level, height and radiator length are synonymous. Not so with the v-h-f or u-h-f antenna which is made up of an element or elements whose length is determined by frequency, and is substantially the same regardless of how high the antenna system is elevated above the earth.

Another conspicuous contrast with medium-frequency ground-wave propagation is the difference in effect produced by ground conductivity and by irregularities in the terrain. Soil conductivity has little effect upon v-h-f and u-h-f propagation except when using vertical polarization at the low frequency end of the v-h-f band, but hills—even low hills—cast "shadows" in which the signal strength is considerably if not greatly attenuated. Medium-frequency waves, on the other hand, pass over low hills almost as though they did not exist, but the propagation of medium-frequency waves is greatly affected by the soil conductivity.

The field strength or "coverage" due to the ground wave of a medium-frequency transmitter can be predicted quite accurately from comparatively simple calculations based upon the effective conductivity of the terrain over the transmission path, the transmitter power, and the known effectiveness of the proposed type of radiator and ground system.

Because of the great effect of the nature of the intervening terrain upon the attenuation of a v-h-f or u-h-f transmission path, and because the effect of irregular terrain is too complex to be reduced to a simple pat formula, determination of the coverage to be expected from a v-h-f or u-h-f transmitter is much more difficult and more liable to error. Oftentimes an intelligent off-hand appraisal of the transmission path, amounting to nothing more than an intelligent "guess" by someone who has had considerable experience with such problems, will give almost as accurate an answer as will the detailed application of formulas and charts based upon the latest theoretical and experimental data. However, it is impossible to acquire the ability to do any "intelligent guessing" without first knowing and appreciating what considerations are involved. The principle ones will now be dealt with.

FACTORS DETERMINING V-H-F/U-H-F COVERAGE OR CIRCUIT PERFORMANCE

Starting at the transmitter, the first factor which determines the maximum communication or broadcast range is the type of modulation. Intelligence requiring a large bandwidth (such as television) results in reduced circuit performance, because the receiver bandwidth must be made correspondingly wide, and more noise energy is included in the pass band.

The next factor is the rather obvious one of transmitter power. Then the transmission line loss. (These two factors combine to determine the power delivered to the antenna.)

Next comes the transmitting antenna directivity, which determines the power gain over an isotropic radiator. Because all commonly used v-h-f and u-h-f antennas are efficient radiators, the radiation efficiency need not be considered. It may be assumed that virtually all of the power delivered to the antenna proper is radiated.

Then comes the factor, which, except under conditions of nearly flat terrain, is so difficult to evaluate accurately: the attenuation of the transmission path. This will be determined or affected by the distance, frequency, antenna heights, contour and conductivity of the intervening terrain, and, in the lower portion of the v-h-f band when the antenna height is low and the earth conductivity high, by the polarization. At extreme distances, particularly beyond the horizon, meteorological conditions become of prime importance.

At the receiving end, the required signal level for minimum acceptable circuit performance above 100 Mc or so seldom is determined by noise (cosmic, atmospheric, or man-made) picked up by the receiving antenna, but is determined by inherent receiver noise. Therefore the intensity of the ambient electrical noise field may or may not be a factor, depending upon frequency, how good the receiver is, etc. We can't simply say, except as a generalization, that a signal field strength of so many microvolts per meter will override the ambient noise field and therefore provide satisfactory service. The efficacy of the receiving system, specifically its signal-to-noise ratio or "noise factor," must be taken into consideration. It is for this reason that the "noise factor" method of rating v-h-f/u-h-f receiver performance has become so important in v-h-f/u-h-f work.

The first factor at the receiving end, as we follow the signal from the transmitter, is the "capture area" or effective intercepting area of the receiving antenna. A factor which is closely related to and helps determine the capture area is the degree to which the energy intercepted by various portions of the antenna or array all adds up in phase at the receiver input. This is simply another way of specifying *directivity*. The directivity of the antenna system (without regard to physical size) determines the equivalent power gain on "transmit," and the effective capture area on "receive," if a fixed value of resistance loss is assumed.

Next comes the factor of transmission line loss. When inherent receiver noise is the limiting noise, as usually is the case above 100 Mc. or so, 10 db transmission line loss will degrade the circuit performance (signal to noise ratio) by 10 db regardless of whether the loss is in the transmitter transmission line or the receiver transmission line.

Then comes the noise factor or noise figure of the receiver.* Receiver bandwidth is not considered a factor because it is assumed that in every case it is just sufficient to pass the required modulation, and modulation bandwidth was previously considered.

Except for the attenuation of the transmission path, all of these factors are amenable to simple and accurate calculation. For instance, it is obvious that the use of a directional transmitting array which increases the field strength in the desired direction by a factor of 10 produces the same improvement in circuit performance as an increase of 100 times in the transmitter power, the improvement in either case amounting to 20 db at the receiver. The factor of transmission path attenuation, however, does not necessarily fit inflexible formulas except in certain simple, idealized cases. In most practical cases this factor involves qualitative considerations and recognition of experimental data.

ATTENUATION OF THE V-H-F/U-H-F TRANSMISSION PATH (REGULAR TERRAIN)

Within the *radio path horizon* for a "standard" atmosphere (determined by applying a correction factor of 1.33 to the earth's radius), the variation in signal strength over *perfectly regular terrain* (such as sea water) follows certain simple laws (within limits) which can be utilized to give a "starting point" in the determination of range, field strength, etc. under practical conditions of irregular terrain and man-made shadowing structures. These simple relationships apply to a *normal* space wave only; *they do not apply when strong tropospheric refraction or reflections are present.* Also, they apply only when the length of the transmission path is large compared to the product of the antenna heights (both expressed in terms of wavelength). And they apply to vertical radiators only when located more than a few wavelengths above earth; otherwise they apply to both horizontally and vertically polarized waves.

These relationships may be expressed as follows: *Within the limits and observing the limitations just discussed, the field strength is proportional to frequency, proportional to the product of the antenna elevations, and inversely proportional to the square of the distance.* (It should be noted that if one antenna elevation is fixed, the product of the elevations varies as the elevation of the other antenna).

Beyond the radio path horizon, different conditions obtain. The field produced by the space wave falls off much more rapidly than the square of the distance; it also increases faster than the product of the antenna elevations. And for points well below the radio path horizon the field strength *decreases* with increasing frequency, instead of increasing.

When the product of the elevations is an appreciable proportion of the distance between the radiators in terms of wavelength, to the extent that the difference in path length between the direct and the ground reflected components of the space wave no longer is a very small fraction of a wavelength but is a matter of wavelengths, then the relationship between elevation and field strength, the relationship between distance and field strength, and the relationship between frequency and field strength, will all exhibit an oscillatory character.

The reason is apparent from the discussion of space wave patterns earlier in this chapter. In one sense, it may be said that the relationship between frequency and field strength always exhibits an oscillatory character over smooth earth, because for common numerical heights and distances it always is possible to produce a path difference of a matter of wavelengths *simply by going to a sufficiently high frequency.*

*The *noise figure* or *noise factor* of a receiver is a measure of how closely its signal-to-noise ratio approaches that of an ideal receiver in which all of the noise present in the receiver output is contributed by the thermal agitation noise generated in the resistive component of the antenna system or the internal resistance of the signal generator. It may be expressed as a numerical ratio or in decibels. A receiver of good design, incorporating the latest developments in circuits and tubes, may show a noise figure of approximately 5 db at 30 Mc., 8 db at 60 Mc., and 12 db at 600 Mc. These figures are typical, but are not the best that can be obtained, especially if other performance features and operating conveniences are sacrificed and if cost is not important.

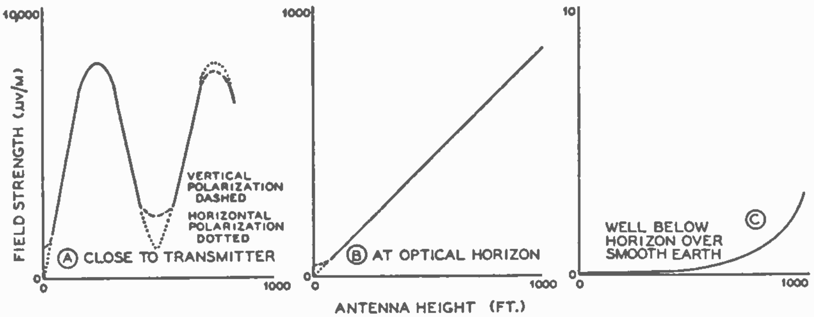


Figure 1-14.

ILLUSTRATING MANNER IN WHICH THE GROUND WAVE FIELD STRENGTH OF A V-H-F TRANSMITTER VARIES WITH ELEVATION AT A POINT NEAR THE TRANSMITTER (A), AT A POINT ON THE OPTICAL HORIZON (B), AND AT A POINT WELL BELOW (BEYOND) THE HORIZON (C). (TROPOSPHERIC REFLECTIONS ARE NOT CONSIDERED A COMPONENT OF THE GROUND WAVE. SMOOTH TERRAIN IS ASSUMED).

However, this is an academic point, because the oscillatory character is not observed at distances beyond 2 or 3 miles for frequencies below 1000 Mc. (the region in which we are concerned), except when very high antennas are employed. Also, the oscillatory effect tends to be "wiped out" at microwave frequencies under typical conditions of terrain, due to the fact that small ground irregularities become large in terms of wavelength and produce random or "scattered" reflections.

In any event, the effect is seldom of great importance in ground-to-ground work on frequencies below 1000 Mc., because at distances close enough for the effect to obtain (for all practical antenna heights) the direct and ground reflected components of the space wave are so strong that even when they are 180 degrees out of phase the resultant field strength is quite high, due to the slight difference in the amplitude of the two components.

When one or both ends of the circuit is an aircraft, however, the oscillatory effect can assume considerable importance, because of the possibility of its occurance even at extreme ranges when the aircraft is flying at great altitude. The variation in field strength with antenna elevation for the various conditions just discussed is depicted graphically in figure 1-14.

While not of much direct use except in the case of v-h-f/u-h-f propagation over the ocean, prairie, or other terrain which is "flat" except for the curvature of the earth, the following formulas give the distance to the optical and radio path horizons (the latter under conditions of standard atmosphere) as a function of antenna height (elevation):

$$H = \frac{D_o^2}{1.51}$$

$$\text{or } H = \frac{D_r^2}{1.99}$$

$$\text{or } D_o = 1.23\sqrt{H}$$

$$\text{or } D_r = 1.41\sqrt{H}$$

where H is the height (elevation) of the antenna, D_o is the distance in miles to the optical horizon, and D_r is the distance in miles to the radio path horizon *based on normal or "standard" refraction.*

When both antennas are elevated, as is the usual case, the horizon distance for each must be calculated and the two added in order to determine the maximum "line of sight" transmission. When the space wave just grazes the earth in travelling from one antenna to the other, the two antennas have a common horizon at the grazing point, or "point of tangency."

An idea of the ground wave coverage to be expected from a v-h-f transmitter radi-

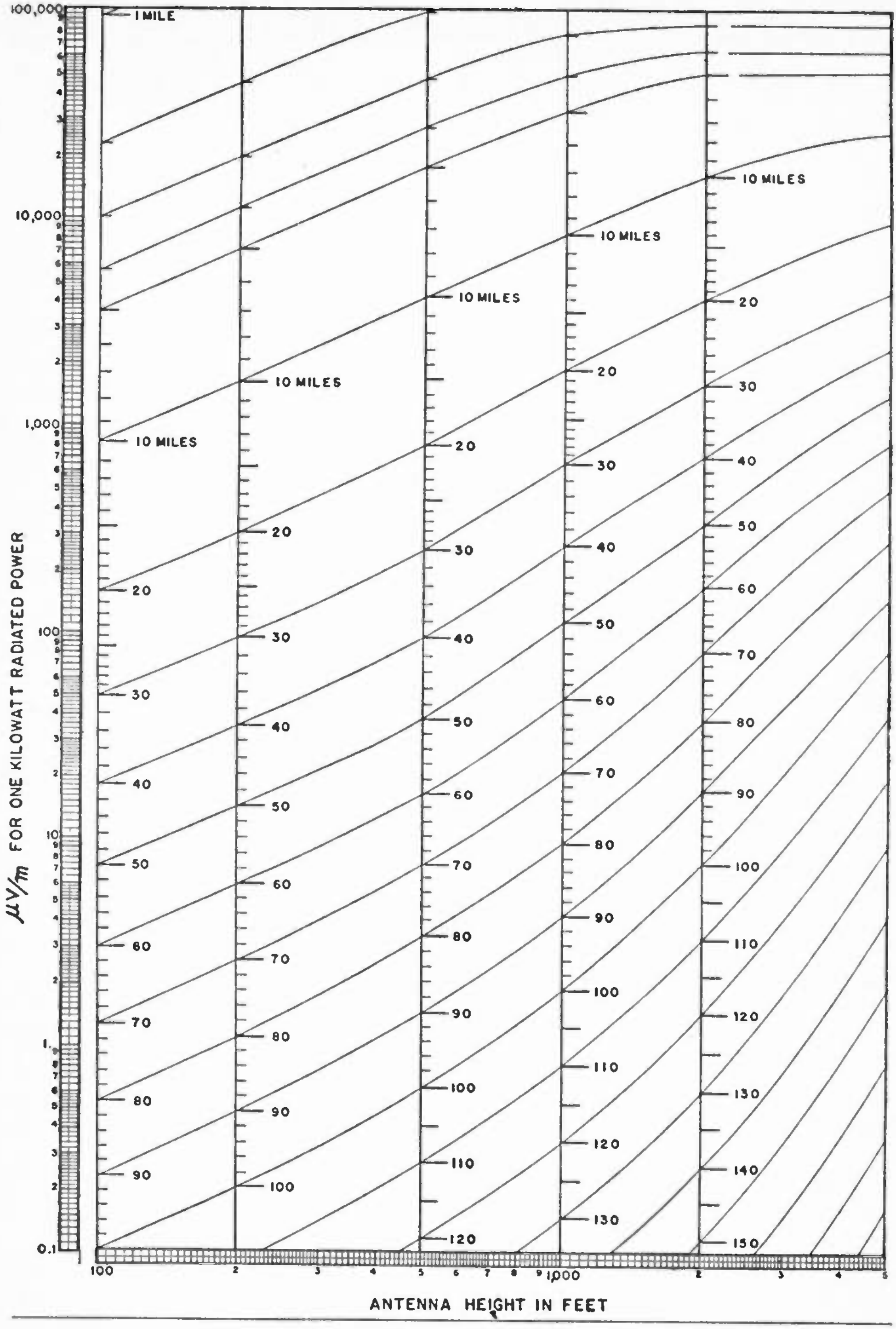


Figure 1-15.

GROUND WAVE SIGNAL RANGE FOR FM BROADCASTING (FROM F.C.C.)

Curves are for horizontal polarization, average soil, regular terrain, a frequency of 98 Mc., a receiving antenna elevation of 30 feet, and "standard" atmosphere. Vertical polarization will give insignificant difference in field strength.

ating through standard atmosphere over smooth land from an antenna having no directivity in the horizontal plane may be had from figure 1-15. This chart is from the Federal Communications Commission's publication, "Standards of Good Engineering Practice Concerning FM Broadcast Stations."

Beyond approximately half the distance to the optical horizon (still assuming smooth earth), the field strength encountered in actual practice will deviate considerably from the values given by the curves, due to variations in the lower atmosphere from average or "standard" conditions. The deviations take the form of random fluctuations and variations which more often increase the field strength above the intensity for standard atmosphere, but also decrease it a good portion of the time. At distances out to twice the optical horizon the maximum deviation in intensity, expressed in db, is about the same above "standard" as below, though, as noted above, at any given moment the intensity is more likely to be above "standard" than below. At the optical horizon the maximum deviation ordinarily encountered is about plus-minus 8 db from "standard."

As a matter of interest, the field intensity considered necessary for "program quality" or "primary" broadcast service is 1000 μv/m. for business or factory areas of cities, and 50 μv/m. for rural areas.

The field strength required for minimum satisfactory intelligibility on a *communication* circuit under conditions of low ambient electrical noise (single voice channel) is much lower for the same frequency, usually being on the order of 0.5 to 5 microvolts per meter, depending upon the noise figure of the particular receiver employed and the effective intercepting area of the particular array or antenna system employed.

It should be kept in mind that while the field strength at distances just inside the optical horizon is roughly proportional to frequency (for the same transmitter power, antenna heights, etc.), a resonant half wave receiving dipole antenna becomes proportionately smaller at higher frequencies and "picks up" proportionately less voltage

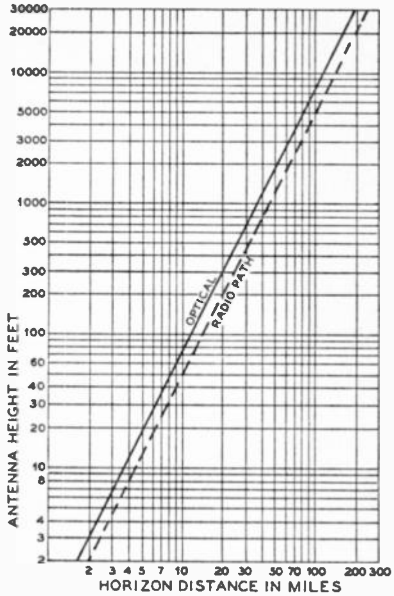


Figure 1-16.

CHART SHOWING DISTANCE IN STATUTE MILES TO OPTICAL AND RADIO PATH HORIZONS AS A FUNCTION OF ANTENNA ELEVATION FOR PERFECTLY SMOOTH EARTH AND "NORMAL" OR "STANDARD" REFRACTION

for a given field strength. This puts a practical limit upon the frequency which can be used for *non-directional reception*.

By using a directional receiving array and maintaining the effective intercepting area of the array constant as the frequency is raised, approximately the same receiver input can be maintained for a given field strength regardless of frequency. However, as the array becomes larger, the directivity becomes greater (assuming the phase relationships in the array are such as to provide maximum gain). This may be an advantage or a disadvantage, depending upon the particular application. In actual practice the effective area of the array must be made somewhat larger for the same performance as the frequency is raised beyond about 100 Mc., because both the noise figure of typical receivers and transmission line losses go up appreciably with frequency above this point, and greater input is required for the same circuit performance.

It should also be kept in mind that some of the generalizations which have been made are not valid in the lower portion of the v-h-f band for propagation of vertically polarized waves over sea water or salt marsh. Under these conditions the surface-wave component has considerable magnitude, and is effective for considerable distance when the elevation of the antennas is low at both ends (less than 100 feet or so).

Out to a distance of about 25 miles, the strength of a 100 to 1000 Mc. ground wave usually will be strong enough to dominate any tropospheric waves sufficiently to restrict fading to a tolerable value at all times, provided the two antennas are within optical line of sight. When the antenna heights are barely sufficient to provide an optical path, fading effects usually become noticeable beyond this distance, and increase in intensity with both frequency and distance.

The fading effects can be reduced, usually to a negligible magnitude for frequencies in the v-h-f range at distances out to 50 miles or so, *simply by utilizing considerably more antenna elevation than the minimum required for an optical path.*

It will be noted that reference is made to an *optical* path. The reason is that the "radio path horizon" for the regular ground wave varies with the weather, and under some conditions may be no farther than the optical horizon. In rare cases it may be even less. Therefore, when the utmost in circuit reliability is desired and deep fading cannot be tolerated, the only safe procedure is to base everything on optical line of sight. Then the circuit performance will be substantially independent of any vagaries in the refractive index of the lower atmosphere.

It should be kept in mind that the foregoing remarks apply to fairly smooth or "regular" terrain (horizon determined primarily by the bulge of the earth rather than by terrain features).

PROPAGATION OVER IRREGULAR TERRAIN

Propagation over irregular terrain does not conform to any hard and fast formulas, and it will be found that when irregularities of considerable proportion are present, the profile of the terrain separating the transmitting and receiving antennas has more effect upon the field strength than does the distance between then. The effect of the profile upon field strength will be discussed later in this chapter.

Even when the intervening land is quite flat, reflections from ground irregularities, trees, buildings, and other nearby man-made structures cause interference patterns or "standing waves" which produce variations of as much as plus-minus 10 db (total variation of 20 db) in field strength between two points of comparatively low elevation only a few feet or yards apart. Because the interference pattern is highly sensitive to frequency, these "multipath" effects cause severe distortion in the case of certain complex types of modulation covering a wide band, such as television. The effects can be minimized by using a directive array, high and in the clear, so as to make the direct path the dominant path, and by the use of horizontal polarization. Because the spurious reflections from buildings usually are reflected from surfaces lying in a vertical plane, the reflection coefficient can be reduced by the use of horizontal polarization.

Multipath reflections become more troublesome as the frequency is raised, because the flat reflecting surfaces of buildings and other structures then are larger in terms of wavelength and constitute a more effective "zone plate," or mirror. "Shadow" effects also increase with frequency. An object large enough to produce a strong shadow at 1000 Mc. might cast only a negligible shadow at 30 Mc.

The effect of residential or low business buildings in an area which is otherwise fairlv level is to produce variations in field strength at low elevations of approximately plus-minus 10 db as previously discussed, and to lower the *mean* intensity by about 10 db. This means that the signal usually will be anywhere from 0 to 20 db weaker than if the buildings, trees, power lines, etc. were not present (unless, of course, the receiving antenna is well elevated.).

The wide variations in field strength with slight changes in location are particularly noticeable in the case of a v-h-f receiver in a moving vehicle and an average signal level near the a.v.c. or limiter threshold.

PROPAGATION BEYOND THE RADIO PATH HORIZON

Ground-wave propagation beyond the optical horizon and out to the variable limit of the radio path horizon is explained by refraction. Ground wave propagation beyond the radio path horizon is due to diffraction. As a result the attenuation beyond the radio path horizon due to the bulge of the earth is quite high in the v-h-f/u-h-f range, the attenuation being greater with increasing frequency. Below (beyond) the radio path horizon, fading of the ground wave becomes very severe; the high rate of attenuation over the diffracted path causes wide variations in field strength as the diffracted path length to a certain point changes in conformance with the varying radio path horizon.

When the horizon is limited by a hill or mountain rather than by the normal bulge of smooth earth, the attenuation beyond the shadowing object is difficult to predict accurately, and may even become negative. This is explained by the fact that as the receiving antenna is moved beyond the shadowing object (assuming the same antenna elevations are maintained), less diffraction of the wave is required, and the reduction in the required diffraction may more than offset the attenuation due to the increased path length. In fact, if the transmitting antenna is appreciably higher than the shadowing structure, moving the antenna sufficiently beyond the shadowing object may bring it back within line of sight.

TROPOSPHERIC WAVES— ANOMOLOUS PROPAGATION AT V.H.F. AND U.H.F.

Tropospheric waves and the mechanism whereby they are produced were described briefly in general terms earlier in this chapter under "Radio Wave Paths." Because their existence and effects are highly pertinent to a discussion of v-h-f and u-h-f propagation, it is recommended that the reader review that portion of the section which deals with tropospheric waves before proceeding to the following elaboration and supplementary information.

SUPERREFRACTION

When the atmosphere or a layer of atmosphere has a refractive index gradient which is substantially lower than that of "standard" atmosphere or is reversed, the atmosphere or layer of atmosphere is said to be *sub-standard,* and the refraction produced thereby likewise is called "substandard."

When a layer of atmosphere has a refractive index gradient enough higher than standard to produce a curvature of grazing-incidence waves *which exceeds the curvature of the earth,* the layer is said to be *superrefractive,* and the refraction produced thereby is called *superrefraction.* If the layer is sufficiently thick to give the wave time to execute enough of an arc while within the superrefractive region, the wave will not "perforate" the layer but instead will be turned downwards and emerge from the side of the layer it entered.

GUIDED PROPAGATION—"DUCTS"

When a superrefractive layer exists over smooth earth, a wave leaving a ground-level transmitter along a path tangent to or nearly tangent to the earth will be refracted downward from the layer and strike the surface of the earth again, in spite of the earth's curvature, if the frequency is high enough. Most of the energy striking the earth will be reflected upward again towards the layer, to be refracted downwards again, and so on out to the limit of the superrefractive layer. Some of the energy leaving the antenna below a certain angle will never "perforate" the layer. It will be guided around the bulge of the earth and will spread only in one dimension: laterally. Under certain conditions the absorption and "leakage" will be very low, and the attenuation therefore quite low.

This type of propagation is called *guided propagation, duct transmission,* or *trapping,* for obvious reasons. The region between the earth and the top of the superrefractive layer is called a *duct,* in this case a *ground based duct* because the earth forms the lower boundary. If the frequency is sufficiently high and the height of the ground-based duct is low, say a few hundred feet or less, it will be found that there are no "skip zones" in the coverage by guided propagation except perhaps at distances well within the maximum range covered by the regular ground wave. There will be a "cutoff" frequency below which the duct does not produce effective propagation by means of superrefraction. Thus it can be seen that the duct in many respects resembles a hollow, metallic wave guide of the type used for microwave power transmission, except for spreading in the horizontal plane and the fact that there is no reflection from sharply defined boundaries or walls in the case of the duct.

The frequency below which "trapping" no longer occurs is determined by the height of the duct and the total variation or differential in refractive index represented by the superrefracting layer. The greater the total variation in the atmosphere within the duct, and the thicker (higher) the duct, the lower will be the "cutoff" frequency of the duct.

Just as in the case of a hollow wave guide, there will be many modes of propagation within the duct at frequencies considerably in excess of the cutoff frequency. This explains the absence of "skip zones." However, when propagation is primarily or exclusively by means of guided propagation, there is an optimum elevation within the duct for the transmitting antenna and for the receiving antenna.

In the case of ducts ordinarily encountered in actual practice, the cutoff frequency may be anything between about 100 Mc. and 3000 Mc. or so, the latter case applying to a particular type of ground-based duct which hugs the ocean in parts of the trade wind belt and is only a few feet high. Cutoff frequencies considerably above 100 Mc. are the general rule, and for this reason guided progagation is of greatest importance in connection with frequencies in the u-h-f band and above. However, duct propagation should not be completely ignored in connection with propagation in the v-h-f band, particularly the upper portion. In fact, duct propagation along the sea coast is frequently encountered down to 50 Mc.

The ground-based duct just described is an example of *surface trapping*. The transmitting antenna must be below the refracting layer in the case of a well defined duct, or there will be no trapping of the wave within the duct. The receiving antenna must likewise be within the duct or no "guided" wave will be received.

ELEVATED DUCTS

It also is possible to have a duct comprising two elevated "layers," in which case the duct is referred to as an *elevated duct*. The lower layer or boundary of an elevated duct is a layer of *substandard* atmosphere, the variation in refractive index being such as to bend a near-grazing, downcoming wave *upwards*. The lower boundary may in some cases be as effective as the upper one in refracting the trapped wave. Then a wave leaving a transmitting antenna placed within the well defined elevated duct will "perforate" the upper layer and be "lost" if it strikes the lower layer at too steep an angle and is reflected from the earth rather than being refracted from the lower layer. This means that at an appreciable distance from the transmitter, virtually all of the energy (ignoring absorption) will be either trapped between the two layers forming the elevated duct, or else radiated into space either directly or by reflection from the earth. The field strength below the elevated duct will be weak or negligible at distances beyond the regular ground wave.

When the latter condition obtains, little or no signal will be received beyond the regular diffraction region of the ground wave unless the receiving antenna is at such an elevation as to be *within the elevated duct*. When a signal temporarily drops below usability due to the formation of a transitory elevated duct or a transitory layer of substandard atmosphere which causes the signal to pass over the receiving antenna, it is called a "tropospheric blackout." While this condition occasionally is encountered in actual practice, it is not nearly so common as that in which the signal strength is greatest with the receiving antenna placed at an optimum height within the limits of an elevated duct, but is also usable at elevations above and below the elevated duct.

COMPLEX PROPAGATION BY SUPERREFRACTION

The simple cases of trapping just cited, that of a ground based duct and that of an elevated duct, each existing alone and with rather idealized refracting layers or boundaries, do sometimes exist in actual practice, even being rather common in a few certain geographical locations or areas. However, the more usual situation is a *complicated structure of stratification which is exactly the same only over a rather limited lateral area and is continually changing,* due to air turbulance and the general effects of weather. The ducts may not be sharply defined, and the vertical pattern or structure may vary considerably within a lateral distance of less than the maximum transmitter range. Transitions from a ground-based duct to an elevated duct, or vice versa, are not uncommon over a long path. This may produce a "skip zone" at ground level.

Under these conditions the chief manifestations of guided propagation are wide and erratic variations in signal strength with both time and antenna elevation, the latter being superimposed on any variations due to lobes in the vertical space wave pattern. The variation in field strength at a distant point over a considerable period of time, say 12 hours, may exceed 100 decibels for fixed antenna installations, the field strength at the receiving antenna often varying from a maximum value having the same order of magnitude as or considerably above the "free-space" value, to a minimum value so far below the free-space value as to be unusable. The fading or variation in field strength usually tends to be greater with increasing frequency, particularly over very long *optical* paths.

In addition to the day-to-day or hour-to-hour variations in signal strength observed under conditions of guided propagation, there usually are two types of fading present: a relatively slow and deep fade having an average period of from perhaps 10 to 20 minutes, and a rapid, random, "scintillating" fade which produces small, irregular fluctuations in the signal.

Duct transmission is not appreciably affected by the polarization of the wave, the long distance circuit reliability being about the same (usually rather poor) for either vertical or horizontal polarization. The fading is not in phase for the two types of polarization, however, and some improvement in the reliability of a communication circuit subject to the effects of superrefraction can be effected by employing circular polarization at the transmitter and "polarization diversity" at the receiving end.

TROPOSPHERIC REFLECTIONS

Under conditions sometimes encountered in certain geographical areas, a duct forms which has an upper boundary with an unusually abrupt and pronounced discontinuity. Under such conditions there may be considerable *reflection* from the layer of waves striking it at near-grazing angles, and waves may be propagated by such means well beyond the regular ground wave range, just as in the case of a superrefractive layer. But when the wave is bent downward by partial reflection as in this case, rather than by superrefraction, the duct acts in a somewhat different manner. The cutoff frequency observed in the case of duct transmission by means of superrefraction still is observed, but for a given duct height the cutoff frequency is very much lower in the case of reflection. In fact, the reflecting layer becomes a more effiecient reflector as the frequency is lowered, and strong reflections may be noted at the lower end of the v-h-f band.

When the reflecting boundary is high, say between 4000 and 8000 feet, it acts more like a "small scale ionosphere" than it does a wave guide, exhibiting skip zones and other effects observed with ionospheric sky-wave propagation. For this reason the mechanism of such propagation is called "tropospheric reflection" rather than "guided" or "duct" transmission.

CIRCUIT PERFORMANCE PROVIDED BY TROPOSPHERIC PROPAGATION

As noted earlier in this chapter, superrefraction and tropospheric reflections within the range limit of good ground-wave field strength serve no useful purpose and may cause serious interference fading, thusly requiring that more transmitter power be used, and, with some types of modulation, causing objectionable distortion.

At ranges beyond the useful limit of the regular ground wave, anything in the way of a signal at all is just so much largess; therefore waves arriving at such distances cannot be considered objectionable unless they produce interference to other circuits using the same frequency.

Whether or not tropospheric propagation coverage serves a useful purpose or provides a useful service depends upon the minimum acceptable circuit performance. For a service where interference fading produces serious distortion, such as in television, tropospheric propagation will not be satisfactory even for casual entertainment reception, except possibly for rare, short periods. On the other hand, the possibility of occasional v-h-f and u-h-f communication over distances of several hundred miles by means of superrefraction and tropospheric reflections affords the radio amateur a wonderful time, because he thrives on competition, speculation and uncertainty.

EFFECTS OF WEATHER UPON V-H-F-/U-H-F PROPAGATION

V-h-f and u-h-f coverage via tropospheric propagation, as well as the intensity and character of fading effects within the regular ground wave range, are greatly affected by weather conditions in the lower atmosphere. The effect becomes more pronounced as the frequency is increased.

Wide variations in performance may be noted as weather conditions change over the transmission path unless the path is relatively short. These variations are due to the critical dependence of the refractive index of the atmosphere upon temperature and humidity, especially the latter. Pressure also affects the refractive index, but compared to the effects of temperature and humidity the effects of pressure are insignificant.

THE TEMPERATURE INVERSION AND STEEP MOISTURE LAPSE

In normal or "standard" atmosphere the temperature decreases slowly with height. This tends to refract a horizontally traveling wave very slightly *upwards*. However, the decrease in humidity with height under standard conditions has a much greater bending effect, *downwards*, and the net result is a slight downward bending under conditions of normal or standard atmosphere.

Under certain conditions of weather and terrain the temperature of the atmosphere may, over a limited range of height, *increase* with height. This anomalous condition is called a *"temperature inversion."* If the increase is sufficiently rapid and extensive, superrefraction can occur from temperature effects alone. However, seldom is a temperature inversion encountered which is pronounced enough to support superrefraction under humidity conditions approximating those of standard atmosphere. When one does occur, the duct so formed is referred to as a *"dry duct."*

A rapid decrease in humidity with height, called a *"steep moisture lapse,"* produces a refracting layer with much greater bending properties than one produced by a simple temperature inversion, especially when the steep moisture lapse occurs in a warm atmosphere. The latter effect is explained by the fact that a difference in humidity represents a greater difference in refractive index in the case of warm air, because warm air contains a greater amount of water vapor for a given humidity.

Even though the refractive index and therefore the bending properties of the atmosphere are much more dependent upon humidity than upon temperature, temperature inversions are nevertheless of prime importance, because *a temperature inversion usually is conducive to the production of a steep moisture lapse,* for reasons which will be discussed in some detail later. Not only do temperature inversions tend to cause or, more accurately, permit a steep moisture lapse, but they also are responsible for a minor direct contribution to the variation in refractive index with height. Quite often this slight contribution makes the difference between superrefraction and no superrefraction when the moisture lapse alone is not quite sufficient to support superrefraction.

When the steepness of the moisture lapse rate is the major contributor towards superrefraction, the duct so formed is designated as a *"moist duct."*

It should be kept in mind that while a pronounced temperature inversion almost invariably produces or permits a steep moisture lapse (except perhaps in arid or desert areas), a steep moisture lapse can and often does exist *without* the contributing factor of temperature inversion.

Generally speaking, superrefraction is more common in warm than in cold climates; is most likely to occur below a few hundred feet in warm climates and between a few hundred and a few thousand feet in cold climates; and usually is due to a moist duct in warm climates and to a dry duct in very cold climates.

METEOROLOGICAL CONDITIONS PERTINENT TO SUPERREFRACTION

The meteorological mechanisms which are primarily responsible for non-standard variations in temperature and moisture with height are *advection, subsidence,* and *nocturnal cooling.* These will be briefly described before proceeding to a discussion of how they act to produce anomalous propagation.

The term *advection* refers to the lateral movement of a layer or body of air. A "breeze" is an example of advection. The effects of advection upon the vertical variations in temperature and humidity are greatest at low altitudes.

The term *subsidence* refers to a gradual, downward movement or sinking of air over an extensive area. The sinking air spreads out horizontally over the underlying air, usually at moderate altitudes, to produce pronounced stratification (usually a temperature inversion) in the atmosphere at altitudes somewhat above those at which advective effects tend to produce stratified gradients in the refractive index. Some degree of subsidence always accompanies a barometric "high," though not always of sufficient intensity to produce a temperature inversion. Subsidence seldom occurs in the region of a barometric "low."

Nocturnal cooling refers to the loss of heat by the ground during the night as a result of infra-red radiation skyward. When the cooling of the ground surface by radiation is sufficiently great, the air coming in contact with it is cooled enough to produce a low-level temperature inversion. Because there is little change in the surface temperature of a large body of water at night, nocturnal cooling does not occur over the ocean (excepting, of course, very shallow bays).

Certain other meteorological factors come into play in the generation of superrefractive ducts, either directly or indirectly, to impede or otherwise affect the three basic mechanisms mentioned. Chief among these are *convection* and *frictional turbulence.*

Convection refers to the tendency of the atmosphere to seek the equilibrium conditions of standard atmosphere as regards temperature lapse. More specificially it applies to the rising of air as it is heated from contact with the ground during that portion of the day the surface of the ground is being warmed by the sun and is hotter than the air above it. The rising of the warm air is not smooth and uniform, but somewhat turbulent, especially under conditions of a steep temperature lapse over irregular terrain.

When the temperature lapse rate is *lower* than that of standard atmosphere, or else is reversed (temperature inversion), then there will be no tendency for the lower air to rise, and negligible convection will take place. *This is an important characteristic of a temperature inversion,* for reasons which will become apparent.

Frictional turbulence refers to the turbulence produced in the lower portion of the atmosphere as a result of friction between the wind and the earth's surface. Naturally the turbulence becomes more intense and vertically extensive as the velocity of the wind increases, and is more intense and vertically extensive over land than over water (due to the greater friction).

Both convection and frictional turbulence tend to produce a rapid, thorough mixing of the lower atmosphere and therefore act to maintain standard atmosphere and refraction. When either convection or frictional turbulence is *very pronounced,* superrefraction is quite unlikely in that area.

The atmosphere also becomes mixed by seeking the equilibrium conditions of standard atmosphere through *diffusion*. However, compared to the "stirring up" of the atmosphere as a result of convection and frictional turbulence, the process is very slow. In fact it is so gradual that the time taken for a thorough mixing of a substantial layer of air by diffusion often is greater than the time it takes for the "weather" to change.

Formation of Advective Ducts

When warm, dry air flows over a cooler, moist surface as a result of advection, the following condition is created. A temperature inversion is formed as a result of the lowest air being cooled by contact with the ground. This process is conducive to the evaporation of moisture from the wet or moist surface, because a condition of warm, dry air in contact with a moist or wet surface is highly favorable to rapid evaporation. This causes the air just above the surface to acquire a very high humidity. Because of the temperature inversion, there is little mixing of the atmosphere by convection. The only rapid mixing is due to frictional turbulence, and this type of turbulence produces a rather well-defined upper boundary, the height of which is determined by the wind velocity. The end result is not only a temperature inversion but a steep moisture lapse at the upper boundary of the frictional turbulence. If there is but little frictional turbulence, such as in the case of a very light breeze over the ocean, then the duct will be very low, sometimes not more than 100 feet or so in height.

Probably the condition most favorable to the formation of an advective duct is that of a land breeze carrying warm, relatively dry air out over a cooler sea. Off-shore advective ducts formed in this manner are quite common in coastal regions, especially in temperate or warm climates.

Where the land along the coast is moist, and particularly if it is covered with lush vegetation (as in certain parts of the tropics), strong advective ducts often are formed over the land along the coast as a result of sea breezes. These ducts are formed during that part of the day when the land is cooler than the ocean, or in other words, at night.

Advective ducts formed along the coast, either off shore or over the coastal land, often are quite limited in extent in a direction normal to the coastline. In such a case the duct takes the shape of a long, narrow, coastwise strip.

The presence of nearby land is not essential, however, to the formation of an advective duct over the ocean. When the wind is steady and moderate, the frictional turbulence is low except just above the water. If the air is not appreciably cooler than the water, convection will be negligible, and under these conditions the only turbulence will be right near the surface of the water. The result is a steep moisture lapse within the first few yards of height and a duct from 25 to 75 feet high, the height depending somewhat on wind speed. Because a duct of this type is not only shallow but also rather weak, it will produce superrefraction only at microwaves, the cut-off frequency varying from roughly 3000 to 6000 Mc. (making the duct of academic interest so far as v-h-f and u-h-f propagation is concerned). The duct usually will be strongest at a wind speed of 15 to 20 m.p.h.

Advective ducts also are formed over inland areas under favorable conditions, but conditions are much less likely to be favorable and therefore such ducts are rather uncommon.

Ducts Resulting From Subsidence

By a complex thermodynamic process which is beyond the scope of this book, subsidence can produce strong temperature inversions. As previously related, the inversion may occur at an elevation of from a few hundred up to several thousand feet. If the inversion is high, and it produces a steep moisture lapse, a duct will be formed which may be ground-based or elevated, depending upon the nature of the atmosphere near ground level. If the atmosphere at ground level is sufficiently *sub*-standard, and the inversion layer is strong enough and high enough, an elevated duct will be formed.

Unless aided by advection or nocturnal cooling, the temperature inversion created by subsidence will not ordinarily produce a sufficiently strong moisture lapse for superrefraction *except over a large body of water*. Over inland areas the chief role played by subsidence is to facilitate the formation of ducts by means of advection and nocturnal cooling.

Formation of Ducts by Nocturnal Cooling

Ducts formed by nocturnal cooling of land surface as a result of radiation resemble advective ducts in many respects. In each case there is a temperature inversion caused by the land surface being cooler than the air above it, and the tendency of the air just above the land surface to be cooled by a combination of contact and turbulence. Nocturnal cooling, like advection, seldom will result in a duct unless it occurs over moist earth, because otherwise no steep moisture lapse will be created.

Nocturnal cooling is greatly retarded by the presence of heavy fog or clouds, because then much of the heat radiated by the ground is reflected back towards the ground, and the ground does not cool off nearly as rapidly as it does in the case of a clear sky. Even a highly humid clear atmosphere slows down the surface cooling by reflecting some of the radiated heat back towards the ground, though the effect is not so pronounced as in the case of fog or clouds. Clear, cool, dry atmosphere is the condition most favorable to formation of pronounced temperature inversion from nocturnal cooling. A dry atmosphere also facilitates surface evaporation and thereby increases the steepness of the moisture lapse.

"Composite" Ducts

While a duct may be due primarily to one of the three basic mechanisms just discussed (advection, subsidence, nocturnal cooling), a common situation is one in which two or all three factors make a substantial contribution. The analysis of such a situation often is highly complex and requires a comprehensive knowledge of meteorology.

Effect of Fog upon Refraction

For a given water vapor content, the refractive index of atmosphere is much lower when the water is contained in droplets than it is when the water is contained in the form of vapor. Thus a layer of fog can have a much lower index of refraction than might be expected, and even result in a layer of substandard atmosphere under conditions which might otherwise be favorable to duct formation. The extent of the effect depends somewhat upon the nature of the fog. Generally speaking, the type of fog usually encountered (radiative or advective) is deterimental to formation of a superrefracting layer within the fog layer. However, under some conditions a duct may be formed with the upper boundary located coincident with or above the *top* of the fog layer. When this is the case, the duct quite often is of the elevated type.

Scattered Reflections From Water Droplets

We have seen how fog can reflect infra red heat waves (under the discussion of nocturnal cooling). The same thing happens in the case of rain and radio waves, provided that the water drops are large enough and the wavelength short enough. Very heavy precipitation (either rain or snow) is necessary in order to produce perceptible reflection of 1000 Mc. waves, but scattered reflections begin to be quite noticeable at frequencies above approximately 5000 Mc. under conditions of moderately heavy precipitation.

Clouds do not produce reflection of microwaves unless they contain rain drops, or at least this is true of the highest radio frequencies used for practical purposes at the present time.

PATH PROFILE VS. V-H-F/U-H-F CIRCUIT PERFORMANCE

As stated previously, the most important consideration in evaluating reliable v-h-f-u-h-f coverage, or the circuit performance between two points, is the nature of the *profile* along the transmission path. In the case of smooth earth such as sea water, the profile is determined only by the earth's curvature, and therefore is a simple function of distance, as illustrated in figure 1-16. The curves permit instant determination of the distance to the optical or the radio path horizon (standard refraction) for a given distance, and vice versa.

Except for over-water paths, the horizon usually is determined more by irregularities in the terrain than by the bulge of the earth, though in some cases each may be of the same order of importance.

A convenient representation of a profile of the transmission path can be constructed from a suitable contour map or maps. The United States Geological Topographical Quadrangle Sheets are excellent for those portions of the U. S. A. which they cover. These may be obtained from the United States Geological Survey, Department of the Interior, Washington D. C. by enclosing 10 cents in coin for each sheet. For areas not covered by these maps, the Sectional Aeronautical Charts available for 25 cents each (in coin) from the United States Coast and Geodetic Survey, Department of Commerce, Washington D. C. are also good.

To facilitate interpretation of the profile or "contour strip" which is constructed from data taken from these maps, the horizontal scale should be compressed from 12 to 30 times more than the vertical scale. Unless such an expanded vertical scale is employed, the variations in elevation with respect to the distance along the path will be so small as to be difficult to work with.

While no longer a true representation or relationship, the expanded vertical scale emphasizes the shadow areas, and still gives a true picture as to whether or not a certain path is above or below line of sight. However, it should be kept in mind that the expanded vertical scale distorts the picture in so far as depicting the amount of diffraction required to reach a point which is in shadow. Such a countour strip may show that the wave must make an acute angle to reach a point several miles behind a shadowing mountain, whereas in actuality the required amount of bending may be quite small. Rather than a disadvantage, this distortion really is advantageous, as it makes it easier to see which points are in shadow. However, when comparing various "indirect" paths by means of contour strips, it is important that the same scales be used on all contour strips; otherwise the visual comparison is deceptive.

When irregularities of the terrain are of large magnitude as compared to the bulge of the earth for the distance in question, the curvature of the earth may be ignored in constructing the contour strip. This would apply, for instance, in constructing a contour strip of a 15-mile path which embraced a 300 foot rise, but not to a 60 mile path embracing a hill of the same height. In the latter case the base line should be drawn to correspond to the earth's curvature (taking the vertical and horizontal scales into consideration).

Before making a countour strip to evaluate an indirect path, the basic map or chart should be referred to in order to determine whether or not the profile includes any important peaks which are sheer in a lateral direction. Sometimes it will be found that if the profile passes near a gap or "pass," less diffraction or refraction is required to bend the wave *around* the peak than is required to bend the wave *over* it. Furthermore, in such a situation the far side of the gap or pass may be sheer enough to act as a reflector and give a surprisingly high field strength at the far end of the path. This points up the fact that, ignoring tropospheric reflections and refraction, there are two kinds of "indirect" path: that which relies entirely upon diffraction to get into the "shadow" region, and that which is aided and abetted by reflection from irregularities in the terrain existing to one side of the geometrically direct path.

When the profile shows a clear path (optical line-of-sight, no intervening hills), ordinarily there is no question of whether communication will be possible if the transmitter has a fair amount of power, the receiver has good sensitivity and the antenna systems are efficient. The only question concerns the ratio of ground wave intensity to tropospheric wave intensity. Increasing transmitter power, antenna directivity, or receiver sensitivity will not help in this respect; the only thing that will do any good if the distance is so great as to provide an objectional proportion of tropospheric wave is to *increase the product of the antenna elevations,* even though they already may be within line of sight.

When the receiving point is below line of sight, the amount the circuit performance is degraded by the presence of the intervening object is determined not only by the amount by which it rises above the "geometrical line of shoot," but also by such things as the configuration of the intervening object and its distance from the mid point of the path. For instance, a sharp, transverse ridge rising appreciably above line of optical sight will not degrade the performance nearly as much as a slightly lower obstruction at the same location but sporting a plateau having a longitudinal depth of several hundred feet. Also, any shadowing object (such as an ordinary hill for instance) will reduce the field strength of the regular ground wave least when it is located midway along the path, and most when it is at one end of the circuit, provided things are not complicated by the presence of a still higher shadowing object.

However, it is difficult to make even qualitative statements about the effect of the profile without making qualifications or exceptions. For instance, in the last case, suppose that the power were great enough and the distance short enough that diffraction of the regular ground wave would put in a strong signal regardless of whether the hill or other obstruction was at the mid point on the path or near the receiving end. Then with the obstruction at the mid point, the ground wave would be reduced much more than any tropospheric waves present, while with the obstruction very near the receiving end, they would be reduced in about the same proportion. Thus, less interference fading would be encountered and the circuit *performance* therefore would be better with the obstruction at the receiving end, even though the field strength would be less under these conditions.

It is even necessary to qualify this qualification, by saying that it would apply only in locations where strong superrefractive effects are common. Thus it can be seen that it is not practicable to take these factors into consideration in a pat formula. A qualitative approach, based on experimentally determined information, is as good as any. *Generally speaking,* it may be said that the sharper the edge of a shadowing object, the more readily a wave will be diffracted over the edge into the region that is in optical shadow, and that the height of intervening shadowing objects is important only in so far as it determines the amount of bending or diffraction required to reach the other point. The less the bending required, the better the circuit.

This means that a site at the base of an intervening hill, particularly a hill surmounted by considerable level area, is highly undesirable, and the circuit probably can be improved by locating the antenna back from the hill even though it increases the path length and decreases the antenna height. Unfortunately, nothing can be done about the detrimental effect of the flatness of the hilltop.

When the antenna location is well below optical line of sight due to the presence of highly irregular terrain consisting of *high* hills or peaks in the *immediate vicinity,* a very interesting condition obtains which at first appears to conflict with previous statements to the effect that antenna height is of prime importance at v-h-f and u-h-f. Communication over a 5 or 10 mile path is possible under such conditions by virtue of diffraction (provided the transmitter power and receiver sensitivity are reasonably high) and sometimes also bv scattered reflections which tend to pass *around* the obstructions. In such cases it may be found that increasing the antenna height from, say, 20 feet to perhaps 100 feet will offer little or no improvement, or even degrade the performance.

The reason for this is that the downward diffracted signal acts as though it were radiated from the summit or brow of the shadowing hill, and the horizontal distance between "antennas" no longer is great compared to the product of the antenna heights. Therefore, if the frequency is high enough, raising the low antenna may simply put the summit or brow in a null, thus reducing the signal strength. And if the signal is being reflected from the sides of other hills, raising the antenna may or may not increase the signal strength.

Generally speaking there is nothing to gain by raising the antenna under such conditions, assuming the antenna height in all cases is at least several wavelengths above earth yet much lower than the summit of the shadowing hill; on the contrary, the longer transmission line required for the higher antenna usually will introduce enough loss to more than offset any slight gain due to increased antenna height. Rather than attempt to obtain as much height as possible, it is better to keep the height reasonably low yet take care to avoid a height which produces bad interference cancellation in the pertinent direction or directions at the effective vertical angle. A furthe advantage of a reasonably low antenna height is that it more readily permits use of a rather elaborate directional array.

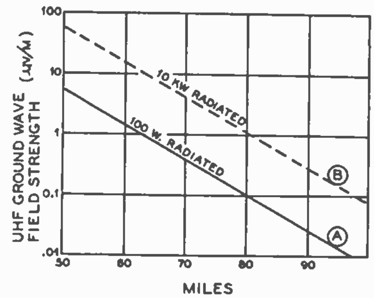


Figure 1-17.

ILLUSTRATING HOW A CONSIDERABLE INCREASE IN POWER RESULTS IN A COMPARATIVELY SMALL INCREASE IN GROUND WAVE RANGE OVER RELATIVELY SMOOTH EARTH AT DISTANCES WELL BELOW (BEYOND) LINE OF SIGHT.

The slope of the solid curve A represents typical attenuation of a u-h-f ground wave at distances well beyond the horizon. If the radiated power is increased 100 times, the curve simply is displaced upwards by a field strength factor of 10, without altering its shape. Observe that when this is done, the distance for a given field strength is extended outward by a comparatively small amount. It should be kept in mind, however, that this condition does not necessarily apply when the horizon is dictated primarily by irregularities in the terrain rather than the bulge of the earth. Also, in a practical case the actual extension of communication range ordinarily will be even less than that indicated for the ground wave, due to the fact that interference fading due to tropospheric effects tends to become worse as the distance is increased.

POWER VS. DISTANCE

Increasing the radiated power (either by substituting a more powerful transmitter or a more directional antenna array) will in every case increase the range of the transmitter. However, in many cases the increase in range for a moderate increase in radiated power will be surprisingly small, for reasons which will be explained.

In free space the field strength falls off directly as the first power of the distance, or in other words, is inversely proportional to distance. Thus, for a certain receiving apparatus requiring a certain field strength for satisfactory operation, increasing the transmitter power by four times (field strength by two times) would double the range. However, in the case of ground-to-ground communication, intereference effects resulting from the presence of the earth upset this relationship. Thus in the case of u-h-f reception well within the line of sight, but considerably beyond the maximum range at which the oscillatory characteristic is noted, the field strength over *flat earth* falls off approximately as the inverse square of the distance, assuming fixed antenna heights. The attenuation over actual, smooth *curved* earth is slightly greater but not much so within optical line of sight. This means that under these conditions and within these limitations, about 16 times the power is required in order to double the range.

But when we get to a point well below line of sight, particularly at the upper part of the u-h-f band, we find that the attenuation of the regular ground wave over smooth earth is so high that an increase of 100 times in radiated power may increase the ground wave range by less than 50 per cent. This is illustrated in figure 1-17, in which the solid line A depicts a typical field strength curve for a radiated u-h-f power of 100 watts. Assuming that a field strength of 1 microvolt per meter is required for satisfactory operation of the receiver, it is seen that the ground wave range of the 100 watt signal is about 65 miles. Now if curve A is simply displaced upward by a factor of 10, which is equivalent to raising the power 100 times, it is seen that the ground wave range has been extended only to approximately 90 miles. In a practical case the extension of actual communication range will be somewhat less, because, as the range is increased, interference fading becomes worse due to a higher ratio of tropospheric wave to ground wave, and still more power is required in order to ensure satisfactory signal strength at the bottom of a fade.

In the case of tropospheric waves beyond the significant ground wave range, it is difficult to deal in anything other than broad generalizations and rough approximations. However, in one common type of tropospheric propagation the signal is received with good *average* strength from a point just beyond the "skip" zone for the tropospheric wave out to a distance of somewhere between 75 and 200 miles, depending upon frequency, duct height, transmitter power, and other factors.

An increase in power does increase the circuit reliability over moderate distances, and permits reception in locations which might be considered unfavorable from the standpoint of shadowing objects or high ambient electrical noise. It also permits simpler and less expensive receiving apparatus to be employed out to the optical horizon.

The foregoing comments with regard to power vs. ground wave signal range apply to propagation over substantially smooth earth. Under conditions of irregular terrain they do not necessarily apply. For instance, in the case of a sharp ridge running laterally, beyond which there is an extended stretch of flat or slightly rising terrain, it may be found that the attenuation characteristics are such that while the field strength falls off rapidly beyond the crest of the ridge for a short distance, such is not the case farther on. In fact the field strength may even *increase* with distance for a ways. Obviously under some such conditions a moderate increase in transmitter power may result in a considerable increase in range.

SITING RULES OF THUMB

Expanding upon or adding to the suggestions, comments, and considerations concerning v-h-f/u-h-f siting mentioned up to this point, the following "rules of thumb" should prove quite helpful. They apply to both transmitting antennas and receiving antennas.

Choosing the best sight within a given area is a fairly simple problem when omni-directional general coverage is desired. The crest or summit of the highest hill or peak usually will provide the answer, barring freak conditions of terrain.

When the problem is to pick the best site within a given area for transmission to or reception from a given distant point, the considerations are greatly altered. For instance, the brow of a hill may be as good as or better than the summit or crest (assuming, of course, the side of the hill towards the other station). If a bidirectional or omnidirectional antenna is employed, the brow quite often will be the better; if a unidirectional array is used, little difference should be noted, assuming a profile of the general type shown in figure 1-18. If the hill has a rather sheer slope in the direction of the other station but has a very extensive and gradually rounded crest or summit, then the brow usually will provide better point-to-point service regardless of whether or not a unidirectional array is employed.

When the path is long or the power very low, much greater care in siting is required than for short distances and moderate or high power.

For a proposed fixed station, it is highly desirable to compare the most promising sites by actual test, using portable equipment.

Especially avoid low areas where the ground starts rising gradually and smoothly for considerable distance starting right at the antenna location.

If mast height limitations make it impossible to get the antenna up fairly well above surrounding trees or buildings, it is much better to take a site in an open area a few hundred yards back of the woods or buildings, even though it increases the path length slightly.

In "flat" country, the antenna should be elevated well above both nearby *and* distant objects such as trees, buildings and low hills along the transmission path whenever possible.

When checking a terrain map, do not

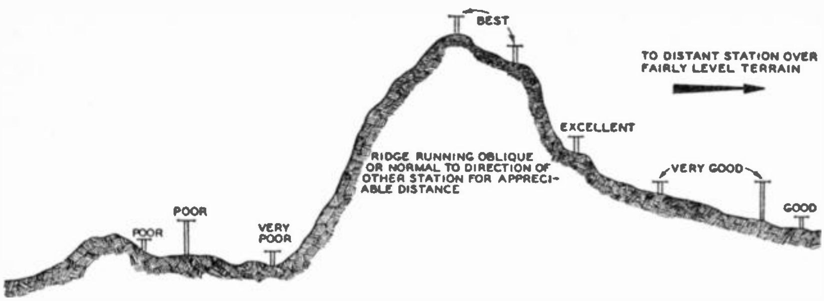


Figure 1-18.

GRAPHICAL ILLUSTRATION OF THE RELATIVE MERITS OF DIFFERENT TYPES OF V-H-F/U-H-F SITE FOR POINT-TO-POINT WORK.

overlook opportunities or possibilities which will permit the wave to go *around* or *through* rather than *over*.

When highly directional antennas are employed in regions of very irregular terrain, keep in mind that the wave may be taking a devious route; do not orient the antenna entirely by compass, without checking to see which orientation gives the best signal strength.

When using frequencies in the lower portion of the v-h-f band for inter-island or other communication over a salt water path from a location near the edge of the water, increased height is of little benefit with vertical polarization unless a site having an elevation of several hundred feet is available. If a very high site is not available, use vertical polarization, use the lowest possible frequency, and place the antenna as near the edge of the water as practicable, without regard to elevation. With horizontal polarization, maximum possible height is desirable regardless of frequency, for salt water paths as well as over land.

HORIZONTAL VS. VERTICAL POLARIZATION AT V-H-F/U-H-F

At *extreme ranges*, regardless of whether the receiving antenna is within or below (beyond) line of sight, and regardless of whether propagation is via ground wave, tropospheric wave, or both, the field strength for a given radiated power and antenna directivity will be approximately the same for either horizontal or vertical polarization (except for the special case of propagation of 30 to approximately 70 Mc. waves over sea water). Experimental data gives a slight edge to horizontal polarization, but the difference is insignificant.

Over an optical path, when no large objects with flat, oblique surfaces are close to either end of the circuit, it will be found that the use of opposite (mutually perpendicular) polarization at the two ends of the circuit will cause a terrific loss in signal strength.

When an appreciable proportion of the signal is reaching the receiving antenna via diffraction, or by reflections from large, flat, obliquely-oriented surfaces or from highly irregular surfaces, the necessity for matching polarization usually is not so important. In most of such cases, little or no difference in signal strength will be noted between matched and unmatched polarization. The same applies in most cases to *very long range* tropospheric propagation; usually there is little necessity for matching the transmitter polarization at the receiver.

These characteristics often can be exploited profitably to minimize interference from a nearby station when difficulty is had in receiving a weak, very distant station on an adjacent frequency because of interference from the nearby station. By using a receiving antenna of polarization opposite to that of the interfering station, oftentimes the interference will be greatly reduced without greatly affecting the signal strength of the distant station, regardless of the polarization of the latter.

A simple horizontal dipole has two "nulls" in two opposite compass directions, while a simple vertical radiator is uniformly effective in all compass directions. In a certain application, one or the other of these characteristics may be of advantage. Of course, it is possible to have an omnidirectional pattern with horizontal polarization or a figure 8 pattern with vertical polarization, but it requires two or more dipoles, or some configuration of comparable complexity.

Distortion in wide band services (such as television) due to interference patterns caused by reflections from buildings, etc., is reduced by the use of horizontal polarization, for reasons previously discussed. Also, because most man-made interference originates near street level, and because it is more effectively propagated to a low receiving antenna under such circumstances by vertical polarization, the use of horizontal polarization improves signal-to-noise ratio in metropolitan areas. It is for these reasons that horizontal polarization has been adopted as standard for FM broadcasting and for television. In the case of police and similar services where a simple, vehicular antenna having omnidirectional characteristics is required, vertical polarization is widely employed.

Except at very low (near grazing) and very high (near vertical) angles, the magnitude of the ground reflection coefficient is much lower with vertical polarization than with horizontal polarization. This means that with vertical polarization the field strength amplitude does not go through such violent oscillations as the angle of elevation is increased. This is a desirable characteristic in the case of communication with aircraft, because the pronounced "nulls" which exist with horizontal polarization over farily regular terrain greatly restrict the reliable range. Also, in the case of "indirect" transmission or reception over a high nearby hill or cliff, the use of vertical polarization precludes the possibility of an almost complete "null" occuring at the angular elevation of the skyline in some direction.

CHAPTER TWO

Transmission Lines

RADIO FREQUENCY TRANSMISSION LINES

While transmission lines are widely used in ordinary power and telephone work, and while basic transmission line theory is valid regardless of application or frequency, the considerations involved in the case of *radio-frequency* power transmission are in many respects more exacting and in some respects unique. Therefore, transmission lines will be treated in this chapter only from the standpoint of radio-frequency use. If a statement is made to the effect that such and such is or is not important, the implied reference is to *radio*-frequency applications, usually above 100 Kc. or so. While a certain *effect* will exist for all frequencies from d.c. (zero) to infinity, the *magnitude* may be considerably different at, say, 60 cycles or 1000 cycles from what it is at 100 Kc. or 100 Mc., and standard engineering practice may be considerably different.

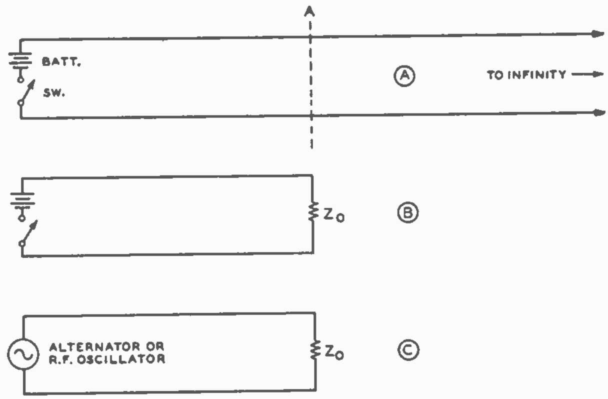
The most familiar use of radio-frequency transmission lines is to transfer energy from "here" to "there," usually from an antenna or array to a receiver, or from a transmitter to an antenna or a directional array. However, they also are used to perform the functions of wave filters, tank circuits, phase inverters, and so on. In the v-h-f and u-h-f band, sections of transmission line are widely used as resonant tank circuits, which function they are able to serve most ably for reasons which later will be explained. In such cases, the element probably is more correctly termed a "linear circuit element," as contrasted to a "lumped" element such as a coil or capacitor, even though physically and electrically it qualifies as a section of transmission line. In many cases a section of line is called upon to serve a dual function. In a directional array, for instance, a section of transmission line may feed power to a radiating element and at the same time act as a phasing section or as an impedence transformer. However, it should be kept in mind that regardless of the application, anything that qualifies as a section of transmission line obeys certain fixed laws and can be expected to possess certain basic characteristics regardless of how the section of line is employed or what it is called.

THE DISSIPATIONLESS INFINITE LINE

Before taking up practical lines, it is well first to have an understanding of the basic idealized transmission line: a line uniform throughout its length, having zero radiation, zero resistance and zero dielectric loss. In case the reader might feel that the behavior of such a hypothetical transmission line may be of only academic interest, it should be pointed out that actual radio-frequency transmission lines can be constructed having losses so low that their performance in many cases may be considered for all practical purposes the same as that of an absolutely lossless line. Furthermore, knowledge of behavior of an idealized lossless line facilitates determination of the conditions under which an actual line having appreciable losses operates, because usually the dissipation can be taken care of in simple correction factors.

Figure 2-1.

A lossless, uniform, infinite line, or a finite section of such line terminated in a non-reactive impedance equal to the surge impedance of the line, will present to the battery or alternator a non-reactive load equal to the surge impedance of the line. For an explanation refer to the accompanying text.



While transmission lines have various configurations, such as parallel wire, coaxial (concentric), symmetrical (balanced), asymmetrical (unbalanced), etc., their *general* behavior is the same regardless. Let us consider for the moment an idealized two-wire "open" line consisting of two parallel conductors, having zero resistance, stretching for an infinite distance through a vacuum or empty space, and having uniform cross section and spacing throughout their infinite length. Let us also assume for the moment that the line will not radiate. In actual practice such an open line always radiates slightly, though in practical open wire lines the radiation usually is so low as to be negligible except at frequencies so high that the line spacing becomes a significant fraction of a wavelength. The factor of radiation is best ignored for the moment; we can take up the question of why a line radiates or doesn't at a later and more appropriate point.

D-C OR "SURGE" CONDITIONS

Let us connect across the one end of our infinite line a switch and battery in series, as in figure 2-1A. Let the battery be a very special and hypothetical one having zero internal resistance, in order to simplify the discussion.

When the switch is closed, the line *progressively* becomes charged up along its length to the potential of the battery. It should be noted that the line does *not* become instantaneously and simultaneously charged throughout its length. The dividing point or boundary between the charged and uncharged portions of the line *starts* the instant the switch is closed and travels down the line at the characteristic velocity of space (about 186,000 miles per second) in the case of conductors separated by empty space, or somewhat less in the case of dielectric insulation. The boundary represents the point at which, at any given instant, a displacement current is flowing in the space between the two conductors. To supply this displacement current, a conduction current flows all along the line from this point back to the battery. The result is a "leading edge" of voltage and current which travels down the line towards infinity at the speed of light. In back of the leading edge the voltage and current are uniform all the way to the battery, while ahead of the leading edge the voltage and current are zero.

It is apparent that to charge up, say, a 186,000 mile section of this infinite line in one second requires a finite and specific amount of current out of the battery. Because the line characteristics do not change with applied voltage or current flow, it is obvious that the current flow required to charge the line is proportional to the applied voltage. It is also apparent, though perhaps less so, that the current flow starts *instantaneously* when the switch is closed and remains *constant* as long as the switch is closed, because displacement current flows only at the "leading edge," and the latter has constant characteristics except for its position on the line. This means that, to the battery, the lossless, infinite, uniform line looks like an impedance with zero net reactance, or like a *resistance*. It is true that the line has inductance and capacitance uniformly distributed throughout its length, but the two always are such as to cancel each other. So long as they are uniformly distributed, rather than lumped, the only effect on the generator of changing the ratio of inductance to capacitance per unit length is to change the amount of current flow for a given voltage. Thus, increasing the capacitance and/or lowering the inductance per unit length, as would occur if larger conductors or closer spacing or a higher dielectric constant insulation were employed, simply increases the amount of current flow for a given applied voltage; under no conditions will it introduce a net reactance.

By now it should be pretty obvious that our lossless, infinite, uniform line looks to the battery exactly like a pure resistance, and that the magnitude of the equivalent resistance depends upon the ratio of inductance to capacity per unit length, which in turn depends upon the diameter and spacing of the conductors and upon the dielectric constant of the medium separating the conductors. The magnitude of the equivalent resistance is called the *characteristic impedance* or *surge impedance* of the line, and is abbreviated Z_0. Practical transmission lines ordinarily have a surge impedance of from 20 to 800 ohms, or most commonly, from about 50 to 650 ohms. A chart showing the surge impedance as a function of conductor dimensions and spacing is shown for the two most common configurations of line later in the chapter (figures 2-15 and 2-26) in connection with a discussion of characteristics of practical transmission lines.

To get back to our hypothetical line of figure 2-1, it should be noted that if an infinite line of 100 ohms surge impedance is removed from the battery and a 100-ohm resistor substituted, exactly the same amount of current will flow through the battery. Also note that for any particular line, there is only one value of resistor which can be substituted without the battery's "knowing the difference."

If the line is broken at point A, the portion of the line to the right will look like a resistance equal to the surge impedance of the line, because it still is infinite in length. (Something finite subtracted from infinity still leaves infinity.) Therefore, if we break the line a A (or any other point) and substitute a resistor having a resistance equal to the surge impedance of the line, as in figure 2-1B, the remaining portion of the line "won't know the difference," and will behave exactly the same as though the line were infinite in length. Thus, we might say that when the switch is closed, the battery can't tell whether (1) it is working into a lossless, infinite, uniform line, or (2) the line has finite length and is terminated in its characteristic or surge impedance, or (3) a resistor equal to Z_0 is connected directly across the battery and switch terminals; the surge of current *is the same in every case.*

If an alternating current generator (alternator) having zero internal impedance is substituted for the battery and switch as illustrated in figure 2-1C, the line still looks like a pure resistance equal to Z_0, regardless of frequency, and regardless of whether the line is infinite in length or is of finite length and terminated in Z_0. This is readily understandable when it is recalled that when the switch was closed in figures 2-1A and 2-1B, the current flow out of the battery was instantaneous and remained constant, indicating a net reactance of zero.

CONDITIONS WITH SINE WAVE GENERATOR

With the alternator, however, conditions are somewhat different as regards voltage and current distribution along the line. At any given *instant* the current and voltage no longer are uniform from the "leading edge" or from the terminating resistance back to the source; instead they vary sinusoidally along the line. Because of the absence of any net reactance, the current and voltage are in phase along the line; at any instant a point of zero current on the line will coincide with a voltage node and vice versa, and at any point on the line the ratio of voltage to current is equal to the characteristic impedance. This is another way of saying the line has *unity power factor*. Likewise at any given *point* on the line, the voltage and current amplitudes vary sinusoidally with time, and are in phase. To summarize, the voltage and current vary sinusodially with distance along the line at at any given instant, with time at any given point on the line, have a constant ratio equal to Z_0, and always are in phase.

If the output of the generator is not sinusoidal, the line still looks to the generator like a pure resistance, because the behavior of the lossless line is not sensitive to frequency. The wave form at the generator terminals will appear at any given point on the line at a later time determined by the velocity of the wave and the distance from the generator. Thus we see that a transmission line can be used as a phase shifter, to produce any desired delay (assuming in a practical case that the required length of line is not prohibitive).

If the two conductors are separated by a lossless dielectric other than empty space. then both the characteristic impedance and the velocity of propagation are reduced, the reduction becoming greater as the dielectric constant is increased. However, the net reactance still is zero, so long as the dielectric is lossless and uniformly distributed.

Before proceeding further, it is desirable that one be familiar with certain transmission line terminology pertaining to the two ends of a terminated line. The *generator end* is variously known as the *sending end, input end,* or *source;* and the *load end* is known also as the *receiving end, output end,* or *sink.* When a line is terminated in pure resistance equal to Z_0, it is said to be *matched.* The speed at which a wave travels on the line is called the *velocity of propagation* (abbreviated V), and the ratio of this velocity to that of light is called the *velocity constant* or *velocity factor* of the line.*

*In a uniform line in which the losses are low enough to be insignificant, the characteristic impedance, the velocity constant, and the wave length are all inversely proportional to the dielectric constant of the insulation or medium separating the conductors.

THE GUIDED FIELD CONCEPT

It should be apparent to the reader by now that there is a close resemblance between the propagation of a wave down an idealized, lossless transmission line, and the propagation of a wave through space as discussed in the preceeding chapter.

For instance, the relationship between frequency, wave velocity and wavelength applies to propagation along a transmission line just as patently as it does to a wave being propagated through empty space or a dielectric. In fact, the current in and voltage between the conductors of a transmission line produce magnetic and electric fields in the vicinity of the conductors which store or contain the energy which has left the generating source but has not yet arrived at the load end. In one sense, it is more accurate to say that an r-f transmission line guides and at the same time confines or restricts an electromagnetic wave containing energy than it is to say that the line actually conducts the energy in the form of a current. However, unless the reader intends to pursue the subject considerably beyond the scope of this book, it will suffice and is somewhat simpler to consider that the energy flows down the conductors to the load in the form of a conduction current. But while the current concept is permissable and adequate for an understanding of what follows in this book, the electromagnetic field concept will be discussed briefly for the benefit of the curious, the rigorous, and those who expect eventually to concern themselves with wave guides.

In simple terms a wave guide in practical form is essentially a hollow conductor used to prevent spreading or dispersion of radiated r-f energy, thus permitting r-f energy to be transmitted from one point to another with very low losses by the simple expedient of squirting the "radiation" into one end and extracting it at the other. Because the minimum permissable cross-sectional dimensions of a wave guide are limited in terms of wavelength from an electrical standpoint, and the maximum are limited in terms of absolute dimensions from a practical or economical standpoint, wave

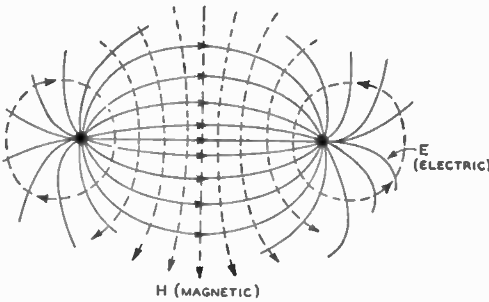


Figure 2-2.

END-ON OR CROSS-SECTIONAL VIEW OF TWO-WIRE OPEN TRANSMISSION LINE CARRYING ENERGY TOWARD THE READER.

This is an instantaneous representation of the electromagnetic field surrounding the line at a particular point on the line at a particular instant. It is assumed that the line is infinite in length or is terminated in Z_0. A half cycle of time earlier or later the direction of BOTH fields of force would be reversed. Reversing ONE field would send the wave in the opposite direction, away from the reader.

guides seldom are used below about 1000 Mc.

Obviously the concept of power transmission by means of current in "go" and "return" circuits does not fit the case of the wave guide. Therefore if we are after a single concept which can be applied to all types of r-f or a-c power transmission, the concept of *guided* travelling electromagnetic lines of force is the answer.

Referring to the end-on view of our two wire line shown in figure 2-2, the current flowing in the wires will produce magnetic lines of force as represented by the dashed flux lines, while the voltage difference between the wires will produce electrostatic lines of force (solid lines). It is impossible to have current and voltage at a point on the line without the existence of a corresponding electromagnetic field, and vice versa. They are so interrelated that it really is immaterial whether at a point somewhere along the transmission line we say that the r-f current in and voltage on the wires are due to the electromagnetic field, or vice versa, or that they simply are two manifestations of the same phenomenon.

Even if we stick to the current concept and consider the electromagnetic field simply as a by-product of the voltage on and current in (or on) the conductors of the transmission line, it should be kept in mind that wherever current is present on a line, there always is an associated and corresponding magnetic field, and wherever there is voltage on a line, there is an associated and corresponding electrostatic field.

WAVE MOTION ON A LINE

If we should connect an infinite line (or a line that is terminated in Z_0) to a source of sine wave r-f voltage, the conditions existing at a particular *instant* could be represented by figure 2-3, which is a graphical picture of a "frozen" wave. In this case the horizontal axis represents distance along the line. One cycle later, or an integral number of complete cycles later, the "frozen" wave would look exactly the same. At fractions of a cycle later, the wave train and therefore the whole picture would be moved a corresponding fraction of a wavelength to the right, in order to depict the conditions on the line at that instant. To get a *moving* or true picture of what actually happens, it would be necessary to run the picture past the observer at a velocity which caused F cycles to pass the observer per second. At any frequency over a few cycles per second the result would be a blur; hence the necessity for showing a "frozen" wave.

If the horizontal axis is used to represent *time* instead of *distance along the line*, the same graphical representation then shows the variation in voltage and current with time at a fixed point on the line. This is the familiar representation of ordinary 60 cycle a.c. wherein voltage is plotted against time in Cartesian rather than polar coordinates. It can just as well apply to 20 kilocycles or 20 megacycles as to 60 cycles.

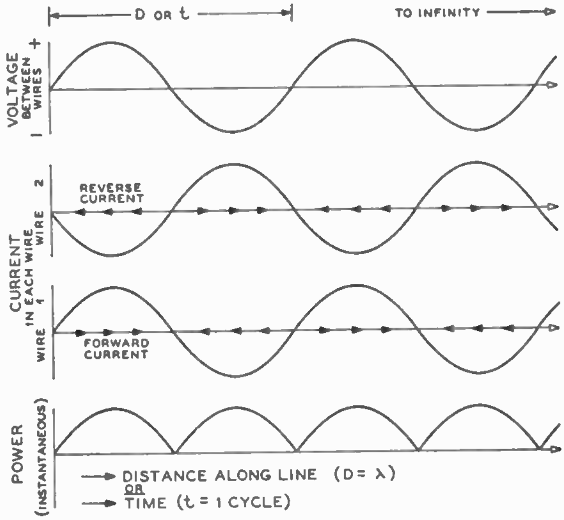
REFLECTION AND STANDING WAVES

In Chapter 1 under *Reflection* it was stated that when a radiated wave strikes an abrupt change in medium or a sharp boundary, some of the wave is reflected, and that all of it is reflected in the limiting case

Figure 2-3.

VOLTAGE, CURRENT, AND POWER RELATIONS IN A LOSSLESS, UNIFORM, INFINITE LINE OF TWO-WIRE, OPEN CONFIGURATION.

Distance along the horizontal axis can represent distance along the transmission line in the case of a "frozen" travelling wave, or it can represent time in the case of a fixed point on the line; in either case the representation is valid. It is also for a finite line terminated in Z_0, its characteristic or surge impedance.



of a conducting sheet consisting of an infinite plane of zero resistance. It was also explained how reflection causes *standing waves* in the medium. The same considerations apply to waves propagated along transmission lines. When they encounter a discontinuity in the line such as a variation in surge impedance, a portion of the energy is reflected. When the discontinuity is the limiting or extreme case of an open circuit or a dead short, *all* of the energy is reflected. In any case, standing waves will be present if there is reflection; the variation of r-m-s voltage or current amplitude with distance along the line will be a function of the amount of reflection or *reflection coefficient.*

Because there is no attenuation of the wave on a line due to spreading (as there is in the case of radiation and propagation through space), a lossless line terminated in an open circuit or short circuit or pure reactance will produce a reflected wave having the same amplitude as the incident wave. This is true because such terminations absorb no power. In such a case the two waves will combine to produce a standing wave of infinite *standing wave ratio;* in other words, the voltage along the line will vary from zero to infinity, and the current will do likewise but at intermediate points (for reasons explained under *Standing Waves* in Chapter 1).

A "standing wave" should not be confused with a "frozen wave" of the type previously discussed; a standing wave is simply an interference effect which causes the voltage or current as read on an r-m-s meter to vary cyclically with distance along the line, while a "frozen" wave is simply an instantaneous picture of a traveling wave.

Comprehension of the mechanism whereby standing waves are produced is prerequisite to a clear understanding of other transmission line phenomena; therefore some deliberate reiteration and recapitulation pertaining to reflection and the generation of standing waves immediately follows, in order to make absolutely certain that the reader is able to visualize what transpires.

GENERATION OF A STANDING WAVE

Perhaps the simplest termination for a transmission line is just to have it stop—in other words, an open circuit. Let us see what happens when a travelling wave encounters such a termination.

When energy consisting of in-phase voltage and current waves approaches the "end of the line," the abrupt ending is not anticipated by the travelling electromagnetic wave, and it acts up to that point just as though the line were infinite in length. But upon reaching the open end, the current must disappear, because a current can't exist unless it keeps moving, and at the very end of the line it is stopped dead in its tracks. This causes the magnetic field at that point to collapse, but in so doing it generates an e.m.f across the open end of the line which adds *in phase* to the e.m.f of the arriving or *incident* wave, causing the e.m.f. across the extreme end of the line to double.

When the wave arrived at the end, the energy represented by the incident in-phase current and voltage waves existed half as an electric field and half as a magnetic field. But upon reaching the end, the magnetic field of necessity collapsed, and in so doing converted itself into an electrostatic field which adds to the already existing electrostatic field. This reinforced electrostatic field, represented by a doubled e.m.f. across the extreme end of the line, is maintained by the continuing collapse of the incident magnetic field, and serves in effect as a generator which launches a new electromagnetic wave in the only direction it can travel: the reverse direction. Because the e.m.f. across the end of the line (or electric field) is not reversed in phase, it is obvious that in order to have a wave travelling in the opposite direction, or a *reflected* wave, the current (or magnetic field) *must* be reversed in phase. This condition is satisfied by the absence of current at the extreme end of the line; the current component of the reflected wave is equal in amplitude and opposite in phase to that of the incident wave, giving a resultant of zero. This is true at the reflection point at all portions of the cycle. But as we move away from the reflection point the reflected wave has travelled increasingly farther than the incident wave, and the finite time taken for the additional travel represents a phase shift.

At a point a quarter wavelength from the end of the line, the reflected wave always has travelled a half wavelength farther than the arriving wave. This means that at this point the reflected current, which is reversed in phase (shifted 180 degrees) upon reflection, is *in* phase with the arriving current, because the 180 degree phase angle due to difference in path length plus the 180 degree shift due to inversion at reflection adds up to 360 degrees, which is the same as 0 degrees or no phase difference. This condition exists at all times, and not just at a certain point on or portion of a cycle. Therefore, the resultant current at this point, while oscillating in amplitude, has a higher r-m-s value than at points a slight distance to either side where the phase angle is not zero.

At a point a half wavelength from the end of the line, the reflected wave has travelled a full wavelength, or 360 degrees, farther than the arriving wave. This is the same as 0 degrees. Therefore the effective

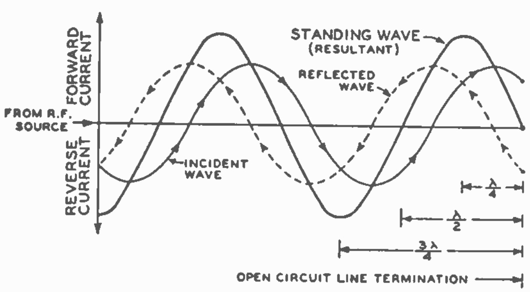


Figure 2-4.

ILLUSTRATING DEVELOPMENT OF A STANDING WAVE.

The example illustrated is a standing wave of current on one wire of a lossless, two-wire open line terminated in an open circuit. The other wire will have an identical standing wave except for reverse polarity. This is an instantaneous picture; at an earlier or later time the standing wave will have different amplitude but will OCCUPY THE SAME LOCATION ON THE LINE. It oscillates in amplitude at a sine rate but does not travel laterally like the constant-amplitude incident and reflected waves. This same standing wave pattern also applies to the VOLTAGE wave on a line with a SHORT CIRCUIT termination.

phase difference is due entirely to the 180 degree shift upon reflection, which puts the incident and reflected currents in *phase opposition* at this point. So the resultant current at this point is zero at all times, just as it is at the extreme end of the line.

Moving back to a point ¾ wavelength from the end simply adds 360 degrees to the condition existing at ¼ wavelength from the end, because again the additional 360 degrees of travel (180 degrees each way) has no effect upon the phase difference, or the resultant. Hence, at a point ¾ wavelength from the end, the resultant current is maximum again at all portions of the cycle.

Without carrying things further it should be obvious that these two conditions will repeat alternately at all higher multiples of a quarter wavelength from the termination. Thus we see that at zero and all even multiples of quarter wavelengths (0, 2, 4, etc.) from the open circuit termination, the resultant r-m-s current is zero, while at all odd numbers of quarter wavelengths, (1, 3, 5, etc.) from the termination the resultant r-m-s current is maximum.

Now it is known from elementary alternating current theory that two sine waves of the same frequency always add to produce either zero (a special case) or a resultant sine wave, and the same still applies to two sine waves of the same frequency travelling in opposite directions. Therefore the resultant or "standing wave" also will be a sine wave. Inasmuch as we have the zero and the crest (double current) points, all that is necessary is to construct a sine wave which intersects or includes these points. Or, more laboriously, the instantaneous magnitude of the incident and reflected waves can be added algebraically at many close, successive points for several cycles along the line. The result will be the same: a sine wave which takes in the already determined zero and crest points on the horizontal (distance) axis.

GRAPHICAL REPRESENTATION OF A FROZEN WAVE AND OF R-M-S AMPLITUDE

Figure 2-4 shows how a reflected current wave combines with the incident current wave to produce a resultant as explained. This is an instantaneous or "frozen" picture. If a similar instantaneous representation were constructed for every few degrees throughout a cycle, it would be observed that while the *amplitude* of the sine resultant varies in a sine fashion, it never moves laterally. It oscillates in amplitude, but is fixed in space. That is why it is called a *standing wave.*

While useful in determining or proving the characteristics or configuration of a standing wave, such a series of graphical representations is rather ineffective in assisting the reader to visualize the actual, continuous formation and maintenance of a standing wave. Unfortunately the simultaneous existence on a transmission line of an incident travelling wave sweeping to the right, a reflected travelling wave sweeping to the left, an algebraic sum or resultant which "stands still" yet oscillates in magnitude is best illustrated in actual motion, with the aid of moving picture film.

The amplitude oscillation of a standing wave on a two-wire line terminated in an open circuit is shown in figure 2-5. The instantaneous amplitude and polarity of the current are shown for every 30 degrees of time over a complete cycle for both wires. The bottom representation shows the r-m-s current at different points along the line as read on a thermocouple or other r-f ammeter. As such a meter has no "sense" as regards polarity, and because a standing wave has only *instantaneous* polarity, it is permissible to show all current readings or r-m-s values *above* the base line. When such a curve is shown, it is implied without further notation that it is a curve of r-m-s readings as might be observed on a meter, and not a curve of instantaneous values.

If the standing wave of *voltage* is plotted for a line with open circuit termination, it will be found that the lack of phase reversal upon reflection causes the resultant or standing wave of voltage to be displaced along the line by 90 degrees or a quarter wavelength from the current curve. The voltage peaks will coincide with current nodes, and vice versa.

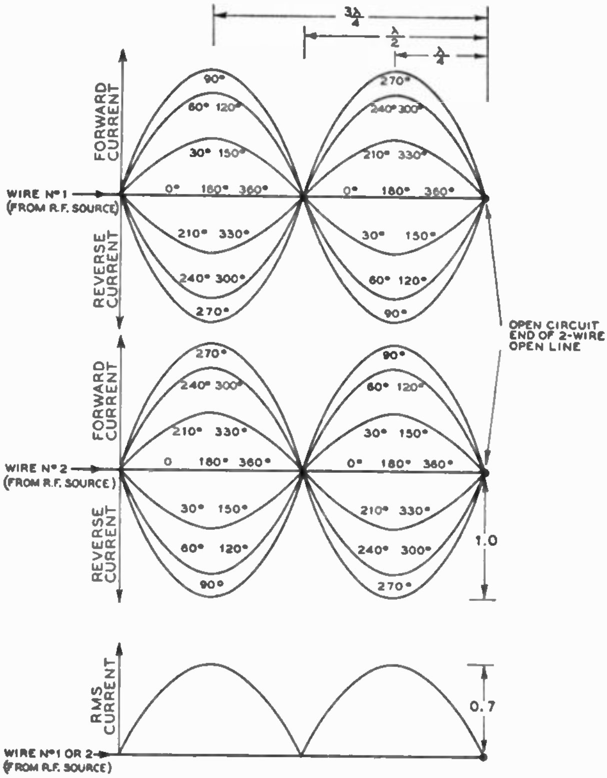


Figure 2-5.

INSTANTANEOUS VALUES AND POLARITY OF STANDING WAVE ON TWO-WIRE LINE (ABOVE), AND R-M-S VALUES AS READ ON METER (BELOW).

Instantaneous amplitude and polarity of the resultant current is shown for every 30 degrees of time over a complete cycle for both wires (top), while below is shown the conventional representation of the r-m-s value as read on an r-f ammeter in either wire. Note that the r-m-s value is 0.707 of the maximum instantaneous peak amplitude.

EFFECT OF LINE TERMINATION ON STANDING WAVE PATTERN

If a line is terminated in a dead short, instead of an open circuit, the *voltage* "collapses" (because a potential difference cannot exist across zero resistance) and the current is doubled, the current and voltage roles being just reversed from the case of the open circuit termination. The *voltage* now undergoes the phase reversal upon reflection, and not the current. The result is that the *voltage peaks* of the standing wave which is formed by combination of the incident and reflected waves occur at the same distance from the termination as the *current peaks* did in the case of an open circuit termination, and the *current peaks* occur at the same distance from the termination as the *voltage peaks* did in the case of an open circuit termination. Thus, the standing wave shown in figure 2-4 not only holds for current in the case of an open circuit termination, but also for voltage in the case of a short circuit termination.

If the termination is a pure reactance, the standing wave pattern will be somewhere intermediate to that of a short circuit termination and an open circuit termination. The voltage and current nodes in the standing wave pattern still will be 90 degrees apart, but the standing wave will contain *both* voltage and current at the termination, instead of only one. If a very small capacitance or a very large inductance is used for a termination, the standing wave will resemble or approach that obtained with an open circuit termination, except that a small amount of resultant current will flow at the termination. If a very large capacitance or a very small inductance is used for a termination, the standing wave pattern will resemble or approach that obtained with a short circuit termination, except that a small amount of resultant voltage will appear at the termination.

If we should start with a reactance having a value approaching infinity (open circuit conditions) and lower it gradually (increase C or decrease L), we would find that in the case of capacitive reactance the whole standing wave pattern is effectively "slid along" *towards* the termination in the case of *capacitive* reactance and *away* from the termination in the case of *inductive* reactance, the maximum shift approaching but never reaching 90 degrees so long as the reactance still has a finite value.

A more convenient way to remember this relationship is that *decreasing either L or C moves the pattern away from the termination.* This is explained by the fact that decreasing the value of L lowers the reactance and makes the termination approach a short circuit, while decreasing the value of C increases the reactance and makes the termination approach an open circuit.

EFFECT OF GENERATOR ON STANDING WAVES AND VICE VERSA

In order to facilitate the explanation of reflection and standing waves, certain simplifying assumptions were made. They do not affect the basic mechanism of the generation of standing waves, but they do enter into the picture when we start considering practical lines.

Most important, we have assumed in the foregoing explanations that there was only one incident wave and one reflected wave: in other words, a single reflection point. But unless the generator is so far away as to be just slightly this side of infinity, some of the returning energy or reflected wave will "bounce off" the generator and start down the line again as an incident wave unless the generator happens to have an internal impedance which is non-reactive and exactly equal to the surge impedance of the line. In any case the returning waves (those reflected from the line termination) will have considerable effect upon the apparent load on the generator, the exact effect depending upon the relative magnitude and phase of the arriving reflected wave as compared to the voltage delivered by the generator. The net result is that a line which is not terminated in Z_0 looks to the generator like a pure resistance of Z_0 only for the length of time it takes for the "leading edge" of the first surge down the line to arrive back at the generator. Then it can look like most anything, as we shall soon see: a reactance, a resistance equal to some-

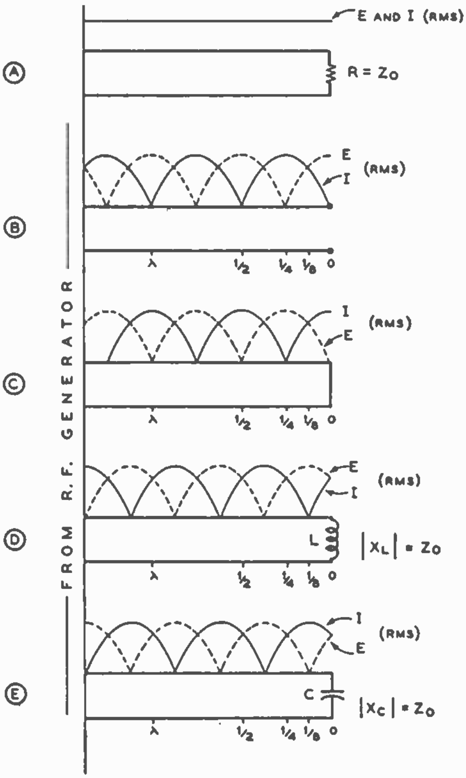


Figure 2-6.

EFFECT OF LINE TERMINATION ON STANDING WAVE PATTERN.

In the case of a lossless line and lossless termination, the standing wave pattern is the same in all cases except for displacement along the line with respect to the termination. The reactances in (D) and (E) are equal in ohms to the surge impedance. Decreasing the value of either L or C moves the pattern away from the termination, to approach a short circuit in the case of L and an open circuit in the case of C.

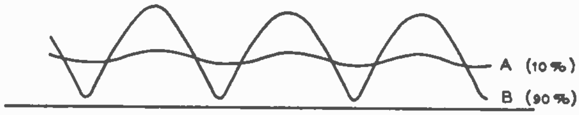


Figure 2-7.

ILLUSTRATING EFFECT OF REFLECTION COEFFICIENT UPON STANDING WAVE RATIO (SWR) AND SHAPE

As the magnitude of the reflection coefficient is increased from zero to 1.0, the standing wave increases in amplitude and the shape gradually changes from a sine wave at zero amplitude to a "full wave rectified" wave form at maximum amplitude. The two curves shown are for a reflection coefficient of 10 per cent (A) and 90 per cent (B). A curve showing reflection coefficient vs. SWR is given in figure 2-8.

thing other than Z_0, or a complex impedance, depending upon the nature of the termination and the length of the line in terms of wavelength.

The reader need not be unduly concerned for the moment with the reference to multiple reflections on a line as a result of reflection at the generator end. The important end is the *receiving end* in so far as reflections are concerned, because all the incident waves form a resultant incident wave and all the reflected waves form a resultant reflected wave, and the conditions at the point of the first bounce, or receiving termination, determine the points along the line at which nodes, etc. will occur. For instance, if there is cancellation of incident and reflected voltage waves at a point a quarter wavelength from the receiving end of a long line, the same condition will hold true as the length of the line is varied. However, the phase relationships existing at a point on a long line a quarter wavelength from the *generator* will vary cyclically as the length of the line is increased or shortened. It is for this reason that *the receiving end always is used as the reference point.*

UNMATCHED TERMINATION WITH RESISTIVE COMPONENT

In illustrating the principles of reflection and the generation of standing waves, the discussion was confined to lossless terminations, in order to simplify the explanation. When a transmission line is used for the purpose of feeding energy to a practical load circuit, the termination no longer is lossless. It either is purely resistive or has a resistive component. And if the load is not exactly equal to Z_0, there will be reflections, and when there are reflections there are standing waves. Such an "unmatched" or "mismatched" load may consist of a resistance not equal to Z_0 (assuming Z_0 is purely resistive), or of a complex impedence comprising both resistance and reactance.

Regardless of the exact condition of mismatch, as long as the mismatched termination has a resistive component *part of the arriving energy will be absorbed and part will be reflected.* This is different from the conditions previously considered in which reflection always was total, or complete. When the incident wave is neither totally reflected nor totally absorbed, the standing wave of current or voltage no longer will have zero amplitude at recurring points along the line.

STANDING WAVE RATIO (SWR) AND REFLECTION COEFFICIENT

The ratio of maximum r-m-s voltage (or current) to minimum r-m-s voltage (or current) as measured along the line is called the *standing wave ratio,* abbreviated *SWR.* Obviously when the line is non-resonant or "flat" (no reflection), the SWR is 1.0, and when the termination is dissipationless (total reflection), the SWR is infinite. In the case of a mismatched load having a resistive component (partial reflection), the standing wave ratio will have a value somewhere between 1.0 and infinity, the exact value being determined by the magnitude of the *reflection coefficient*, which is simply the ratio of the voltage or current of the reflected wave to that of the incident wave. The greater the reflection coefficient, or percentage of reflection, the greater the SWR. The relationship between the two is illustrated in figure 2-8.

In addition to magnitude, the reflection coefficient has a *phase angle*, indicating the phase difference between the incident and reflected waves at the point of reflection. If the load is purely resistive, or is an open or short circuit, the phase difference is zero or 180 degrees. But in the case of a complex load impedence the picture is different; the phase angle of the reflection coefficient is somewhere betwen 0 and 180 degrees, the exact phase angle depending upon both the magnitude and power factor of the load impedance.

The phase angle of the reflection coefficient simply fixes the *position* of the standing wave pattern with respect to the load. For instance, it is possible to substitute other values of resistance and reactance in a complex termination and maintain the same SWR. However, the power factor of the load will be different, the ratio of load impedance to surge impedance will be different, and the standing wave pattern will be "slid along" the line. Even though the magnitude of the reflection coefficient (and therefore the SWR) remains unchanged, the phase angle of the reflection coefficient has been changed.

The term "SWR" always refers to voltage or current unless power is specifically mentioned. The term "VSWR" (for *voltage* standing wave ratio) oftentimes is used, particularly in tables and charts, to indicate specifically that the voltage ratio and not the power ratio is being given. The VSWR must be squared to get the SWR in terms of power. The SWR in terms of current is of course the same as the VSWR.

SWR also may be expressed as a reciprocal fraction. Thus a VSWR of 5 is the same as a VSWR of 0.2, and a VSWR of 3 is the same as a VSWR of 0.33.

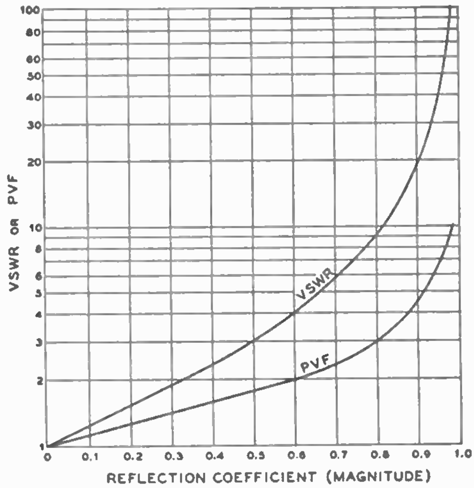


Figure 2-8.

MAGNITUDE OF REFLECTION COEFFICIENT VS. STANDING WAVE RATIO (VSWR), AND PEAK VOLTAGE FACTOR (PVF).

The peak voltage factor is the ratio of maximum voltage on the line to the voltage on a flat line delivering the same power to a load. Unless the line loss is insignificant, these curves are valid only near the load end of the line.

Getting back to figure 2-8, the relationship between reflection coefficient and SWR holds regardless of whether or not the generator is matched to the transmission line. When the generator is mismatched and there actually are multiple reflections between the generator and the line termination, the incident wave is simply the resultant of all the forward travelling waves, and the reflected wave is simply the resultant of all the backward travelling waves.

EFFECTS OF SWR UPON SHAPE OF STANDING WAVE

If the reflection coefficient is say 10 per cent, meaning that 10 per cent of the incident voltage or current wave is reflected, then 10 per cent of the incident wave will combine with it to form a resultant or vector sum in the form of a standing wave similar to B, C, D, or E of figure 2-6. The other

90 per cent of the wave is absorbed in the termination and produces a curve of r-m-s voltage or current vs. distance which is a straight line, such as shown in A of figure 2-6. The resultant or total r-m-s current or voltage at any point along the line is due to the combination of this unidirectional travelling wave and the standing wave, which, when relative phases are taken into consideration, produces an r-m-s curve similar to that of A in figure 2-7.

If the reflection coefficient is high, say 90 per cent, the resultant r-m-s curve or combination of the travelling and standing waves gives a curve similar to that of B in figure 2-7. The interesting thing to note is that when the standing wave ratio or SWR is high, due to a high reflection coefficient, the curve resembles a full-wave rectified sine wave, with sharp nulls,* while in the case of a low SWR, the curve approaches that of an unrectified sine wave, with relatively broad nulls. This is somewhat analogous to the case of linear dection or demodulation of two beating waves; the demodulated signal resembles the output of a full wave rectifier when the two waves are of the same magnitude, but approaches a sine wave as the ratio of the two is increased or decreased from unity. The same condition applies here; as the ratio of the incident to the reflected wave is increased (lower reflection coefficient) the resultant approaches a sine wave.

EFFECT OF LINE DISCONTINUITIES

So far we have in every case assumed a transmission line which was uniform throughout its length. But what happens when it is not?

Whenever there is any *sudden* change in the characteristic impedence of the line, partial reflection will occur at the point of discontinuity. Some of the energy will be transmitted and some reflected, which is essentially the same as having some of the energy absorbed and some reflected in so far as the effect upon the line from the generator to that point is concerned. The discontinuity can be ascribed a reflection coefficient just as in the case of an unmatched load.

In a simple case, such as finite length of uniform line having a Z_0 equal to 500 ohms feeding into an infinite length of uniform line having a Z_0 equal to 100 ohms, the behavior is easily predicted. The infinite 100-ohm line will have no standing waves, will accept the same power from the 500-ohm line as would a 100 ohm resistor, and the rest of the energy will be reflected at the discontinuity to produce standing waves from there back to the generator. However, in the case of a discontinuity which is not straightforward and occurs an odd distance down a line which is terminated in an unmatched, complex impedance, things begin to get somewhat complicated, especially when the discontinuity is neither sudden nor gradual but intermediate between the two.

In any case, when a discontinuity exists somewhere on a line and is not a smooth, gradual change embracing several wavelengths, it is not possible to avoid standing waves throughout the entire length of the line. If the discontinuity is sharp enough and is great enough to be significant, standing waves must exist on one side of the discontinuity if not both.

LINE LOSSES AND THEIR EFFECTS

So far it has been assumed that the transmission line in every case was entirely lossless. This was done because in actual practice good transmission lines have such low losses that in general their behavior is similar to that of a lossless line except for gradual attenuation of the transmitted power, and also because the assumption of a dissipationless line simplified the discussion. Practical lines do have losses, however, and it is desirable to know what effects these losses produce besides attenuation, even though the effects may be small.

Losses in transmission lines are of two kinds: dielectric losses due to an imperfect dielectric, and ohmic losses due to imperfect conductors. The first is a shunt loss, and the latter is a series loss. The shunt leakage,

*The term "null" is used to indicate a minimum which may either be zero or have a finite value.

or conductance, usually is designated as G; the series or ohmic resistance as R.

The dielectric loss per unit length of line varies directly as the frequency, assuming the power factor and dielectric constant remain constant, while the ohmic loss varies as the square root of the frequency. The variation in the relationship of the dielectric loss is readily understandable; the amount of energy dissipated during one cycle will be the same regardless of the duration of the cycle. Therefore the longer the cycle (lower the frequency) the longer it takes to dissipate a given amount of energy, which is simply another way of saying that proportionately less power is dissipated.

The ohmic loss relationship is not so obvious. It might first appear that the ohmic loss should be the same at all frequencies. The explanation is a higher effective resistance as the frequency is raised, due to skin effect.

Because of these relationships, in practical lines the ohmic or conductor losses predominate at the very low frequencies and the dielectric or insulation losses predominate at the very high frequencies. While ordinary air is an almost perfect dielectric, line conductors have to be spaced and supported, and the best available insulation begins to look bad if we go high enough in frequency. Thus it readily can be seen that in very low frequency lines the design must be such as to minimize ohmic losses, while in very high frequency lines especial care must be taken to minimize the dielectric losses. If the ohmic loss is, say, 100 times the dielectric loss, the thing to work on is the ohmic loss, because in such a case reducing the dielectric loss to zero would make but an insignificant reduction on the total losses or *attenuation* of the line.

The relative losses at different orders of frequency for various types and configurations of actual transmission line will be taken up later under a discussion of practical transmission lines. For the moment let us examine the nature and magnitude of the effects of dissipation upon the line.

EFFECT OF LINE LOSSES ON CHARACTERISTIC IMPEDANCE

An infinitely long, lossy line will have a characteristic impedance that contains negligible reactance (or is for all practical purposes purely resistive) provided that either (1) the attenuation per wavelength is small or (2) the ohmic loss is equal to the dielectric loss per unit length and the resistance and conductance do not approach the inductive and capacitive reactances per unit length.

Condition (1) obtains for *well-designed, high quality* practical lines at all radio frequencies except the very lowest. However, in the case of a radio frequency line having high attenuation per wavelength, the characteristic impedance still can be maintained substantially resistive by balancing the ohmic and dielectric losses for that particular frequency. But note that because of the different frequency vs. loss ratio for the two types of losses it is possible to make the losses exactly equal *at only one frequency*.

When the frequency becomes so low (usually in the lower audio frequency range) that the series resistance and/or the shunt conductance per unit length approach the distributed reactances per unit length, then simply making the ohmic and dielectric losses equal will no longer result in a resistive characteristic impedance. However, this is of only academic interest in the case of lines operating above 100 kc.

Thus we see that it is quite possible to have a transmission line with a characteristic impedance which has a reactive component and changes with frequency. But both the ohmic and dielectric losses per wavelength are sufficiently low for practical lines of good design at the radio frequencies with which we are concerned to result in a characteristic impedance which approaches a pure resistance and is substantially constant with frequency. As a matter of interest, however, it should be noted that in the case of a very lossy line with substantially all of the loss in one branch (ohmic or dielectric), the line must be terminated in a complex impedance containing both reactance and resistance in order to avoid reflection and standing waves. It should also be noted that the generateor "sees" a complex impedance when looking into such a line when the line is matched and no standing waves are present.

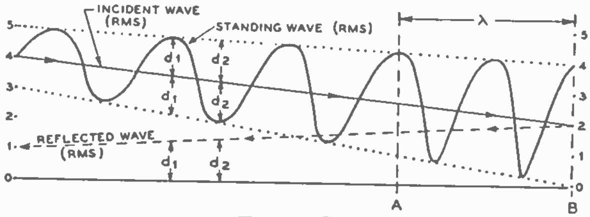


Figure 2-9.

ILLUSTRATING EFFECT OF LINE ATTENUATION UPON STANDING WAVE

The line is assumed to have a "normal" attenuation of approximately 2 db loss per electrical wavelength for a unidirectional travelling wave. Reflection from point B is total; the termination is dissipationless and the reflection coefficient is 1.0. The effect of an unmatched resistive termination such as to give a reflection coefficient of approximately 0.65 (partial reflection) may be had by assuming that the line is terminated at point A instead of point B. Observe that the r-m-s value of the standing wave oscillates about the amplitude of the incident wave with a maximum displacement (d) equal to the amplitude of the reflected wave. The result is a decrease in average r-m-s amplitude and an increase in both the standing wave ratio and absolute r-m-s amplitude variation as the distance from the generator is increased.

EFFECT OF LINE POWER FACTOR ON SWR

The effect of a lossy line upon the standing wave ratio is illustrated in figure 2-9. An arbitrary attenuation of approximately 2 db per electrical wavelength *for a unidirectional travelling wave* is assumed. This attenuation (that occuring when the line is matched, with reflection coefficient equal to zero and SWR equal to 1.0) is called the *normal* attenuation of the line.

The incident wave becomes weaker as it travels toward the load, while the reflected wave becomes weaker as it travels toward the generator. The result is a standing wave which oscillates about the amplitude of the incident wave with a maximum displacement equal to the amplitude of reflected wave. The "peaks" of the standing wave are equal to the amplitude of the incident wave plus that of the reflected wave, while the "nulls" are equal to their difference.

It is apparent from the curves that the effect is to produce both a higher standing wave ratio and a higher absolute r-m-s amplitude variation or displacement as the distance from the generator is increased.. Thus it may be said that when line losses are significant, the effect of increasing the length of the line for any given load is to reduce the standing wave ratio at the generator end of the line.

The relationship between normal attenuation and reduction in SWR is illustrated in figure 2-10. Note that when the normal attenuation of the line is considerable (yet not too high per wavelength), a high SWR at the generator end is impossible even with a reflection coefficient of 1.0 (infinite SWR) at the load end. A lossy line tends to "smooth out" the standing waves at the generator end.

EFFECT OF SWR ON TOTAL LINE DISSIPATION

The efficiency of a transmission line may be expressed as the ratio of the power dissipated in the load to the total power delivered by the generator. Obviously it is important in the transmission of radio frequency power to minimize the percentage of power dissipated in the line itself, especially when (because of the frequency, length of line, or type of line) the losses tend to be appreciable.

The attenuation constant or "normal" attenuation of a line at a certain frequency is expressed in db per electrical wavelength or per absolute unit length, and refers to a *unidirectional* travelling wave: in other words, to the attenuation of the line per unit length when it is infinite or else terminated in its characteristic impedance, and no standing waves exist on the line.

The presence of reflection, as indicated by the existence of standing waves, increases the percentage of total power dissipated in the line and thereby decreases the efficiency of the line. This is explained as follows:

When standing waves exist, reflection is present, and there is both an incident wave component and a reflected wave component, the standing wave simply being a manifestation of the resultant of these two travelling waves. 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. If the reflection coefficient is high, there is considerable energy flowing in both directions along the line, yet only a comparatively small amount of power is dissipated in the load. This increases the ratio of power dissipated in the line as compared to that dissipated in the load and reduces the efficiency of the line. For instance:

If the line has a characteristic impedance of 100 ohms, the termination a reflection coefficient of 0.8, and the r-m-s voltage of the incident wave a value of 100, then the voltage of the reflected wave will be 80, producing a resultant or standing wave of r-m-s values having voltage nodes and loops of 20 and 180 respectively (representing an SWR of 9) near the load end.

The incident current will be 1 amp. and the reflected current 0.8 amp., determined in each case by the voltage and surge impedance. Thus we have 100 watts arriving and 64 watts returning, leaving 36 watts to be dissipated in the load. But the transmission line is carrying 164 watts: 100 in one direction and 64 in the other. So actual line dissipation will be the same as though 164 watts were being dissipated in a perfectly matched load and no reflection and standing waves were present. In effect, the average power factor along the line has been increased considerably.

From figures 2-9 and 2-10 and the previous discussion relating to them we know that when the normal or flat-line attenuation of a line is considerable, a high SWR can exist only near the load end. Therefore,

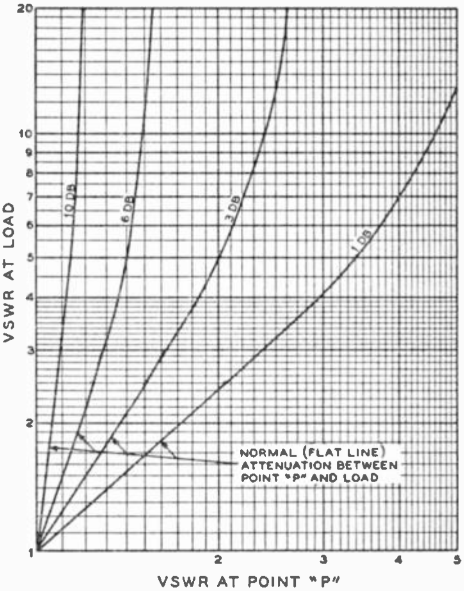


Figure 2-10.

VSWR AT POINT "P" VS. VSWR AT LINE TERMINATION FOR INDICATED AMOUNTS OF "NORMAL" OR FLAT-LINE ATTENUATION BETWEEN POINT "P" AND TERMINATION.

It is assumed that the attenuation per wavelength of line is small. Note that a high SWR at point "P" is impossible when the normal attenuation of the line is high. Even with a reflection coefficient of 1.0 (infinite SWR at the termination) a high SWR cannot exist at point "P" if the normal attenuation amounts to at least several db. When standing waves exist, the actual line loss or total line dissipation will be higher than the "normal" or flat-line attenuation. The increase in total line dissipation with SWR is illustrated in figure 2-11.

figured percentage wise the increase in total line loss due to a mismatched load is not so great as is the increase in a section of line adjacent to the receiving end. Except for this factor, the increase in total line dissipation with degree of load mismatch would be even worse than it is.

The increase in transmission line attenuation with increasing SWR at the load end is illustrated graphically in figure 2-11. An SWR equal to 1.0 indicates the normal or

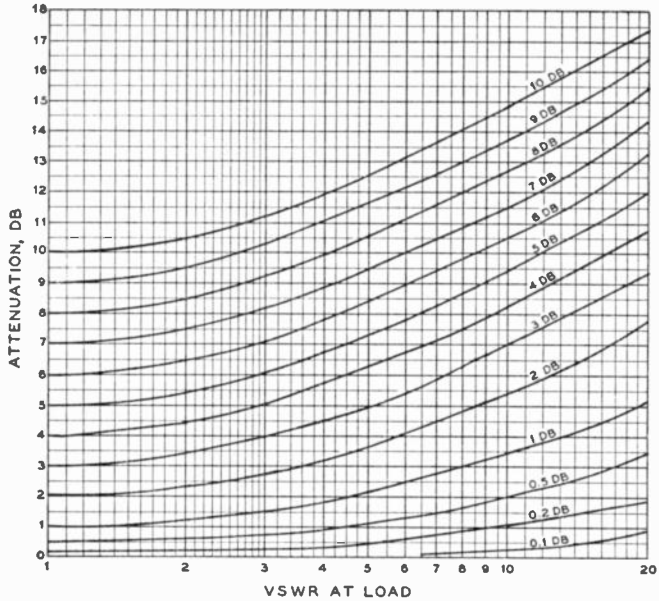


Figure 2-11.

ILLUSTRATING INCREASE IN TRANSMISSION LINE ATTENUATION WITH INCREASING SWR.

It is assumed that the attenuation per wavelength is small. The "normal" or flat-line loss (VSWR equal to 1.0) is the rated loss or attenuation of the line. When this is known, the actual loss present under conditions of an unmatched load may be found by picking the closest curve for a VSWR of 1.0 and following it to the right to the point of intersection with the value of VSWR existing at the load, interpolating between curves where necessary.

flat-line attenuation for the transmission line to which each particular curve applies.

It is apparent from the curves that unless a very good line is employed, one having a low "normal" attenuation, the load must be closely matched to the line in order to avoid a serious loss in power. However, if the normal loss is negligible, then a fairly high SWR can be tolerated from an efficiency standpoint because even though the presence of the moderately high SWR increases the line dissipation several times, it is still low compared to the load dissipation. For instance: if the normal attenuation is such that the line efficiency is 99.5 per cent, then it is possible to increase the line loss 20 times without having the line efficiency drop below about 90 per cent. But if the normal loss is high to begin with, we cannot very well afford to increase it 20 times.

Even when the normal attenuation is low and moderate standing waves can be tolerated from the standpoint of line efficiency, they nevertheless are objectionable in the case of practical lines when the power is sufficiently high. The peak voltage across the line conductors or to ground increases as the square root of the VSWR. (Refer to figure 2-8.) In the case of a line carrying high power, the presence of standing waves might mean that the physical dimensions of the line have to be increased in order to avoid flashover, thus increasing the cost. There is still another reason why a high SWR may be undesirable even though the line efficiency is tolerable. This has to do with the *frequency sensitivity* of the line, which will be taken up later.

Lines which are operated with a load which is exactly matched or even *nearly* matched to the characteristic impedance of the line often are called *non-resonant* lines, for reasons which will become apparent in the following discussion. There is no hard-and-fast line of demarkation, but usually a line is classified as a "non-resonant" line if the VSWR is less an 2.0.

INPUT IMPEDANCE OF UNMATCHED LINES

It is readily apparent that when reflection exists and some of the energy delivered by an r.f. generator "bounces off" the load and arrives back at the generator, the line will not look like a pure resistance equal to Z_0.

The impedance the generator "looks into" is determined by the ratio of voltage to current at the sending end, and from our study of standing waves we know that this ratio depends upon the nature of the line termination and the length of line (assuming for the moment a uniform line with no attenuation). The sending end impedance will be maximum when the line length is such as to put a voltage loop (current node) at the generator, and minimum at a current loop (voltage node).

It should be noted that the impedance at the sending end is not necessarily nonreactive in the case of a mismatched or "resonant" line. The power factor of the impedance at the sending end is determined by the phase angle between the voltage and current, which likewise is determined by the nature of the line termination and the length of line. Only at certain critical line lengths will the relationship between incident and reflected waves be such as to result in unity power factor. At all other lengths the load on the generator will be reactive. This is illustrated graphically in figure 2-12.

It will be observed that as the standing wave ratio is increased from unity or flat line conditions, the phase shift of the voltage or current no longer is uniform; it tends to become concentrated at recurrent points along the line. Also, the phase angle between the voltage and current approaches 90 degrees at points other than the critical or "resonant" points as the SWR is increased from unity to infinity. It will be observed that in every case where the SWR is not unity, the line may be divided up into quarter wavelength sections, with the current leading the voltage in one section, lagging in the next, leading in the next, and so on down the line, and also that the current is in phase with the voltage (unity power factor) at the points where the current changes from leading to lagging and vice versa.

If the SWR is small, the maximum phase angle is small, which means that the maximum reactive component is small and the power factor is high regardless of where the generator is connected. If the SWR is large, then the maximum phase angle approaches 90 degrees, the power factor approaches zero, and the line looks like a substantially pure reactance (either capacitive or inductive, depending upon whether the current is leading or lagging) *except* when the generator is connected at one of the recurring resonant "crossover" points.

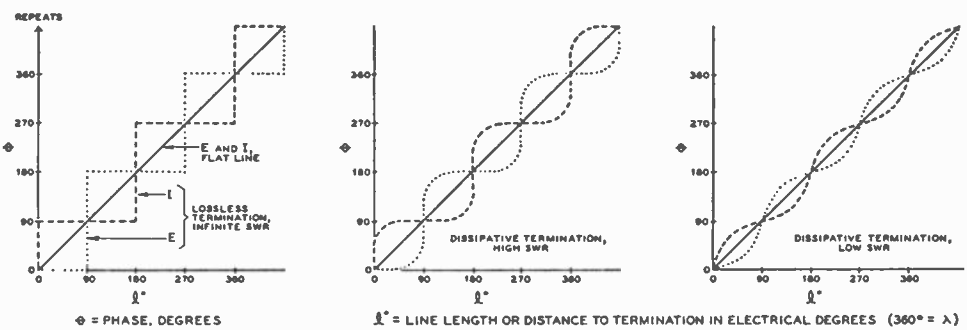


Figure 2-12.

ILLUSTRATING NON-UNIFORMITY OF PHASE SHIFT OF VOLTAGE AND CURRENT ALONG A LINE WHEN STANDING WAVES ARE PRESENT

It will be noted that when the line is flat, the voltage and current are in phase at all points, and that the phase shift is uniform. When standing waves are present, the voltage or current phase shift tends to become more concentrated at half-wave points as the SWR is increased, with the phases of the voltage and current becoming more and more nearly in quadrature except at the critical or "resonant" lengths. The line termination or "receiving end" is taken as the phase reference. It should be kept in mind that the recurring pattern may START at any portion of the cycle, depending upon the character of the termination.

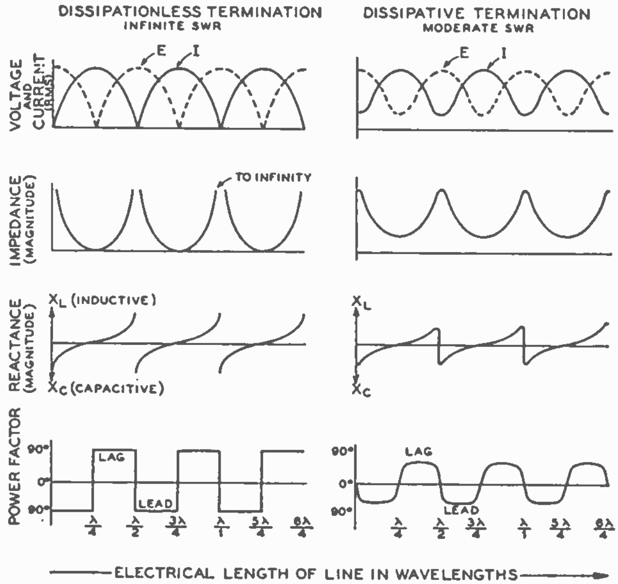


Figure 2-13.

ILLUSTRATING VARIATION IN IMPEDANCE, REACTANCE, AND POWER FACTOR AT SENDING END OF A MISMATCHED LINE AS THE ELECTRICAL LENGTH IS VARIED.

It should be noted that the electrical length of a line (its length in wavelengths) can be changed either by altering its actual physical length and maintaining the frequency constant, or by altering the frequency and maintaining the physical length of the line constant. The curves not only illustrate the change in sending end characteristics, but also the cyclic variation in impedance, reactance, etc. along the line. As the SWR is reduced and approaches 1.0, all curves approach straight, horizontal lines.

Thus we see that as the physical line length or the frequency is varied, thus effectively varying the *electrical* length of the line in either case, the magnitude of the impedance and the magnitude and sign of its reactive component will vary cyclically *whenever reflection is present.* The exact manner in which they vary is illustrated in figure 2-13, which represents (left) conditions existing on a hypothetical lossless line with dissipationless termination (infinite SWR), and (right) conditions existing on an actual low loss line with dissipative termination sufficiently mismatched to give a moderate SWR.

It is apparent that, regardless of the SWR, at certain critical lengths (namely, multiples of an electrical half wavelength) the sending end impedance is the same as the load impendance, both as to magnitude and as to power factor. And figure 2-13 also shows that for all other lengths of line the sending end impedance differs from the load impedance. It is also apparent from the curves that a lossless line terminated in a short or open circuit looks like a pure reactance which can be either capacitative or inductive and have any magnitude, depending upon its length, except at the critical "resonant" lengths just mentioned. The whole gamut is included when the electrical length is varied between zero and a half wavelength, or between zero and one quarter wavelength if the option of short or open-circuited termination is exercised.

Because an actual line of such short length can be constructed to have extremely low losses, in actual practice it is possible to realize a reactance having very high "Q", much higher than can be obtained with ordinary "lumped constant" coils and capacitors. It will also be noted from the curves that in the regions of line length approximating an impedance maximum, the characteristics resemble those of an ordinary parallel or anti-resonant (tank) circuit made up of lumped C and L. In the vicinity of impedance minima, the curves resemble the characteristics of an ordinary series resonant circuit made up of lumped C and L.

Now let us see how these various characteristics of the "resonant" line are utilized to advantage and how in some applications of transmission lines they are detrimental.

USE OF "RESONANT" LINES FOR POWER TRANSMISSION

First let us consider how the behavior of a resonant line as just described can be detrimental. We already have seen that the standing waves cause an increase in the total line dissipation and the peak voltage on a line, reducing the efficiency of the line, and also in some cases requiring a line of larger physical size and greater cost in order to avoid flashover or excessive temperature rise. In addition to these characteristics, which may or may not be objectionable depending upon the circumstances, a "resonant" line has one feature which can be a very serious disadvantage as compared to a "non-resonant" line when the line is used simply for transferring power from "here" to "there" and operation on more than a single frequency is involved.

From figure 2-13 we see that as the frequency is varied, effectively changing the electrical length of the line, the load on the generator varies, both as to magnitude and power factor. A resonant line may be said to be *frequency sensitive.* The change in input impedance with frequency is a function of the length of the line and the reflection coefficient. If the line is very long and the SWR high, only a slight change in frequency can make a great difference in the magnitude and power factor of the input impedance, even though the load may be a pure resistance. Usually the picture is modified somewhat by the fact that the impedance of the load itself is not a pure resistance and is frequency sensitive, but seldom in such a fashion as to alleviate the overall condition and oftentimes in such a fashion as to aggravate the overall condition.

As an example, let us suppose a mismatched load which comprises nothing but pure resistance and produces an SWR of 4.0 terminating a line which is 20 wavelengths long (which is not an unusually long line at the higher frequencies). This is not an idealized or academic case, because in practical applications it is quite common or possible to have a load which has a fairly constant resistance and comparatively a very low reactance over a frequency range of 5 per cent or more.

Assuming the line has negligible losses, which is another valid assumption in the case of a good quality line, the standing wave ratio at the sending end will be approximately 4.0 and the input impedance will be the same as the terminating impedance. Increasing the frequency (or decreasing the frequency) by only 1.25 per cent will effectively increase or decrease the electrical length of the line by a quarter wavelength, because *frequency* and *line length in terms of wavelength* are *directly proportional.* This will increase or decrease the *impedance* by the *square* of the VSWR, depending upon whether the terminating impedance is higher or lower than Z_0. (At a voltage loop the voltage will be 4 times as high and the current ¼ as great as at a voltage node.) The input impedance will still be resistive, but it will be increased or decreased by 16 times. If the length is altered by less than a quarter wave, the change in impedance will not be so great, but the input impedance *then will have a reactive component.*

In the case of a c-w telegraph transmitter operated on various frequencies this condition is undesirable but operation is nevertheless still possible, even though it entails "tuning out" (or otherwise eliminating) the reactance and matching the impedance each time the frequency is changed (except when by sheer accident two frequencies just happen to be so related that this is unnecessary). In the case of a "wide-band" service which transmits simultaneously in the form of sidebands a spectrum of frequencies having considerable width, such as television, a "frequency sensitive" line can produce distortion which is intolerable.

TROMBONE SECTION OR "LINE STRETCHER"

When a transmission line is used under resonant conditions, so that standing waves of appreciable magnitude exist throughout the length of line, it is necessary to make provision for either adjusting the line precisely to a resonant electrical length or else "tuning out" any reactive component at the sending end, so that the generator will "look" into a non-reactive load. This can be accomplished either electrically, by means of adjustable compensating reactances connected at the generator end, or physically by inserting in series with the line a section of line having adjustable length. The latter consists of a telescopic section of line formed in the shape of a "U" which may be slid in or out as required to achieve the desired total line length. Such a telescopic section is known as a *"trombone section"* or *"line stretcher."*

Trombone sections are practicable only in the v-h-f and u-h-f range, where the physical size does not become formidable. Because of the difficulty in constructing such a device for a coaxial line, due to problems encountered in supporting the inner conductor and obtaining low contact resistance, trombone sections are used almost exclusively with parallel-conductor open lines. Another reason why trombone sections are seldom used with coaxial lines is that resonant operation of coaxial lines tends to become impractical when the SWR is high. Under conditions of high SWR, the losses become excessive in solid dielectric lines, and the power handling capability of air dielectric lines becomes quite low because of flashover difficulties due to the relatively close conductor spacing as compared to "open" lines.

One advantage of the resonant line and trombone section combination is that a particular installation can be used over a wide frequency range with a wide variety of terminations, and the trombone section may be inserted at the most convenient point, anywhere along the line, without restriction. To be most useful, the trombone section should permit a variation of one half wavelength in the total line length at the lowest frequency employed.

SECTIONS OF TRANSMISSION LINE USED AS CIRCUIT ELEMENTS

The properties of low loss resonant lines are such as to make them very useful as high Q circuit elements, either (1) as parallel resonant circuits, (2) as series resonant circuits, or (3) as reactances having very low power factor. Sections of transmission line are widely used for such purposes when frequencies become so high that circuits using lumped constants are either inefficient or impossible to attain. Because of their configuration, circuit elements consisting of sections of transmission line are called *linear* circuit elements, as contrasted to the term *lumped* circuit elements which is applied to ordinary inductors and capacitors. It should be kept in mind that the reactive effects produced by linear circuit elements are not due to the distributed C or L of the line but to reflection effects which are determined by the relationships between C, L, line length, line termination, and other factors.

The three basic applications as circuit elements enumerated above will be discussed separately because of the wide application and importance of linear circuit elements.

LINEAR PARALLEL-RESONANT CIRCUITS (ANTI-RESONANT OR "TANK" CIRCUITS)

Perhaps the characteristic of a resonant line section which is most exploited in the v-h-f and u-h-f bands is its ability to act as a parallel resonant or "tank" circuit. At extremely high frequencies ordinary tank circuits utilizing coils and condensers become inefficient and finally cease to function in the intended manner, because they cannot be made sufficiently small physically. A conventional capacitor of the smallest practical physical size begins to show considerable inductive effect when the frequency becomes sufficiently high. And if a u-h-f coil is so designed as to have a high Q and low distributed capacitance, we find that it begins to resemble a short section of transmission line rather than a conventional solenoid.

It will be noted in figure 2-13 (left), that when a line having negligible losses is terminated in an open or closed circuit, points on the line in the immediate vicinity of a current node exhibit the characteristics of a parallel resonant "tank" circuit. At exact "resonance" the linear tank circuit looks like a pure resistance of very high value (infinite in the case of a hypothetical lossless line), the voltage across the line is the voltage of the source, considerable energy is flowing along and is stored or contained in the line, and little power is required from the generator to maintain this energy. Thus at exact resonance a linear tank circuit acts as a "flywheel" having high inertia and very low friction. In the immediate vicinity of exact resonance a curve of impedance vs. frequency resembles that of a conventional tank circuit consisting of lumped capacitance and lumped inductance in parallel. However, at points substantially removed from exact parallel resonance (maximum impedance), the curve departs from the conventional parallel-resonance curve obtained with lumped constants.

Harmonic Response

The behavior of a linear tank circuit differs from that of a tank using lumped constants in another respect. From figure 2-13 it can be seen that for maximum impedance or parallel resonance, an open circuited line can be any even number of quarter wavelengths long, while a short circuited line may be any odd number of quarter wavelengths long. This means that a section of transmission line theoretically has an infinite number of resonant frequencies, all harmonically related. In actual practice the usual integral relationship of harmonic response is upset by *end effect,* which causes an open end of a resonant line to act as though the line were a very small fraction of a wavelength longer than it actually is. When a section of line is resonant on a harmonic, quarter wavelength sections which do not include an open end are not so affected, and are a full quarter wavelength long (allowing for the velocity factor of the line. The net result is that the parallel resonant frequencies of a section of line are not in exact intergral harmonic relation, the frequency of resonance on a harmonic being somewhat higher than an integral multiple of the fundamental resonant frequency.

The relationship between fundamental and harmonic frequencies is further upset when lumped reactance is present across the line. For instance, a short-circuited quarter-wavelength parallel-resonant line can be made considerably shorter physically than a quarter wavelength as calculated from the velocity factor of the line if lumped capacity is connected across the open or high impedance end of the line section. However, as the physical "shortening effect" (or electrical "lengthening effect") in terms of wavelength is determined by the magnitude of the capacitive reactance, and as the reactance of the lumped capacitance is inversely proportional to frequency, it can be seen that if enough capacitance is employed to provide appreciable shortening of the line at the fundamental frequency, then the near-integral relationship of the harmonic resonant frequencies will be completely upset.

The amount by which a fundamentally resonant quarter-wave section of line is electrically lengthened or physically shortened by a given amount of lumped, shunt, capacitive reactance *depends upon the surge impedance of the line.* The higher the characteristic impedance of the line, the greater the effect of a given lumped, shunt, capacitive reactance upon the resonant length (or upon the resonant frequency when the length is constant).

For resonance, the characteristic impedance of the line in ohms multiplied by the tangent of the length of the line in electrical degrees must equal the reactance of the capacitor in ohms. From this it is apparent that doubling the surge impedance of a capacity-loaded quarter-wave linear tank has the same effect upon the resonant frequency as doubling the lumped shunt capacitance, and vice versa.

From figure 2-13 it can be seen that in the hypothetical case of a parallel resonant section of line which is absolutely lossless, the impedance and therefore the Q are infinite in magnitude, and the line looks like an open circuit. In practical linear tank circuits the Q will of course be finite, because of losses in the line section and in any capacitor which may be connected across the circuit.

Tank Circuit Q

If no lumped capacitance is employed, the Q of a parallel-resonant quarter-wave section of coaxial line utilizing copper conductors and air as the dielectric will be as shown in figure 2-14, provided that the ratio of the diameters of the effective sur-

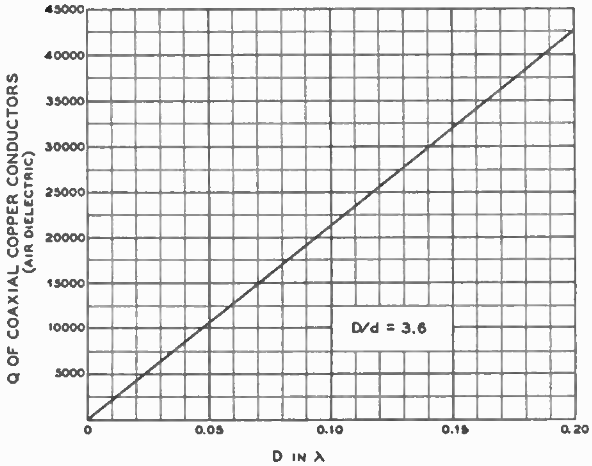


Figure 2-14.

"Q" OF AIR SPACED COAXIAL LINE VS. INSIDE DIAMETER OF OUTER CONDUCTOR WHEN RATIO OF DIAMETERS IS OPTIMUM.

The curve is for copper conductors and a ratio of D/d equal to 3.6, which is optimum. D refers to the inside diameter of the outer conductor and d refers to the outside diameter of the inner conductor. The Q may be obtained for other values of D/d by multiplying by a factor taken from the Q curve of figure 2-15.

faces is 3.6. The ratio of 3.6 gives the highest Q for a given outer diameter. The Q for other diameter ratios may be obtained from the Q curve of figure 2-15.

With metals other than copper, the Q varies inversely as the square root of the effective resistivity of the metal used for the conductors. It should be kept in mind, however, that due to skin effect the effective resistivity is not the same as the d-c resistance of the metal, and that at a given frequency the ratio of d-c resistance to effective resistivity is not constant but varies with the resistance of the metal.

It is obvious that for good Q, only metals of high conductivity such as aluminum, copper, and silver should be used. Solid copper or silver-plated metal ordinarily is employed, the latter having a slightly lower effective resistivity than solid copper if the plating is done properly.

When a section of line used as a linear tank circuit is shunted by a capacitor, it is important that the capacitor have the highest possible Q. The Q of a "quarter wave" section of low loss line is not greatly reduced by a moderate amount of physical "shortening" by means of shunt capacitance if the capacitor is a pure reactance, but in actual practice it is difficult to achieve a capacitor having a Q in the v-h-f and u-h-f range of the same order as an air dielectric line unless the plates of the capacitor are entirely supported by the line conductors, so that no insulation other than air appears across the "hot" end of the line. When a linear tank circuit must be made tunable, the use of a variable lumped shunt capacitor often is circumvented by using instead a laterally adjustable shorting plunger to change the effective length of the line. However, the heavy currents flowing in the shorting plunger require that it make very low resistance contact with the inner and outer conductors if an appreciable reduction in Q is to be avoided.

Impedance Versus Q

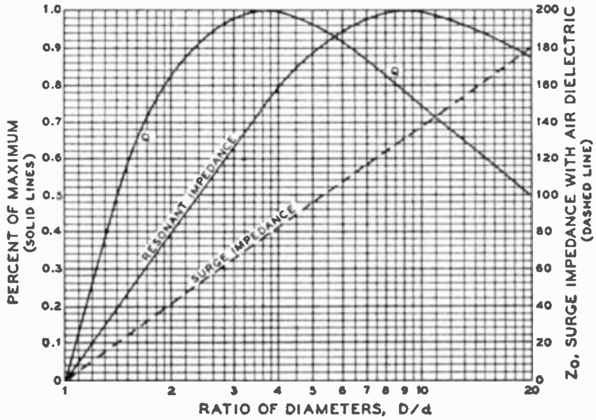
It will be noted from figure 2-15 that maximum open-end impedance of a coaxial linear tank is obtained with a considerably higher ratio of D to d than that giving maximum Q (assuming a given inside diameter of outer conductor), but that both curves are fairly broad.

So far we have not mentioned any loading on the linear tank other than its inherent losses. But a tank circuit is not useful unless we connect it to something or connect something to it, and in either case there will be additional loading of the tank circuit and lowering of the effective Q. In fact, if the loading is fairly heavy, it may be found that the *loaded* impedance of a linear tank designed for maximum unloaded impedance may be but little higher than a tank using

Figure 2-15.

COAXIAL LINE CHARACTERISTICS.

Illustrating how the line characteristics vary with the outside diameter of the inner conductor (d) when the inside diameter of the outer conductor (D) is held constant, thus changing the ratio D/d. Relative line attenuation may be determined from the Q curve; the attenuation is proportional to the reciprocal of the Q. Within certain limitations and with certain qualifications discussed in the text, the curves can be applied also to parallel conductor open lines.



the same outer conductor but with the inner conductor proportioned for maximum unloaded Q rather than maximum unloaded impedance.

Coupling

Coupling to a linear tank circuit usually is made either by tapping directly on to the circuit element at an appropriate impedance point or by means of a small, inductively coupled loop which sometimes is called a "hairpin loop." Because the Q of a linear tank usually is quite high, and heavy current is flowing in the conductors, a high coupling coefficient is obtained.

It should be kept in mind that a generator or load can be tapped on to a quarter-wave linear tank circuit at any distance from the shorted end and the tank will appear resistive, just as in the case of tapping on to the coil of a tank circuit using lumped constants, so long as the total length of the line is resonant at the frequency involved. For instance, it is true that the open end of a short-circuited one-eighth wavelength line "looks" predominantely reactive (purely reactive in the case of a lossless line); but if we tap on to a resonant quarter-wave linear tank circuit an eighth wavelength up from the shorted end it looks like a pure resistance, the magnitude being determined by the Q of the circuit, which in turn is determined by the inherent losses and external loading.

Quarter Wave vs. Longer Tank Elements

Linear parallel-resonant tank circuits ordinarily are made with an electrical length of either a half wavelength or a quarter wavelength, usually the latter. The Q of the tank is not improved by making the line longer than a quarter wavelength, even though the energy storage is greater in a longer line. The line losses and energy storage go up together as the length is increased, keeping the Q of the unloaded line the same. A longer line will stand more external loading, however, for the same reduction in Q, or conversely the reduction in Q is less for a given external loading. But the usual practice is to obtain a higher loaded Q, when such is required, by making the line only a quarter or half wavelength long and using larger diameter conductors, tapping down on the tank to maintain the desired impedance.

One advantage of a quarter-wavelength tank over a half-wavelength tank is that the former is easily "tuned" by means of a movable shorting plug or bar, or by a variable capacitor. Also, a quarter-wavelength tank is responsive only to odd harmonics, while a half-wavelength tank is responsive to both odd and even (for reasons which are apparent from study of standing wave patterns). The response of a high Q linear tank to the second and third harmonics (usually the only ones to which attention

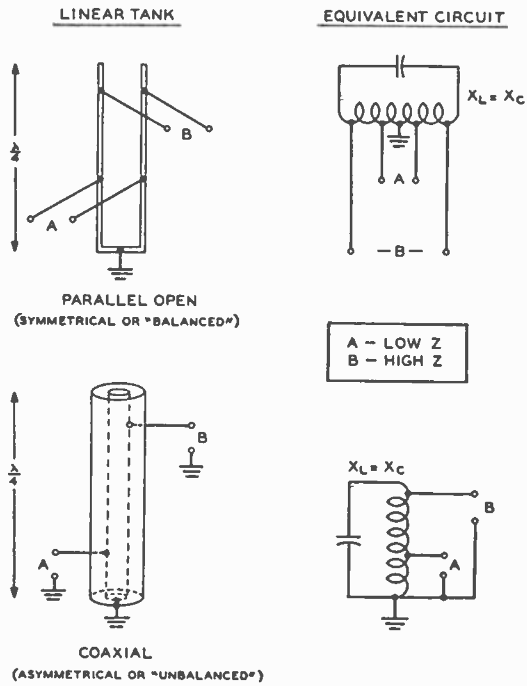


Figure 2-16.

ILLUSTRATING THE TWO BASIC FORMS OF LINEAR TANK CIRCUIT AND USE AS AN IMPEDANCE TRANSFORMER

Because the voltage on a linear tank follows a substantially sine function and the voltage between turns of a coil is substantially uniform (in a high Q tank using lumped constants), and also because the effective or equivalent mutual coupling is not the same for both linear and lumped circuits, the points of attachment or taps for the two types of circuit will not have quite the same physical relationship for the same impedance transformation.

must be given) can be greatly reduced by using a small amount of capacity loading, causing the second and third harmonic response frequencies to be removed by several per cent from the actual harmonic frequencies. Another common method of virtually eliminating the third harmonic response of a quarter-wavelength tank is to tap down on the tank to a point which corresponds to a voltage node at the harmonic (approximately 1/3 of the way down from the open end).

Parallel Conductor Open Line vs. Coaxial Line Tank Circuits

While the foregoing discussion of linear tanks has assumed a section of coaxial line, the same considerations apply to sections of balanced, two-conductor open line except for minor qualifications. Coaxial line is used almost exclusively for linear tank circuits above approximately 200 Mc. because the radiation from and stray coupling between open lines of practical configuration becomes excessive at such frequencies. Sections of two-conductor open line oftentimes are used on lower frequencies, because adjustments are more easily made than with coaxial line, due to the fact that both conductors are exposed throughout their length.

Except at close spacings (ratios less than 3 or 4), the coaxial line curves of figure 2-15 apply *approximately* to two-conductor open lines if the *ratio of center-to-center spacing to conductor radius is substituted for the radio D/d* and the *values of surge impedance are doubled.*

Proximity Effect

At comparatively close spacings the effect of one conductor upon the other results in a current concentration along the inner sides of the two parallel conductors, and the current distribution over the surface of the conductors no longer is uniform. Also, undesirable eddy currents are produced. The overall effect is called "proximity effect," and it acts to cause both Q and surge impedance to be lower at very close spacings than would otherwise be the case.

Practical Applications of Linear Tanks

Linear tank circuits commonly are used in connection with v-h-f and u-h-f antenna systems as impedance matching devices (figure 2-16) and as selective filters to suppress undesired frequencies, just as lumped-constant tank circuits are used for such purposes at lower frequencies. They also are widely used as transmitter final plate tank circuits and as receiver input circuits in the v-h-f and u-h-f range. Although it is open to argument as to whether the latter circuits can be considered part of the antenna system, the function of such tank circuits is so closely related to the antenna system to which they are coupled that they cannot very well be ignored in a study of overall antenna systems.

LINEAR SERIES-RESONANT CIRCUITS

It will be noted in figure 2-13 (left) that when a line having negligible losses is terminated in an open or closed circuit, points on the line in the immediate vicinity of a voltage node exhibit the characteristics of a series resonant circuit. At exact "resonance" the linear series-resonant circuit looks like a short circuit having a very low, non-reactive impedance (zero in the case of a hypothetical lossless line). In the immediate vicinity of exact resonance a curve of impedance vs. frequency resembles that of a conventional series resonant circuit consisting of lumped capacitance and lumped inductance in series. However, at points substantially removed from exact series resonance (minimum impedance), the curve departs from the conventional series-resonance curve obtained with lumped constants.

From figure 2-13 it can be seen that for minimum impedance or series resonance, an *open* circuited line may be any *odd* number of quarter wavelengths long, while a *short* circuited line may be any *even* number of quarter wavelengths long.

From figure 2-13 it also can be seen that in the hypothetical case of a series-resonant line which is absolutely lossless, the impedance looking into the line is zero, and the voltage across the line at an electrical distance of ¼ wavelength from the zero impedance point is infinite when such a line is connected across a generator. In the case of practical linear series-resonant circuits the line losses cause both the input impedance and the resonant rise in voltage across the line at the current node to assume finite values, though in the case of very low loss lines the input impedance will be virtually zero and the voltage at the voltage loop will be quite high. This condition is compatible with that existing in a series resonant circuit using lumped constants: the input impedance is low and the voltage across the input is that of the generator, but the voltage across the coil *or* the capacitor is many times greater.

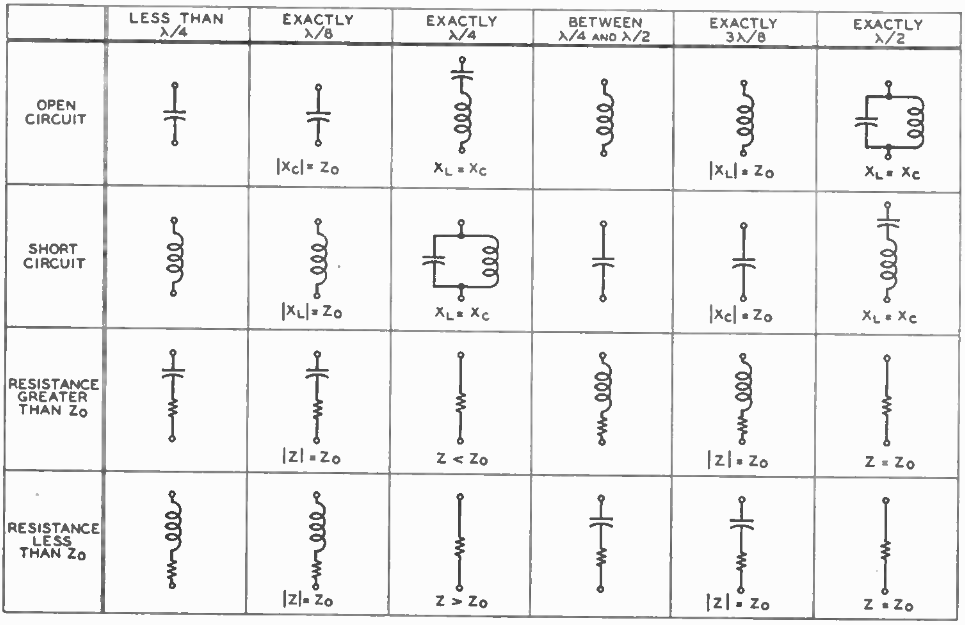
Series-resonant linear circuits are not so widely used in connection with antenna systems as are parallel resonant or "tank" circuits using linear elements, though they are found very useful for such purposes as the suppression of harmonics, and the incorporation of "compensating reactance" in certain antenna systems to make them less frequency sensitive.

LINEAR ELEMENTS AS LOW LOSS REACTANCES

From figure 2-13 it can be seen that a section of line having negligible losses can have almost any magnitude of reactance, either capacitive or inductive, depending upon the electrical length of the line and whether the far end is open circuited or short circuited. This is illustrated graphically in figure 2-17. In the hypothetical case of a section of lossless line, a pure reactance having zero power factor would result. Losses present in practical lines result in a small resistive component, and therefore a finite power factor. However, the losses in a section of well constructed line are so low that such a section of line can serve as a reactance having a power factor much lower than can be realized at the higher frequencies in conventional "lumped constant" capacitors and inductors.

Line Matching Stubs

Sections of line called *stubs* commonly are used as low loss reactances to "cancel" or eliminate undesirable reactive components in antenna systems and provide an imped-



CHARACTERISTICS ARE UNCHANGED WHEN ANY MULTIPLE OF AN ELECTRICAL HALF-WAVELENGTH IS ADDED.

Figure 2-17.

ILLUSTRATING BY MEANS OF EQUIVALENT LUMPED CIRCUITS HOW VARIOUS LENGTHS OF TRANSMISSION LINE "LOOK" TO THE SOURCE WHEN VARIOUS TERMINATIONS ARE EMPLOYED.

ance match. They function in the following manner.

When standing waves exist upon a line, there are recurring points on the line (four every wavelength) where the resistive component is equal to the characteristic impedance of the line. If the line is shunted at this point with a reactance which is equal in magnitude and opposite in sign to the reactive component existing at that point on the line, then the line will be perfectly mached at that point, and, assuming a uniform line, no standing waves will exist from that point back to the generator. (See figure 2-18A.)

Naturally when this arrangement removes standing waves only betwen the stub and source, the stub should be placed as close as practical to the mismatched load or to the line discontinuity responsible for the standing waves, even though there are many recurring points on a long line where a stub having the appropriate amount and kind of reactance will eliminate the standing waves. Increasing the distance between the load and stub not only increases the total line loss, but also makes the overall antenna system more frequency sensitive.

While the characteristics and point of attachment of a stub which will remove standing waves can be predetermined with a fair degree of accuracy (see figure 2-19), it often is necessary to "touch up" the adjustment to obtain a perfect match, because of various unpredictable factors which usually enter into any practical installation. This final adjustment by experimental means is easily accomplished in the case of a two-conductor open line where both conductors are exposed, but is rather difficult in the case of a coaxial line. In the latter case the stub length can be made variable by means of a shorting plunger, but varying the point of attachment to the line is not so easily accomplished.

Double Stub Matching

This difficulty is circumvented by the use of *two* stubs of adjustable length, located anywhere on the line, with a predetermined spacing between stubs (commonly either 1/8 or 3/8 wavelength). Such an arrangement (figure 2-18D) is limited as to the range of complex impedances which may be matched, while a single stub arrangement is not. However, where the standing wave ratio is not unusually high, no particular attention need be paid to the exact point of attachment of the two stubs, or to the exact value of fixed stub spacing to be employed.

The functioning of the dual stub arrangement is not as simply analyzed as is the single stub arrangement, even though the adjustment is no more difficult. The two stubs, together with the resonant portion of the line to which they are connected, act as a combined impedance transformer and reactance cancelling circuit just as in the case of the single stub arrangement. The additional variable-length stub produces somewhat the same effect as being able to change the position of the single stub.

The Stub as a Transformer

It should be appreciated that a matching stub is basically an impedance matching device, and a very good one because it can transform or "match" any complex impedance to the characteristic impedance of the line. The fact that standing waves can be eliminated by means of a stub near the load end simply indicates that the stub, in conjunction with the resonant section of line between stub and load, in effect transforms the mismatched load impedance into a purely resistive load equal to Z_0.

Now if the stub arrangement actually is nothing more than an impedance transformer, it is apparent that the arrangement can be used not just to transform a mismatched load impedance into an impedance equal to the characteristic impedance of the line, but also to transform the line impedance to some other value at the generator end, should this be necessary or desirable. For instance, a transmitter designed for operation only into a 300-ohm load could be used in conjunction with a 600-ohm non-resonant line and an antenna having an

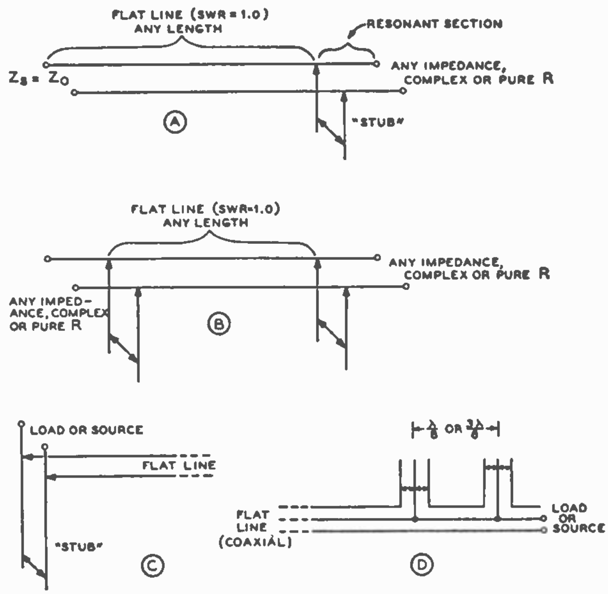


Figure 2-18.

"STUB-LINE" IMPEDANCE MATCHING

It is possible to match any load to a transmission line, or to make a line look like any load to a generator, by shunting the line at a suitable point with a suitable reactance. This is best accomplished by means of an adjustable, short length of line known as a "stub." For open two-conductor lines the configuration may be as shown at either A or C. The use of two stubs, as shown at D, makes it possible to make experimental adjustments without varying the stub position.

input or feed point impedance of 73 ohms by utilizing a matching stub at each end of the line, as in figure 2-18B. Impedances are matched throughout the system, and the line works under "flat line" conditions.

When studying or considering the use of matching stubs in connection with antenna systems used for receiving, it should be kept in mind that the role of the antenna proper is reversed: it no longer is the load, but rather the *generator*. To remove standing waves from the line, a stub at the *receiver end* is required. The ratio of antenna impedance to line impedance will have no bearing upon the standing wave ratio. However, if the antenna impedance is appreciably different from the characteristic impedance of the line, and the line is matched at the receiver end, then a stub can be used to advantage at the antenna end to enable the antenna to deliver maximum power into the line.

Stub construction may follow either the configuration of figure 2-18A, or that of figure 2-18C, depending upon which physical arrangement best suits the particular installation.

Stub Design and Adjustment

While it is possible to obtain a perfect match and equally good performance with either a shorted stub or an open stub by observing appropriate dimensions, a shorted stub is much more readily adjusted. Therefore, the following discussion of practical application of single stub matching to an open parallel conductor line will be confined to that type.

If the transmission line is so elevated that adjustment of a "fundamental" shorted stub cannot be accomplished easily from the ground, then the stub length may be increased by exactly one or two electrical half wavelengths, without appreciably affecting its operation.

While the correct position of the shorting bar and the point of attachment of the stub to the line can be determined entirely by experimental methods, the fact that the two adjustments are interdependent or "interlocking" makes such a cut-and-try procedure a tedious one. Much time can be saved by determining the approximate adjustments required by reference to a chart such as figure 2-19 and using them as a starter. Usually only a slight "touching up" will produce a perfect match and flat line.

In order to utilize figure 2-19, it is first necessary to locate accurately a voltage node or current node on the line in the vicinity that has been decided upon for the stub, and also to determine the SWR.

Stub adjustment becomes more critical as the SWR increases, and under conditions of high SWR the current and voltage nulls are more sharply defined than the current and voltage maxima, or loops. Therefore, it is best to locate either a current null or voltage null, depending upon whether a current indicating device or a voltage indicating device is used to check the standing wave pattern.

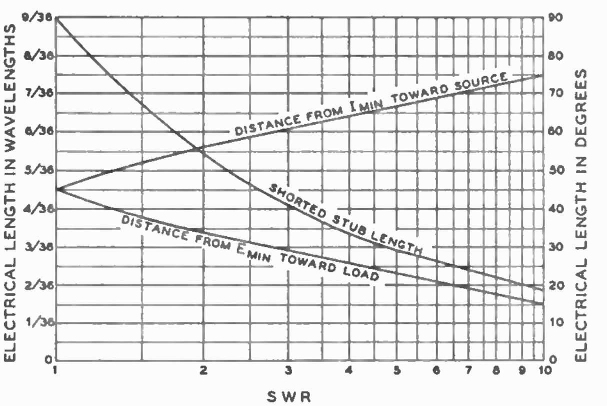


Figure 2-19.

SHORTED STUB LENGTH AND POSITION CHART.

From the standing wave ratio and current or voltage null position it is possible to determine the theoretically correct length and position of a shorted stub. In actual practice a slight discrepancy usually will be found between the theoretical and the experimentally optimized dimensions; therefore it may be necessary to "touch up" the dimensions after using the above data as a starting point.

The SWR is determined by means of a "directional coupler," or by noting the ratio of E_{max} to E_{min} or I_{max} to I_{min} as read on an indicating device. The correct method of making such measurements will be treated later in this chapter.

It is assumed that the characteristic impedance of the section of line used as a stub is the same as that of the transmission line proper. And it is preferable to have the stub section identical to the line physically as well as electrically.

The procedure is somewhat more complicated in the case of single stub matching of a coaxial line. The SWR can be determined by means of either a directional coupler or a section of *slotted line*, which will be described later under the subject of measurements. However, unless the calculated position for the stub happens to turn out to be exactly right, the lack of flexibility of the single stub arrangement makes it impossible to do a simple "touching up" job on the stub position. With the double stub arrangement previously mentioned, the "touching up" is facilitated, but the mathematical or graphical determination of the theoretically correct stub positions and lengths is somewhat more complicated, and will not be dealt with here. It is recommended that if the reader has or expects to have need for solving numerous transmission line problems of this and similar nature he familiarize himself with the workings of the "circle diagram," a very useful device which permits rapid graphical solution of a wide variety of transmission line problems.*

Electric Wave Filter Elements

Sections of low loss transmission line are useful as reactive elements in electric wave filters of the low pass, high pass, and band pass types in the v-h-f and u-h-f range. They are especially useful as substitutes for lumped capacitance, as capacitors having very low power factor in this frequency range are quite expensive. The characteristics of the filter are modified somewhat when sections of line are substituted for some of the "lumped" reactors, because the reactance of linear elements changes with frequency in a somewhat different manner than does that of conventional capacitors and inductors.

Filters of this type are useful in the suppression of spurious transmissions and receiver responses, and have the advantage over highly selective resonant circuits of not requiring readjustment for a moderate change in frequency.

*For an excellent treatment of the circle diagram and its application to common transmission line problems see King, Mimno, and Wing, *Transmission Lines, Antennas, and Wave Guides*, McGraw Hill Book Co., Inc., 1945.

MISCELLANEOUS APPLICATIONS OF LINEAR ELEMENTS

In addition to such fundamental applications as parallel resonant tank circuits, series resonant circuits, and as low loss reactances, sections of transmission line are used for various other important applications in connection with antenna systems, applications which do not fall strictly within any of these three categories. In some such applications the line section serves a composite function and is difficult to classify. Some of the more important of these applications will be described without any attempt at classification.

QUARTER-WAVE LINEAR TRANSFORMER ("Q" SECTION)

From our previous study of standing waves, reflection coefficient, and SWR, we know that if a 100-ohm line is terminated with a resistive load of 50 ohms, the SWR will be 2.0 and a current loop (voltage node) will appear at the load and every electrical half wavelength back towards the generator. Voltage loops (current nodes) will occur every half wavelength at intermediate points, the closest being a quarter wavelength from the load.

This means the voltage at the closest loop will be twice that at the load, and the current at the same point will be half that at the load. Thus, at a point a quarter wavelength from the load the voltage is doubled and the current halved, which represents the same energy or power but four times the impedance.

If the termination were 25 ohms, the

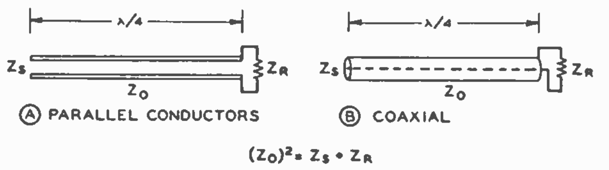


Figure 2-20.

QUARTER-WAVE LINEAR TRANSFORMER ("Q" SECTION).

A section of line an electrical quarter wavelength long "inverts" the load, thus making it useful as an impedance transformer. It functions properly only when the load is non-reactive.

same relationships would apply except that the SWR would be 4.0, the voltage quadrupled, and the current reduced to a quarter, representing the same energy or power but 16 times the impedance.

If the load impedance were made 200 ohms and then 400 ohms, the SWR still would be 2.0 and 4.0 respectively, but the current and voltage relationships would be reversed and the impedance at a distance of a quarter wavelength would be reduced to 1/4 or 1/16, respectively, instead of being increased to 4 times or 16 times as high.

Thus we see that a section of line which is one or any odd multiple of quarter wavelengths long "inverts" the load and can be used as an impedance transformer, and furthermore that the relationship between the input and load impedances is such that their geometric mean is equal to the characteristic impedance of the linear transformer section, or, referring to figure 2-20,

$$(Z_0)^2 = Z_s Z_r$$

where Z_0=surge impedance of quarter-wave transformer

Z_s=reflected or sending end impedance

Z_r=load impedance

While the linear transformer may be any odd number of electrical quarter wavelengths long, assuming that the losses are negligible, the transformer section usually is made one quarter wavelength long for reasons of economy and convenience as well as to minimize the frequency discrimination. The quarter-wave linear transformer also is known as a *"Q Section."*

In practical applications the Q section usually is made with a fixed surge impedance. However, when it must be adjustable to allow for precise adjustment, or for use where the load impedance is not known accurately, the surge impedance of the Q section can be made variable within limits in the following manner. In the case of parallel conductors (figure 2-20-A) the spacing between conductors is made adjustable. In the case of a coaxial linear transformer, the conductors can be made eliptical, with provision for rotating the inner conductor with respect to the outer one. Obviously the latter scheme is practicable only at v-h-f and u-h-f, where the length of a linear quarter-wave section is short.

Unlike the stub matching arrangement previously discussed, the use of a linear quarter-wave transformer requires that the load impedance be nonreactive for proper operation. Fortunately this is not a serious disadvantage in many antenna system applications.

As is the case with any "resonant" impedance transformer, the frequency discrimination becomes high for high ratios of impedance transformation. That is, the overall antenna system becomes more frequency sensitive as the impedance transformation ratio goes up. This may or may not be desirable, and may or may not be a disadvantage, depending upon the particular application. For broad-band or multi-channel operation where it definitely is detrimental, the effects may be greatly reduced by employing two quarter-wave linear transformers in tandem, so that each Q section handles a transformation ratio equal to the square root of the total impedance transformation ratio.

Development of Exponential Line From Tandem Q Sections

If the number of articulated Q sections should be increased indefinitely, the impedance transformation ratio handled by any one section would be minutely small, and the adjacent quarter-wave sections would differ only very slightly in characteristics. Such an arrangement would resemble a transmission line having a gradual taper in the conductor spacing or diameter, such as to give an exponentially varying characteristic impedance. As would be expected, an "exponential" or "tapered" line of this type would have negligible frequency discrimination or "Q", provided that the taper was not too great per wavelength at the lowest frequencies involved.

METALLIC INSULATORS

From previous discussion of the quarter-wave shorted line (resonant tank circuit), we know that if the line losses are low in the quarter-wave section the impedance across the open end is extremely high. At the higher frequencies, where a quarter wavelength is short in terms of physical length, this permits the use of a stiffly constructed quarter-wavelength section as a mechanical support for the main transmission line. Such supports are called *insulating stubs* or *metallic insulators.* The basic types are shown in figure 2-21. Obviously this restricts the frequency range over which the line may be employed, but for many applications this is not disadvantageous.

Contrary to the conditions existing in the case of conventional dielectric insulators, the losses in a metallic insulator become less as the conductivity of the metal is increased, because the resonant impedance goes up with the conductivity. However, unless the line impedance is very high, such as could exist at certain points under conditions of a very high SWR, the conductivity of the metallic insulators will be sufficiently high even in the case of unplated steel to give sufficiently high resonant impedance to keep "insulator" losses negligible.

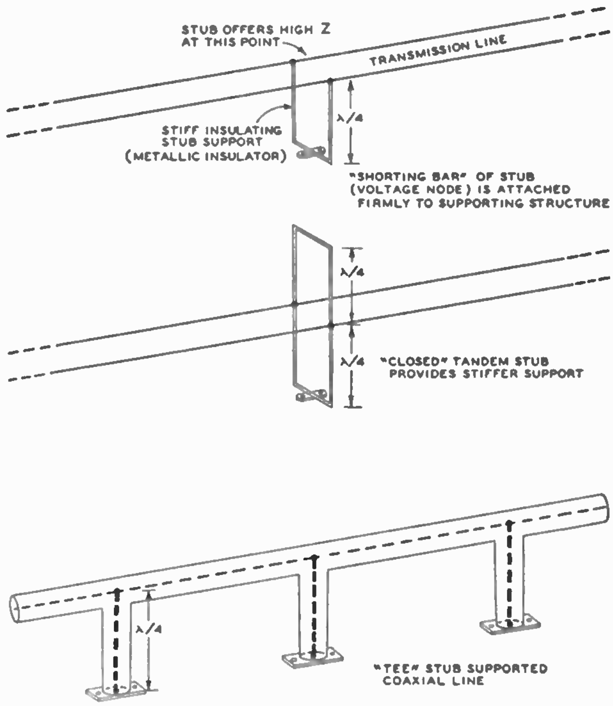


Figure 2-21.

INSULATING STUB SUPPORTS FOR TRANSMISSION LINES ("METALLIC INSULATORS").

The "U" shaped resonant supports for the open, parallel conductor line can be used at v-h-f and higher frequencies. Stub-supported coaxial lines are practical only at u-h-f and microwave frequencies.

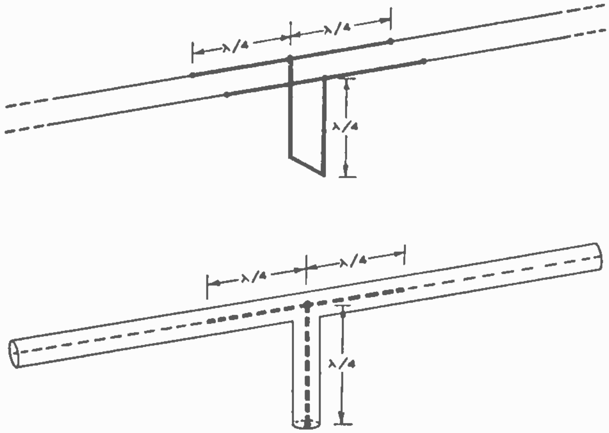


Figure 2-22.

BROAD-BANDED STUB SUPPORTS FOR OPEN AND COAXIAL LINES.

The line is made less frequency sensitive by incorporation of a reactance-cancelling section of line having a somewhat lower surge impedance, usually about 80 or 85 per cent of Z_0. By itself the reactance-cancelling section has no effect upon the line, because any half-wave section of line "repeats" the load. But combined with the stub, as shown here, it produces a "triple hump" resonance characteristic which is much broader than the resonance curve of the loaded stub alone.

A stub-supported line can be made less frequency sensitive by spacing the stubs an odd number of quarter wavelengths, so that the reflections more or less cancel. However, this is effective only when the line is short in terms of wavelength.

A highly effective and widely used method of "broad banding" a stub-supported line is the reactance-cancelling arrangement shown in figure 2-22, wherein the characteristic impedance of the line is made lower for a distance of a quarter wavelength in either direction from the point of attachment of the insulating stub. A properly designed, stub-supported line using this type of broad banding can be used over a frequency range of at least plus-minus 15 per cent, while a simple stub-supported line ordinarily is useful over a frequency range of about 1 per cent.

"BAZOOKA" LINE BALANCE CONVERTER

In some applications of transmission lines it is desirable or necessary to feed a symmetrical or "balanced" load with an asymmetrical or "unbalanced" line, or vice versa. If the line is simply attached to the load in such cases, either the load may be unbalanced, or else the operation of the line may be disturbed, or both. The condition usually is typified by the presence of current on the outer surface of the outer conductor of a coaxial line (causing line radiation), and by an unsymmetrical distribution of voltage to ground and an inphase current component on a two-wire open line (causing increased line radiation). In addition to aggravating line radiation or "antenna effect" as it is called, the load on the generator may no longer be strictly balanced or strictly unbalanced, but intermediate between the two. Obviously these conditions are to be avoided in the interest of optimum performance of the overall system.

It is possible to transfer from a symmetrical (parallel conductor) line to an asymmetrical (coaxial) line or vice versa without detrimental effects by incorporation of a device called a *line balance converter* or *balun*. One widely used type which accomplishes this without effecting an impedance transformation is called a "bazooka."

The basic principle is illustrated in figure 2-23. The resonant "detuning sleeve" acts with the outer conductor of the enclosed section of line to form a shorted quarter-wave coaxial section and permits a high impedance to exist between points A and B, which means that point B is thusly isolated and need not be at ground potential even though the outer conductor of the coaxial line is grounded at a point or points below the detuning sleeve.

In effect the detuning sleeve acts as the outer conductor of the coaxial line for the last quarter-wavelength, and the regular outer conductor (which is an inner con-

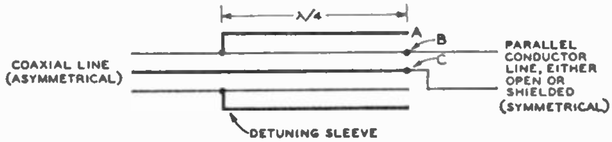


Figure 2-23.

"BAZOOKA" TYPE OF LINE BALANCE CONVERTER

A quarter-wavelength detuning sleeve is used to isolate the end of the outer conductor of a coaxial line so that it can be connected to a symmetrical (balanced) line or load without unbalancing it, and without causing undesirable currents to flow on the outside of the outer conductor of the coaxial transmission line (except within the detuning sleeve). No impedance transformation takes place; the ratio of line impedances is 1 to 1.

ductor so far as the sleeve is concerned) is left "floating," so that the impedance to ground will be determined by the nature of the load or line to which points B and C are connected. If they are connected to a symmetrical line, as illustrated, then they will be "nailed down" by the symmetrical line and will have equal impedance to ground.

The simple type of line balance converter shown provides very good performance over a narrow band of frequencies. Where broad band characteristics are required, more complicated converters are employed.

One important characteristic of this type of line balancer is that it provides a 1 to 1 ratio of impedance transformation. The system is perfectly matched when the impedance across the balanced parallel conductor line or load is equal to that across the unbalanced coaxial line or load.

PHASE INVERTER TYPE OF COMBINED LINE BALANCE CONVERTER AND IMPEDANCE TRANSFORMER

The fact that a half wavelength of line "repeats" the load but with an 180 degree phase reversal, and thus acts as a phase inverter, permits the simple form of line balance converter shown in figure 2-24. However, unlike the "bazooka" formed by the addition of a detuning sleeve, the phase inverter arrangement does not provide a 1 to 1 impedance ratio, but rather a 4 to 1 ratio. The latter is due to the fact that the same voltage appears across the asymmetrical or "unbalanced" line as appears between each element of the symmetrical or "balanced" line and ground, and vice versa, giving a voltage ratio of 2 and an impedance ratio of 4. As is the case with the simplest form of "bazooka," the phase inverter type of line balance converter exhibits optimum performance over but a narrow band of frequencies.

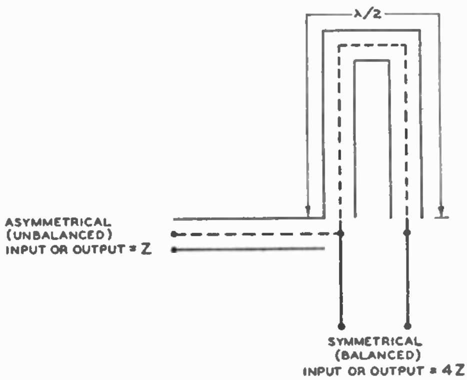


Figure 2-24.

PHASE INVERTER TYPE OF LINE BALANCE CONVERTER AND IMPEDANCE TRANSFORMER.

In this arrangement the phase inverting properties of a half wavelength of line are used to effect a line balance converter, but this time with an impedance transformation of 4 to 1. The latter is due to the fact that on the balanced line the same voltage appears between each conductor and ground as appears between the "hot" coaxial conductor and ground, thus giving a voltage step-up or step-down of 2, or an impedance transformation of 4. When computing the physical length of the phase inverting section, the velocity factor of the line must be taken into consideration.

"BAZOOKA" EQUIVALENT   PHASE INVERTER EQUIVALENT

Figure 2-25.

EQUIVALENT LUMPED CONSTANT CIRCUITS OF THE LINE BALANCE CONVERTERS ILLUSTRATED IN FIGURES 2-23 AND 2-24.

Figure 2-25 illustrates in terms of equivalent lumped constant circuits the difference between the two types of line balance converter discussed.

PRACTICAL TRANSMISSION LINES

When it comes to practical transmission lines and their physical characteristics, we find that there are numerous variations of the radio frequency transmission line differing in minor and major respects, but all exhibiting the same basic behavior. And usually when convenience, performance, overall economy, and other pertinent factors are considered it will be found that one particular type of line is best suited for a particular application.

All transmission lines can be classified as either *shielded* or *unshielded.* They also may be classified as *symmetrical* (also known as "balanced", or "push pull") or as *asmmetrical* (also known as "unbalanced" or "single ended"). The only qualification that must be made with this system of general classification is made necessary by the *semi-shielded* multi-wire line or "cage," in which a center conductor is surrounded by several wires which are grounded. While the line resembles an asymmetrical coaxial line more closely than it does a symmetrical parallel-conductor line, the shielding is poor compared to a true coaxial line.

While two-conductor parallel open lines are referred to as "balanced" lines and coaxial lines commonly are called "unbalanced" lines, this terminology is equivocal, because "balanced" and "unbalanced" also are employed (and more correctely so) to indicate *how a line is functioning* rather than a certain *type* of line. "Unbalance" is more correctly applied when used to indicate abnormal operation of a line. For this reason it is recommended that, widespread usage to the contrary, the reader adopt "symmetrical" and "asymmetrical" to designate the *type* of line and "balanced condition" and "unbalanced condition" to indicate how the line is *functioning* with regard to the presence of undesirable "antenna effect" (which is described elsewhere in this chapter).

SYMMETRICAL UNSHIELDED LINES

Lines in this category consist of parallel conductors, usually two but sometimes four or a higher even number, spaced sufficiently close to keep radiation and induction effects within tolerable limits, and symmetrically disposed and cross-connected with regard to polarity when four or more are employed.

THE TWO-WIRE OPEN LINE

One very common type of line in the "symmetrical, unshielded" category consists of two round conductors, usually wire but sometimes tubing, spaced from 1½ to 12 inches by means of the fewest number of ceramic, polystyrene, or other low loss "spacers" or "spreaders" which will maintain the spacing with satisfactory rigidity. This type of line can be designed to have the lowest losses of any practicable transmission line at frequencies below about 50 Mc., and it is inexpensive to construct. At higher frequencies the radiation becomes appreciable, but this type of line still finds some applications at frequencies up to approximately 300 Mc.

An unshielded line using two, identical, parallel conductors is generally known as a *two-wire open line,* even when the conductors happen to be tubing instead of wire. For "flat line" power transmission, conductors larger than no. 6 B&S gauge ordinarily are used only to provide mechanical rigidity or to permit a lower value of surge impedance for a given spacing, and not to decrease line losses (which are already very low) or to increase the power handling capability (which is already high). For a high Q resonant tank circuit, however, larger conductors are more appropriate. They also are more appropriate to a "Q section" transformer when a low value of surge impedance (175 to 300 ohms) is required.

The two-wire open line almost invariably is operated under balanced conditions, or at least an attempt is made to secure the best possible balance. When the line is long and runs near the earth or other object for considerable distance, the two wires sometimes are *transposed* every few feet in order to minimize unbalance caused by unequal capacity effects. Transposition of the conductors also is advisable when several open lines run parallel for some distance in close proximity; otherwise "cross talk" (mutual coupling) and line unbalance are likely.

Commercial installations in locations subject to sleet and icing often provide for running a heavy 60-cycle current through the line from a low voltage transformer, the current being high enough to heat the line sufficiently to melt the ice.

Surge Impedance of Two-Wire Open Line

The surge or characteristic impedance of a well constructed two-wire open line using a practical minimum of insulation (low loss) for support, and operated under balanced conditions (currents in the two wires equal in amplitude and 180 degrees out of phase), is given in figure 2-26. Very close spacing (center to center spacing only slightly greater than the conductor diameter) causes a non-uniform current distribution over the outer surface of each conductor, the current being greater on the adjacent portions of the conductors. The increased losses due to this effect (described earlier in this chapter under "proximity effect"), together with mechanical difficulties and voltage breakdown considerations, make it impracticable to obtain extremely low values of surge impedance with this type of line by resorting to very close spacing.

The values of surge impedance given in figure 2-26 are the "free space" values. Proximity to large objects such as the earth will lower the surge impedance, but the reduction in impedance is negligible unless the separation approaches the magnitude of the line spacing. It is assumed that in every case the line balance is maintained by preserving symmetry between the two line conductors and the juxtaposed object.

Attenuation of Two-Wire Open Line

The attenuation of a well constructed two-wire open line using low loss insulation and a practical minimum of it for support runs about as follows *when the line is operating under balanced conditions and does not run near any large objects.* For a 600-ohm line, operated under nonresonant conditions (VSWR not much over 1.0), the decibel loss per 100 feet will run about .018 times the square root of the frequency in megacycles for no. 8 B&S copper wire, and about .034 times the square root of the frequency in megacycles for no. 14 B&S copper wire.

These figures are based on actual measurements, and are somewhat higher than the theoretical losses. The figures are typical for clean insulators and dry weather; with damp weather and a dust and soot deposit on the insulators, particularly near the seashore, the losses will increase considerably. It will also be observed that above approximately 100 Mc. the losses go up faster than the square root of the frequency, because of radiation losses, unless very close spacing is used.

It is interesting to note that at frequencies low enough that radiation losses are not appreciable, only a well constructed, gas dielectric, large diameter coaxial line has losses as low as a well constructed, two-wire, open line using no. 12 copper wire.

Large commercial installations commonly use a 600-ohm line consisting of two no. 6 B&S copper wires spaced 12 inches to transmit many kilowatts of power at frequencies up to approximately 20 megacycles. Radio amateurs commonly use no. 12 or no. 14 B&S copper wire spaced either 6 or 4 inches for frequencies below 15 Mc., and spaced either 4 or 2 inches above 15 Mc.

Effect of Spacing Insulators; "Periodicity"

Since the spreader insulators and stand-off

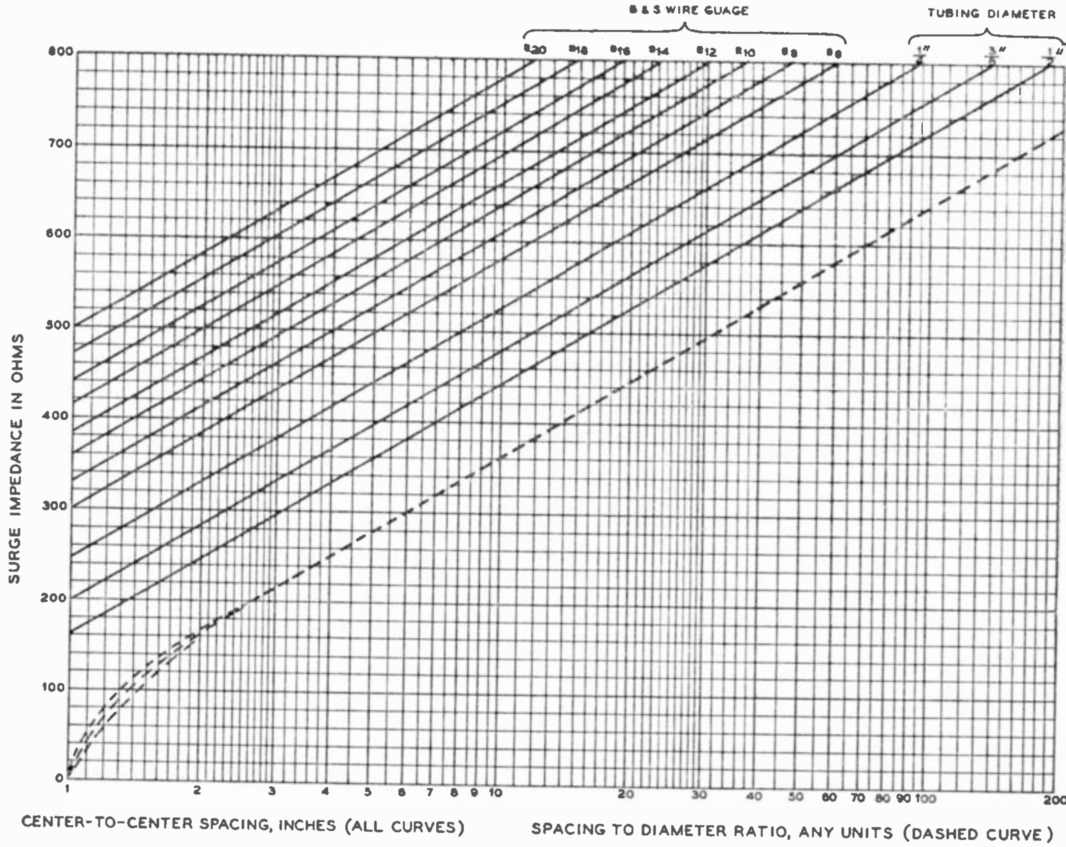


Figure 2-26.

SURGE IMPEDANCE OF TWO-WIRE OPEN LINE AS A FUNCTION OF CONDUCTOR DIAMETER AND CENTER-TO-CENTER SPACING. THE DASHED CURVE NOT ONLY APPLIES TO 1 INCH DIAMETER TUBING, BUT ALSO IS A UNIVERSAL CURVE APPLICABLE TO ANY UNIT OF MEASUREMENT, IN WHICH CASE IT APPLIES TO THE RATIO BETWEEN CENTER-TO-CENTER SPACING AND THE CONDUCTOR DIAMETER.

insulators have a higher dielectric constant than air, they constitute irregularities in the line constants. The higher the frequency, the bigger the "bump" in the line looks to the travelling wave. In other words, the reflection coefficient of a single insulator increases with frequency. At the lower frequencies, where there are many insulators per wavelength and the reflection coefficient of each insulator is negligible, the only effect of the insulators (other than to introduce a very slight dielectric loss) is to lower the velocity factor and impedance of the line slightly.

However, when the frequency is high enough that the reflection from one insulator no longer is insignificant, and the insulator spacing begins to be comparable to the wavelength, then other factors enter the picture. Under these conditions, reflection from the insulators makes it impossible to achieve a perfectly "flat" line, regardless of what is done at the termination. This increases the line loss and makes the line frequency sensitive.

A particularly bad condition is encountered when the insulators are spaced exactly an electrical half wavelength or multiple thereof apart. Under these conditions the reflections all add up in phase at the sending end of the line, producing the effect of a single discontinuity or irregularity of serious proportions. Another way of explaining the condition is as follows. We know that a section of line a multiple of an electrical half wavelength long "repeats" the load. Thus, under the aforementioned conditions all of the insulators are effectively in parallel across the line at the same point.

The avoidance of, or remedy for this condition is obvious. The insulator spacing should be so arranged that the reflections do not add up in phase. When the line must be used over a wide range of frequencies, the safest procedure is to stagger the spacing in random fashion. When the line is to be used at a single frequency an improvement over random spacing can be obtained by spacing the insulators by an electrical quarter wavelength, so that the reflections cancel. However, the latter arrangement need not be resorted to unless the reflection from each individual insulator is for some reason unusually high and a very flat line is necessary.

The undesirable condition whereby reflections from individual line discontinuities all add up in phase to produce a high SWR because of their regular spacing along the line is aptly called *periodicity*. It can occur in any type of transmission line at one or more frequencies unless precautions are taken to avoid it.

Effect of SWR on Attenuation

From figure 2-11 it is apparent that a line with a normal or flat line attenuation as low as that of a well constructed two-wire open line can be operated at a rather high SWR without the total line loss becoming serious, particularly if the total line length is short enough and the frequency low enough to keep the normal or flat line loss less than 0.1 db or so. And because a typical line consisting of no. 12 B&S copper wire spaced 6 inches will handle lots of current without overheating and lots of voltage without arcing over, it is possible to transmit considerable power over such a line even with a VSWR as high as 15 to 1.

Thus we see that a two-wire open line can handle fairly high power with tolerable attenuation even when the SWR is high. In fact, the two-wire open line is the best suited of all common transmission lines for use as a "resonant" line when both attenuation and power handling capability are considered.

Velocity Factor of Two-Wire Open Lines

Because the supporting and spacing insulators have a dielectric constant considerably higher than that of air; because of line losses; and because of various other factors; the velocity factor of a two-wire open line is slightly less that 1.0, though it generally is higher than that of any other type of transmission line. For typical lines the velocity factor will run between 0.95 and 0.98. With lines having large conductors, close spacing, and numerous insulators it will run close to 0.95; and with lines having small wires, large spacing, and few insulators it will run close to 0.98.

Radiation From and Pickup by a Two-Wire Open Line

The radiation from or pickup by a two-wire open line which has a spacing not exceeding a small fraction of a wavelength and is long in proportion to the spacing is, under *perfectly balanced*, flat line conditions, roughly equal to that of a doublet antenna having a length equal to twice the line spacing. Thus it can be seen that the amount of power radiated or picked up by the line is very low compared to that picked up by or radiated by the antenna to which the line is connected. In the transmitting case it can be ignored as an insignificant increase in line attenuation, but in the receiving case the small amount of pickup by the line may significantly degrade the discrimination of a highly directional receiving array, or significantly degrade the noise reduction obtained by placing the receiving antenna out of a high ambient noise field.

The radiation or pickup of the two-wire line will increase considerably if it is operated under "resonant" conditions with a high SWR, but still will not be excessive for most applications so long as the line is accurately balanced.

Radiation or pickup usually will be increased to a prohibitive magnitude, regardless of the SWR, if the line is operated under conditions of appreciable unbalance (instantaneous current not exactly equal and 180 degrees out of prase in the two conductors at every point on the line). It is for this reason that the condition is known as "antenna effect." The result of line unbalance is to produce an "in phase" component in the two wires, and this in-phase current causes the two wires to act as an antenna.

It should be noted that when the unbalance is due entirely to the phase relationship being other than 180 degrees, the current or voltage on the two wires will *read the same on a meter*, because while the *instantaneous* values will be different at a given point on the line, the r-m-s, average, and peak values will be the same in the two wires. Therefore, just because the current reads the same in the two wires on a meter it cannot be assumed that no in-phase component exists. However, phase unbalance is not so common as amplitude unbalance, and fortunately the latter type of unbalance can be detected by comparing the r-m-s, peak, or average current in the two wires at the same point on the line as read on the meter.

Some idea as to the importance of avoiding unbalance can be had from the fact that an in-phase component equal to only 25 per cent of the current flowing in one wire may, under some conditions, cause the line to radiate or pick up as much power as the antenna proper. The exact magnitude of the effect is dependent not only upon the degree of line unbalance, but also upon line length, spacing between the line and grounded objects or ground, and various other factors. In almost any practical case, however, the effect will assume objectionable proportions if the line is not accurately balanced.

Line unbalance can be caused by an unbalanced condition at either the load or the generator, or by not preserving symmetry of the two wires with respect to ground and other objects.

THE MULTI-WIRE OPEN LINE

It is possible to reduce both the surge impedance and the radiation (or pickup) considerably by going to four, six, or a greater *even* number of parallel conductors, so disposed and polarized as to observe geometrical symmetry. The most common version is the four-wire arrangement with diagonally opposite wires cross connected as shown in figure 2-27. The surge impedance for various combinations of conductor diameter and spacing is given in the same figure. The chief disadvantage of this type of line is the complication of maintaining uniform tension on all four wires when a change in direction is required.

MOLDED PAIR OR "RIBBON" PARALLEL LINE

Instead of using spacers to hold the two wires of an open wire line in position, it is possible to construct an open line using

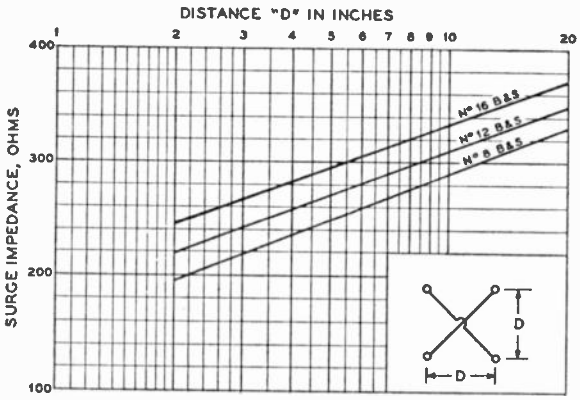


Figure 2-27.

SURGE IMPEDANCE OF FOUR-WIRE OPEN LINE FOR COMMON WIRE SIZES AND SPACINGS. SURGE IMPEDANCE USING NO. 10 OR 14 B&S MAY BE OBTAINED BY INTERPOLATION.

a continuous, flexible dielectric to position the wires. One widely used line of this type consists of two parallel conductors of stranded wire moulded into a spacing strip or ribbon of polyethylene. Polyethylene is a flexible, low-loss plastic which is described in detail later in this chapter under "Solid Dielectric Coaxial Cable." However, in the case of "ribbon" line it often contains an anti-oxidant which gives it a dark brown color.

Moulded pair line using polyethylene dielectric is sold in numerous power and impedance ratings, and under various trade names. In many cases the same or nearly the same style of line is available under two or more trade names. While ribbon line is available in 75, 100, 150, 200 and 300 ohms characteristic impedance, the following discussion will be devoted to what, at the time of this writing, is the most widely available and commonly known line of this description, namely, "Twin-Lead."

Twin-Lead is available in a "receiving" size in three surge impedances: 300 ohms, 150 ohms, and 75 ohms, the only significant physical difference in the three lines being in the spacing between conductors. It also is available in a "transmitting" size in 75 ohm surge impedance. Though the transmitting size uses much larger conductors than the receiving size, the attenuation characteristics are improved only slightly. However, it will handle much more power without damage (over 1 kilowatt continuously at 30 Mc.).

The 300 ohm receiving Twin-Lead line is widely used by radio amateurs for transmitting purposes, and stands up satisfactorily under intermittent service at powers up to about 500 watts at 30 Mc. and about 200 watts in the upper v-h-f range *when the VSWR is kept low.* 150-ohm Twin-Lead will stand slightly less. The 75-ohm receiving Twin-Lead will stand about 100 watts up to approximately 14 Mc. if the VSWR is low.

The chief disadvantage of using the 300-ohm receiving type line for transmitting purposes is that the surge impedance and velocity factor are lowered considerably when the line becomes wet, causing variations in the loading at the transmitter. The normal or "flat-line" loss goes up only slightly when the line is wet if there is no soot or salt deposit on the line, but under some conditions the lowered surge impedance may upset the functioning of the antenna system to the extent of an appreciable increase in the total attenuation of a *long* line, due to increased SWR. If the line is comparatively short and is clean, the only significant change noticed in operation will be a change in loading at the transmitter, and this can be taken care of by a slight adjustment in tuning.

If the range in frequency over which the transmitter works is narrow and the distance between transmitter and antenna is short in terms of wavelength (as often is the case below 4 or 5 Mc.), then the effects of moisture on the transmitter loading usually can be considerably reduced by making the line exactly one, two, or three (the fewer the better) electrical half wavelengths long. Because such a line "repeats" the load, a moderate change in surge impedance will have but little effect at the generator end of the line, assuming of course that the characteristics of the antenna proper do not change appreciably when it becomes wet. The only reason that this procedure does not virtually eliminate the change in loading due to shifting surge impedance is that moisture on the surface of the line affects not only the surge impedance but also the velocity factor and electrical length of the line as well.

The effects of moisture on 300 ohm Twin-Lead can be considerably reduced by applying a film of Simoniz Wax or Silicone waterproofing compound. The 75-ohm type line is not significantly affected by water clinging to it; the surge impedance and velocity factor will remain substantially constant. However, a heavy coating of ice will produce some change in surge impedance. The effect of water upon 150-ohm Twin-Lead is not nearly so pronounced as in the case of 300-ohm Twin-Lead, but waterproofing compound still can be employed to advantage.

The velocity factor of Twin-Lead transmission line (dry) runs about 82% for the 300-ohm line, about 77% for the 150-ohm line, about 68% for the receiving type 75-

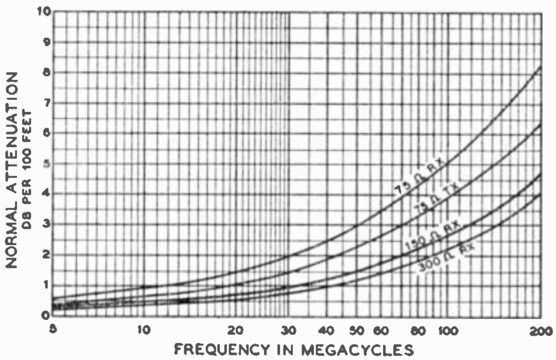


Figure 2-28.

NORMAL OR "FLAT-LINE" ATTENUATION CURVES FOR "TWIN-LEAD" PARALLEL LINE WHICH IS REASONABLY CLEAN AND DRY.

The loss for any VSWR of less than 2.0 is not appreciably greater than the "normal" loss obtained from the curves. When the VSWR is greater than 2.0, determine the "normal" attenuation and then refer to figure 2-11.

ohm line, and about 71% for the transmitting 75-ohm line.

The normal or "flat-line" attenuation for the four types of Twin-Lead when reasonably clean and dry is given as a function of frequency in figure 2-28. For a VSWR appreciably above 1.0 the attenuation may be determined by applying a conversion factor obtained from figure 2-11.

The 300-ohm Twin-Lead matches the input impedance of many television receivers and is ideally suited to the purpose because of the small field of the closely spaced wires and the lack of "bumps" in the line. The latter prevents a surge impedance which is frequency sensitive at v-h-f and u-h-f, thus avoiding one possibility of distortion on a television screen.

Twin-Lead parallel conductor line qualifies as an "open wire" line in that it is unshielded. However, through loose usage the term "open wire" often is employed to indicate specifically a line which uses air as the dielectric (except for well-separated spacers), as contrasted to a closely spaced line which depends upon a continuous rubber or plastic dielectric to maintain the spacing. Therefore, to avoid ambiguity, it is recommended that the ribbon or "Twin-Lead" type of line be classed as a "moulded pair" line. Thus any closely-spaced, balanced line using a continuous, solid dielectric for spacing can be classed either as a "moulded pair," a "twisted pair," or a "shielded pair" as appropriate. The latter two types will be dealt with subsequently.

TWISTED PAIR LINE

Ordinary rubber covered twisted wire, such as "yellow and green" lamp cord, or outdoor telephone wire, can be used as an r-f transmission line at the lower frequencies. But the losses are high, particularly if the line is wet, unless the line length is kept quite short. The loss for such wire (dry) is about 3 db per hundred feet at 5 Mc. and about 12 db per hundred feet at 30 Mc. The loss may be tolerable for some *receiving* applications below 15 Mc., because usually the atmospheric noise delivered by a good antenna is much greater than the set noise at these frequencies. The surge impedance for such lines is about 100-150 ohms, the exact value depending upon the particular type of line and whether it is wet or dry.

The losses and the surge impedance both can be reduced somewhat by improvising a line of no. 12 rubber covered wire (of the type used for house wiring) twisted together by hand.

A commercially manufactured line (commonly known as EO-1 cable), using no. 12 solid copper conductors and a high quality of rubber dielectric, twisted together and covered with an outer weatherproof wrap, is useful for many transmitting applications where neither frequency nor line length is excessive. This cable has a characteristic impedance of 70 ohms, and a normal (flat line) attenuation of about 0.7 db per hundred feet at 5 Mc. and about 3 db per hundred feet at 30 Mc. If the ends are carefully sealed, the attenuation of this line goes up only very slightly and the surge impedance down only slightly when the line becomes wet. It is relatively inexpensive, easy to install (as it need not be spaced from metallic objects), and ideally suited for such applications as link coupling to a nearby antenna tuner or matching device at frequencies below 30 Mc.

Any type of rubber-dielectric twisted-pair line should be operated with the lowest possible SWR, as the efficiency goes down rapidly as the SWR is increased. The power handling capability also is greatly reduced by a high SWR. Type EO-1 cable will handle approximately 1 kilowatt at frequencies up to 30 Mc. if the VSWR is kept below 1.5.

Because of the close spacing of the conductors, radiation or pickup from a twisted pair line operating under perfectly balanced conditions is negligible at any frequency at which the use of such line is otherwise practical.

The velocity factor of twisted pair lines runs from about 0.55 to 0.65, depending upon the characteristics of the dielectric material and upon the method of fabrication (number of twists per foot, etc.).

ASYMMETRICAL SHIELDED LINES

Lines in this category consist of two concentric, coaxial conductors which are spaced by a dielectric, the outer tubular conductor confining the field and therefore effectively shielding the inner conductor. Such lines are called *coaxial lines* for obvious reasons. While there are several specialized types which are used for special applications (such as video, delay, or pulse work), there are only two basic types which are commonly used as antenna "feeders." These are the gas dielectric (bead or stub supported) type, and the semi-flexible, solid-dielectric type.

Before proceeding to practical examples of and various considerations pertaining to the two types, one general consideration regarding proper and improper operating conditions which is applicable to both types of line will be discussed.

Line Balance in Asymmetrical Lines

When a coaxial line is operating properly, the r-f current flowing on the outer portion of the inner conductor is equal to the r-f current flowing on the *inner surface* (!) of the outer conductor. Because the two currents are equal, the line may be said to be operating under balanced conditions. Note that under this condition there is *no current flowing on the outside of the outer conductor;* in other words, the field is completely contained within the line, and there is no radiation from or pickup by the line. There is no "antenna effect."

Unless this condition obtains, one of the chief advantages of the coaxial line—its ability to operate with no significant external field—is not being exploited. It is possible to reduce antenna effect to an insignificant magnitude for one particular frequency, regardless of the load and generator characteristics, simply by the judicious application of detuning stubs and line balance converters (described previously in this chapter), and by careful orientation of the line with respect to the antenna. If the line is effectively detuned at the antenna end, and if the inductive coupling between the line and radiating portions of the antenna is kept low, the "antenna effect" will be negligible. To minimize inductive coupling between the outer conductor of the line and the radiating portions of the antenna, the line preferably should be brought away from the radiating system in a symmetrical manner for as great a distance as possible, so that the significant forces acting on the line to produce antenna effect by induction will tend to cancel.

Because "balanced" is synonymous with "symmetrical" to many people, it is safer to refer to such a coaxial line as operating under "normal" conditions, or "without antenna effect," rather than to refer to the operation as "balanced."

Antenna effect, when present, produces exactly the same undesirable effects as does an in-phase component on an unshielded two-wire line. Its presence can be detected in transmitting applications by inductively coupling a sensitive current indicating device to the outer conductor of the line at various points along the line. In the case of receiving applications, its presence can be detected by inductively coupling a high-output signal generator to the outer conductor of the line at various points.

Coaxial lines, operating without antenna effect, are better suited than two-wire unshielded lines for applications where juxtaposed lines may give spurious coupling or cross talk trouble; for use at frequencies so high that an open line does not provide adequate restriction of the field; for applications requiring complete impunity to all types of weather (including sleet and ice) or requiring that the line be run under ground; and to all applications utilizing a low impedance, asymmetrical or "single-ended" generator and load.

It should be noted that, in the receiving case, it is not possible to make the outer conductor immune to induced currents resulting from intercepted radiation unless the line is grounded continuously throughout its length. Such currents, however, do not constitute antenna effect *unless they are coupled into the receiver.* No matter how well the outside of the outer conductor is isolated from the antenna proper, the line still will act as a receiving antenna unless it is buried or continuously grounded throughout its entire length, an obviously impracticable procedure. If proper isolation is used at the receiver in the method of coupling the line to the receiver, these induced currents are not transferred to the input circuit of the receiver and therefore produce no detrimental effects, and are not considered a manifestation of antenna effect.

COAXIAL LINES USING GAS DIELECTRIC

The ideal coaxial line would employ no spacing insulators to position the inner conductor, thus permitting a continuous vacuum or gas dielectric. But in a practical gas-filled line some means must be employed for positioning the inner conductor, the usual arrangement being a succession of captivated "spacers," "washers," or "beads" of glazed, low-loss ceramic material. Dry air or nitrogen under pressure is better than a vacuum from a practical standpoint, because in the case of a vacuum a minute leak in the line would permit moisture to enter, while in the latter case it would not.

Thus, the lowest loss practical line consists of two coaxial copper tubes, with the inner one positioned by isolantite or similar spacing washers, and the whole line made gas tight and filled with gas under pressure. One or more gas inlet fittings, and, when a long, straight run is used, one or more expansion joints complete the necessary features.

This is the most efficient of all practical coaxial lines, but is much more expensive and difficult to install than the semi-flexible solid-dielectric type line which will be described later. Also, it obviously is practicable only in a permanent installation. In spite of the greater cost, the line is economical in the case of an AM or FM broadcast station, because the carrier power in watts is high, the dollar investment and running expense per watt is high, and a db saving in transmission line loss therefore represents a considerable dollar saving when figured over a period of several years.

It is for these reasons that the gas filled line (including dry air) has become more or less standard practice in the AM, FM and television broadcast station fields. In the FM and television field the characteristic impedance has been standardized at a nominal 50-ohm impedance (actually 51.5 ohms). In the AM broadcast field the standard characteristic impedence long has been 70 ohms, but in some of the new installations 50-ohm line is being employed, particularly where the station operates both an AM and an FM or television transmitter. The one type of line can be used for both installations, thus affording economy and simplification in the stocking of spare fittings and accessories. The new, commercially available 50-ohm line is designed to minimize discontinuities in the line characteristics; it exhibits very low reflection at frequencies up through 200 Mc., and tolerable reflection at still higher frequencies.

The standard line diameters (outer conductor) are ⅞; 1⅝; 3⅛; and 6⅛ inches.

Power Handling Capability

As the line diameter is increased, the

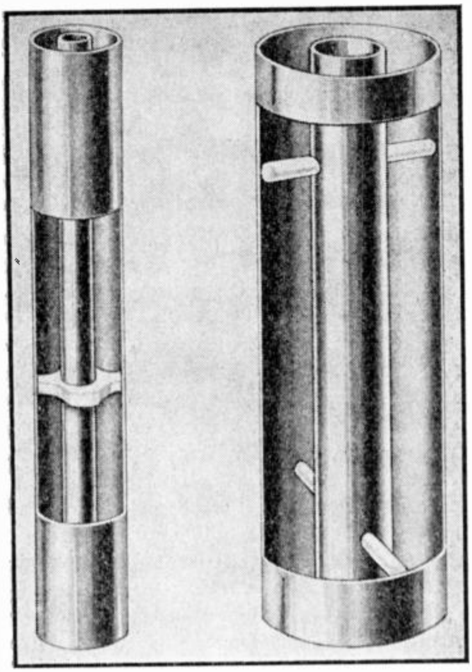


Figure 2-29.

CUT-AWAY VIEW SHOWING INTERIOR CONSTRUCTION OF TWO SIZES OF GAS DIELECTRIC COAXIAL LINE OF COMMERCIAL MANUFACTURE (ANDREW CO.).

power handling capability goes up and the attenuation goes down. The power handling capability for the various sizes of line is aproximately as follows for a frequency of 100 Mc. and a VSWR of 1.5:

⅞" dia.	2 Kw.
1⅝" dia.	7 Kw.
3⅛" dia.	28 Kw.
6⅛" dia.	110 Kw.

For a perfectly flat line (VSWR equal to 1.0) the power handling capability is increased about 50 per cent, while a VSWR of 2.0 reduces the power handling capability to about 75 per cent of the values given above. The power handling capability for a VSWR less than 2.0 is determined by maximum safe temperature rise, and not by breakdown voltage.

Attenuation

The normal or flat-line attenuation of the various sizes of line at a frequency of 100 Mc. is approximately as follows, the loss being expressed in db per hundred feet:

⅞" dia.	0.45 db
1⅝" dia.	0.25 db
3⅛" dia.	0.15 db
6⅛" dia.	0.06 db

The db loss increases approximately 50 per cent at a frequency of 200 Mc. By referring to figure 2-11 it will be seen that the total attenuation in db for typical line lengths goes up only very slightly for any VSWR not exceeding 2.0.

Unless the line length is quite short, the optimum or most economical line diameter for a given power is determined by line attenuation rather than by power handling capability. This is particularly true above 100 Mc.

"Prefabricated" line, including all necessary elbows, solderless couplings, expansion joints, gas inlet fittings, pressurizing equipment, etc. is available commercially at such reasonable prices that there no longer is any justification for the old practice of individual fabrication of the line from basic components such as spacing washers and bulk tubing, using basic metal working procedures such as sawing and soldering. Examples of prefabricated line material are illustrated in figures 2-29 and 2-30.

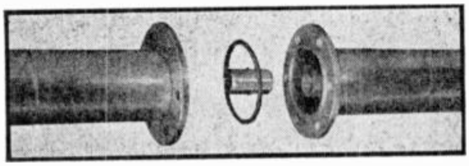


Figure 2-30.

ILLUSTRATING GAS TIGHT, FLANGED CONNECTOR WHICH PERMITS RAPID FABRICATION OF A COAXIAL LINE FROM STANDARDIZED 20-FOOT SECTIONS AND FITTINGS (ANDREW CO.).

Either dehydrated air or oil-pumped nitrogen, under a pressure of 10 lbs. or more, will satisfactorily prevent moisture from getting into the line, or condensing on the inside of the line as a result of a temperature change. A leak is evidenced by a gradual loss in the gas pressure. Higher pressures permit earlier detection of any leak which might develop in service.

Expansion joints can be avoided in long runs of the ⅞ inch size line by arranging it in a sinuous path. Sufficient curvature to avoid expansion troubles increases the length of the line required to reach between two points by only a few per cent. The method is not practicable for the larger diameter lines.

It is very important that a gas filled line never be operated with a short or without a termination, even for an instant, because the Q is high enough to build up terrific voltages under such conditions, even though the power fed to the line may be quite low. The resulting arc usually will "burn" one or more of the spacing washers or beads and make it necessary to repair or replace any section of line in which an arc occurs.

The velocity factor of gas-filled coaxial lines varies widely with the amount of ceramic material used to position the inner conductor, and may be anything from about 0.83 to 0.94.

SOLID DIELECTRIC COAXIAL CABLE

A standardized series of semi-flexible coaxial cables (designated RG/U) using solid polyethylene as a dielectric is commercially available for applications requiring a line which will stand limited twisting or flexing, or for applications where additional attenuation can be tolerated in the interest of ease of installation and virtual elimination of maintenance responsibility and work.

Polyethylene has excellent characteristics for the purpose. Both the dielectric constant and power factor are very low even in the v-h-f/u-h-f range, giving a loss factor that compares favorably with polystyrene. Polyethylene remains flexible at minus 40 degrees F., and will withstand temperatures up to 180 degrees F. safely when used in this application. It is mechanically stable and very inert, there being no known solvent at temperatures ordinarily encountered.

Except for a special "lossy" line which is used as an attenuator, the center conductor consists of either solid or stranded wire which is either solid copper or copper-clad steel. The outer conductor consists of either a bare or plated copper braid, either a single or double layer, woven so as to have minimum r-f resistance.

The braid usually is covered with a Vinylite plastic jacket. This material is virtually immune to sunlight and impervious to and insoluble in any materials such as oil or water which might come in contact with the jacket in ordinary use. It therefore forms a tough, abrasive-resistant jacket which is waterproof and shows good aging qualities when exposed to weather. Like polyethylene, Vinylite is a thermoplastic, and softens at approximately the same temperature as polyethylene. Some cables intended for military use in exposed or vulnerable locations are provided with an additional sheath or jacket of steel braid. These cables are designated as "armored," and otherwise have the same characteristics as the more common, unarmored cables.

Prior to 1944, many r-f cables were manufactured which resemble the present "RG/U" series of polyethlene dielectric cables except that the dielectric material was *not* polyethylene and had inferior characteristics, the most important being that the attenuation increased rather rapidly with the age of the cable. Polyethylene may be distinguished by the fact that the unpigmented variety ordinarily used in coaxial cables is grey, translucent, melts rather sharply at about the temperature of boiling water, and does not become tacky or rubbery at temperatures approaching the melting point. Upon solidifying after being melted, it resembles the original material in every respect.

Because both polyethylene and Vinylite are thermoplastic, a very neat and effective splice can be made in cables using these materials by fusing together the butt ends of first the polyethylene and then the Vinylite by heating to between 225-235 degrees F. Special jigs and moulds are available which greatly facilitate the splicing operation.

Eccentricity of the inner conductor and variations in diameter of the cable components are held to very close tolerances. The result is that a long, uninterrupted run ex-

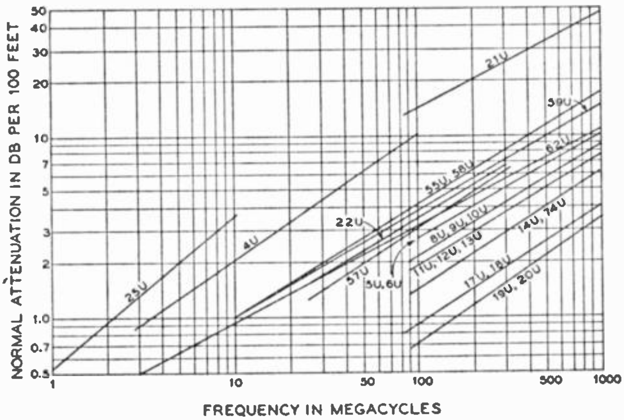


Figure 2-31.

NORMAL OR "FLAT-LINE ATTENUATION OF POPULAR RG/U RADIO FREQUENCY CABLES AS A FUNCTION OF FREQUENCY.

hibits less reflection than does a gas-filled line. The difference is most noticeable in the u-h-f range, where the ceramic spacing washers in a gas-filled line constitute an appreciable discontinuity. It is for this reason that the characteristic impedance of one of the RG/U series solid-dielectric r-f cables is much more constant with frequency than is the impedance of a gas-filled line, the difference of course being most noticeable at the higher frequencies. Because of the care taken in manufacture, "periodicity" effects are insignificant in these solid dielectric lines at all frequencies for which the characteristics are otherwise acceptable.

The basic characteristics of the most common coaxial r-f cables in the RG/U series are given in the accompanying tables (pages 126 and 127) and in figure 2-31.

A wide assortment of connectors, couplings, adapters, elbows, junctions, etc. for the RG/U series of cables is available, many of them in either regular or waterproof type. Those designed for u-h-f use have a very low reflection coefficient, and even at frequencies as high as 1000 Mc. several may be used in a line without producing serious reflection.

The smaller diameter cables (such as RG-8/U, RG-58/U, etc.) are quite flexible, and will stand considerable twisting and flexing without giving trouble. The larger diameter cables (such as RG-19/U) are more properly classified as semi-flexible. They can be bent permanently on a radius equal to 10 times the outer diameter of the cable during installation without damage or risk of future trouble, but should not be used where there is continual twisting or bending of considerable magnitude.

The RG/U series of coaxial r-f cables is divided into two main impedance families: those with a nominal surge impedance of 50 ohms (actually slightly higher) and those with a nominal surge impedance of 70 ohms (actually slightly higher). The proper choice of cable depends upon the frequency, length of line, power to be handled, the desired impedance, and the degree of shielding required. While a single braid is adequate for most applications, it is impossible for a woven braid to provide as thorough shielding as does a solid copper tube. For those few exacting applications for which a single braid does not offer adequate shielding, double braid cable such as RG-9/U, RG-14/U, etc. will be found highly satisfactory.

The velocity factor of solid-dielectric coaxial cables using polyethylene runs close enough to 0.66 that this figure may be used for all purposes not requiring a highly precise figure.

It should be kept in mind that not all of the cables in the standardized RG/U series of cables are coaxial r-f cables using polyethylene as a dielectric and intended for use in conjunction with antennas. This standardized cable series includes special lossy cable, low capacitance cable, twin con-

COMMONLY USED STANDARD ARMY-NAVY RADIO FREQUENCY CABLES

class of cables		Army-Navy type number	inner conductor	dielec mate-rial (1)	nominal diam of dielectric (in)	shielding braid	protective covering	nominal overall diam (in)	weight lb/ft	nominal imped-ance ohms	nominal capaci-tance $\mu\mu f/ft$	maximum operating voltage rms	remarks
50–55 ohms	Single braid	RG–58/U	20 AWG copper	A	0.116	Tinned Copper	Vinyl	0.195	0.025	53.5	28.5	1,900	General purpose small size flexible cable
		RG–8/U	7/21 AWG copper	A	0.285	Copper	Vinyl	0.405	0,106	52.0	29.5	4,000	General purpose medium size flexible cable
		RG–10/U	7/21 AWG copper	A	0.285	Copper	Vinyl (non-contaminating) armor	(max) 0.475	0.146	52.0	29.5	4,000	Same as RG–8/U armored for naval equipment
		RG–17/U	0.188 copper	A	0.680	Copper	Vinyl (non-contaminating)	0.870	0.460	52.0	29.5	11,000	large high power low attenuation transmission cable
		RG–18/U	0.188 copper	A	0.680	Copper	Vinyl (non-contaminating) armor	(max) 0.945	0.585	52.0	29.5	11,000	Same as RG–17/U armored for naval equipment
		RG–19/U	0.250 copper	A	0.910	Copper	Vinyl (non-contaminating)	0.120	0.740	52.0	29.5	14,000	Very large high power low attenuation transmission cab'e
		RG–20/U	0.250 copper	A	0.910	Copper	Vinyl (non-contaminating) armor	(max) 1.195	0.925	52.0	29.5	14,000	Same or RG–19/U armored for naval equipment
	Double braid	RG–55/U	20AWG copper	A	0.116	Tinned copper	Polyethylene	(max) 0.206	0.034	53.5	28.5	1,900	Small size flexible cable
		RG–5/U	16 AWG copper	A	0.185	Copper	Vinyl	0.332	0.087	53.5	28.5	2,000	Small microwave cable
		RG–9/U	7/21 AWG silvered copper	A	0.280	Inner—silver coated copper. Outer-copper	Vinyl (non-contaminating)	0.420	0.150	51.0	30.0	4,000	Medium size, low level circuit cable

Notes:
1. Dielectric materials
   A Stabilized polyethylene

(COURTESY FEDERAL TELEPHONE AND RADIO CORPORATION)

COMMONLY USED STANDARD ARMY-NAVY RADIO FREQUENCY CABLES (Continued)

class of cables		Army-Navy type number	inner conductor	dielec material (1)	nominal diam of dielectric (in)	shielding braid	protective covering	nominal overall diam (in)	weight lb/ft	nominal impedance ohms	nominal capacitance $\mu\mu f/ft$	maximum operating voltage rms	remarks
70–80 ohms	Single braid	RG–14/U	10 AWG copper	A	0.370	Copper	Vinyl (non-contaminating)	0.545	0.216	52.0	29.5	5,500	General purpose semi-flexible power transmission cable
		RG–74/U	10 AWG copper	A	0.370	Copper	Vinyl (non-contaminating) armor	0.615	0.310	52.0	29.5	5,500	Same as RG–14/U armored for naval equipment
		RG–59/U	22 AWG copperweld	A	0.146	Copper	Vinyl	0.242	0.032	73.0	21.0	2,300	General purpose small size video cable
		RG–11/U	7/26 AWG tinned copper	A	0.285	Copper	Vinyl	0.405	0.096	75.0	20.5	4,000	Medium size, flexible video and communication cable
		RG–12/U	7/26 AWG tinned copper	A	0.285	Copper	Vinyl (non-contaminating) armor	0.475	0.141	75.0	20.5	4,000	Same as RG-11/U armored for naval equipment
	Double braid	RG–6/U	21 AWG copperweld	A	0.185	Inner—silver coated copper. Outer—copper	Vinyl (non-contaminating)	0.332	0.082	76.0	20.0	2,700	Small size video and I-F cable
		RG–13/U	7/26 AWG tinned copper	A	0.280	Copper	Vinyl	0.420	0.126	74.0	20.5	4,000	I-F cable
Cables of special characteristics	Twin conductor	RG–22/U	2 Cond. 7/18 AWG copper	A	0.285	Single—tinned copper	Vinyl	0.405	0.107	95.0	16.0	1,000	Small size twin conductor cable
		RG–57/U	2 Cond. 7/21 AWG copper	A	0.472	Single—tinned copper	Vinyl	0.625	0.225	95.0	16.0	3,000	Large size twin conductor cable
	High attenuation	RG–21/U	16 AWG resistance-wire	A	0.185	Inner—silver coated copper. Outer—copper	Vinyl (non-contaminating)	0.332	0.087	53.0	29.0	2,700	Special attenuating cable with small temperature coefficient of attenuation

(COURTESY FEDERAL TELEPHONE AND RADIO CORPORATION)

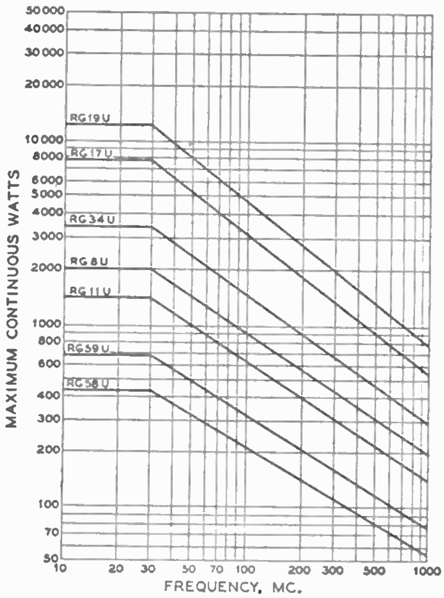


Figure 2-32.

MAXIMUM SAFE CONTINUOUS POWER HANDLING CAPABILITY OF COMMON RG/U COAXIAL CABLES UNDER FLAT LINE CONDITIONS.

The curves are based on an ambient temperature of 140 degrees F. Double shielded cables and armored cables will handle the same power as the above prototypes. When the VSWR exceeds 1.0, the power handling capability of the cable will be reduced approximately in inverse proportion to the VSWR.

ductor shielded cable, highly flexible cable, pulse cable, and other special r-f cables. Only those having antenna applications have been included in the abbreviated list given in the accompanying tables.

SYMMETRICAL SHIELDED LINES

For certain very exacting applications which entail a balanced generator and load, it sometimes is found that neither a symmetrical unshielded line nor an asymmetrical shielded line is entirely satisfactory from the standpoint of antenna effect. This is particularly true in direction finder work, where absolutely no antenna effect can be tolerated and the frequency range involved often is quite extensive. For such work two solid-dielectric coaxial cables sometimes are used as an integral pair. Special "dual coaxial" cables consisting of two coaxial cables in a single outer weatherproof sheath are available in some sizes, making it unnecessary to lash two separate cables together for use as a single symmetrical cable or pair.

Also available are symmetrical, shielded cables variously designated as "Twinax" (a trade name), "twin conductor shielded," "shielded pair," and "dual shielded." These resemble regular solid-dielectric flexible "coax" of the RG/U type except that *two* inner conductors are employed instead of one, the two inner conductors being spaced from each other and symmetrically disposed with respect to the outer shield. When a balanced generator and load are connected to the two inner conductors and the outer shield is grounded, a system is realized which is highly invulnerable to antenna effect over a wide frequency range and therefore exhibits virtually zero line radiation and pick-up except under unusually unfavorable conditions. Typical examples of this type of cable are RG-22/U and RG-57/U.

ASYMMETRICAL UNSHIELDED LINES

While it is possible to operate a two-wire open line with one conductor grounded, so as to permit use with an asymmetrical generator and load without resorting to a line balance converter, the resulting "antenna effect" ordinarily is so great under these conditions as to make such operation unfeasible. However, if instead of one grounded wire *several* wires are disposed around the "hot" wire in such a manner as to form a sort of cage, the radiation or pickup is greatly reduced. While not as good as that of a regular coaxial line utilizing a woven braid or solid tube for the outer conductor, the shielding is sufficiently good for some purposes, and such lines sometimes are employed below 1500 kilocycles for transmitting applications where high power is involved, thus avoiding the expense of a comparatively costly gas-filled line. Such a line is called a "semi-shielded" or "cage" coaxial line.

A single-wire well removed from the

earth and large objects has a characteristic impedance of about 500 ohms when worked "against ground" (meaning that the earth is used as the return circuit). But when such a wire is used as a transmission line, it represents the ultimate in undesirable "antenna effect," the radiation or pick-up often being comparable to or even greater than that of the antenna to which it is connected. Just because it is possible to terminate such a line in a manner which results in a very low SWR on the line does not mean that the line will not serve as an excellent antenna. As will be seen in a later chapter, the rhombic antenna array—one of the most effective ever conceived—uses radiating elements which show very little cyclic change in r-m-s current amplitude along their length.

Although the single-wire line enjoyed considerable vogue at one time as a simple method of feeding a nondirectional high-frequency antenna, it now is seldom used except where a hastily improvised method of feed is required for a temporary installation.

MATCHING THE TRANSMISSION LINE TO THE GENERATOR AND LOAD

Because the considerations involved in coupling or matching a transmission line to an antenna are so closely tied up with the operation of the antenna, a discussion of coupling and matching methods will be deferred until after an elementary treatment of basic radiating systems. In fact, in the case of certain overall systems it is difficult to distinguish between phasing or matching elements of the antenna, and the transmission or feeder line proper. The only logical treatment in such cases is to consider the system as a whole.

For the same reason, transmission line measurements and adjustments will be treated later in the book.

CHAPTER THREE

Basic Antenna Theory

In the preceding two chapters it was shown how any conductor carrying a radio frequency current will radiate energy into space unless the conductor is either shielded or effectively neutralized by a closely-spaced, similar conductor carrying a current which is equal in amplitude and opposite in phase. In the latter case, the radiation never can be *completely* neutralized, but if the spacing between radiators is a sufficiently small fraction of a wavelength, the radiation is insignificant and may be ignored. Obviously to neutralize radiation completely the fields produced by the two wires would have to be exactly 180 degrees out of phase at all points in space, a physical impossibility because the two wires would have to occupy exactly the same position in space.

FICTITIOUS REFERENCE ANTENNAS

The ficticious "isotropic radiator" often referred to in conjunction with antenna problems is not physically realizable, because to radiate equally well in all directions the radiator would have to radiate waves with longitudinal polarization instead of transverse polarization. Because of the very nature of electromagnetic waves, this is not physically possible. The isotropic radiator is simply a convenient reference antenna which, though it cannot actually exist, makes a logical basis of comparison for the directional characteristics of practical antennas.

Another hypothetical radiator (or receptor) often used as a reference antenna or basis of comparison is the *elementary or infinitesimal dipole,* either electric or magnetic. The former may be visualized as simply a pair of conducting spheres (having capacitance) connected by a linear filamentary conductor (having inductance), all dimensions being *minutely small* in terms of wavelength. The elementary *magnetic* dipole may be visualized as a circular loop of wire, minutely small in terms of wavelength. By convention, the unqualified term "elementary dipole" refers to an elementary *electric* dipole.

The *elementary doublet* is another hypothetical radiator (or receptor) often referred to. It resembles the elementary electric dipole with the two spheres removed but their capacity retained; in other words, it is simply a straight wire which is minutely short in terms of wavelength and has uniform current distribution.

The elementary doublet, because of its minute size and negligible radiation resistance, would be found a very inefficient radiator were it possible to construct one and feed power to one. But even though the radiation resistance of a single elementary dipole is negligible, a large enough number working together can acquire the characteristics of an efficient radiating system. Therefore we find that the hypothetical elementary doublet is not only a convenient reference antenna, but also a convenient analysis mechanism or tool in the study of practical radiating systems. If the current distribution of an antenna is known, the radiation field can be determined for any point in space by considering the radiating system to be made up of a series of elementary doublets, then calculating the contribution of each doublet, and finally adding the separate fields by means of graphical integration (with due regard to phase and polarization).

To avoid confusion, it should be stated that the expressions "dipole" and "doublet," without the qualifying "elementary" or "infinitesimal," often are used to refer to actual, practical antenna systems which will later be described in detail.

RECIPROCITY

The directional characteristics of an antenna (ignoring the feed line for the moment) are the same regardless of whether it is used for receiving or for transmitting energy.

The only time an antenna *system* does not exhibit the same directivity when receiving as it does when transmitting is when the antenna consists of phased elements, asymmetrically fed, and the antenna and receiver do not present the same terminal impedance to the transmission line. As this is a special case, seldom encountered, antenna directivity will for the time being be considered solely from the standpoint of radiation, in order to simplify the discussion. The reader need only keep in mind that, with the one exception noted, everything regarding radiation directivity applies "in reverse" to the directivity of a receiving antenna as well.

Although the classical reciprocity laws must be applied with care to practical antenna problems involving electromagnetic radiation, it is safe to assume that the directivity of an antenna system is the same when transmitting as when receiving if the terminal impedances are maintained constant, and that *generally speaking* a good transmitting antenna makes a good receiving antenna system.

"FREE SPACE" CONDITIONS VS. PRACTICAL INSTALLATIONS

The characteristics of any radiator are affected by the presence of the earth and other objects, including other nearby antennas. The characteristics of a simple radiator vary considerably with height above ground, for instance. Therefore, the only logical recourse is to separate the basic charactristics of the antenna from the effects produced by the presence of other objects, and then to evaluate or make allowances for the presence of other objects as fit the circumstances. This is done by defining the basic characteristics of an antenna as those which the antenna would exhibit in *free space.* These are called the "free space" characteristics, or sometimes just "space" characteristics. Obviously this does not represent a condition to be encountered in practice, but it does give us a common "starting point."

In the case of certain antennas which are actually *connected* to the earth, and utilize the earth as part of the oscillatory circuit, the presence of the earth is a prerequisite and there can be no true "free space" pattern. Therefore, the basic characteristics in this case are defined as those which would be obtained with the antenna grounded to a flat plane of infinite extent and conductivity, forming a "half space." Though not completely realized, this condition is approached in actual practice by a low frequency antenna working against a sea water "ground."

The characteristics of a grounded antenna working against an infinitely conducting half-space have an interesting relationship to those of an antenna formed by replacing the conducting plane with a "mirror image" of the antenna, which then works under "free space" conditions. This will become apparent as the discussion develops.

RADIATION CHARACTERISTICS OF A LINEAR CONDUCTOR

The simplest form of effective antenna consists of a rectilinear conductor whose length is comparable to a wavelength and whose thickness is very small in terms of wavelength. When r-f power is fed to such a conductor which is well removed from other objects, some of it will be reflected from the discontinuities at the ends and produce a standing wave of voltage and current in much the same manner as a mismatched transmission line, the distance between nodes being essentially the same as for an open-wire line with air dielectric. In other words, the energy travels at the same speed along the single linear wire as it does along a two wire open line, or at a velocity closely approaching that of a wave in free space.

If the radiator is not of a resonant length, reactance effects will be present just as in the case of an unterminated or mismatched line. With the transmission line the standing wave and any resulting reactive effects are due to the fact that the terminating resistor or load does not absorb all of the incident wave. In the case of the radiator the standing wave and any resulting reactive effects are due to the fact that the conductor does not radiate all of the traveling energy before it reaches the end. If the diameter of the radiator is very small in terms of wavelength, the r-m-s voltage and current amplitudes will vary substantially as a sine function.

The closer the terminating resistor matches the line or the greater the coupling of the radiator to the universe (greater the radiation resistance), the less will be the reflected energy and the less pronounced will be the standing wave and any resulting reactive effects. In both cases the effective "Q" will be reduced, and resonance effects will not be so pronounced. But there are important basic differences which should be noted, and these fundamental differences limit the extent to which transmission line analogies and similarities can be carried over into the field of radiators.

In the case of a transmission line, if we assume negligible resistance losses in the line and a nonreactive load on the generator, then all of the dissipated power is removed from the line at one point: the load termination. In the case of a linear radiator, the entire antenna radiates, and therefore power is removed from it throughout its length.

The second basic difference is that in the case of a transmission line there is little coupling between one part of the line and another part an appreciable fraction of a wavelength distant, because the neutralizing effect of one conductor upon the other tends to restrict the field. In the case of a linear radiator no such cancellation exists, and therefore the coupling between different portions of the radiator cannot be ignored.

It also is important to note that, except for the very ends of the radiator, any elementary section of the radiator is, in effect, at the same time a radiator and a transmission line feeding power to the adjacent section, though the mutual coupling also serves to transfer energy from one section to another.

In spite of these basic differences between a true transmission line and a linear radiator, it is possible to construct what amounts to a cross between an open-wire line and an antenna. If the spacing between the two conductors of a terminated two-wire open line is increased until it is a substantial fraction of a wavelength, then cancellation of the two fields no longer will be complete and appreciable radiation will occur. The characteristics in this case will then resemble both those of an inefficient antenna and an inefficient transmission line; not all of the power will be radiated, and not all of the power will be dissipated in the terminating resistor.

Electrical Length of a Linear Radiator

The speed at which energy is propagated along a vanishingly small, highly conductive, linear radiator in free space closely approaches the velocity at which an electromagnetic wave travels in free space, and therefore the standing wave current nodes (assuming more than one) are almost exactly a "free space" half wavelength apart. However, the *electrical* length of an isolated linear radiator having a *finite diameter* is slightly more than the *physical* length, because of "end effect," mutual coupling between different portions of the radiator, and other factors. This increases the electrical length of the radiator by an amount varying from about 2 per cent to about 8 per cent of a half wavelength for cylindrical radiators of practical configuration. The greater the diameter of the conductor in terms of wavelength, the greater the difference between its electrical and physical lengths.

In practical cases the radiator sometimes is not rectilinear but contains one or more bends. Usually the effect upon the ratio of physical length to electrical length is slight, especially when the radiator makes only one change in direction. In most cases when the electrical length is affected, it will be found that the *electrical* length has been *decreased*, making it necessary to *increase* the *physical* length slightly over what would be correct for a perfectly straight conductor with no bends. Generally speaking, the effect of a bend upon the electrical length is least when the bend occurs at a current loop, and greatest when it occurs near a voltage loop, though there are special situations where this does not apply.

It is possible to determine the electrical length of a completely isolated, rectilinear, cylindrical radiator from its physical length and vice versa when the frequency and the diameter of the radiator are known. The relationship for a resonant, half-wave radiator in free space is shown in figure 3-1.

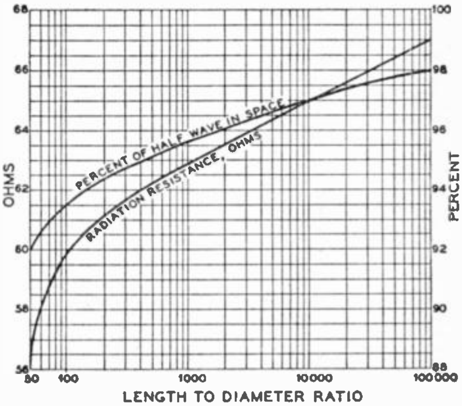


Figure 3-1.

EFFECT OF THE LENGTH/DIAMETER RATIO OF A CYLINDRICAL "HALF-WAVE" RADIATOR UPON ITS RADIATION RESISTANCE AND UPON THE EXACT PHYSICAL LENGTH AT WHICH IT RESONATES.

It is assumed that the radiator is in free space, is rectilinear, and has infinite conductivity. Thus it is seen that a "half-wave" radiator with a length to diameter ratio of 1000 resonates in free space when its physical length is 95.6% of a half wavelength in space. The curves may be applied to a quater-wave vertical radiator working against a perfect ground by halving both the "ohms" and "length/diameter" scales.

Before proceeding further, it is important that the reader fully appreciate the distinction between "physical length," "free space wavelength," "electrical length," and "free space conditions."

Physical length is that quality which may be measured with a yardstick. The only possible confusion that may arise here is that the physical length sometimes is converted to and expressed in terms of *"free space wavelength"* (or just "space wavelength"), which simply is the distance an electromagnetic wave of the frequency involved is propagated through empty space during one complete cycle. Because the velocity of an electromagnetic wave travelling in empty space is independent of frequency, free space wavelength is solely a function of frequency. Thus, for a certain frequency, a "free space wavelength" is just as definite a physical distance as is "so many feet and so many inches." (See figure 3-2.)

Electrical length always predicates a certain frequency, and is a measure of the time, *measured in cycles,* it takes for a wave to travel from one end of the radiator to the other. In a sense, it is a measure of the number of complete standing waves which will exist on a radiator at a specified frequency. It is determined by frequency, radiator length (physical), and by the velocity of propagation of a wave along the radiator. It is expressed in terms of wavelength, degrees, or radians. Thus if a radiator is of such physical configuration that a wave travels from one end to the other during exactly one complete cycle, then the radiator has an electrical length of one wavelength, or 360 degrees, or approximately 6.28 radians.

The expression *"under free space conditions,"* or reference to a "free space" characteristic of an antenna, means "with the radiator located in empty space." Thus the "free space electrical length" of a radiator in wavelengths is equal to the physical length of the radiator divided by the distance a wave would travel along that radiator during one complete cycle if the radiator were located in empty space and not influenced by surrounding objects. As another example, the "free space physical

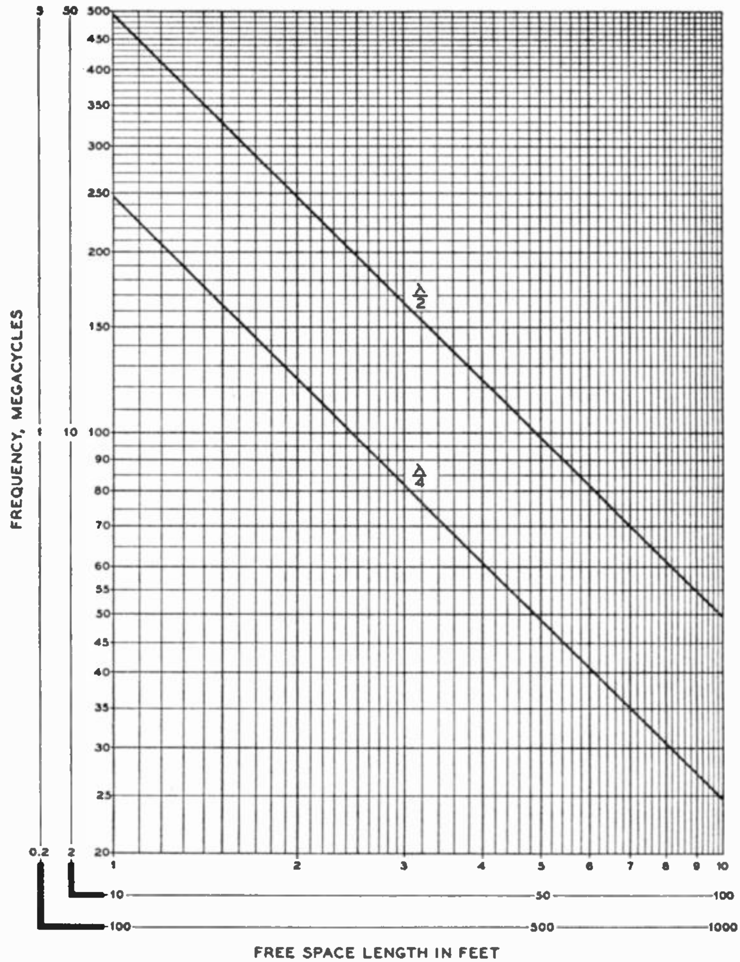


Figure 3-2.

PHYSICAL LENGTH IN FEET OF A QUARTER WAVELENGTH AND A HALF WAVELENGTH IN FREE SPACE, VS. FREQUENCY IN MEGACYCLES.

"Shortening effect" is not considered, and must be allowed for when the curves are used to determine the approximate length of a resonant radiator. Decimal fractions of a foot may be converted to inches by multiplying by 1.2. The chart can be read with sufficient accuracy for all ordinary antenna and transmission line problems, because the probable error will be considerably exceeded by the latitude of the unpredictable factors acting to alter the electrical length of the radiator or line.

length" of a radiator in terms of "free space wavelength" refers to the actual physical length of a radiator in free space, expressed in a unit of measurement equal to the distance an electromagnetic wave of that particular frequency is propagated through empty space during one cycle.

To get back to a discussion of the relationship between the physical length and electrical length of a radiator, the curves of figure 3-1 are applicable only to a half-wave ungrounded or a quarter-wave grounded radiator, the former in free space and the latter on a half-space. In actual practice the presence of supporting insulators, feed systems, and surrounding objects such as the earth and other antenna elements have an aggregate effect upon the electrical length which may even exceed the variation in length caused by practical variations in the conductor diameter. This makes the unknown length (either physical or electrical) difficult if not actually impossible to predict accurately under practical conditions.

Therefore the usual procedure is to cut or adjust the radiator to a length equal to or slightly less than the correct "free space" physical length, check the characteristics of the antenna experimentally, and then alter the physical length as necessary. If a half-wave radiator is supported from the ends by means of insulators, it is advisable to start with a length equal to about 0.97 of the length determined from figure 3-1.

Radiation Resistance of a Linear Radiator

In order to keep the efficiency of a radiating system high, the ratio of radiated energy to the energy dissipated in the radiating system must be high. This means that the antenna must be extensive enough physically in terms of wavelength to effect a fair amount of coupling to the universe, thus minimizing the amount of circulating current in the radiator for a given radiated power. As previously noted, the ratio of radiated power to the square of the circulating current at a maximum current point or "loop" in a simple radiator defines the *radiation resistance,* a fictitious term used to designate how closely the radiator is coupled to the universe. From the standpoint of efficiency, the value of radiation resistance is not too important so long as it is much higher than the loss resistance of the radiating system. However, a higher radiation resistance than the minimum required for good efficiency sometimes is desirable for other reasons, which will be covered later on.

When the radiating portion of the antenna consists of several phased elements with complex current distribution and complex mutual impedances, difficulty arises in defining the radiation resistance of the system in terms of the currrent at a current loop. It is for this reason that there exists considerable confusion and ambiguity with regard to the ficticious term "radiation resistance," and explains why the expression "feed point resistance" or "input resistance" often is used instead in conjunction with a complex array. "Feed point resistance" or "input resistance" refers to the resistive component of the impedance which the antenna presents to the feed line, or to the generator if no feed line is used. In the

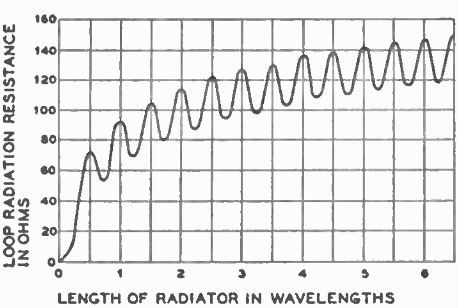


Figure 3-3.

LOOP RADIATION RESISTANCE VS. LENGTH OF A RECTILINEAR RADIATOR OF INFINITE CONDUCTIVITY AND VANISHINGLY SMALL CROSS SECTION WHEN LOCATED IN FREE SPACE.

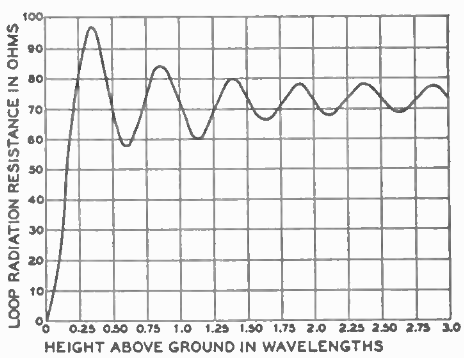


Figure 3-4.

LOOP RADIATION RESISTANCE OF A HALF-WAVE HORIZONTAL RADIATOR VS. ELEVATION, A VANISHINGLY SMALL CROSS SECTION AND A PERFECTLY CONDUCTING GROUND BEING ASSUMED.

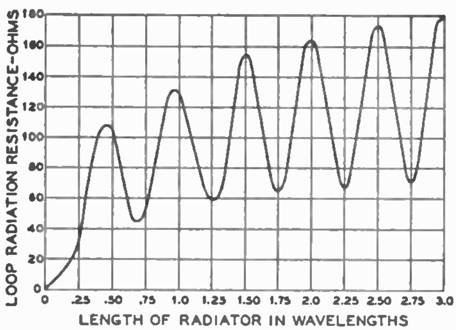


Figure 3-5.

LOOP RADIATION RESISTANCE OF A GROUNDED VERTICAL RADIATOR

A rectilinear radiator of vanishingly small cross section and perfectly conducting ground are assumed. It is also assumed that the radiator is grounded through whatever reactance is necessary to resonate it. For radiator lengths equal to an odd number of quarter wavelengths this reactance will be zero, and for radiator lengths equal to an even number of quarter wavelengths this reactance will be infinite.

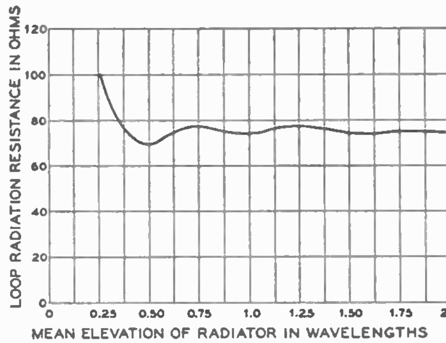


Figure 3-6.

LOOP RADIATION RESISTANCE OF A HALF-WAVE VERTICAL RADIATOR VS. ELEVATION OF THE CENTER ABOVE GROUND.

A rectilinear radiator of vanishingly small cross section and a perfectly conducting ground are assumed. Substitution of imperfect, average ground does not affect the curve as much as in the case of a horizontal radiator.

case of a simple antenna, the feed point resistance may be equal to the radiation resistance, if the antenna is fed at a current loop. But it should be kept in mind that the two are not always synonymous.

The loop radiation resistance of a rectilinear conductor of infinite conductivity and vanishingly small cross section is shown in figure 3-3 as a function of length, free space conditions being assumed. In the case of practical radiators having finite conductivity and cross section, and affected by the earth and surrounding objects, the values can be taken as only very approximate, especially if the radiator is not removed from other objects by at least several wavelengths.

While the effect of the proximity of a perfectly conducting earth is not of direct practical interest except possibly in the case of applications involving antennas located over sea water, a *general idea* of what happens to the radiation resistance of a horizontally oriented linear radiator a half wavelength long when located over average soil may be had by referring to figure 3-4, which shows the variation in radiation resistance with elevation above perfect ground. The effect of an imperfect ground is to reduce the amplitude of the oscillation and to slide the entire curve slightly to the left by an amount which is dependent upon the frequency and the ground constants. When the radiator is vertically polarized or is substantially longer than a half wavelength, the curve shown no longer is even approximately correct, even for perfect earth. When the radiator is horizontal but shorter than a half wavelength, the effect is to reduce the ordinate values without appreciably affecting the shape of the curve.

The corresponding curves are shown for a vertically oriented radiator over a perfectly conducting half space for purposes of comparison. Figure 3-5 shows the variation in loop radiation resistance (not feed point resistance) with height of a vertical radiator of vanishingly small cross section and infinite conductivity, with the lower end grounded to the conducting half space. Figure 3-6 shows the loop radiation resistance of a vertical, half-wave radiator as a function of elevation of the mid point above a perfectly conducting ground, the lower end being ungrounded in this case.

In all cases the effect of a finite cross section causes the radiation resistance to be very slightly lower than for the case of a vanishingly small cross section, but the effect is insignificant except when the cross section is an appreciable fraction of a wavelength. In the latter instance, as for example the use of large diameter tubing as a u-h-f antenna, the radiation resistance may be several per cent lower than the "vanishing cross section" value indicated by the curves. The effect of conductor diameter upon the radiation resistance of a radiator having an electrical length of one half wavelength is shown in figure 3-1.

Feed Point Impedance of a Center Driven Radiator

When a linear radiator is series fed at the center, the resistive and reactive components of the driving point impedance are dependent upon both the length and diameter of the radiator in wavelengths. The manner in which the resistive component varies with the physical dimensions of the radiator is illustrated in figure 3-7. The manner in which the reactive component varies is illustrated in figure 3-8.

Several interesting things will be noted with respect to these curves. The reactive component disappears when the overall physical length is slightly less than any number of half waves long, the differential increasing with conductor diameter. For overall lengths in the vicinity of an odd number of half wavelengths, the center feed point looks to the generator or transmission line like a series-resonant lumped circuit, while for overall lengths in the vicinity of an *even* number of half wavelengths, it looks like a parallel-resonant or anti-resonant lumped circuit. Both the feed point resistance and the feed point reactance *change more slowly* with overall radiator length (or with frequency with a fixed length) *as the conductor diameter is increased,* indicating that the effective

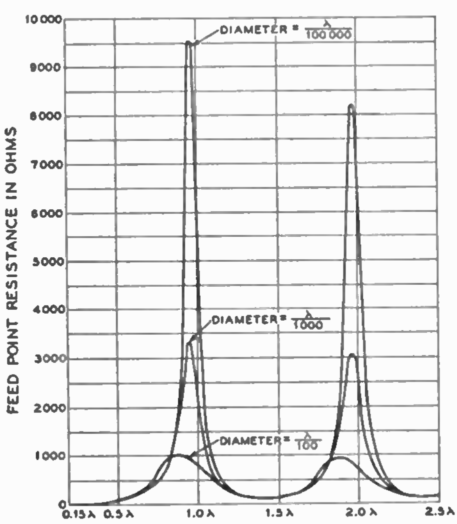


Figure 3-7.

FEED POINT RESISTANCE OF A CENTER DRIVEN RADIATOR AS A FUNCTION OF PHYSICAL LENGTH IN TERMS OF FREE SPACE WAVELENGTH.

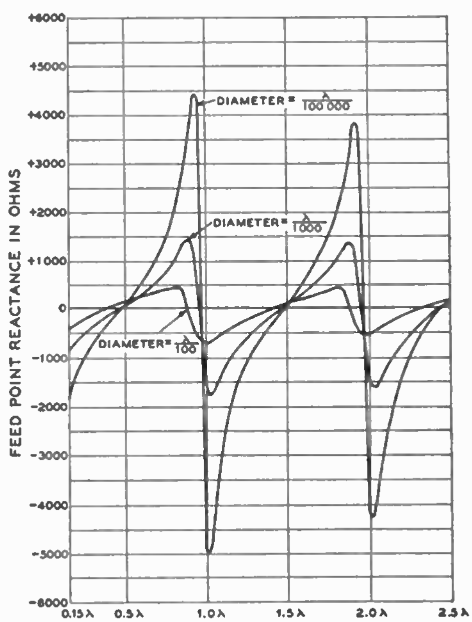


Figure 3-8.

REACTIVE COMPONENT OF THE FEED POINT IMPEDANCE OF A CENTER DRIVEN RADIATOR AS A FUNCTION OF PHYSICAL LENGTH IN TERMS OF FREE SPACE WAVELENGTH.

"Q" is lowered as the diameter is increased. However, in view of the fact that the damping resistance is nearly all "radiation resistance" rather than loss resistance, the lower Q does not represent lower efficiency. Therefore, the lower Q is desirable, because it permits use of the radiator over a wider frequency range without resorting to means for eliminating the reactive component. Thus, the use of a large diameter conductor makes the overall system less frequency sensitive. If the diameter is made sufficiently large in terms of wavelength, the Q will be low enough to qualify the radiator as a "broad-band" antenna.

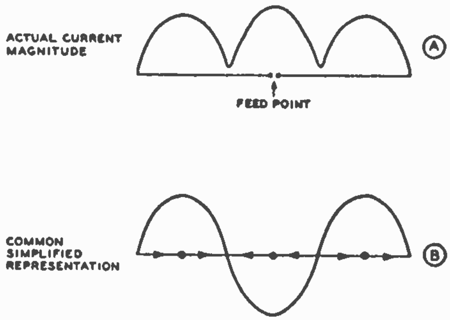


Figure 3-9.

ACTUAL AND SIMPLIFIED REPRESENTATION OF CURRENT DISTRIBUTION ON A CENTER-FED, RESONANT, LINEAR RADIATOR OF SMALL, UNIFORM CROSS-SECTION.

The upper representation, showing the current magnitude without regard to phase, is a true representation. Note that the current does not drop to zero except at the ends, and that the center section (the fed section) has somewhat greater loop current than have the outer sections. The phase shift is not uniform, being concentrated near the current nodes. Because at a node the phase shifts very rapidly (though not instantaneously) through 180 degrees and the magnitude of current is quite small compared to the value at a current loop, the lower idealized, simplified representation is commonly employed. A dot indicates a current loop, while the arrows designate the direction of current flow in each half-wave section. This assumption of sinusoidal current distribution, zero current magnitude at the nodes, and an instantaneous phase shift of 180 degrees at each node gives results which are sufficently accurate for many practical applications. No particular significance should be attached to the fact that the illustration happens to show a radiator having an electrical length of exactly 1.5 wavelengths.

Voltage and Current Distribution on a Linear Radiator

When a linear radiator has a comparatively small and uniform cross section, is symmetrically disposed with respect to surrounding objects, and exhibits considerable reflection (as does any filamentary linear radiator which is not terminated), the current is, to a first approximation, sinusoidally distributed. By making certain simplifying assumptions and ignoring certain second order effects, we may illustrate the current distribution by the common representation of figure 3-9. As in the transmission line case, both voltage and current are distributed sinusoidally and are displaced 90 degrees with respect to each other, causing a voltage node to occur at a current loop, and a voltage loop to appear at a current node. Because of this fixed relationship, it is common practice to show only the current distribution.

As in the transmission line case the current never is zero at a node under practical conditions except at an open end, because of the finite Q of the system. However, if the Q of an antenna is high, the current distribution resembles that of a transmission line with a high SWR. Under these conditions the current will drop to a very low value at a node and the phase shift will be concentrated near the nodes, just as in the case of the transmission line illustration, figure 2-12. This justifies the simplified schematic representation of figure 3-9, wherein the current at a node is assumed to be zero and the phase of the current in adjacent half-wave section to be exactly 180 degrees out of phase (as indicated by the small arrows representing direction of current flow). The dots indicate the position of current loops. These assumptions give acceptably accurate answers in computations involving directivity patterns and distant fields, but lead to appreciable error in computations involving feed point impedance.

RADIATORS LOADED WITH LUMPED SERIES REACTANCE

The electrical length of a radiator, as well as the relative phase of currents flowing in different parts of the radiator, can be greatly modified by the insertion of lumped inductance or capactiy or both, at one or more points along the radiator. The effect produced by a lumped reactance upon the current distribution depends upon the magnitude of the reactance, whether it is inductive or capacitive, and the point at which it is inserted. In the case of a transmitting antenna, the current distribution is significantly modified only from the point at which the reactance is connected back towards the generator or transmission line; the current distribution between the lumped reactance and the free end of the radiator is not appreciably affected.

Ordinarily, lumped series reactances are used as follows: (1) at the feed point terminals to cancel out the reactive component of the feed point impedance, and (2) at a point a small fraction of a wavelength from one end of a short, grounded antenna to increase the effective height of the radiator, thus increasing the radiation resistance. Because the use of a series reactance at the feed point permits efficient use of an antenna on a frequency which may be considerably different from its self-resonate or *natural* resonant frequency, the insertion of reactance at the feed point sometimes is considered as altering the electrical length of the radiator, thus "loading" it to a different frequency. Actually this simply is a matter of "tuning out" the reactive component presented at the feed point by a *radiator* which is *not* resonant, thus making the "antenna system" resonant. The insertion of the reactance makes the radiator "look" to the generator or transmission line like an antenna having a different natural resonant frequency.

As an example, insertion of a small amount of inductive reactance at a current loop will *effectively* increase the electrical length (lower the natural frequency) of a radiator, while insertion of the same amount of capacitive reactance will *effectively* decrease the electrical length. Insertion of the same amount of series reactance at a current *node* would have much less effect. But remember that in any case it is all a matter of terminology, and that when series reactance is inserted at the feed point the real purpose is to match the generator or line to a complex impedance, or at least to eliminate the reactive component.

VOLTAGE AND CURRENT FEED

If the electrical length of an ungrounded radiator is an exact integral number of half waves long, either a voltage loop or a current loop will occur at the exact center. If the radiator is center fed, the feed point will present either a high impedance or a low impedance to the feed line or impedance matching device, depending upon whether the feed point is at a voltage loop or current loop. If the feed point is at a voltage loop, the radiator is said to be *voltage fed,* while if the feed point is at a current loop, the radiator is said to be *current fed.* The same terminology is used when the feed point coincides with *any* current loop or voltage loop, regardless of whether it is at the center of the radiator.

When the feed point does not coincide with either a current loop or a voltage loop, as will be the case for a center fed radiator which is not an integral number of half wavelengths long electrically, the distinction no longer is sharply defined and is not strictly applicable, though a feed point impedance of 600 ohms or less generally is referred to rather loosely as "current feed," while a feed point impedance of 1000 ohms or more usually is designated as "voltage feed." The important thing to remember is that, regardless of terminology, what is fed to the antenna in every case is *power.* The expressions "high impedance feed" and "low impedance feed" are much more applicable, but are not widely employed.

FACTORS AFFECTING FREQUENCY SENSITIVITY

To minimize the frequency sensitivity of an overall antenna system (reduce the variations in the reactive and resistive components appearing at the sending end of the transmission line as the frequency is varied), it is desirable that the characteristic impedance of the line be such that no impedance transforming device is necessary between the line and the radiating system, thus realizing a "matching ratio" of 1.0. When this is impractical, the matching ratio should be kept as low as possible. If it cannot be kept below approximately 4.0, considerable improvement can be obtained by accomplishing the transformation in two or a greater even number of steps.

Other expedients can be resorted to which will provide a substantial reduction in the frequency sensitivity when "broad-band" characteristics are a requirement. The basic principles involved can be understood from the following explanation.

When a line is terminated in a purely resistive load which differs from the surge impedance, there will be recurring points along the line where the power factor is zero and the impedance therefore purely resistive, as explained in chapter 2. As the frequency is lowered, the distance between these points of zero reactance increases, causing them to move away from the line termination. As the frequency is raised, they move toward the termination. This is called "line effect."

If a series-tuned circuit is substituted for the terminating resistor, the resistive points on the line still move in the same direction as the frequency is changed, but at a faster rate. But if a parallel-tuned (anti-resonant) circuit is substituted, the non-reactive points tend to move in the opposite direction.

Now the feed point of an antenna looks to the transmission line very much like a resistance-loaded tuned circuit, series resonant if fed at a current loop and parallel resonant if fed at a voltage loop. From the foregoing discussion of line effect, it would appear that a multi-element antenna system would be less frequency sensitive if fed at a voltage loop, because then the "normal" line effect and the effect of the reactance slope would tend to cancel and cause the impedance presented at the sending end of the line to change more slowly with frequency. Unfortunately the beneficial effects of this cancellation are offset by the fact that, when "voltage feed" is employed, it usually is impractical to employ a line having a surge impedance comparable to the feed point resistance, and some kind of impedance transformer must be used. The latter has an adverse effect upon the effective "Q" of the overall system as noted previously.

A current-fed antenna system can be made considerably less frequency sensitive simply by connecting a parallel-resonant circuit using lumped L and C across the feed point, the optimum L/C ratio being most easily determined by experimental means because of its dependence upon so many factors.

When it is realized that the various resonant lengths of line used as phasing stubs, etc. in a complex antenna system can produce effects similar to those of resonant or anti-resonant circuits using lumped L and C, it helps to explain why some antenna arrays are much more frequency sensitive than others, though of course other factors are involved.

One simple antenna which owes in part its "broad-band" characteristics to reactance cancelling effects similar to those just described is the "folded dipole" which will be treated in chapter 5.

"RESONANT" ANTENNAS

Usually an attempt is made to operate an unterminated transmitting antenna under "self-resonant" conditions when it is used only on a single frequency or over a relatively narrow frequency band. When practicable, the radiating portion of the antenna is made of such length that at the frequency of operation (or center frequency of operation) the feed-point impedance is nonreactive. This means that the radiator must be cut to a "resonant" length. The simplest resonant radiators are the half-wave resonant dipole or "Hertz," which has an electrical length of a half wavelength, and the quarter-wave resonant unipole or "Marconi," which has an electrical length of a quarter wavelength and is grounded at the lower end. However, any number of electrical half wavelengths can be added without destroying the inherent resonance of the radiator. Several such

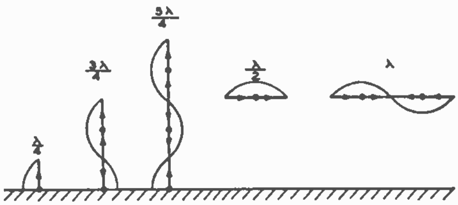


Figure 3-10.

SIMPLIFIED REPRESENTATION OF CURRENT DISTRIBUTION ON VARIOUS FUNDAMENTAL AND HARMONIC TYPE RESONANT RADIATORS.

No curve of voltage is shown in this common simplified representation because in all cases the curve would be the same as the current curve except for being displaced by 90 degrees (or, in other words, a voltage node always corresponds to a current loop).

antennas, with a simplified representation of the current distribution, are illustrated in figure 3-10.

Harmonic Operation

When a radiator is cut to the shortest resonant length (an electrical half wavelength for an ungrounded antenna and an electrical quarter wavelength for a grounded antenna), it is said to be operated on its *fundamental*, or *first harmonic*. An ungrounded resonant radiator two electrical half wavelengths long is operating on its *second harmonic*, and is often referred to as a *full wave* antenna. A radiator three electrical half wavelength long operates on its third harmonic, and so on.

In the case of a grounded resonant radiator the numerical designation of the harmonic is somewhat ambiguous, because the order of the harmonic depends upon whether the earth is considered to constitute a duplicate or "mirror image" of the radiator, or whether it is just supposed to substitute for the last quarter wavelength of an ungrounded harmonic radiator. A grounded, resonant radiator having an electrical length of 5/4 wavelengths is operating on its 5th harmonic when considered on the first basis, and on its 3rd harmonic when considered from the second. The complication arises due to the fact that under certain conditions it is difficult to classify the antenna as either grounded or ungrounded. This condition exists when the radiator is worked "against ground" but is approximately an integral number of half wavelengths long. The current flowing in the ground connection in such cases is practically negligible, and it is debatable whether the radiator should be classed as a "grounded" radiator.

Because of the foregoing confusion regarding terminology, it is becoming more and more common to refer simply to a "harmonic radiator so many wavelengths long," with the qualifying expression "worked against ground" appended if the latter is applicable.

NONRESONANT ANTENNAS

While the term "resonant" antenna or radiator usually implies a *naturally* resonant radiator, or one which presents a pure resistance at its feed point, the term "resonant" in its broadest sense applies to any antenna whose end or ends are open or short circuited, causing the radiator to exhibit considerable reflection and standing waves. The term "nonresonant" does *not*, though it might seem appropriate, apply to an unterminated radiator which lacks inherent resonance and therefore presents a reactive component at the feed point. Instead it refers by common usage to a radiator which is terminated in an impedance of such magnitude and character as to eliminate or virtually eliminate reflection from the end of the radiator.

The magnitude and phase of the current along a nonresonant radiator corresponds closely to that of a lossy transmission line which is terminated in its characteristic impedance. The current is gradually attenuated as it travels towards the terminating resistor, no standing waves being apparent, and the phase shift along the radiator is substantially uniform instead of being concentrated at half wavelength points. Obviously such a radiator is not very frequency sensitive.

It is readily apparent that the nonresonant radiator, with its effective SWR of 1.0 and linear phase shift, does not lend itself to the simplified representation of current

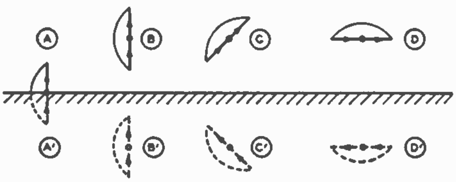


Figure 3-11.

ILLUSTRATING HOW AN IMAGINARY "IMAGE" ANTENNA MAY BE SUBSTITUTED FOR A PERFECTLY FLAT GROUND OF INFINITE CONDUCTIVITY IN MAKING CERTAIN ANTENNA CALCULATIONS.

Observe that while the configuration of the imaginary antenna is a mirror image of the actual radiator, the direction of current flow is not. This is explained by the fact that the phase of a horizontally polarized wave is reversed upon reflection from the perfect ground, while the phase of a vertically polarized wave is not. The result is that corresponding vertical components of current flow always are in the same direction, while the corresponding horizontal components of current flow always are in opposite directions.

magnitude and phase distribution commonly employed for resonant radiators (figure 3-9B).

IMAGE ANTENNAS

If an absolutely flat, perfectly conducting ground is assumed, the effect of the earth always can be duplicated by an "image" antenna which is a mirror image of the actual radiator, regardless of whether or not the actual radiator is connected to earth. This assumption of a "mirror image" is useful in calculating or comparing directivity patterns. The "mirror image" representation is illustrated in figure 3-11.

MUTUAL IMPEDANCE

The reader no doubt is familiar with the mechanics of coupling between two lumped circuits by means of mutual impedance in accordance with ordinary electric circuit theory. As was mentioned previously in this chapter in a discussion of the radiation characteristics of a linear conductor, one important distinction between a transmission line and an antenna is that in the case of an antenna each portion of the antenna is coupled to every other portion, the coupling assuming considerable magnitude for portions of the antenna system which are spaced less than several wavelengths.

As will be discussed later under antenna directivity, antenna systems or "arrays" often are employed which utilize a multiplicity of radiating elements in a complex configuration. Whether we consider all elements in the system as a single, elaborate radiator, as in the case of a harmonically operated long wire, or whether each composite element is considered as an individual radiator, it is apparent that mutual coupling exists between them and that the current in one element affects and is affected by the current flowing in every other element.

One significant difference between the mutual coupling associated with inductively coupled lumped circuits and the coupling between different antenna elements is introduced by the factor of *time*. In the lumped circuit case the interchange of energy is assumed to be instantaneous, and in fact is virtually instantaneous, while in the case of neighboring radiating elements the time taken for the current in one element to affect the other element and react back on the first element must be taken into consideration, because the spacing between them no longer is an insignificant fraction of a wavelength. The mutual interaction, instead of being effectively instantaneous, is retarded because of the finite velocity of propagation and the significant electrical distance between the elements. In figuring the effect upon the impedance and distribution of current in the elements, the reaction time must be taken into consideration. The mutual impedance has a *phase angle*.

In a very simple case such as two, identical elements which are either parallel or colinear, it is possible to compute the mutual impedances from the theory with fair accuracy, but in complex situations involving a multiplicity of elements with complex orientation and dimensions the problem is either formidably difficult or actually impossible of accurate solution except by experimental means. It is for this reason that the only accurate data available on mutual impedance calculations is restricted to very simple cases such as the two just mentioned.

In view of the foregoing, it is sufficient for the purposes of this book if the reader simply appreciate the basic facts presented in the following very brief qualitative summary.

The interaction or "coupling" between two antenna elements which are separated by not more than several wavelengths can produce a mutual impedance which is significant or even considerable when compared to the self-impedance exhibited by either element under isolated conditions.

Because of the significant effect of reaction time, the mutual impedance has a *phase angle* which varies with the spacing between the radiators. It also varies with the shape, size, conductivity, and orientation of the conductors.

The magnitude of the mutual impedance between two conductors decreases with increased spacing, and the phase angle varies cyclically through 360 degrees as the spacing is increased. When the phase angle is between 90 and 270 degrees, the mutual impedance has a negative resistive component.

Because the resistive component of the mutual impedance may be either positive or negative, the presence of a neighboring radiator may either raise or lower the feed point resistance of a given radiator, depending upon the spacing, the phasing of the feed system, and other factors.

The reactive component of the mutual impedance will cause the resonant frequency (or length) of a radiator to be appreciably or considerably affected by the presence of another radiator except at certain critical spacings.

The effect of a perfectly flat, infinitely conducting ground upon the impedance characteristics and current distribution of a radiator is the same as that produced by the mutual impedance afforded by an "image" antenna which is substituted for the perfect earth. The effect of the actual earth (including trees, buildings, etc.) is not easily determined in most cases except by experiment.

Except in very simple cases, the only safe method of ascertaining accurately the characteristics of a complex array of radiators is the experimental method. No array should be so designed that alteration of the pertinent physical dimensions is impracticable, unless the same effect can be achieved by the adjustment of compensating reactances (coils, condensers, or "stubs"), or the array is an exact duplicate of one known to be optimum and it is possible to simulate the same environment characteristics (terrain, method of support, etc.).

EFFECT OF FEED POINT POSITION ON PHASE RELATIONSHIPS

As will be discussed in the next section, the directional characteristics of an antenna of certain configuration are dependant upon the phase relationships of the currents flowing in the different radiating elements or portions of the antenna. Therefore it is important that we examine the effect which is produced upon these phase relationships by the attachment of a feed line or generator to the antenna.

Regardless of where the generator or feed line is connected, provided that it is at a point which does not destroy the line balance, the current at any given instant always is flowing into one terminal and out the other. When the feed point is at a current loop, the "free" phase relationships are not disturbed, because the current in one half-section of a half-wave element naturally flows towards the current loop while the current in the other half-section naturally flows away from the current loop, as illustrated in figure 3-12A. Therefore the phase relationships are not upset when "current feed" is used, as in figure 3-12B.

When the feed line or generator is inserted at a voltage loop, the antenna still is fed current (even though it is shown as vanishing in the simplified representation). And it still flows into one terminal of the generator or line and out the other at any given instant. However, because the terminals are at a voltage loop or current node, there is a 180 degree phase shift at each terminal, causing the current to reverse direction. This is illustrated in figure

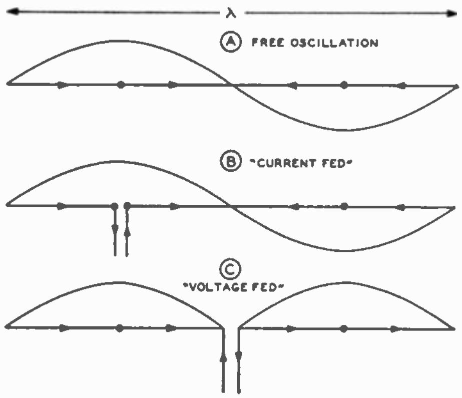


Figure 3-12.

ILLUSTRATING EFFECT OF FEED POINT LOCATION UPON PHASE RELATIONSHIPS.

When the balanced feed point is at a current loop, the phase relationships are the same as for a "free" radiator. When the feed point is at a current node, the phase relationships are reversed.

3-12C. The important thing to note is that the current is caused to flow in the same direction in the two half-wave elements of the antenna, rather than in opposite directions.

This effect is not unique to the particular antenna used as an example, but applies to any antenna system. When the feed point is at a current node ("voltage feed"), the natural phase relationships are reversed, causing the current in the adjacent half-wave elements to be in phase instead of out of phase. When the feed point is at a current loop ("current feed"), the natural or "free" phase relationships are not disturbed.

While the foregoing discussion is not rigorously accurate, because of the simplified representation of current distribution and phase relationships on the radiator, it is a sufficiently true picture for the purpose of explaining the basic principles involved.

In many directional arrays, "dummy" sections of transmission line, called "phase reversing stubs," or "phase inverting sections," are inserted at current nodes to reverse the phase relationship between two adjacent half-wave elements. These will be described in detail in a later section.

ANTENNA DIRECTIVITY

As previously explained in this chapter under "Fictitious Reference Antennas," all practical antennas radiate better in some directions than others. This characteristic is called "directivity." The more "directive" an antenna is, the more it concentrates the radiation in a certain direction, or directions. The more the radiation is concentrated in a certain direction, the greater will be the field strength produced in that direction for a given amount of total radiated power. Thus the use of a directional antenna or "array" produces the same result in the favored direction as an increase in the power of the transmitter.

The increase in radiated power in a certain direction with respect to an antenna in free space as a result of inherent directivity is called the "free space directivity power gain" or just "space directivity gain" of the antenna (referred to a hypothetical "isotropic radiator" which is assumed to radiate equally well in all directions). Because the ficticious isotropic radiator is a purely academic antenna, not physically realizable, it is common practice to use as a reference antenna the simplest ungrounded resonant radiator, the half-wave "Hertz," or resonant doublet. As a half-wave doublet has a space directivity gain of 2.15 db over an isotropic radiator, the use of a resonant dipole as the comparison antenna reduces the "gain figure" of an array by 2.15 db. However, it should be understood that power gain can be expressed with regard to *any* antenna, just so long as it is specified.

As a matter of interest, the directivity of an "infinitesimal dipole" provides a free space directivity power gain of 1.5 (or 1.76 db) over an isotropic radiator. This means that *in the direction of maximum radiation* the infinitesimal dipole will produce the same field strength as an isotropic radiator which is radiating 1.5 times as much total power.

A half-wave resonant doublet, because of its different current distribution and significant length, exhibits slightly more free space power gain as a result of directivity than does the infinitesimal dipole, for reasons which will be explained in a later section. The space directivity power gain of a half-wave resonant doublet is 1.63 (or 2.15 db) referred to an isotropic radiator.

DIRECTIVITY GAIN VS. PRACTICAL SIGNAL GAIN

In the foregoing discussion of free space directivity gain the only consideration was the increase in field strength resulting from concentrating the radiation on (or confining it to) a restricted solid angle. Resistance losses, and the effect of the ground on directivity, were ignored. However, what we are interested in from a *practical* standpoint is the *effectiveness of an array in increasing the field strength at a distant receiving location*. This increase will be called the *practical dx signal gain*, and unless otherwise stated will assume as a comparison antenna a half-wave dipole with identical polarization, suspended at the same average elevation above earth as the array.

While the practical dx signal gain is influenced primarily by the amount of free space directivity gain possessed by an antenna, it is also influenced by loss resistance (either earth losses or antenna heat loss) and by the effect of the earth upon the vertical directivity pattern. The directivity gain of an array over actual earth usually varies with the vertical "wave angle," even though the comparison antenna is at the same average elevation.

In most cases the effect of resistance losses (earth losses and antenna losses) will not degrade the gain by more than about 2 db. With many arrays the reduction will be less than 1 db, and with some it will be negligible or even negative.

Also in most cases the difference between free space directivity gain and the directivity gain at the optimum wave angle in the presence of the earth will not exceed about 2 db if the optimum wave angle is low, as is the case for long distance, high-frequency, sky-wave comunication, and for v-h-f and u-h-f ground-wave communication at all distances.

Thus we may generalize and say that *ordinarily* the difference between the space directivity gain and the practical dx signal gain of an array (both referred to a half-wave dipole) will not exceed 4 db, and that increased space directivity gain *ordinarily* means increased practical dx signal gain. However, it is important that the reader appreciate the limitations that must be placed upon these generalizations with regard to high-frequency sky-wave communication.

If the effective or optimum wave angle is above 5 or 10 degrees, the practical dx signal gain of an array having sharp vertical directivity may be as much as 10 db greater than the space directivity gain, or as much as 10 db less. This is explained by the fact that the nulls and maxima of an array *do not necessarily* occur at the same vertical angles as for a reference antenna at the same average elevation above ground.

In view of the fact that the practical dx signal gain of a high frequency array depends upon the optimum wave angle, and because the latter depends upon frequency, distance, and variable ionosphere conditions, the practical dx signal gain must represent an average and must necessarily be an approximation (unless the optimum wave angle is specified, or unless the free space vertical directivity is rather broad).

ASPECT DIRECTIVITY AND INTERFERENCE DIRECTIVITY

The intensity of radiation from an infinitesimal dipole or doublet expressed in terms of field strength is proportional to a projection of the doublet in any particular direction. Thus the radiation is maximum in directions lying in a plane which passes through the center of the radiator at right angles (a median plane), is zero in the direction of the conductor, and for other points is proportional to the cosine of the angle which the direction makes with the radiator.

The infinitesimal dipole or doublet is so short that the radiation from all parts of it arrives at any point in space with no significant phase difference. Because electromagnetic waves are transverse, the relative

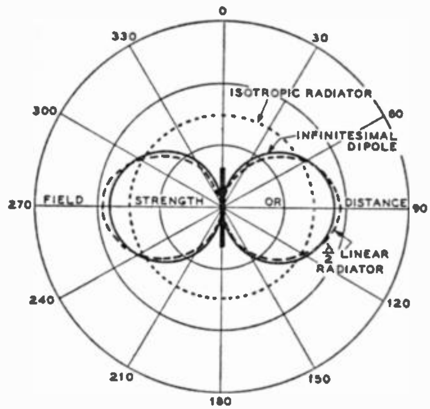


Figure 3-13.

POLAR DIAGRAM OF RADIATION PATTERN OF AN ISOTROPIC RADIATOR (DOTTED), AN INFINITESIMAL DIPOLE (SOLID), AND A HALF-WAVE, RESONANT, LINEAR RADIATOR (DASHED), FOR THE SAME AMOUNT OF TOTAL RADIATED POWER.

The curves are for free space conditions, and represent either distance for a given field strength, or field strength for a given distance. A two dimensional diagram such as this can illustrate the pattern for only one plane, but for the antennas under consideration the same pattern applies to any plane containing the radiator. The distance to any point on the solid curve is proportional to the cosine of the angle made with a plane bisecting the radiator at right angles. It will be observed that while the resonant half-wave radiator has slightly more directivity, the difference is insignificant from a practical standpoint, and only a very small error is introduced if the resonant radiator is assumed also to have a cosine relationship.

magnitude of radiation therefore is determined entirely by the magnitude of the projection of the infinitesimal dipole or doublet upon an imaginary sphere centered on the radiator. This is *aspect directivity.*

To project our imaginary dipole or doublet we must assign a finite length, even though minutely small. Therefore such a ficticious reference antenna is probably more correctly referred to as an *elementary* dipole or doublet.

The free-space directivity pattern of an infinitesimal or elementary dipole is illustrated in figure 3-13. The pattern shows the relative magnitude of *field strength* for points at constant range in any plane which includes the radiator. This magnitude is proportional to the square root of radiated power, the relationship being pertinent because "power gain" figures naturally are based on power, rather than field strength. For instance, if a directivity pattern shows that the field strength is doubled in a certain direction, then the power gain in that direction is 4 times, or 6 db.

For the reader who may have difficulty in visualizing the radiation pattern in three dimensions from the two-dimensional polar diagram of figure 3-13, figure 3-14 should facilitate an understanding of how a two-dimensional polar diagram can give a complete picture of the pattern. It is apparent that because of the axial symmetry of the radiation pattern of a dipole radiator, all cross-sectional views of the pattern which include the radiator will be identical.

Directivity patterns sometimes are drawn to a power scale, but are more commonly drawn so that the curve represents field strength, because this is the type of pattern obtained when the directivity of an antenna system is solved graphically, and the curve must be converted if a power scale is to be used. When a directivity pattern does not indicate whether the curve represents field strength or power, field strength is implied. However, to avoid possible misinterpretation, it is customary to indicate whether field strength or power is used as the unit of measurement.

Another good reason for using a field strength scale for a *free-space* polar diagram is that the same curve then can be used to represent either relative field strength vs. direction at a constant range, or range vs. direction for a constant field strength. The reason for the interchangeability is that field strength along any constant direction from the radiator is, in free space, inversely proportional to distance from the radiator. Thus, if the field strength at a given point is halved, it can be restored to its original value by halving the distance to the radiator. So if we plot a curve of all points in the plane which have the same field strength, the curve automatically becomes a polar diagram of direction vs. relative field strength at a constant range.

Polar diagrams of antenna directivity also are commonly drawn to a *decibel* scale, because a logarithmic relationship gives a more useful or convenient presentation for certain purposes. When a db scale is used, the center of the polar diagram cannot represent zero, but rather has some finite value, because a decibel is simply a ratio. As the scale represents a nonlinear relationship and is arbitrary, the geometrical shape of a field strength pattern drawn to a db scale can vary considerably, depending upon the total range assigned to the scale. The smaller the total db range encompassed by the scale, the "fatter" the pattern will appear in a polar diagram of given size.

Getting back to figure 3-13, the field strength produced by the infinitesimal dipole at any particular angle with respect to the radiator is proportional to the cosine of the angle made with a plane bisecting the radiator at right angles.

This property of *aspect directivity* is a property of all linear radiators. It is not unique to the elementary dipole or doublet. However, the elementary dipole or doublet is the only radiator whose significant directivity is solely due to aspect. Unless a radiator is very short in terms of wavelength, another factor contributes to directicity: *interference* due to the radiation arriving from different portions of the radiator at a given point in space at different times. In fact, wave interference is the basic mechanism of antenna directivity control. Directivity obtained in this manner may be referred to as "interference directivity" in order to distinguish it from simple aspect directivity.

In the case of any complex combination of radiating elements, called a *directional array*, interference directivity contributes much more to the total directivity than does simple aspect directivity. The latter is no greater in the case of an elaborate array than in the case of a single radiating element, while the interference directivity of a properly designed array increases with the size of the array.

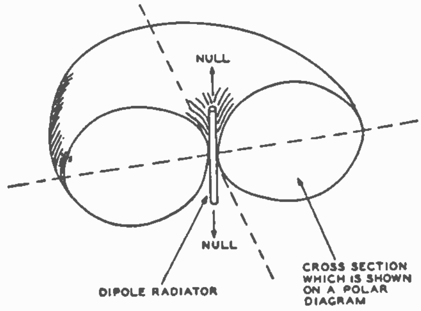


Figure 3-14.

SURFACE PATTERN OF RADIATION, PERSPECTIVE VIEW, OF A DIPOLE RADIATOR IN FREE SPACE.

This illustration shows how axial symmetry permits use of a single polar diagram such as figure 3-13 to indicate the relative field intensity at all points in space which have the same range.

While interference calculations are quite complex, there is one simple empirical relationship which is useful in understanding and appraising directional systems. Two identical radiating elements or arrays in free space, having identical orientation and currents, will in certain directions produce a resultant field strength equal to twice the field strength produced in that direction by one element or array, *provided the spacing between them is sufficient to reduce the mutual impedance to an insignificant magnitude*. This means that doubling the total power (by the addition of the second element or array with the same current) increases the field strength in certain directions by 6 db. The *net* increase, or power gain of the combination, therefore is 3 db in those directions.

If the two elements or arrays are sufficiently close that the mutual impedance is appreciable, then the maximum net "free space" directivity gain may be somewhat more or somewhat less than 3 db over a single element or array, depending upon the magnitude and phase angle of the mutual impedance.

The radiation pattern of any practical antenna, whether it consists of simply a linear conductor having a length of an appreciable fraction of a wavelength, or whether it consists of a multiplicity of radiating elements having complex orientations and spacings and current distributions, may be determined by resolving the whole antenna system into a very large number of articulated elementary doublets, then calculating the contribution of each doublet (which is determined by current magnitude and conductor aspect) and finally adding the separate fields by means of graphical integration (with due regard to phase and polarization). In the case of a simple antenna such as a thin, half-wave resonant doublet this is easily done, and gives the dashed pattern of figure 3-13. In the case of highly complex antennas the solution by this method is much more difficult and laborious, but is still more practicable than attempting to calculate the gain on the basis of mutual impedances.

DISCRIMINATION VS. POWER GAIN

All practical antennas radiate better in some directions than in others, and usually have two or more complete "nulls." As just shown, even the simple half-wave resonant dipole exhibits pronounced directional qualities, producing no significant radiation at all directly off the two ends. However, the term "directional antenna" by convention ordinarily implies a more complicated radiating system than a single half-wave resonant dipole; it usually refers to an antenna or "array" which is expressly designed for the purpose of concentrating the radiation into a relatively narrow "beam," though it also is sometimes employed to designate an antenna which is expressly designed to *prevent* or *minimize* radiation in one or more certain directions.

But the two are more or less interrelated, because if the antenna configuration is designed to reduce radiation in certain directions, the radiation in certain other directions will be increased, assuming that the loss resistance is at all times insignificant. It is axiomatic that if radiating elements are disposed in such a manner that their radiation adds or aids in preferred directions, there will be other directions in which the radiation opposes or cancels. However, it is obvious from the directivity pattern of a simple dipole, figure 3-13, than an antenna can exhibit pronounced nulls in the radiation pattern without providing much power gain.

Thus it is seen that directivity control can be used to "protect" a certain direction or directions, or to provide the maximum possible power gain in a given direction or directions, or to accomplish both simultaneously, depending upon the application. "Protection" of a certain direction can be accomplished with a simple array, but a high power gain in a certain direction requires a comparatively elaborate array.

EFFECT OF GROUND ON "SPACE" DIRECTIVITY PATTERNS AND GAIN

As previously noted, the radiation pattern of an array is affected by the ground, just as in the case of the simplest radiator. The directivity contributed by ground reflection is interference directivity.

The pattern of an array spaced several wavelengths above flat earth can be determined with good accuracy by multiplying the "free space" pattern of the array by the ground interference pattern of an isotropic radiator having the same average height and polarization. Another method is to use the concept of the "image" antenna, in which case the earth is ignored temporarily and the image is considered as part of the array. Only the top hemisphere of the pattern so obtained is utilized, because obviously there can be no radiation into the lower hemisphere when the earth is reinstated in place of the image. Confining the radiation to one hemisphere adds 3 db to the field strength at all points, but this is exactly offset by the additional power required to excite the image. As these two factors cancel, they can be ignored: the radiation below the horizontal and the power required to excite the image.

As a simple example of applying the image concept, the pattern of a quarter-wave radiator working against perfect ground is the same as the "top half" of the pattern of a vertical half-wave radiator in empty space *carrying the same current,*

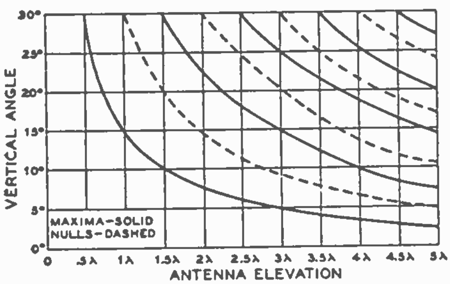


Figure 3-15.

VERTICAL ANGLES AT WHICH MAXIMA AND NULLS OCCUR FOR A HORIZONTALLY POLARIZED ISOTROPIC RADIATOR OVER PERFECT EARTH.

The curves also give a close approximation for a horizontal dipole over actual earth of average resistivity and dielectric constant, in a direction broadside to the radiator, provided the earth is reasonably flat. In the case of a horizontally polarized array having sharp "free space" vertical directivity, the angles at which maxima and nulls occur depend on the composite directivity (free space directivity and ground reflection factor) and may deviate considerably from the above curves. If the free-space vertical directivity of the array shows a very broad "nose," however, the above curves are applicable as a close approximation.

because the latter antenna corresponds to the grounded quarter-wave radiator plus its image. However, it takes only half as much power to produce a given current in the quarter-wave grounded radiator, because its radiation resistance is half that of a half-wave radiator in free space. So the grounded quarter-wave radiator may be said to have 3 db gain over a half-wave radiator in free space.

In the case of an ungrounded antenna located above earth, the mutual impedance between the antenna and its image must be taken into consideration. When a horizontal radiator is sufficiently close to a perfect ground that the mutual impedance is appreciable, then the ground interference pattern may exhibit lobes having a maximum amplitude up to approximately 8 db above the free space amplitude, assuming the same total radiated power in each case.

When the array is sufficiently removed from the ground that the mutual impedance between the array and its fictitious image is negligible (in other words, having a spacing exceeding several wavelengths), then the maximum increase over the "free space" field strength will be 6 db. This figure has been widely used without regard to antenna elevation, but it is not rigorously correct when the effective elevation of the radiator is not great enough to prevent the ground from having an appreciable effect upon the impedance of the antenna.

The effect of an actual, imperfect ground, *particularly one which is irregular,* is difficult to predict in the case of either an array or a simple radiator. In fact, the presence of highly irregular terrain under and around an array can so distort the radiation pattern that the practical signal gain cannot be predicted with even fair accuracy.

The angles of the maxima and nulls below 30 degrees due to the ground reflection factor are shown for horizontal polarization and perfect ground in figure 3-15. For horizontal polarization and within the vertical angles included in the chart, "perfect ground" curves will give a close approximation in the case of actual ground *if the ground is flat.* They are not rigorously accurate because the *effective* elevation is greater than the *actual* elevation except at grazing angles. The disparity increases with radiation angle.

VERTICAL AND HORIZONTAL PATTERNS

When the radiation pattern of an antenna or array does not exhibit axial symmetry, a single cross-sectional view of the "solid" or three-dimensional pattern cannot be used to show the relative magnitude of radiation at all points in space. However, this is not too serious. It isn't necessary to show the relative magnitude of radiation at all points in space in order to give a useful picture of the directional properties of an antenna.

The pertinent data in the case of ground wave propagation is a plot of the field strength versus compass direction at constant range. This gives a polar diagram of

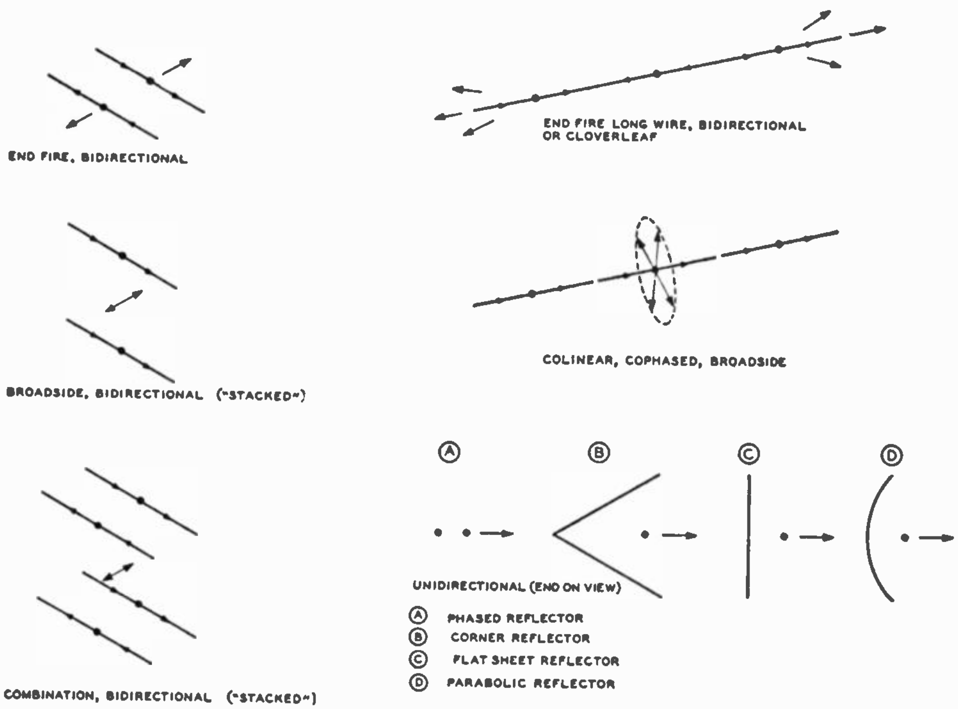


Figure 3-16.

ILLUSTRATING VARIOUS DESCRIPTIVE TERMS OFTEN USED IN CONNECTION WITH DIRECTIONAL ANTENNA ARRAYS.

To simplify the illustration, the sections of transmission line used to feed and phase the various radiating elements have not been included. The large arrows indicate the directions of maximum radiation; the small arrows on the radiators indicate the instantaneous direction of current flow, and therefore the relative time phasing of the elements.

the basic *horizontal directivity* (at ground level) and is referred to as the "horizontal pattern."

In the case of sky wave propagation, where information on the relative field strength and response at various vertical angles of departure and arrival is pertinent, a single vertical cross section of the solid pattern, taken along the direction of maximum "zero elevation" or ground level radiation, often will give the needed data. This is explained by the fact that directional sky wave propagation often is employed for point-to-point work, and with typical h-f arrays the compass direction in which radiation is maximum usually corresponds very closely for all vertical angles below 15 degrees or so (the angles useful for long distance sky-wave transmission). Patterns of this type are called *vertical patterns*, depicting *vertical directivity* in a certain horizontal direction.

If one or two vertical cross section patterns do not give the required information, horizontal patterns may be shown for various pertinent vertical angles, rather than for zero elevation only, on the same polar diagram. This is somewhat analagous to giving the "profile" data for all of a geographical area on a map by means of contour lines, rather than by an infinite number of profile drawings or "contour strips."

The foregoing description and explanation should enable the reader to interpret polar diagrams of vertical and horizontal directivity when encountered, if he is not already familiar with them.

DIRECTIVE ARRAY TERMINOLOGY

The directivity characteristics realized with certain radiator configurations or relationships have descriptive designations with which the reader should be familiar.

End fire as applied to directivity or an array refers to an arrangement where the currents in adjacent radiating elements are considerably out of time phase (usually in time phase opposition or quadrature), but are so disposed in space that the radiation from the different elements is caused to travel different distances to reach a distant point in the desired direction, thus compensating for the difference in time phase at the elements and arriving at the distant point in phase.

In the stricter sense, end-fire directivity always is either (1) in a direction which lies in the plane containing the radiating elements, or (2) along the *approximate* direction of a single, long, "harmonic" radiating element. This is illustrated graphically in figure 3-16.

When the directivity shoots off in two opposite compass directions (maximum gain occuring in two opposite directions), the array is said to be *bidirectional*. When radiation in one of these two directions is suppressed, the array is said to be *unidirectional*.

Broadside directivity is that which is obtained when the currents in a number of coplanar radiating elements are in time phase and the elements are disposed in such a manner that all are equidistant from a distant point in the desired direction of maximum radiation. Such an array is called a broadside array. Broadside directivity always is in a direction normal to a plane containing the radiating elements. This is illustrated graphically in figure 3-16. In the case of the *colinear* broadside array* (figure 3-16), the cophased radiators are in-line and there are an infinite number of planes which will include the radiating elements. Therefore the radiation is uniformly maximum in all directions at right angles to the line of the radiators.

A harmonically operated long wire also might be described as "colinear," in view of the fact that the radiating elements are in a straight line. However, by usage the term "colinear" has come to imply in-line *cophased* elements exclusively.

Some arrays employ both broadside and end-fire directivity to enhance the overall directivity, and they therefore qualify as "combination" arrays.

When the elements of a broadside or combination broadside and end-fire array are arranged one over the other, rather than side by side, the array is said to be "stacked," and the array of radiators sometimes is referred to as a "stack."

When an array is designed to concentrate radiation in one direction and to minimize it in the opposite direction, it is said to be *unidirectional*. The basic types are shown in figure 3-16. The phased reflector is an element resembling the driven radiator, either parasitically excited (as a result of mutual impedance) or directly driven in such a manner as to produce the phase difference required for a unidirectional interference pattern. The various sheet reflectors act in somewhat the same manner as an artificial ground, with the parabolic and "corner reflector" arrays being designed to give a "beaming" effect and greater directivity than a flat sheet.

A coplanar broadside array consisting of several tiers of colinear elements or "strings" often is referred to as a *broadside curtain*, or just *curtain*, and also is known as a *panel*. A "curtain" backed up by a flat sheet or grid reflector to give a unidirectional characteristic commonly is called a *mattress*.

Parasitic and Driven Elements

In a complex array consisting of a multiplicity of radiating elements, there are sev-

*The colinear array sometimes is placed in a separate classification, but throughout this book it will be considered simply as a special case of the broadside array.

eral methods of exciting each element with current having the desired phase and magnitude. When an element is not physically connected to the main transmission line, either via other elements or through a section of line, it is fed solely through its mutual impedance. An element coupled and fed in such a fashion is said to be "parasitically excited." An element so excited is called a "parasitic element."

When an element is connected to the main transmission line in any physical manner, either (1) directly, (2) through a section of line, (3) via another element, or (4) by any combination of line and radiator(s), it is referred to as a *driven* element, even though it may have considerable mutual coupling to other elements which are more intimately or more directly connected to the main transmission line.

CHAPTER FOUR

Low and Medium Frequency Antenna Systems

VERY LOW FREQUENCY ANTENNAS

The main object in the design of very low frequency transmitting antenna systems can be summarized briefly by saying that the general idea is to get as much wire as possible as high in the air as possible, and to use excellent insulation and an extensive ground system.

Most of the very low frequency or v-l-f transmitting installations now in use have been in service many years, because the limited number of channels available in that portion of the spectrum long have been fully utilized, and also because most long distance circuits can be provided more economically by the use of high frequency installations. In view of the fact that the number of v-l-f stations is comparatively small and virtually no new v-l-f transmitting installations are being made, the author does not feel justified in devoting more than brief mention to v-l-f antenna systems.

Physical and economical limitations on radiator height result in a very low radiation resistance in the v-l-f range. Therefore, various means are employed to make the radiation resistance as high as possible and to keep the loss resistance (ground resistance, tuning or "loading" coil resistance, etc.) as low as possible. Even so, the total loss resistance usually is several times the radiation resistance.

In some installations a "multiple" antenna system is employed to increase the radiation resistance and lower the ground resistance, thus increasing the radiating efficiency of the system. The "multiple" arrangement employs several vertical conductors spaced an appreciable distance, each working against its own ground and all effectively working in parallel. By manipulation of the phase angle between vertical elements, various directional effects are possible, and sometimes are employed to give a further effective power gain.

V-l-f radiating elements invariably are "top loaded" or are "capacity loaded" by means of a *flat top* (a number of wires parallel with each other and ground). This makes the current in the vertical or main radiating portion of the antenna substantially constant, effectively increasing the radiation resistance and substantially lowering the peak voltages on the antenna system. The latter is an important consideration when high power is employed in the v-l-f range, because of difficulties with corona. Even with a large flat top it is necessary to employ corona shields and take other precautions to prevent corona losses when the power is sufficiently high and the frequency sufficiently low.

LOW-FREQUENCY ANTENNA SYSTEMS

The foregoing remarks on v-l-f antenna systems apply equally well to antenna systems in the lower portion of the "l-f" or low frequency range. The upper portion of this range falls in a somewhat different category, however.

Marine transmitting antennas in the l-f range usually consist of an inverted L or T radiator using a single conductor, though some of the older installations still in use employ a "cage" or "flat top" instead of a single wire. The antenna is made as large as the physical limitations of the ship permit, and is kept as much in the clear as possible. On small ships the same antenna often is used on a harmonic for h-f operation, but on large ships separate half-wave resonant doublets usually are employed for h-f operation.

Large fixed-station installations in this range usually employ "loaded" vertical radiators with extensive ground systems. High-power foreign broadcast stations in this frequency range employ vertical radiators which sometimes approach a quarter wavelength in height, making them considerably higher than the average AM broadcast band radiator used in this country.

Special navigational services in the l-f range (such as radio range beacon stations) employ specialized l-f transmitting antenna systems which in general are not especially complex, but because of their many types and variations they will not be described here. The basic elements of such systems are the Adcock antenna (two, anti-phased vertical radiators), the single-turn loop antenna, and the vertical radiator. Because the maximum range required of such services is about 200 miles, a high order of radiator efficiency is not as important as the directional characteristics, and radiation efficiency is often sacrificed in this service in order to permit economical realization of a certain directional characteristic.

V-L-F AND L-F GROUND SYSTEMS

Because much of the heat loss experienced with v-l-f and l-f transmitting antenna systems occurs as a result of soil resistance, rather elaborate ground systems are the general rule. In the case of good soil, the usual arrangement is a network of buried wires extending somewhat beyond all elevated portions of the radiator. When the antenna is very extensive or is of the multiple type and several ground networks are used, the usual procedure is to connect the various networks together and to the transmitter via an overhead conductor or transmission line. Lumped reactances sometimes are inserted in the overhead ground returns in order to obtain a desired phasing characteristic.

In the case of very rocky or poorly conducting soil a *counterpoise* often is substituted for a buried network of wires. A counterpoise is a network of wires placed *above* the earth a slight distance and insulated from it, so arranged as to produce a very high capacity to the earth.

V-L-F AND L-F RECEIVING ANTENNAS

While v-l-f and l-f transmitting antenna systems usually are quite large physically, the separate receiving antennas which are normally employed are, in many cases, quite unpretentious. This is explained by the fact that any v-l-f or l-f receiver having good sensitivity will get down to the atmospheric noise level with a comparatively short antenna or modest loop. Under such conditions an increase in antenna directivity will provide an effective power gain, even though no more signal is delivered to the receiver, by improving the signal-to-noise ratio. However, to obtain substantially more directivity than is achieved with a simple loop or small Adcock antenna requires an array which is large in terms of wavelength, and wavelength at these frequencies represents many thousands of feet. In spite of their great physical size, v-l-f and l-f receiving arrays are used in some long-distance point-to-point circuits. But they differ from the transmitting arrays used at these frequencies in that the *height* is quite modest, because receiving antenna efficiency is not of great importance at these frequencies (for the reasons previously discussed).

The various types of loop antennas which are commonly employed in the v-l-f and l-f frequency ranges are discussed briefly in a later chapter devoted to the basic principles of direction finder antennas and other navigational antenna systems.

MEDIUM FREQUENCY ANTENNAS; AM BROADCAST ANTENNAS

Antenna systems employed in the lower end (between 300 and 550 kc.) of the m-f range fall in the same category as those in the upper portion of the l-f range and the same considerations apply.

The most important portion (or at least the best known portion) of the m-f band is the AM broadcast band between 550 and 1600 kc. Modern practice is to use one, sometimes two, and occasionally three, four or five steel masts which serve as vertical radiators, the physical height ranging from somewhat less than a quarter wavelength to slightly more than a half wavelength. A single tower is used for nondirectional radiation, and two or more when a certain directional characteristic is desired.

Tower Types

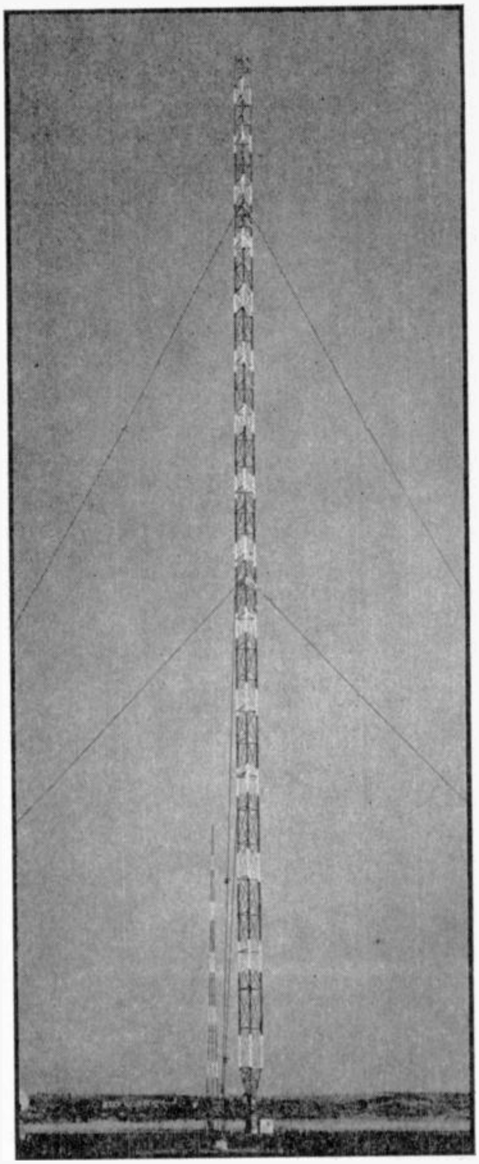
Towers of late design usually are one of two main types: a slender tower of uniform cross section, sometimes guyed and sometimes self-supporting; or a tapered (wide base) tower having considerable variation in cross section, usually self-supporting but sometimes guyed.

A tower of uniform cross section possesses the best electrical characteristics, provided they are not compromised by poor guying, and the characteristics usually can be predicted quite accurately by quantitative methods. A tower of uniform cross section can be erected to a certain height with assurance that it will have (within acceptable tolerances) a certain electrical height.

Top Loading

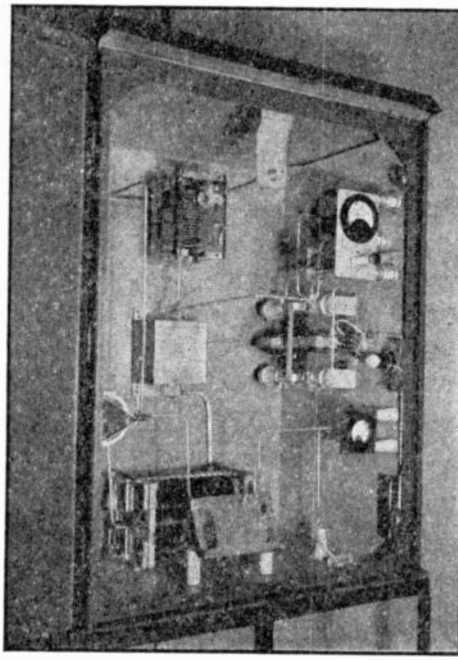
The electrical characteristics and performance of a tapered tower can be made to approach those of a uniform tower by incorporation of a lumped capacity called a "hat" at the top. However, the electrical length of a top-loaded, tapered tower cannot be predicted with any degree of certainty, and usually the top loading must be adjusted experimentally to give the desired electrical length. An exception, of course, is where experimental data is available for a similar tower used on the same or a slightly different frequency.

A top loading "hat" is more effective for a given size if it is insulated from the tower proper and connected to it via an inductance. Also, variation of the inductance provides a convenient means of adjusting the electrical length of the radiator. However, the mechanical construction is not so simple as in the case where the "hat" is connected directly to the tower both physically and electrically.



GUYED TOWER RADIATOR OF UNIFORM CROSS SECTION (TRUSCON) AT WNAX, 927 FEET HIGH.

This tower is typical of the taller AM broadcast station towers of uniform cross section. In the background is a tapered, self-supporting tower 325 feet high.



TYPICAL COUPLING UNIT (E. F. JOHNSON CO.) FOR AN AM BROADCAST RADIATOR, WITH FRONT PANEL REMOVED TO SHOW INTERIOR CONSTRUCTION.

A coupling unit of the general type shown often is placed at the base of an AM broadcast radiator to match the radiator to the transmission line, protect the line in case of antenna shorts or open circuit, and to provide lightning protection.

The current distribution on a tower of uniform cross section is substantially sinusoidal, but the distribution departs considerably from sinusoidal when the tower has a cross section which varies appreciably. The distribution on a tapered tower becomes very nearly sinusoidal, however, when a significant amount of top loading is employed. But the current no longer aproaches zero near the top.

The electrical length of an unloaded, *tapered* tower usually is appreciably *less* than its physical length, while the electrical length of an unloaded, *uniform* tower is slightly *greater* than its physical length. The electrical length of either type can be increased as desired by incorporation of top loading, though there is an economical limit to the amount of top loading than can be employed. A point is reached where it is cheaper to increase the height of the tower than it is to increase the size of the "hat" or its series inductances and insulation. However, when a station is close to an airport and is limited as to physical height by regulations, considerably more top loading may be employed than would otherwise be the case.

Optimum Electrical Length of Tower

The optimum length of a vertical tower radiator for a particular installation depends upon many factors when economy and sky-wave interference problems are considered. An electrical length of about 230 degrees (or approximately 0.63 wavelength) gives maximum surface wave amplitude for a given transmitter power when the base of the radiator is at ground level, but produces a strong high-angle lobe which is capable of causing objectionable sky-wave interference inside the limit of ground wave coverage. When the tower base is at ground level, this high-angle lobe occurs for any electrical length exceeding 180 degrees (one half wavelength), but it does not have sufficient amplitude to be objectionable until the electrical length exceeds 190 degrees. As the surface wave field strength is 90 per cent as strong for an electrical length of 190 degrees as it is for an electrical length of 230 degrees, and because a length greater than 190 degrees requires recourse to rather elaborate techniques to suppress the high-angle lobe, an electrical length exceeding 190 degrees seldom is employed.

If an efficient, low resistance ground system is employed, the surface wave field strength drops less than 25 per cent as the electrical length of the radiator is decreased from 190 degrees to 90 degrees (one quarter wave length). For this reason it is more economical for low power "community" stations in the higher frequency portion of the AM broadcast band to employ radiators having an electrical length of approximately one quarter wavelength. High power "regional" stations at the lower end of the band can more profitably employ a radiator having an electrical length approaching 190 degrees.

Because of the foregoing considerations, the F.C.C. prescribes the antenna height to be employed by a particular station, thus ensuring the most efficient and practical utilization of facilities.

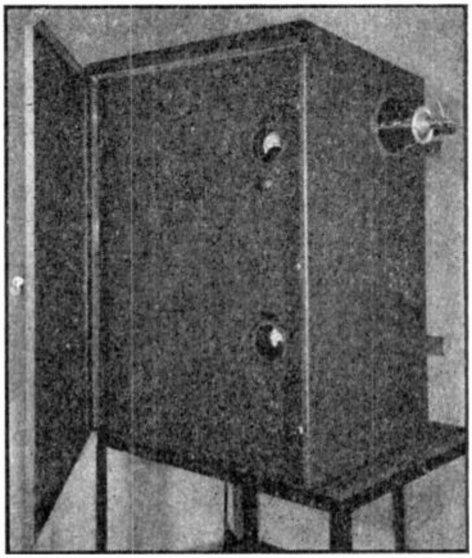
It was assumed in the foregoing discussion of optimum electrical length that there is not a great difference between the electrical and physical lengths. For economical amounts of top loading this condition obtains, and the surface wave field strength and space wave pattern will be very nearly the same for a given electrical height with or without the top loading. However, if an unusually large amount of top loading is employed, as in the case of a radiator near an airport, this no longer is true.

Feed Methods

Vertical tower radiators may be either series fed or shunt fed. With series feed the tower base is insulated from ground. With shunt feed the base of the tower is grounded. The two methods give comparable performance, but require different methods of matching the radiator to the feed line. In either case the radiator is matched by experimental variation of one or more lumped reactances, because the impedance offered by the tower can be calculated only aproximately. However, such calculations are used in determining what order of variable impedance to employ in the coupling network. The small structure used to house and weatherproof a coupling network at the base of a tower radiator is referred to as a "dog house."

The radiator most amenable to accurate calculation is one which has a uniform cross section, no top loading, and is exactly one quarter wavelength long electrically. But even with a radiator of this type, provision should be made for experimental adjustment.

Both coaxial and open type lines are used to feed tower radiators in the AM broadcast band. The former is to be preferred from a performance standpoint, but the latter is much less expensive. When an open wire line is employed it often is of the multi-wire type.

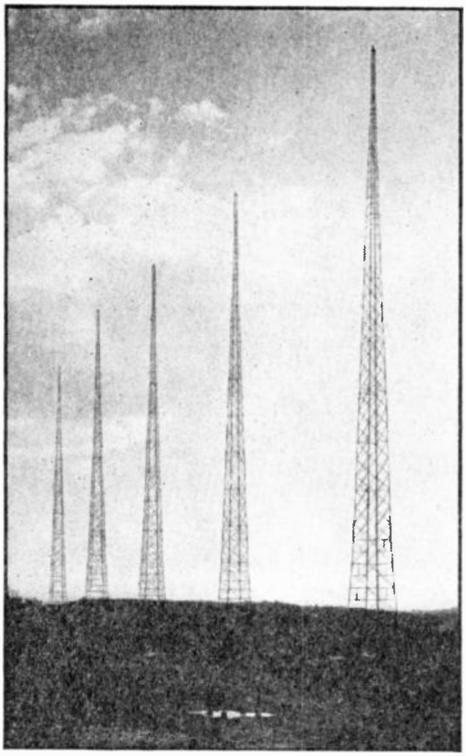


ILLUSTRATING WEATHERPROOF CONSTRUCTION OF TYPICAL ANTENNA COUPLING UNIT OR "DOG HOUSE" EQUIPMENT FOR AN AM BROADCAST STATION (E. F. JOHNSON CO.).

Insulation

The base insulator of a series-fed tower radiator should possess great mechanical and electrical strength, low loss characteristics, and not introduce an excessive amount of shunt capacity between the base of the tower and ground. The electrical requirements are most severe when the electrical length of the tower is one half wavelength, and least severe when the electrical length is one quarter wavelength.

Guy wire insulators should have a low loss factor, high mechanical strength, and be of the "strain" type so that in case of insulator failure the guy will still function. Guys usually are broken up at least every 0.1 wavelength and often every .05 wavelength or so. An insulator always is inserted between a guy and its point of attachment to the tower, and oftentimes an insulator with a longer leakage path is employed at this particular point, especially when the



FIVE TAPERED SELF-SUPPORTING TOWER RADIATORS (BLAW-KNOX) AT KQV, 350 FEET HIGH.

These base-insulated radiators are used in a phased directional array to give a desirable coverage pattern. In many cases two or three radiators will give the desired pattern, but under unusual conditions more are necessary.

point of attachment is a high impedance point.

Ground Systems

The most desirable ground system for a tower radiator which has its base at ground level is a buried network of wires consisting of a large number of "radials" extending out from the base of the tower at least a quarter wavelength and preferably a half wavelength. A worthwhile increase in field strength is obtained with an increase in the number of radials up to approximately 120, beyond which little further improvement is noted. The improvement in performance with increased number and length of radials is most pronounced with a very short radiator (considerably shorter than a quarter wavelength), but is still very pronounced and substantially uniform for practical radiator lengths of from one quarter to one half wavelength. Also, the improvement is more pronounced in the case of poor soil than in the case of good soil, but still is quite pronounced in the case of good soil.

Current general practice is to make the radials about ½ wavelength long physically, to bury them just deeply enough to afford sufficient mechanical protection, and to "ground" the far end of each radial with a rod driven as deeply as is practical into the subsoil when the latter is substantially more highly conductive than the topsoil.

When high voltages are developed between the lower end of the radiator and ground, as in the case of a half-wave vertical radiator, a "ground screen" or "ground plate" oftentimes is placed directly under the tower in order to minimize earth losses at this point. The screen or plate is much smaller in extent than a system of radials, and also differs in that it is either a solid sheet or effectively of much finer mesh. It is used in addition to, rather than in place of, the system of radials.

Tower Radiators on Buildings

The foregoing discussion assumed that the base of the AM transmitter tower was at or near ground level, as this condition usually prevails. However, the radiator of a low power metropolitan or "downtown" station sometimes is for practical reasons located atop a high office building. To minimize r-f currents flowing in and r-f voltage appearing on the plumbing and other metallic components of the building, the top of the building usually is covered with a closely spaced network of wires which are insulated from the building to form a "counterpoise." Unless the counterpoise is quite extensive in terms of wavelength, it will be found that the physical length of a self-resonant "quarter wave" must be considerably greater than for the case of a perfect ground of infinite extent (a condition which is approached in practice by a system of buried radials).

It is for these reasons that the term "length" rather than "height" has been used in the foregoing discussion of tower dimensions. The terms are interchangeable *only when the base of the radiator is at ground level.*

Directional AM Broadcast Arrays

In order to obtain a more advantageous coverage pattern or to "protect" another station on the same channel, or to accomplish both at the same time, an antenna system possessing horizontal directivity oftentimes is desired. This requires the use of more than one tower radiator. Directivity control is by choice of tower spacing and orientation, and by manipulation of the magnitude and phase of the current in each tower.

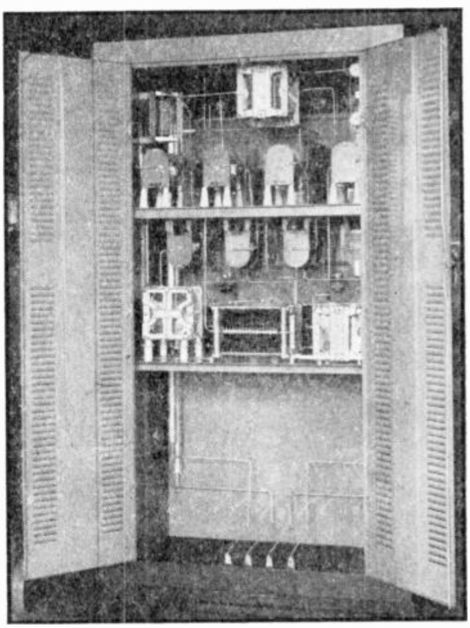
Various mechanical and electro-mechanical "plotters" have been developed which facilitate the determination of the pattern obtained with a given combination of parameters. However, the engineer usually is more concerned with analysis than synthesis, as the desired pattern usually is the known and the required parameters the unknown. A very useful device which greatly facilitates either synthesis or analysis is the "Antennalyzer" developed by Brown and Morrison. The action is entirely electrical. Its use is described briefly as follows.

Four sets of four potentiometers enable the operator to set up the various tower factors of four radiators. These are: the distance from the reference (fifth) antenna; the azimuth angle; the magnitude of the current; and the phase angle of the current in each radiator. Thus, the device will handle any conventional array having not more than five elements. Because the resulting pattern is displayed instantaneously on a cathode ray tube, it is a practical matter to manipulate the dials until the desired pattern is obtained. The dial settings then indicate the correct tower location as well as the required magnitude and relative phase of the current in each tower. At the option of the operator, the pattern is displayed in either rectangular or polar coordinates.

When the desired horizontal pattern can be approximated closely with either an end-fire or a broadside arrangement, the end-fire combination usually is to be preferred. An end-fire array provides a slight increase in vertical directivity along the line of maximum radiation, thus increasing the distance at which interference fading is first encountered.

Oftentimes it will be found that the desired current relationship can be realized or closely approached by the use of one or more parasitic elements. When this is not the case, the elements must be driven.

Design of a variable coupling and phasing network which will permit experimental adjustment of the various tower currents to the exact magnitude and phase desired, requires a thorough knowledge of the methods used to calculate and measure antenna



TYPICAL DIRECTIONAL ARRAY PHASING EQUIPMENT FOR AN AM BROADCAST STATION (E. F. JOHNSON CO.).

A bank of reactive networks permits the magnitude and phase of the current fed to each radiator in an array to be adjusted to the values required to produce the desired directivity pattern.

impedances. These techniques are described in detail in a pamphlet "Impedance Measurements on Broadcast Antennas" issued by the General Radio Co., Cambridge, Mass.

Receiving Antennas for the AM Broadcast Band

The more modern AM broadcast receivers (except for the very cheapest "midgets") usually are equipped with a self-contained loop antenna which gives adequate pickup on all except perhaps the weakest stations. Provision is made, however, for attaching an external antenna when the pickup of the loop is inadequate or it is desired to employ a special noise-reducing antenna. The latter is simply a combination of antenna, transmission line, and impedance transformer which permits placing the antenna proper as far from a-c power wires as posible, because usually most of the electrical interference picked up by a simple AM receiving antenna is conducted to the vicinity of the antenna via the power wires.

ANTENNAS FOR 1600-3000 KC

Transmitting antennas used in the lower end of the 1600-3000 kc. range for applications requiring maximum practicable surface wave coverage (such as police radio systems) resemble those used by low power AM broadcast stations at the upper end of the AM broadcast band, as the same considerations apply.

Transmitting antenna systems used in the upper end of the 1600-3000 kc. range for short and medium distance skywave transmission are similar to the h-f antenna systems described in the following chapter for use at the lower end of the h-f range. The basic antenna is a half-wave horizontal doublet supported at a height of from 1/16 to 1/2 wavelength above ground, and the difference between the various systems is simply a matter of how the radiator is fed.

CHAPTER FIVE

High-Frequency Antenna Systems

> **Note on Antenna Schematics**
>
> When the electrical length, height, etc., is not necessarily the same as the physical, the subscript E is used in the schematics throughout the following chapters to indicate an "electrical" dimension, and the subscript P is used to indicate a physical dimension, whether in terms of feet or in terms of free space wavelength. Thus, L_P refers to physical length, L_E to electrical length. When the two are the same, as in the case of the spacing "S" between two radiators, no subscript is used.

With the exception of certain portable military equipment and except for a few applications at the extreme ends of the range, communication at frequencies between 3 and 30 Mc. normally is via sky-wave propagation. The range 3 to about 7 Mc. is utilized primarily for short and medium distance sky-wave communication via the F_1 or F_2 layer, usually the lower end of the frequency range for nighttime work and the higher end for daytime work. Highly directional antenna systems seldom are employed for short and medium distance sky-wave communication; the antenna most commonly used is a half-wave horizontal radiator or receptor. Such an antenna shows little directivity when used for high-angle (short and medium range) sky-wave communication.

The range from approximately 7 to 30 Mc. is employed primarily for long distance sky-wave communication (including international "shortwave" broadcasting), and highly directional antenna arrays are the general rule. At the higher frequencies it is practical and economical to resort to arrays which are large in terms of wavelength for point-to-point work or for broadcasting to a comparatively narrow sector, thus realizing gains up to approximately 20 db over a half-wave dipole at the same average elevation.

As there is an almost infinite variety of combinations of radiating elements which will provide a wave interference pattern which results in useful directivity and power gain, space can be devoted to a description of only the more popular and more representative of the types in general use. First, however, the simple half-wave antenna will be considered in detail.

THE HALF-WAVE ANTENNA

A resonant radiator or receptor an electrical half wavelength long commonly is referred to as a "half-wave dipole," a "half-wave doublet," or a "half-wave Hertz." Such an element serves as the basic unit of many directional arrays. It also is widely used alone, either horizontally, or sloping, or substantially in one of these two configurations, for short and medium distance sky-wave communication. In fact, it may be considered as the fundamental form of h-f antenna.

There are dozens of methods of feeding a half-wave antenna, and all will give about equally good results when properly employed. Because in certain arrangements the feed system is so closely tied up with the operation of the antenna proper, the feed system will be treated as an integral part of the antenna system, and discussed concurrently.

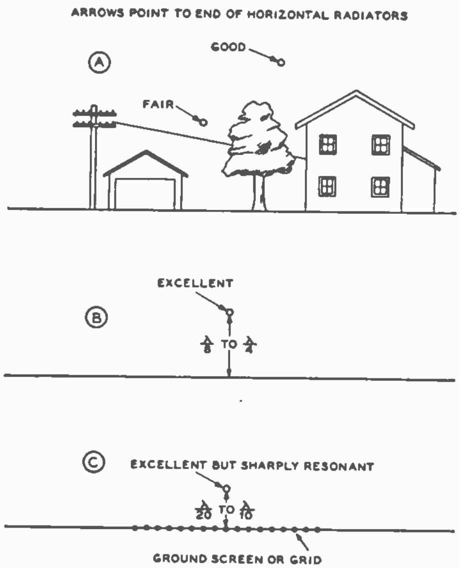


Figure 5-1.

ILLUSTRATING RELATIVE MERITS OF VARIOUS INSTALLATIONS OF HORIZONTAL RADIATOR FOR WORKING OVER DISTANCES AT WHICH HIGH ANGLE RADIATION IS EFFECTIVE.

The representation applies to F_1 or F_2 refracted frequencies between 3 Mc. and 7 Mc. and distances up to approximately 200 or 300 miles (depending upon the height of the controlling ionosphere layer). The B and C installations will in general be about 5 db better than the "fair" installation at A, and about 3 db better than the "good" installation at A.

Communication Via High-Angle Sky Wave

For short and medium distance work via sky-wave propagation over ranges of 0 to 300 miles, the effective radiation (or response) is that which occurs at elevations above about 45 degrees. For such work the optimum height of a horizontal half-wave radiator is between ⅛ and ¼ wavelength above the surface of the earth, and little difference in signal strength will be noted between the broadside and end-on aspects of the radiator, if level terrain and propagation via the F_1 or F_2 layer are assumed. When the antenna is located over or among irregular "scattering" objects such as overhead power and telephone wires, trees, buildings, etc., the optimum height is difficult to predict, though in general it may be said that the radiator should preferably be sufficiently high so as to be well-removed from surrounding objects, in order to minimize absorption losses.

The signal strength at distances less than 150 or 200 miles can be improved by placing the radiator over a perfectly flat piece of ground which is clear for a distance of at least ½ wavelength in every direction. A radiator so located and elevated ⅛ wavelength above the surface of the earth will deliver a signal at 150 miles which is, on the average, several db better than from a radiator which is much higher but is located over "scattering" terrain. (See figure 5-1.)

As the height of a horizontal radiator over smooth terrain of average conductivity and dielectric constant is reduced below ⅛ wavelength, the ground losses become appreciable as a result of the increased strength of the induction field. Were it not for this factor, maximum signal strength over short distances could be obtained with very low antenna elevations, on the order of 1/16 wavelength.

When it is impossible or impractical to elevate the antenna as much as 1/10 wavelength above flat ground, the short-range signal strength can be made to approach that obtained under optimum conditions by installing a ground screen under the radiator, the screen consisting of numerous wires buried in or laid on the ground, parallel to each other and to the antenna, as in figure 5-1C. The wires should be somewhat more than a half wavelength long, and be placed about every 2 degrees of arc for slightly over 90 degrees as viewed from the antenna.

Making use of this expedient, the radiation efficiency can be kept high even with antenna elevations on the order of 1/20 wavelength, but the Q of the antenna tends to become objectionably high, restricting use of the antenna to a very narrow frequency range, unless provision is made for tuning the system to resonance, or measures are taken to broaden the response. Fortunately the sharpness of resonance can be minimized somewhat by resorting to one of several schemes which will be discussed later.

The arrangement of figure 5-1C is the only one to which idealized "perfect ground" patterns apply even approximately. Therefore, unless this type of installation is employed, "perfect ground" curves are more or less of academic interest.

It should be kept in mind that in all compass directions except perpendicular to the radiator, the radiated wave has a vertically polarized component, and that at all vertical angles directly off the ends only the vertical component exists. When used for reception, this also applies to the response. This is an important consideration when such an antenna is employed for reception, because when the antenna is located over either average or poor ground the antenna will have considerable response to vertically polarized noise waves arriving end-on at grazing angles, due to the fact that the imperfect earth causes a wave tilt, as explained in chapter 1 in connection with surface wave propagation. At low angles above approximately 3 degrees another mechanism contributes to the end-on radiation and response in the following manner, particularly when the horizontal radiator is not so close to the earth that low angle sky waves merge with surface waves.

Below the pseudo-Brewster angle the phase shift of a vertically polarized wave on reflection exceeds 90 degrees, causing the direct and earth reflected waves from a *vertical* radiator to cancel. As zero elevation is approached the reflection coefficient acquires a magnitude approaching 1.0 and a phase angle approaching 180 degrees, resulting in virtually complete cancellation of the direct and ground reflected components of the space wave (though the antenna may still be responsive to grazing space waves which are effectively converted to surface waves close to the earth, and likewise may transmit surface waves which excite low angle space waves).

In the case of *end-on* radiation from a *horizontal* antenna, however, there is an extra phase reversal which puts the direct and ground reflected components of the low-angle space wave back in phase. This is due to the fact that the direction of the vertical component of the electric vector of the ground reflected wave is opposite to that of the direct wave *as they leave* the horizontal radiator. Thus we see that if it were not for aspect directivity, a low horizontal radiator over typical soil would provide much better low angle space wave transmission or response in the h-f range (in an end-on direction) than would a vertical radiator at the same height above the same soil.

Utilizing this appreciable end-on response and radiation at low vertical angles, it is sometimes possible to hear and transmit to distances over 2000 miles using modest power in a low, horizontal antenna over flat terrain, though there is a common misconception that such an antenna always exhibits greatest low-angle radiation and response in a direction normal to the radiator rather than end-on. The effect is most pronounced with low antenna heights (less than ¼ wavelength) and poor soil, the former tending to minimize the low-angle broadside response and radiation, and the latter tending to increase the low-angle end-on response and radiation.

The polar diagrams of figure 5-2 illustrate the end-on and broadside vertical directivity of a low horizontal radiator in the range 3 to 7 Mc. over perfect ground and typical soil. The curves are for a constant amount of power fed to the radiator; they take into account the impairment of radiation efficiency due to earth losses. They also assume that the terrain is flat for considerable distance in every direction, a condition seldom encountered in actual practice. The effect of scattering terrain such as figure 5-1A is a moderate reduction in the average amplitude of the high-angle radiation and response in all horizontal directions, and a great increase in the low-angle broadside response.

For instance, the antenna of figure 5-1A marked "fair" may be much more effective in a broadside direction at vertical angles below about 20 degrees than the installation of 5-1B, provided that the type of terrain shown under each radiator continues for considerable distance. This is explained by the fact that with a low radiator over flat terrain the ground reflected waves buck the direct radiation at angles below about 30 degrees, instead of reinforcing it, and

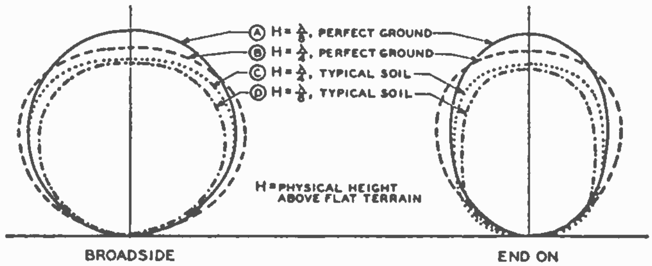


Figure 5-2.

BROADSIDE AND END-ON VERTICAL DIRECTIVITY OF A HORIZONTAL HALF-WAVE ANTENNA OVER PERFECT AND TYPICAL GROUND FOR FREQUENCIES BETWEEN 3 AND 7 MC. AND HEIGHTS COMMONLY USED FOR SHORT AND MEDIUM DISTANCE SKY-WAVE COMMUNICATION AT THESE FREQUENCIES.

The curves represent field strength for the same amount of power fed to the antenna in each case, allowance having been made for resistance losses in imperfect ground. The important things to note are that a low horizontal antenna shows little horizontal directivity at the high angles useful for sky-wave communication via the F_1 or F_2 layer over distances up to about 300 miles, and that over typical soil of average conductivity the antenna actually is more effective end-on than broadside at very low or grazing angles. Curves C and D are affected only slightly by differences in soil ranging from poor to good, but for highly conducting earth like salt marsh the perfect ground curves are more appropriate.

by scattering and dissipating much of the ground directed energy the resultant amplitude of the space wave is actually increased at low angles. At the very low angles effective for propagation over extreme distances, the installations of figure 5-1A may be as much as 10 db better in a broadside direction than the arrangement of figure 5-1B or 5-1C.

In an end-on direction over typical flat earth the ground reflected wave from a low horizontal radiator *aids* the direct wave at very low angles, and therefore scattering and dissipating the ground reflected component usually will *decrease* the amplitude of the resultant space wave, though this depends somewhat on the nature of the scattering terrain.

Receiving Considerations

The low-angle response of a "high angle" antenna intended for short and medium distance reception is of considerable importance from a receiving standpoint because most of the *atmospheric* noise existing in the lower portion of the h-f range in the continental U. S. A. tends to arrive at very low angles from equatorial regions (excepting, of course, noise due to local thundershowers).

When a horizontal antenna is being used for high-angle, short or medium distance reception, it is advisable to orient the antenna so that the antenna is broadside to the distant noise sources. This means that the antenna will be erected in a more or less east-west direction. When this orientation is impossible or impracticable, and it is necessary to run the radiator in a north-south direction, then end-on response to low-angle waves arriving from the equatorial regions can be minimized by tilting the antenna at an angle of from 5 to 15 degrees, with the high end towards the noise source. The optimum amount of tilt is rather critical and depends upon the frequency and the ground characteristics. It is best determined experimentally for each installation.

In some cases it will be found profitable to install a small "anti-static" antenna for reception purposes only, for short distance sky-wave reception on frequencies in the 3 to 5 Mc. range, rather than attempt to employ the same antenna for both transmission and reception. Such an anti-static antenna can be installed as a standby antenna to be used only when signal strengths are strong but atmospherics are bad. Or, if the received signal strength is sufficient at all times to override the inherent receiver background noise when using the anti-noise antenna, the latter antenna may be permanently connected to the receiver, the transmitting antenna in this case serving exclusively as a transmitting antenna.

An anti-noise antenna of this type is illustrated in figure 5-3. It will deliver considerably less signal to a receiver than will a good transmitting antenna, but when the transmitting antenna cannot be installed in such a manner as to minimize equatorial static, the improved signal-to-noise ratio afforded by the anti-noise antenna oftentimes will permit reception of signals which otherwise would be unintelligible because of atmospheric noise. Also, if the anti-noise antenna is made comparatively short it may be easier to keep it away from power wires and other conductors and radiators of local electrical noise than is the case for a half-wave antenna.

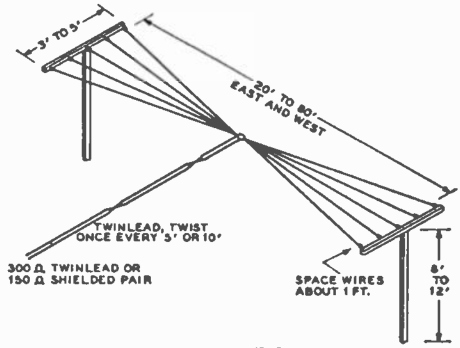


Figure 5-3.

"ANTI-STATIC" RECEIVING ANTENNA FOR SHORT AND MEDIUM DISTANCE SKY-WAVE RECEPTION AT FREQUENCIES BETWEEN 3 AND 7 MC.

If the anti-static antenna is used only when atmospherics are bad, it may be short (20 to 40 ft.) and the feed line need not be resonated at the receiver input. If it is used at all times, it should be longer (40 to 80 ft.) and the overall system (antenna plus feed line) should be resonated to the received signal at the receiver input. For maximum reduction of both tropical static and local man-made electrical noise the feed line should leave the antenna at approximately a right angle for some distance and the receiver input circuit should be symmetrical ("balanced"). When the receiver input is not perfectly balanced and free of stray capacity coupling, a worthwhile improvement in performance can be obtained by substituting for the Twin-Lead a line of 300-ohm 3-wire (neutral ground) ribbon or a line of 150-ohm shielded pair (of the type employed for television receiver lead in).

Half-Wave Antennas for Moderately Long Distance Communication

For some military and commercial applications such as portable work, or where the necessity for simultaneous omnidirectional transmission or reception makes a high degree of horizontal directivity undesirable, half-wave antennas sometimes are employed effectively for communication over moderately long distances (say, 500 to 1500 miles) at frequencies up to 12 Mc. and even higher. Radio amateurs unable or reluctant to erect a more elaborate antenna system have spanned the globe with low power using half-wave antennas at frequencies up to 30 Mc., though it is much more easily accomplished with a high gain directional array.

For single-hop sky-wave communication via the F_1 or F_2 layer over distances of 500 to 1500 miles, the optimum height of a horizontal half-wave radiator *over flat terrain* is between ½ and ¾ wavelength. Such an antenna will show noticeable horizontal directivity at these distances, but not enough to degrade the end-on effectiveness appreciably. When the terrain near the antenna is such as to cause considerable scattering of the ground-directed wave, the optimum height for communication over these distances is best determined by experiment, though usually there will be little difference in performance once the antenna is elevated enough to top the highest surrounding structure.

Vertical half-wave antennas once were commonly used for omnidirectional h-f communication over medium and long distances but no longer are widely employed.

FEEDING THE HALF-WAVE ANTENNA

The most practical method of feeding

a half-wave antenna depends upon various considerations involved in the particular installation, and usually the considerations involved will permit a choice between several arrangements of comparable suitability. The advantages, disadvantages and peculiarities of each of some of the more popular arrangements will be discussed in sufficient detail that the reader should be able to choose one which is well suited to his particular requirements.

It should be kept in mind that while the following discussion refers specifically to the problem of feeding power from a transmitter to an antenna, it applies also to the feeding of energy from an antenna to a receiver. When a transmission line is involved, only the matching of the line to the antenna and vice versa will be considered at this time, as the problem of coupling a line to a transmitter or receiver is a separate problem which can be treated without regard to the type of antenna in use, and will be so treated in a later chapter.

THE END-FED HALF-WAVE ANTENNA

A low half-wave horizontal antenna may be "voltage fed" simply by bringing one end down to the transmitter. If the overall electrical length is a half wavelength and the flat top runs in a straight line, the reactance presented to the transmitter will be zero and the feed point resistance usually somewhere between 3000 and 10,000 ohms, the exact value depending upon the height of the "flat top," the diameter of the conductor, the soil conductivity, and the presence of surrounding objects.

If the feed point resistance is on the order of or lower than the optimum plate-to-ground load impedance of the final amplfier stage, the antenna may be tapped on the tank coil at a point which gives the desired loading, as in figure 5-4A. A series blocking condenser should be inserted if plate voltage on the final stage is series fed. The radiator may be operated slightly off its natural resonant frequency simply by "tuning out" the reactance by means of C_1 in figure 5-4A.

If the feed point resistance of the antenna is too high to provide sufficient loading even when the antenna is clipped directly to the "hot" end of the tank circuit, either the tube plate may be tapped down on the tank, or the feed point impedance of the antenna may be lowered by using either a much larger diameter conductor or else two parallel wires spaced from one to three feet. The use of two parallel wires has the effect of increasing the diameter of the radiator considerably, and either expedient will lower the impedance of the radiator at a voltage loop. The two conductors should be tied together at the transmitter and may be either shorted or left open at the far end.

The use of two conductors also will increase the frequency range over which the antenna may be operated without an objectionable amount of reactance being presented to the tank circuit. When an excessive amount of reactance is present, it is not practicable to compensate for it by tuning of the final amplifier tank circuit.

At frequencies in the lower end of the h-f range this system of direct excitation is just about as efficient a way of feeding a low horizontal radiator as using a transmission line, but ordinarily is employed only for low-power portable or emergency work because there is no attenuation of harmonics which are present in the final amplifier tank circuit.

Discrimination against harmonics and the ability to work over a wider frequency range can be obtained by the incorporation of an intermediate parallel L/C circuit between the transmitter and antenna. This circuit, which is essentially a tank circuit, is most conveniently coupled to the final amplifier tank by means of link coupling, as shown in figure 5-4B; adjustments in loading are more easily accomplished with link coupling than if direct magnetic coupling were employed between the two tank circuits.

A reduction in low-angle noise pickup can be accomplished by making the overall electrical length of the radiator slightly longer than one half wavelength, so that the voltage loop occurs part way up the vertical portion of the radiator and the vertically polarized noise waves picked up

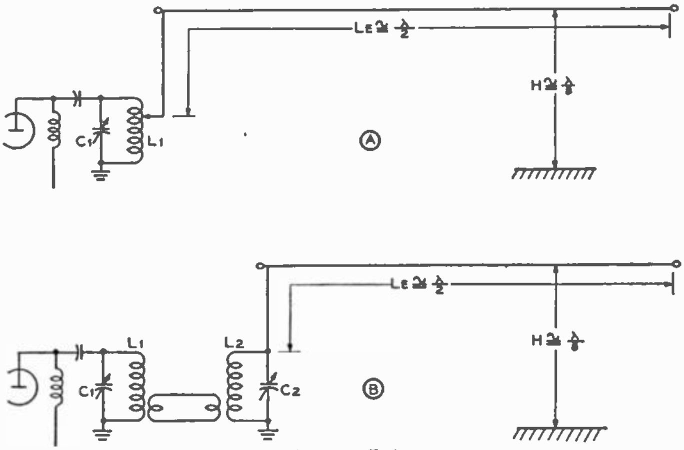


Figure 5-4.

SIMPLE END-FED ("VOLTAGE-FED") HALF-WAVE HORIZONTAL ANTENNAS.

While the arrangement at A is quite efficient when operating under optimum conditions, arrangement B is to be preferred because of its high discrimination against harmonics. If a short, heavy grounded lead is used, its length need not be considered when calculating the required length of the antenna proper. Antennas of the type illustrated seldom are used at frequencies above approximately 7 Mc.

by the various parts of the antenna tend to cancel. The optimum overall electrical length for low-angle noise reduction depends upon the ground constants, radiator height, and other factors, and is not necessarily the same for end-on pickup as for broadside pickup. A compromise value usually will be obtained with an overall length which causes the voltage loop to occur slightly above half way up the vertical portion of the radiator.

The radiator can be matched satisfactorily over a much wider frequency range by substituting a pi or L section network for the tuned tank circuit, but these coupling devices do not afford the foolproof attenuation of harmonics that is obtained with the resonant tank coupler. It is recommended that when the frequency range to be covered exceeds that which can be handled conveniently with a single resonant tank circuit (or roughly 10 per cent), the resonant tank coupler still be employed and the radiator sectionalized with one or more insulators which can be shorted out by means of a switch operated from ground by means of a cord.

The plate spacing of condenser C_2 is determined by the voltage at the end of the radiator, which in turn is determined by the transmitter power and the feed point impedance of the radiator. Though the value usually will be lower, a safe figure to use for the feed point impedance in determining the required condenser spacing is 10,000 ohms for a single conductor of no. 12 wire and 7000 ohms for two parallel wires of the same size. These figures assume that the "flat top" is fairly well in the clear, runs in a straight line, and is not located at a very low height over salt marsh or a ground screen. If the flat top has a pronounced bend or is located very low over a highly conducting surface such as salt marsh or ground screen, the feed point impedance may be considerably higher than these values.

The smallest maximum capacity rating suitable for C_2 is determined by the minimum C/L ratio which will provide sufficient loaded Q to permit the desired coupling to the transmitter to be obtained by adjustment of the coupling links. If too low a C/L ratio is employed, it will be impossible to obtain sufficient coupling without resorting to rather large and unwieldy coupling links, and the loaded Q of the antenna tank will be too low to afford good discrimination against harmonics. Also, there should be some excess capacity available over and above that required to resonate the coil to the operating frequency, in order to permit "tuning out" of a moderate amount of reactive component in the feed point impedance of the antenna. A variable capacitor having a maximum capacity of .0005 μfd. usually will suffice for frequencies in the range of 3 to 5 Mc., the range in which this type of antenna commonly is employed.

When a capacitor is available having considerably more spacing than required for the arrangement of figure 5-4B but with somewhat less maximum capacity than is required for a C/L ratio which will permit satisfactory loading and provide good harmonic attenuation, the antenna may be tapped down on the coil to increase the loaded Q.

The same effect is obtained by moving the ground connection to a tap on the coil, or removing the ground connection and letting the coil "assume" a ground point or node, the location of which is determined by the amount of stray capacity between the "unused" end of the tank and ground and between the link and the coupling coil.

Either tapping the antenna half way down the coil or grounding the center of the tank produces substantially the same effect upon the loaded Q, which is to raise it by an amount corresponding to an increase in the C/L ratio of approximately 4 times. "Fudging" the stray capactiy so that without a direct ground the voltage node occurs at the center of the coil also produces substantially the same result. However, it is preferable to ground the bottom end of the tank and tap the antenna down on the coil, as this permits a cold rotor.

Because of the very high feed point impedance offered by the antenna proper, the ground lead may have appreciable impedance without greatly affecting the tuning of the antenna tank, especially when the overall electrical length of the antenna is exactly a half wavelength. Also, because of the high feed point impedance, the amount of current flowing in the ground wire is so small that its effect upon the radiation pattern may be neglected except when the ground lead is quite long or the antenna proper is operated considerably removed from its natural resonant frequency.

It has been assumed in the foregoing discussion that the antenna tank is right at the transmitter location, and that the length of the link line is an insignificant fraction of a wavelength. When this condition prevails, the characteristic impedance of the short link line and the question of whether it is working under flat line conditions may be ignored. So long as the link line is short and of low loss construction, consideration need be given only to whether the desired loading can be obtained.

Sometimes, in order to minimize the amount of radiator contained within the building housing the transmitter, or to get the "hot" part of the antenna away from the transmitter, the antenna tank is placed some distance from the transmitter. In this case the link line no longer can be considered as simply a part of the coupling mechanism and treated as part of a lumped circuit, but rather must be treated as a transmission line if best performance is to be obtained. 300 ohm "Twin-Lead," rather than low impedance cable, is recommended as the link line except where the power is excessive for this type of line. It is easier to obtain the desired operating conditions when 300-ohm Twin-Lead is employed, and its use minimizes the line loss when a mismatch does exist. When the line is properly adjusted it should be possible to substitute a 300-ohm noninductive resistor at the far end of the link line with negligible effect upon loading or tuning of the final amplifier tank circuit. However, considerable deviation from this optimum condition can be tolerated when Twin-Lead is employed, with no significant effect other than to make the overall system more frequency sensitive (thus requiring that C_2 be retuned for even small changes in frequency).

A reduction in harmonic radiation often

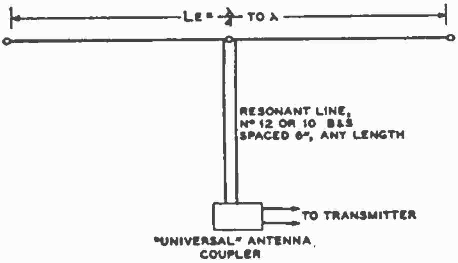


Figure 5-5.

CENTER FED ANTENNA WITH RESONANT FEED LINE.

When a suitable coupling network is employed, this antenna system can be used with good efficiency and substantially the same pattern over a continuous frequency range of about 4 to 1. However, other feed systems are to be preferred when this flexibility is not required. To minimize antenna effect, the feed line should leave the antenna at approximately a right angle for at least one wavelength, or until it is close to the ground.

can be realized by grounding one side of the link line at one or both ends, thus minimizing coupling to the antenna at the harmonic frequency via stray capacity. Also, if one side of the link line is connected to ground at the transmitter and the same side of the link line is connected to the rotor of C_2, the separate ground connection shown at C_2 often may be eliminated with no increase in harmonic radiation, especially if the link line is not unduly long. It is recommended that the combination affording minimum harmonic radiation be determined for each individual installation by experiment, because even a small amount of harmonic radiation from a 3 to 7 Mc. transmitter can cause serious sky-wave interference to other services. Generally speaking, maximum harmonic attenuation will be obtained with the antenna connected to the ungrounded side of the condenser, rather than tapped down the coil, if the loaded Q is made the same in both cases.

There are other methods of high impedance ("voltage") feed for end feeding a half-wave antenna which involve the use of a transmission line, but they will not be dealt with because of the fact that low impedance ("current") feed minimizes feed line insulation requirements and allows a more balanced arrangement, thus minimizing feed line pickup and radiation, or "antenna effect."

While not pronounced enough to be of much practical interest or concern, it should be mentioned in passing that the horizontal directivity pattern of an end-fed half-wave antenna (ignoring radiation from the vertical portion) does not maximize normal to the line of the radiator, but rather is "cocked" slightly *towards the unfed end*, causing the "noses" of the two lobes to be separated by slightly less than 180 degrees. This distortion of the theoretical pattern is a characteristic of all radiating systems which are not symmetrically fed; but because the horizontal pattern of a half-wave radiator is so broad the dissymmetry can in this case be ignored.

CENTER-FED ANTENNA WITH RESONANT FEED LINE

When it is undesirable or impracticable to bring one end of the antenna into the building which houses the transmitter and receiver, it may be fed in the exact center with a resonant feed line having a surge impedance of 500 to 600 ohms, as in figure 5-5. A half-wave radiator fed in this manner is referred to variously as a "center-fed zepp," a "tuned doublet" and a "center-fed Hertz."

The principal advantage of this antenna is that the radiator may be operated well off its natural resonant frequency without introducing antenna effect or appreciable line loss, provided that the radiator is symmetrical with respect to ground and surrounding objects. Its chief disadvantage is that the system is quite frequency sensitive, and a rather elaborate matching arrangement must be used at the transmitter unless the line is cut to a critical length which eliminates the reactive component at the transmitter and the system is used only over a very narrow range of frequencies.

When a suitably flexible matching arrangement is employed, the system may be used at any frequency between one half and twice the natural resonant frequency of the radiating portion without greatly affecting the efficiency or radiation pattern. The efficiency does fall off slightly when the electrical length of the radiating portion is considerably shorter than a half wavelength, and the horizontal directivity in a broadside direction is noticeably increased as the length approaches a wavelength, but when considered on the basis of signal strength in db, the difference is small. As the electrical length of the radiating portion is made longer than one wavelength, the horizontal directivity becomes more pronounced and the pattern acquires multiple "lobes," so that the system no longer can be considered as a nondirectional antenna. Such antennas are discussed later under the subject of directional antenna systems.

Under all conditions of operation the SWR on the open line will be rather high, and to keep line losses negligible the spacing insulators should be of low loss material. Also, if the antenna is used with a transmitter which delivers much power, the line spacing should not be less than about 6 inches or difficulty may be experienced with line flashover.

In determining the nature of the impedance that will be presented at the transmitter end of the line it should be kept in mind that a section of line an electrical half wavelength long "repeats" the load and a section an electrical quarter wavelength long "inverts" the load. However, as the main reason for using this type of antenna is its flexibility with regard to useful frequency range when a suitable matching network is employed, it can be assumed that such an antenna ordinarily will be used in connection with a "universal" coupler at the transmitter, in which case the exact length of radiator and line (within practical limits) are of minor importance.

If antenna tilt or surrounding objects cause dissymmetry, the overall system no longer will be balanced, and some antenna effect will be present on the line. In most applications a moderate amount of antenna effect can be tolerated, but when it cannot, the antenna should be placed over a clear, flat piece of terrain and not be tilted.

The mean SWR over the useful frequency range of the system can be reduced slightly by employing a tuned line having a somewhat lower characteristic impedance, say 200 or 300 ohms. However, because of the much closer spacing required for a line of this impedance, both the losses and the tendency towards flashover can be reduced by utilizing a comparatively wide-spaced line having somewhat higher characteristic impedance than that which would give the lowest average SWR over the useful frequency range (assuming a constant diameter for the line conductors).

Strictly speaking, this antenna is not a half-wave antenna except when operated at or near one frequency in its useful range, because over the useful range the electrical length of the radiator varies from approximately ¼ wavelength to one wavelength. However, for all practical purposes it may be considered in the same category as a half-wave antenna, because the pattern is comparable over the aforementioned frequency range.

THE "HALF-WAVE DOUBLET" WITH SYMMETRICAL FEED

The loop radiation resistance of a practical half-wave horizontal radiator differs considerably from the idealized values indicated in figure 3-4. The loop radiation resistance, which also becomes the feed point resistance when the antenna is center fed, runs between approximately 30 and 80 ohms in typical installations, the exact value depending upon the ground constants, the effective height of the radiator, the presence of surrounding objects, and the effective length/diameter ratio of the conductor or conductors.

If the radiator is cut to a resonant length, it is apparent that when it is fed at the exact center the feed point impedance will be purely resistive and have a magnitude in the range 30 to 80 ohms. If the radiator is fed with a line having a surge impedance of similar magnitude, the SWR will be low and the line will qualify as a "nonresonant" line.

Both "Twin-Lead" type polyethylene dielectric molded pair and "EO-1" type rubber dielectric twisted pair line are available in 70-75 ohm characteristic impedance, and are suitable for direct feed of a half-

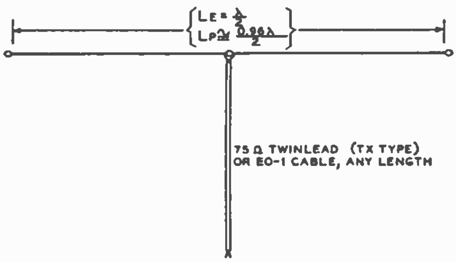


Figure 5-6.

HALF-WAVE DOUBLET WITH SYMMETRICAL FEED.

The feed line should leave the radiator at approximately a right angle for at least a wavelength or until it is close to the ground. The optimum physical length (L_P of the radiator depends upon many factors and is best determined experimentally, though in most cases satisfactory fixed-frequency operation will be obtained if the antenna is cut to 96 per cent of a free-space half wavelength. 300-ohm Twin-Lead can be substituted for the 75-ohm Twin-Lead by using a "Q" matching section (linear transformer) of 150-ohm Twin-Lead exactly an electrical quarter wavelength long. However, the 300-ohm Twin-Lead is more vulnerable to wet weather.

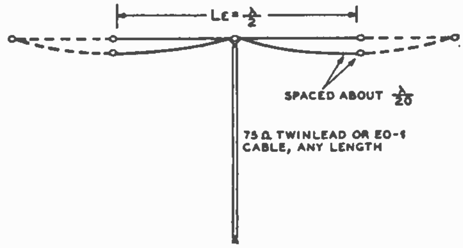


Figure 5-7.

BROAD-BAND HALF-WAVE DOUBLET WITH SYMMETRICAL FEED.

The useful range of a resonant doublet can be increased considerably by lowering the Q of the radiator. This is accomplished here by effectively decreasing the length-diameter ratio. The use of two fanned conductors produces essentially the same effect as does a large increase in the diameter of a single conductor radiator. In some cases a further improvement may be realized by slightly staggering the upper and lower element lengths.

wave doublet without resort to any impedance transforming device at the antenna feed point. Such an antenna is illustrated in figure 5-6.

A characteristic impedance of about 50 ohms would, on the average, provide a closer impedance match, but as of this writing such line is not available in symmetrical configuration. If the antenna is close to highly conducting ground or surrounding objects and the 70-ohm cable overheats or the reactive effects due to the mismatch are found objectionable, the situation can be improved by the use of two identical lengths of 70-75 ohm line connected in parallel (separated at least several times the cable diameter), thus reducing the characteristic impedance to half that of a single length of line. If the two lines are cabled or twisted together, the impedance will be lowered to somewhat less than half.

This type of antenna can be made inherently unresponsive to the second harmonic by making the line an integral number of electrical quarter wavelengths long at the fundamental, and grounding one side of the line (or the center of the coupling coil) at the transmitter. However, this antenna system is an excellent third (or other odd) harmonic radiator regardless of the length of the line. Therefore, when this antenna is used under circumstances requiring that odd harmonic radiation be kept to an absolute minimum, some means of attenuating the harmonics must be provided.

This antenna will work only over a rather limited frequency range, the width of which is determined by the effective Q of the radiator and how well the line impedance matches the feed point resistance at the mid-frequency. Because the physical length required for resonance depends upon so many unpredictable factors, the optimum length of the radiator is best determined experimentally, starting with a physical length equal to 98 per cent of a free space half wavelength and then checking the resonant frequency of the radiator. If it should be 2 per cent lower than the desired band center or mid-frequency, then the ends should be pruned equally to shorten the overall length by approximately 2 per cent.

The range over which the antenna may

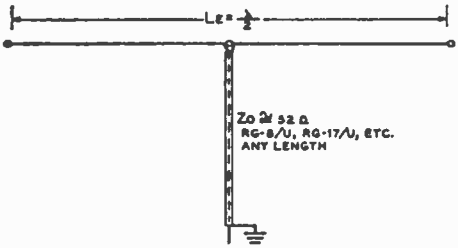


Figure 5-8.

HALF-WAVE DOUBLET WITH COAXIAL FEED.

This antena is similar to that of figure 5-6 except for the use of coaxial line in place of molded or twisted pair line, and the same remarks apply. If the coaxial line leaves the antenna at a right angle for at least a wavelength or until it is close to ground, the antenna effect will not be great in spite of the unbalance at the junction between antenna and feed line, provided that the radiator is operated close to resonance. Antenna effect can be minimized when necessary by cutting the line to a length which detunes the overall system with respect to antenna effect, as described in the text. 52-ohm cable is specified because it comes closer than 75-ohm cable to the mean value of feed point impedances encountered in actual practice.

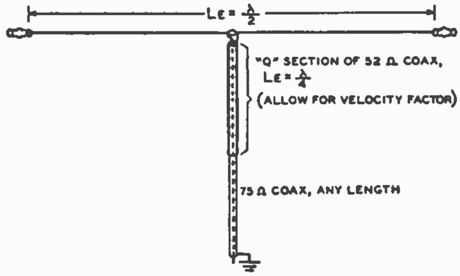


Figure 5-9.

LOW HALF-WAVE DOUBLET WITH COAXIAL FEED, UTILIZING QUARTER-WAVE "Q" SECTION.

When a half-wave horizontal antenna is located close to ground having good conductivity or to a ground screen, the feed point resistance will be considerably less than 52 ohms. In cases where it is less than 30 ohms, the above arrangement will result in a lower SWR than the arrangement of figure 5-8.

be employed without the SWR becoming objectionable and causing excessive line loss or undesirable reactive effects may be increased by the expedient shown in figure 5-7, whereby the Q of the radiator is lowered by decreasing the effective length-to-diameter ratio.

In some cases a further slight improvement can be realized by staggering the element lengths so that the lower wires are slightly longer or shorter than the upper wires, though the possible improvement is dependent upon the maximum tolerable SWR and other factors; in some cases it is negligible and in no case is it great.

HALF-WAVE DOUBLET WITH COAXIAL FEED

Solid dielectric coaxial cable may be substituted for the EO-1 cable or "Twin-Lead" line shown in figures 5-6 and 5-7. Because of the unbalance* at the antenna feed point connection, some antenna effect will be present even if the line is brought away at right angles and the antenna is perfectly symmetrical with respect to ground and surrounding objects. However, *at or close to the resonant frequency of the antenna,* the antenna effect due to the termination unbalance is small compared to the antenna effect that can result from dissymmetry of the antenna with respect to surrounding objects. Therefore, in most practical installations the small amount of antenna effect due to termination unbalance may be ignored. As simple dipoles are not employed for direction-finding installations and other installations requiring a high degree of pattern discrimination, the only exceptions are installations where the lines to several transmitting and receiving anatennas are sufficiently juxtaposed as to cause a "cross talk" problem when antenna effect is present.

In most cases it will be found that 52-ohm cable (such as RG-8/U, etc.) will

*The term "unbalance" refers not to physical symmetry or lack of symmetry, but to an abnormal condition caused by connecting a symmetrical line to an asymmetrical load, or an asymmetrical line to a symmetrical load. Refer to chapter 2 for further explanation.

provide a better match and lower SWR than will 75-ohm cable (such as RG-11/U, etc.). The nominal feed point impedance of 73 ohms commonly used in connection with half-wave antennas applies only to a hypothetical, infinitely thin half-wave radiator in free space.

The antenna effect due to termination unbalance can be minimized by making the feed line an odd number of quarter wavelengths long and grounding the outer conductor at the transmitter. The velocity factor of the line should be ignored in determining the electrical length with respect to antenna effect, because currents flowing on the outside of the outer conductor are not "slowed up" by the dielectric between the conductors and are affected only by an insignificant amount by the plastic outer sheath.

When the proximity of the earth or surrounding objects is such as to bring the feed point impedance down to 35 ohms or less, an improvement in the SWR can be obtained by the use of a 75-ohm line in connection with a 52-ohm quarter-wave matching section or "Q section". Such an arrangement is shown in figure 5-9. If polyethylene dielectric cable is used for the Q section, its length should be made 66 per cent of a free space quarter wavelength. The use of a Q section is recommended in the case of a horizontal antenna suspended one eighth wavelength or less above highly conducting earth or a ground screen.

The same considerations apply with regard to radiator length and harmonic suppression in the case of a half-wave doublet with coaxial feed as were covered under the half-wave doublet with molded or twisted pair feed.

THE FOLDED DIPOLE

If a half-wave radiator is constructed of two, identical, close-spaced parallel conductors which are shorted at the ends, and the combination is excited ("voltage fed") at one end just as though it were a single wire, it is obvious that equal currents will flow in the two conductors. The combination will be equivalent, as regards Q and loop radiation resistance, to a single conductor having a somewhat greater diameter.

With the two conductors connected together at the ends, they are in effect unity coupled in such phase that, even when fed as shown in figure 5-10A, equal currents may be considered to flow in the two identical conductors in the same direction.* When the combination is fed in this manner, the current at the feed point for a given power is only half that which would flow in a single wire alone, which means that the feed impedance is four times as high. Strictly speaking, the feed point impedance is *slightly less* than four times that presented by one of the conductors without the other present, because the effective length/diameter ratio (and therefore the radiation resistance) is lowered appreciably when the second conductor is added.

The *mean* value of the feed point impedance obtained with such an arrangement in actual practice is about 240 ohms for effective heights greater than 1/8 wavelength (assuming horizontal orientation of the radiator). This means that 300-ohm Twin-Lead may be used to feed such an antenna directly, as a low SWR will be obtained without benefit of an impedance transforming device. When the effective height of the radiator is less than ⅛ wavelength and the soil conductivity is high, as is the case with some 3 to 5 Mc. antennas intended for short range sky-wave communication, a lower SWR will be obtained by substituting 150-ohm Twin-Lead.

The feed point impedance can be raised to a value suitable for direct connection of a 450 to 600 ohm open wire line by resorting to either of two expedients. If three

*Actually the current in the driven conductor will be slightly the greater, assuming that the two conductors are of the same diameter, because the energy delivered to the unbroken wire must travel via the driven wire, causing a small "transmission line current" to flow in opposite directions in the two radiators. This component adds to the antenna current in the driven radiator and bucks it in the other, because the antenna current flows in the same direction in both wires. With length-to-diameter ratios encountered in practical radiators, however, the magnitude of the transmission line current component flowing in the radiators is insignificant by comparison with the magnitude of the antenna current, and for practical purposes may be ignored.

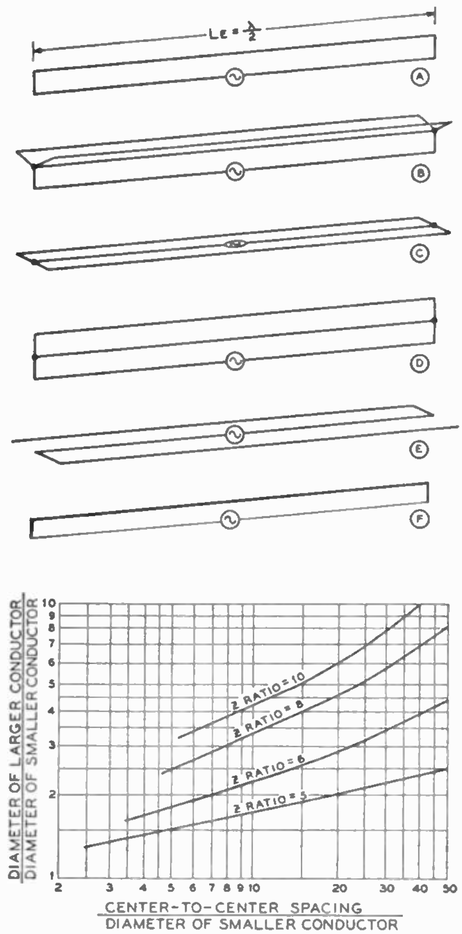


Figure 5-10.

FOLDED DIPOLE (A) AND OTHER FORMS OF THE MULTI-WIRE DOUBLET.

The arrangement at A raises the feed point resistance to approximately 4 times that of one conductor alone, that of B to approximately 9 times. The arrangements of C, D and E are approximately equivalent to that of B, but not exactly because the lack of axial symmetry causes the current to divide unevenly in the three conductors. In practical installations the feed point would be attached to the r-f generator via a transmission line. The impedance transformation obtained with the F arrangement depends upon the ratio of the conductor sizes and upon their relative spacing, and is given in the above chart. Intermediate values can be interpolated. This method is not recommended for impedance ratios higher than 10.

identical, equally-spaced wires are used instead of two (figure 5-10B), the effective impedance transformation is roughly 9 as compared to a single wire alone, instead of approximately 4. This raises the average feed point impedance in practical installations in which the effective height of the horizontal radiator exceeds ⅛ wavelength to about 550 ohms. If the three wires are arranged in the same plane (figure 5-10C, 5-10D) instead of being equally spaced about a common axis, the current in the center wire will tend to be less than that in each of the outside wires, but the effect upon the impedance transformation is not serious from a practical standpoint. When three conductors rather than two are employed, the combination sometimes is referred to as a "multi-wire doublet." The performance of the arrangement shown at E is substantialy the same as that of C, though the latter is more popular.

If in the two-wire arrangement the diameter of the driven conductor is made smaller than that of the unbroken conductor, as at F, the current will divide in such a manner that it is greater in the larger conductor, assuming that both conductors are of the same material. The relationship is not a simple function of the ratio of the diameters or cross-sectional areas of the two conductors, but rather follows a complicated law which involves frequency, the relative wire size, the absolute wire size, the conductor material, the spacing between the wires, and various other factors.

An empirical formula which gives an approximation close enough for most practical applications, provided that the center to center spacing is at least three times the diameter of the larger conductor, is as follows:

$$\text{Feed point resistance} = R_0 \left[1 + \frac{Z_0'}{Z_0} \right]^2$$

where R_0=Radiation resistance of the combination.

Z_0'=Surge impedance of a line having the same center to center spacing as the two radiating elements but with both conductors having the diameter of the smaller conductor.

Z_0=Same as Z_0' except having the diameter of the larger conductor.

Beneath the figure 5-10F representation of a folded dipole with unequal conductor diameters is a chart, based on the above formula, which shows at a glance what diameter ratios and spacing ratios will produce various impedance transformation ratios. Other transformation ratios between 5 and 10 can be interpolated with sufficient accuracy for most practical applications.

This method of increasing the impedance transformation ratio is not recommended for ratios greater than 10, because the broad band features become degraded for extreme ratios. For ratios greater than 10, it is recommended that the desired transformation be obtained by increasing the number of conductors.

The spacing between the elements of a folded dipole or multi-wire doublet is not critical; but the relationship between resonant frequency and element length (ignoring the end jumpers) will be affected somewhat by the spacing. At the higher frequencies 300-ohm Twin-Lead commonly is used as both antenna and feed line in a folded dipole arrangement, the section used as the radiating portion being jumpered at the two ends and again 0.82 of a free space quarter wavelength each side of center. The series and shunt resonant frequencies will differ considerably (due to the polyethylene dielectric) unless these additional shunts are incorporated. One conductor is broken in the center for attachment of the feed line.

However, except for portable or emergency applications, the use of Twin-Lead for the radiating portion of the folded dipole would seem unjustified in the h-f range. The arrangements of figure 5-11 are stronger, cheaper, less inclined to stretch, and more permanent. For reasons of appearance or convenience, Twin-Lead can be used instead of open wire construction for the transmission line, provided its wet weather performance is acceptable. The latter can be improved by "waterproofing" the line with Simoniz Wax or Amphenol 307 Silicone compound.

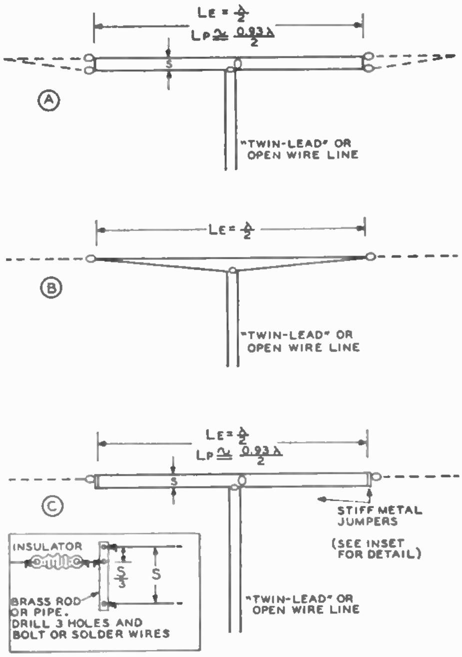


Figure 5-11.

PRACTICAL METHODS OF RIGGING A FOLDED DIPOLE.

Because of the broad resonance peak, the arrangement at A may be cut to a physical length (L_P) equal to 0.93 of a free-space half wavelength and a spacing (S) of approximately 1 per cent of the length with the assurance that it will perform satisfactorily at a given frequency except under unusual circumstances. The arrangement at B is easier to construct but is more frequency sensitive, as the taper increases the Q. Also, the electrical length is dependent upon so many factors that it is difficult to predict accurately, and must be checked experimentally to assure proper operation. The arrangement at C is recommended when the radiator length is short enough to make it practical, as it automatically keeps both wires taut even though differing slightly in length. In typical well-elevated installations a very low VSWR will be obtained with 300-ohm Twin-Lead, and a VSWR of approximately 2.0 will be obtained with a directly connected open wire line. A VSWR comparable to that obtained with 300 ohm Twin-Lead can be obtained by using an open wire line consisting of no. 14 B&S spaced 6 inches in conjunction with a quarter wave "Q" matching transformer consisting of no. 10 B&S spaced 2 inches.

The folded dipole is much less frequency sensitive than a conventional dipole which is fed with either a low impedance line, or an open line in conjunction with a Q matching section. For this reason a "square end" folded dipole (figure 5-11A) with an end spacing equal to about 1 per cent of the length can be cut to an overall length of 0.93 of a free-space half wavelength with the assurance that, except in unusual circumstances with regard to surrounding objects, the system will function satisfactorily at a given frequency within the h-f range. However, if the antenna is expected to work over more than a very narrow range of frequencies, the natural resonant frequency of the radiator should be checked experimentally and the length altered if necessary in order to center the resonant frequency at the middle of the range to be covered. If sufficient sag occurs at the feed point to result in a pronounced taper in the spacing, the current distribution along the conductors will be significantly affected and the length will have to be increased in order to resonate at a given frequency. At A, sag is prevented by the spacing insulator at the center of the elements.

The B arrangement simplifies construction but the inherent sag makes the exact resonant frequency quite unpredictable (unless the installation duplicates one whose electrical length is known). It also causes the system to be slightly more frequency sensitive.

For practical purposes the radiation pattern of the folded dipole may be considered as being identical to that of a conventional half-wave radiator so long as the spacing "S" between the conductors is an insignificant fraction of a wavelength.

THE "T-MATCHED" HALF-WAVE ANTENNA

It will be observed that the two conductors of an experimental folded dipole can be twisted together for some distance from the ends without appreciably affecting the impedance transformation; the two parallel conductors of the split section will still be effectively unity coupled and will still carry substantially equal currents if the two conductors are of the same diameter. However, the resonant frequency and Q of the radiating system will be affected.

When the reduction in the length of the split portion is carried beyond a certain point, the upper and lower elements of the split section no longer are effectively unity coupled, and appreciable difference in the magnitude of current flowing in the two parallel conductors will be noted. Thus, through variation of the length of the split section, control over the impedance transformation is possible.

This effect is exploited in the "T match," two variations of which are shown in figure 5-12. The two are electrically comparable, and the choice is a matter of mechanical considerations. The T match is not widely used for feeding a simple dipole antenna, and when it is, it is almost invariably a self-supporting h-f dipole of tubing rather

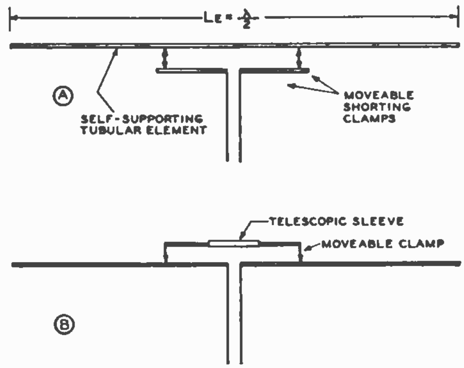


Figure 5-12.

"T-MATCH" METHOD OF TRANSFORMING FEED POINT IMPEDANCE

The "T-matched" dipole is related to the folded dipole. In the case of tubular, self-supporting dipole elements, the T-match permits a more desirable mechanical construction than does a split dipole, and also facilitates experimental adjustment of the impedance transformation. The conductor ratio and spacing are so chosen for the particular application that adjustment of the taps permits variation of the impedance transformation ratio over the range of impedances likely to be encountered. The two arrangements illustrated are electrically comparable; the choice is a matter of mechanical considerations.

than a wire conductor suspended between insulators. However, the T match is widely employed for matching the self-supporting driven element of h-f multi-element arrays which employ parasitic elements for directivity control, because for this application it permits a simpler mechanical structure than does the folded dipole, and also facilitates experimental adjustment of the impedance transformation.

A T-matched dipole is more frequency sensitive than a folded dipole, and is somewhat longer physically for a given resonant frequency (assuming comparable conductor diameters). The impedance transformation obtained with a T-match depends principally upon the spread and spacing of the "T" section, and upon the ratio of the T-section conductor diameter to that of the main element. By controlling these three parameters, a wide range of impedance transformation is possible.

VERTICAL HALF-WAVE ANTENNAS

Though not as popular as they once were, vertical half-wave antennas still are used in some cases for long-distance general coverage work in the h-f range. They show up well by comparison with a horizontal half-wave antenna when the maximum pole height is limited to about a half wavelength, and will provide good signal strength at great distances when the transmitter power is high. But the strong ground wave which is characteristic of such antennas often produces objectional interference to nearby services. For short and medium distance sky wave work the horizontal radiator is to be preferred in every case.

Any of the half-wave antennas shown with horizontal orientation may be used vertically, though some antenna effect is to be expected with all center-fed arrangements, as it is inherent in the asymmetry of the two halves of a vertical antenna with respect to ground. It usually can be kept to a tolerable value by bringing the feed line away from the radiator at a right angle for at least a half wavelength. The end-fed arrangement of figure 5-4B is especially well adapted to feeding a vertical radiator. The antenna tank may be placed in a small "dog house" at the base of the antenna pole (or radiator if a self-supporting radiator is used), and the bottom of the tank grounded to three or four buried radials.

The radiation resistance of a vertical radiator with its bottom end close to the ground is somewhat higher than that of a horizontal antenna close to the ground, and some allowance will usually have to be made in the case of the center-fed antennas illustrated, if a negligible SWR is desired.

DIRECTIONAL H-F ANTENNA SYSTEMS

Space limitations make it impossible to treat more than a small fraction of the various directional antenna systems currently or previously in practical use in the h-f range. Therefore the discussion will be limited to a representative group containing one or more of the most popular, currently employed types in each of the main general categories. In a few cases new modifications or adaptations are suggested, which, in the author's opinion, offer advantages over the conventional version in use at the time of writing.

OPTIMUM ELEVATION OF A DIRECTIONAL ARRAY

The optimum elevation of a high-frequency phased dipole array, assuming that maximum effectiveness at distances exceeding approximately 2500 miles is desired, depends upon numerous factors which may be simplified and generalized briefly as follows:

When an array is located over "scattering terrain" (such as a dense residential or business district) which extends for considerable distance along the azimuthal beam direction, little difference in the 2500-mile field strength or response will be noted for increased array elevation once the array is elevated sufficiently to put the entire array well "in the clear."

When the surrounding terrain is reasonably flat and unobstructed for considerable distance, particularly along the azimuthal direction of the "beam," the recommended minimum effective average elevation of a high-frequency array is approximately 60 feet *for all frequencies between 7 and 30 Mc.* when the array has at least a moderate amount of "free space" vertical directivity. While the condition of the ionosphere and the vertical pattern of the antenna at the other end of the circuit enter the picture, in general the 2500-mile mean field strength or response will fall off noticeably as the effective array elevation is reduced below about 60 feet, and increase only slightly if at all for greater elevations. From an economical standpoint, this calls for an elevation of from 55 to 70 feet.

To determine the effective average elevation of a phased array, use the following emperical rule. Assume a "line of shoot" from the center of the array which makes a depression angle equal in degrees to approximately half the operating wavelength in meters, and determine where this line strikes the ground. The effective elevation of the array is the difference in elevation between this point and the center of the array. Thus, if the array is located in a depression or on a hill, the effective elevation is affected correspondingly.

The optimum elevation of a long wire array such as a "V beam" involves other factors; so the matter of optimum elevation of long wire arrays will be covered concurrently with a discussion of such arrays.

PHASED, RESONANT DIPOLE ARRAYS WITH ALL ELEMENTS DRIVEN

High frequency arrays in this category consist of two or more half-wave dipoles, almost invariably with horizontal orientation, disposed either in broadside relationship, end-fire relationship, or a combination

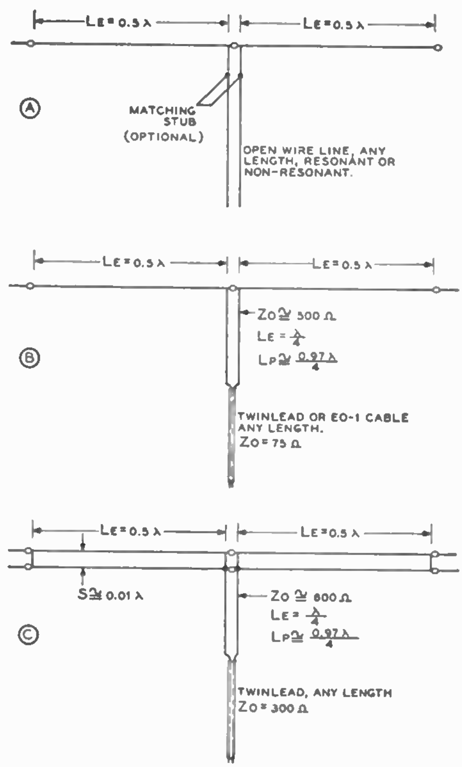


Figure 5-13.

THREE GOOD METHODS OF FEEDING TWO HALF-WAVE COLINEAR ELEMENTS IN PHASE WHEN THEY ARE CONTIGUOUS.

In the arrangement shown at A, the optimum matching stub length and position must be determined experimentally. The stub need not be used unless a flat line is desired. (Compare figure 5-5.) The arrangement shown at B usually will produce a VSWR of less than 2.0 with the constants shown, but for an absolutely flat line the surge impedance of the quarter-wave Q section must be optimized for the particular antenna height, etc. employed. The arrangement at C is considerably less frequency sensitive, due to the lower effective L/D ratio of the elements and the lower impedance transformation ratio. The practical signal gain of all three arrangements is approximately 2 db over a single element at the same height, and is substantially the same at all vertical angles.

of both. In broadside they may be either colinear, stacked one above the other (in a common vertical plane), or arranged in combination. In a broadside arrangement the array always is bidirectional unless a reflector sheet, screen, or phased "curtain" is employed, the latter arrangement qualifying the antenna as a "combination" array. In an end-fire arrangement the elements may be driven in a phase relationship which results in a bidirectional pattern, or they may be driven in a quadrature relationship which gives a unidirectional pattern.

BROADSIDE ARRAYS WITH COLINEAR ELEMENTS

The simplest form of colinear array is "two half waves in phase". The center-fed antenna with resonant feed line, described in a previous section and illustrated in figure 5-5, operates as two colinear half-wave elements in phase when the frequency is such that each leg is an electrical half wavelength long.

The free space directivity gain is not great, being approximately 2 db over a half-wave dipole. Were it not for the appreciable positive resistive component of mutual impedance between the two elements, the gain would be 3 db. While this antenna usually is classed as a directional array, the horizontal directivity does not greatly exceed that of a half-wave dipole, and is pronounced only at low vertical angles. Various recommended methods of feeding two contiguous half waves in phase are illustrated in figure 5-13.

The gain may be increased slightly by increasing the spacing between the elements so that the resistive component of the mutual impedance is lower in magnitude or is actually negative. The arrangement of figure 5-14A gives a free space directivity gain of approximately 3 db, and that of 5-14B approximately 3.3 db. However, when the spacing between the current loops exceeds one half wavelength, minor lobes appear in the directivity pattern. The figure 8 pattern of the major lobe obtained with the figure 5-14B arrangement can be changed to a cloverleaf pattern with slightly less maximum gain by reversing the phase of one feed line. The mutual impedance between elements is low enough that there will be only a moderate change in transmitter loading and line SWR when the polarity of one feeder is reversed. The anti-phase arrangement belongs to the "end fire" family, though in a sense it really is a hybrid.

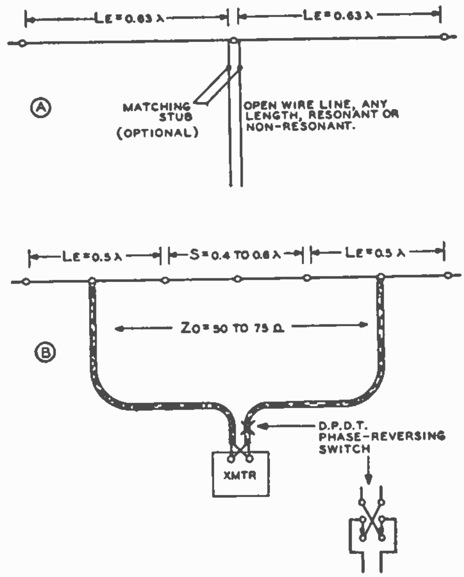


Figure 5-14.

TWO COLINEAR HALF WAVES IN PHASE, EXTENDED (A) AND SPACED (B) TO GIVE GREATER GAIN.

The arrangement at A has a practical signal gain of approximately 3 db and that of B approximately 3.3 db over a half-wave dipole. The arrangement shown at A sometimes is referred to as an "extended double zepp." By means of the phase reversing switch, the arrangement at B can be used with the two half waves either in phase or anti-phase, giving the option of either a figure 8 pattern or a four-lobe cloverleaf pattern. The cloverleaf pattern gives slightly less gain. Both arrays exhibit minor lobes of moderate amplitude.

Because all of the free space directivity provided by colinear disposition of co-phased half-wave elements is in planes containing the radiators, the practical signal gain is substantially the same at all vertical angles and is substantially the same as the free space directivity gain. In other words, colinear disposition of horizontal half-wave

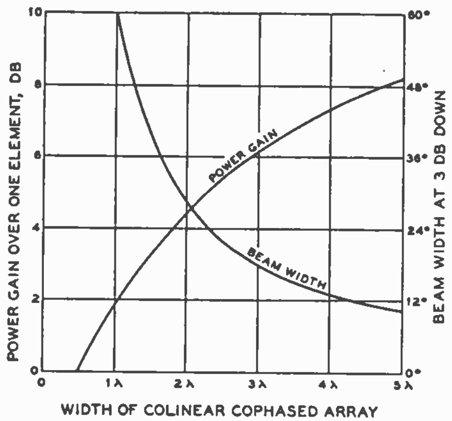


Figure 5-15.

POWER GAIN AND BEAM WIDTH AS A FUNCTION OF OVERALL WIDTH OF A COLINEAR COPHASED ARRAY OF HALF-WAVE DIPOLES WITH UNIFORM LOOP CURRENT AND AN ELEMENT SPACING OF ONE HALF WAVELENGTH BETWEEN CENTERS.

The curves are based on free space gain and directivity, but also apply approximately to the practical signal gain and horizontal directivity obtained with horizontal orientation of elements over typical ground. The curves also apply approximately to any horizontal colinear array having a center spacing exceeding one half wavelength but less than one wavelength, provided the overall array width is two wavelengths or more. The beam width is arbitrarily taken as the angle between half power points.

elements which are excited in phase provides horizontal directivity, and the directivity gain is substantially independent of antenna elevation. The beam width at 3 db down is about 60 degrees for two juxtaposed horizontal elements, and slightly less when the spacing is increased to give greater gain.

Colinear Arrays With 3 Elements or More

It is possible to obtain almost any desired degree of horizontal directivity by using a sufficient number of colinear half-wave elements in phase. When the overall width of the array exceeds approximately 2 wavelengths, the gain is determined primarily by the overall width of the array and is substantially independent of the number and spacing of elements *provided* that the center-to-center spacing of the elements does not exceed approximately one wavelength.

Generally speaking, the power gain is proportional to the width of the array, and doubling the overall width of the array raises the gain by approximately 3 db once the width exceeds one wavelength.

A substantial reduction in the maximum amplitude of the minor lobes can be realized at a slight sacrifice in gain and directivity, or they may be eliminated altogether at a considerable sacrifice in gain and directivity, by utilizing a "tapered" current distribution. The excitation to the various elements is so porportioned that the current is greatest in the center element or elements and least in the outer elements, the distribution being determined by the particular requirements with regard to side lobe amplitude. However, this greatly complicates the feed system and therefore is employed only in applications which require negligible amplitude of the minor lobes, and not where maximum gain is the prime consideration.

With uniform loop current distribution, the number of minor lobes and the relative magnitude of the strongest minor lobe with respect to the main lobe increase with the number of elements in the array and with the center spacing between elements. For the common arrangement in which the current is uniform and the center spacing is one half wavelength, the relative amplitude of the strongest minor lobe increases as the number of elements is increased until it finally levels off at about 12 or 13 db below the main lobe.

For a colinear array with half-wave center spacing and uniform current distribution, the power gain and the beam width at the half power points are shown in figure 5-15 as a function of overall width of the array. The curves also apply *approximately* to colinear arrays having a center spacing exceeding a half wavelength, provided that it does not exceed one wavelength and provided the overall width of the array exceeds approximately 2 wavelengths.

Feeding a Colinear "String"

There are various methods of feeding a colinear "string" of cophased half-wave dipoles, each offering certain advantages and certain disadvantages when the electrical, mechanical, and economical aspects are considered. To facilitate the discussion, three arrangements are shown in figure 5-16 which illustrate the various points to be discussed. While five elements are shown in each array, the same considerations apply in general to any number.

For minimum antenna effect and minimum frequency sensitivity, each element should be current fed with its own individual nonresonant feed line having a surge impedance comparable to the feed point impedance of the element, as illustrated at A in figure 5-16. When an unusually wide band of frequencies is to be covered, the response of the array can be increased still further by making each of these elements a folded dipole.

The arrangement at A is the only one of the three shown which is applicable to the practical utilization of tapered current distribution. When side lobe level is of minor importance and maximum gain is the prime consideration, all feeders may simply be connected in parallel at the transmitter (observing correct polarity) to obtain substantially uniform current distribution. Because the radiation resistance of an end element is not the same as that of an interior element, there will be a slight irregularity in the current distribution, but not enough to affect the pattern greatly.

To make this effect insignificant or even beneficial, instead of detrimental, the center spacing can be increased to between 0.8 and 1.0 wavelength. This not only minimizes the number of elements required for a given gain, but causes the relationship between feed point impedances to be such that the current tends to be either very nearly uniform or slightly greater in the center elements when all feed lines are connected in parallel at the transmitter. This condition is realized because with increased element spacing the resistive component of the mutual impedances becomes zero or negative, rather than positive as in the case where the center spacing is 0.5 wavelength. It should be noted that the arrangement shown at A is the only one of the three which is practicable when the current loop spacing exceeds 0.75 wavelength, because this represents the maximum amount by which elements may be "extended" (see figure 5-14A) without detrimental effects.

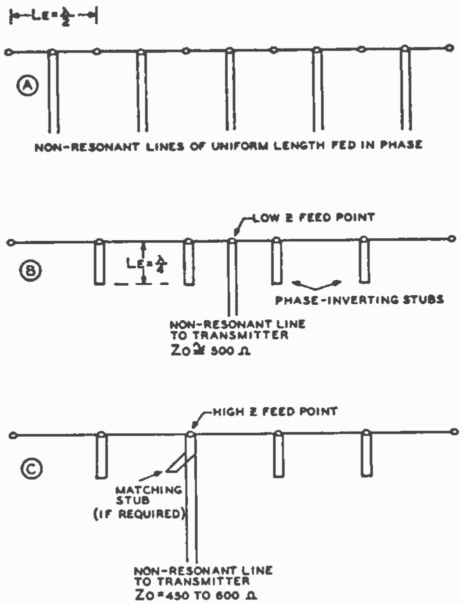


Figure 5-16.

THREE METHODS OF FEEDING A COLINEAR "STRING" OF HALF-WAVE ELEMENTS HAVING HALF WAVE-LENGTH CENTER SPACING.

The relative merits of the above three arrangements, as well as the considerations involved, are discussed in the accompanying text. As articulated elements are added, the feed point impedance at a current loop goes up while that at a voltage loop goes down, until after several elements are connected in series the two are of the same order of magnitude. The impedance at a voltage loop varies more widely with conductor diameter than does the impedance at a current loop.

While possessing many electrical advantages, the arrangement of figure 5-16A is not widely employed in the h-f range for mechanical reasons. It is popular for feeding a colinear string of v-h-f or u-h-f dipoles, however, in which case the elements usually are rigid and each is supported at its center by its own feed line.

For most h-f applications the arrange-

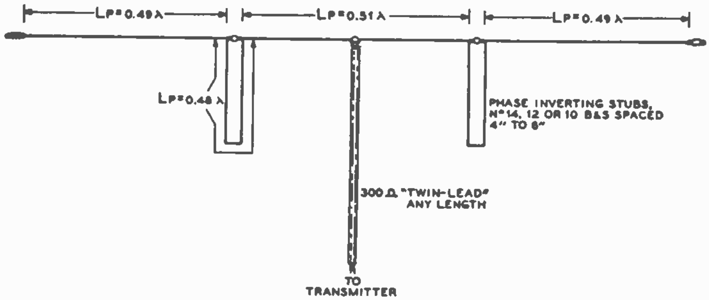


Figure 5-17.

SIMPLE "PRECUT" COLINEAR ARRAY.

This array provides a practical signal gain of about 3 db over a half-wave horizontal dipole at the same height, at all vertical angles, and regardless of the height above ground. The array may be precut to the specified dimensions, if the installation is typical, with the assurance that it will work satisfactorily over a narrow range on either side of the calculated frequency without need for experimental adjustment. The beam width at half power points is approximately 40 degrees. Physical dimensions are specified in terms of free space wavelength. To convert to feet, refer to figure 3-2.

ment at B will be found satisfactory, though its electrical performance is inferior to that of A. With the articulated feed method shown, the currents in the two conductors of each phase inverting stub tend to become unbalanced when the frequency deviates from optimum by even a very slight amount, causing considerable radiation from the stubs unless they are coiled or folded back on themselves in such a manner as to cancel the radiation caused by the in-phase current component. Even at the optimum frequency of the array it is impossible to eliminate completely the in-phase current flowing in the stubs, though for many purposes the stub radiation will be of an acceptible magnitude even though the stubs are not folded or coiled.

It will be noted that at B, as in the arrangement at A, a nonresonant line is employed which matches the feed point impedance without resorting to stubs or other devices, as the latter further increase the frequency discrimination of the array.

The arrangemnt at C is shown only for comparison, and is not recommended. It is inferior to that at B because the inherent dissymmetry produces several undesirable effects. For instance, the asymmetrical feed tends to make the system more frequency sensitive, as well as increasing the radiation from the phase inverting stubs. Also, when a colinear array is not fed in the exact center, a sufficient shift in frequency away from optimum will cause the broadside lobe to shift direction so that it no longer maximizes exactly broadside to the line of the array. However, this effect becomes really pronounced only in the case of arrays which comprise many elements and are fed only via one of the end elements.

It is recommended that, in general, not more than three or at most four colinear elements be fed from one line or branch line unless the array is to be used only over a very narrow range of frequencies. A simple h-f colinear array which has a sufficiently broad response that it may be precut, erected, and used over a narrow frequency range without experimental adjustment of the dimensions is shown in figure 5-17. The specified diminsions are in terms of free space wavelength. (Refer to figure 3-2.) While the feed point resistance depends somewhat upon the height of the elements above ground and other factors, the use of 300-ohm Twin-Lead will result in a low SWR in typical installations in which the effective elevation exceeds a quarter wavelength.

BROADSIDE ARRAYS WITH STACKED ELEMENTS

In the case of a colinear array both the aspect directivity and interference directivity are acting in the same plane. The result is that the directivity of a colinear array is not greatly increased over that of a dipole until enough elements are added to make the interference directivity considerably greater that the aspect directivity. This takes at least three elements.

If each radiating element of a colinear array in free space is rotated 90 degrees so that all are parallel and not staggered, the interference directivity and the aspect directivity no longer are working together in the same plane. In this case the interference directivity acts in a plane in which the antenna otherwise is non-directional, and therefore the interference directivity provides a slightly greater increase in directivity gain over a dipole than is the case for the colinear arrangement. The difference is most marked when only two elements are involved.

Another way of explaining the increased gain provided by the side-by-side (parallel) arrangement as compared to the end-to-end (colinear) arrangement is on the basis of mutual impedances. When the elements are side by side and a half wavelength apart (parallel) the resistive component is negative, while end-to-end disposition (colinear) with a center spacing of one half wavelength results in a positive component of mutual impedance. Thus, the side-by-side (parallel) arrangement requires less total power for a given current in a given number of radiating elements, and therefore provides greater power gain. The difference in free space directivity power gain runs between 2 and 3 db for a given number of elements with half-wave center spacing.

In practical h-f installations the elements of a broadside array with parallel coplaner disposition of the elements usually are arranged with horizontal orientation. The practical directivity of such a "stacked" array differs from that of a colinear array in that both vertical and horizontal directivity are pronounced, and the practical signal gain over a horizontal dipole at the same average elevation varies appreciably with both the antenna elevation and the vertical angle (wave angle) under consideration.

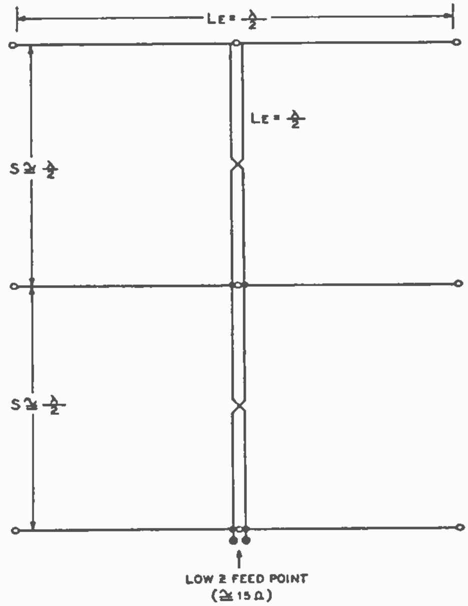


Figure 5-18.

ILLUSTRATING A "STACK" OF DIPOLES AS A FORM OF BROADSIDE COPLANAR ARRAY.

A single stack ordinarily is not employed by itself as an h-f array; the more common arrangement is two or more stacks placed side by side to form a broadside "curtain" or "panel." One stack of such an arrangement is called a "bay." The free space directivity power gain of the three-element stack illustrated is about 5.6 db over one element. The practical dx signal gain over typical soil as compared to a half-wave dipole depends upon average height and the vertical angle under consideration, but usually is comparable to the space directivity gain.

It is also important to note that for a given pole height the average elevation of the antenna is lower when a stack is used than when a colinear arrangement is employed. For low vertical angles of wave departure and arrival, it will be found that with low pole heights the loss in directivity gain due to the decreased effective height when a stack is used tends to offset the increased free-space power gain of a stack as compared to a colinear arrangement. Thus, at the low-wave angles most effective for long distance communication, the broadside stack shows up best when the pole height is great enough that the decrease in effective elevation due to stacking is small when considered percentage wise.

A simple stack of dipoles finds little practical application in the h-f range, and therefore will not be treated as a practical antenna. However, as will be seen shortly, two or more stacks often are employed side by side to constitute a broadside "curtain" or coplanar array having considerable directivity in both horizontal and vertical planes. Therefore the basic properties of a single stack are pertinent and will be discussed. A three-element stack is illustrated in figure 5-18 in case the reader has difficulty in visualizing the configuration under discussion.

As in the case of colinear elements, doubling the number of elements (for a given spacing) very nearly doubles the free space directivity power gain, once the number of elements exceeds three or four. However, as the gain of a two-element stack with half-wavelength spacing is about 4 db, as compared to less than 2 db for two juxtaposed colinear elements, the gain is roughly 2.5 db greater for any given number of elements when the elements are stacked. A free space directivity gain curve for stacked half-wave elements with uniform current and half-wave spacing is given in figure 5-19.

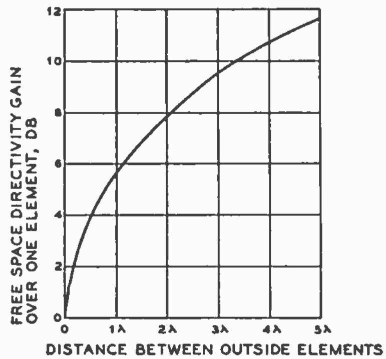


Figure 5-19.

FREE SPACE DIRECTIVITY POWER GAIN OF A BROADSIDE "STACK" OF HALF-WAVE DIPOLES WITH UNIFORM LOOP CURRENT AND AN ELEMENT SPACING OF ONE HALF WAVELENGTH, AS A FUNCTION OF DISTANCE BETWEEN THE TWO OUTER ELEMENTS.

The curve applies specifically to half-wave spacing, but also applies approximately to any spacing less than 0.75 wavelength. Half wavelength spacing is most commonly used because it facilitates feeding the elements in the proper phase.

The element spacing can be increased to 0.75 wavelength without a significant change in power gain for a given overall stack height or "spread," though the magnitude of the minor lobes will increase, particularly if the spread is less than three or four wavelenths. However, half wavelength spacing is the most common, because it simplifies feeding and phasing problems.

The remarks made in preceding sections in connection with colinear arrays with regard to the number and symmetry of feed points applies to a broadside stack as well. However, compromises sometimes must be made between electrical performance and mechanical practicability, and the asymmetrical feed shown in figure 5-18 is quite common.

The average radiation resistance of the elements in a broadside stack in free space is somewhat less than that of a single element when half-wave spacing is employed, and the radiation resistance of the outer elements tends to be slightly higher than that of the inner elements. The absolute value in practical h-f installations is dependent upon many factors, but usually will run in the neighborhood of 40 ohms for half-wave spacing.

From a practical standpoint, comparison between a broadside stack and a reference antenna (half-wave dipole) should be made on the basis of pole height rather than average antenna elevation when the effective power gain at very low vertical angles is involved. When figured on this basis the practical dx signal gain of a broadside stack usually is somewhat less than the free space power gain, though this is not

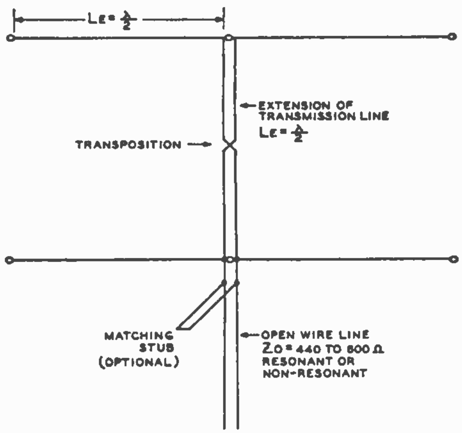


Figure 5-20.

THE "LAZY H" BROADSIDE ARRAY.

This popular bidirectional array, also known as "four half waves in phase," has a free-space directivity gain of approximately 6 db over one element. The optimum element and phasing section physical lengths are best determined by experiment, but satisfactory results usually will be obtained if all are cut to 0.98 of a free-space half wavelength (assuming thin, wire elements, and the minimum practical number of spacing insulators). The correct position of the matching stub, if used, should be determined by experiment. At low wave angles the beam width is approximately 60 degrees at the half power points.

always the case. As with any array having considerable vertical directivity, the practical dx signal gain over a single element (whether at the same average elevation or at the same maximum elevation) depends upon the elevation of the array, the optimum wave angle, and other factors.

The number of tiers of elements that can be employed advantageously in a high-frequency broadside stack is dependent upon frequency. At the higher frequencies (say, 20 to 30 Mc.), where a sharp, low-angle lobe can be tolerated, as many as five or six tiers can be used to advantage with half-wave spacing. At lower frequencies (say, 7 to 12 Mc.), however, such a high degree of vertical directivity ordinarily is undesirable. Three or four tiers are about as many as can be used to advantage in this frequency range in an array with fixed vertical directivity.

THE BROADSIDE CURTAIN

By combining colinear and stacked horizontal half-wave dipoles in the form of a coplaner "curtain," a broadside array having pronounced directivity in both the vertical and horizontal planes is realized. Such an array may be considered as an "array of arrays," either as a stack of colinear arrays, or as colinear stacks. Increasing the overall vertical spread of such an array increases the vertical directivity, while increasing the overall width increases the horizontal directivity. A close approximation of the gain can be had by adding the db gain due to stacking one bay to that obtained from one colinear tier. Because of complex diagonal mutual couplings between the elements in the combined array the gain figure thus obtained for the composite array is not exact, but is sufficiently accurate for most practical purposes.

THE "LAZY H" BROADSIDE ARRAY

The simplest form of broadside curtain comprises four horizontal elements stacked two colinear over two colinear. A simple and widely used method of feeding such a configuration of radiators is shown in figure 5-20. Dubbed the "Lazy H" by the author many years ago, it also is known as "four half waves in phase." The trouble with the latter designation is that, from a descriptive standpoint, it also could apply to a string of four colinear elements.

The Lazy H makes an excellent "dx" antenna for applications where moderate gain and directivity are desired, provided that the pole height is sufficient to elevate the lower elements enough to put them well "in the clear" and at least ½ wavelength above the effective earth level.

The free space directivity power gain of this array is approximately 6 db over one element. The practical dx signal gain over one element depends upon the antenna elevation and upon the wave angle under consideration, but usually it will be comparable to the free space directivity gain if the aforementioned stipulations regarding elevation are complied with.

300 ohm Twin-Lead may be substituted for the open wire line, provided the per-

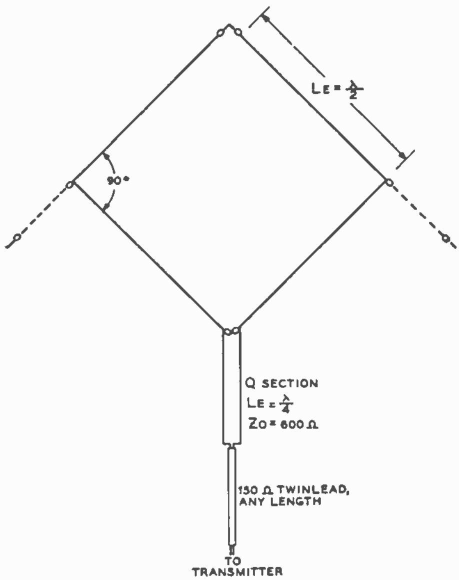


Figure 5-21.

THE "BI-SQUARE" BROADSIDE ARRAY.

This bidirectional array is related to the "Lazy H," and in spite of the oblique elements, is horizontally polarized. It has slightly less gain and directivity than the Lazy H, the free space directivity gain being approximately 4 db. Its chief advantage is the fact that only a single pole is required for support, and two such arrays may be supported from a single pole without interaction if the planes of the elements are at right angles. A 600-ohm line may be substituted for the Twin-Lead, and either operated as a resonant line, or made non-resonant by the incorporation of a matching stub. Usually, satisfactory performance will be obtained in the h-f range if the radiating elements are cut to 0.99 of a free-space half wavelength and the Q section to 0.98 of a free-space quarter wavelength. When tuned feeders are used, the array may be used on half frequency as a vertically polarized end-fire array.

formance of such line is acceptible, by using a quarter-wave "Q" section between the Twin-Lead and the antenna feed point. Usually a 600 ohm Q section will give a tolerable SWR. The feed system will resemble that of figure 5-21 except for the substitution of 300-ohm Twin-Lead for 150 ohm. No stub is used with this arrangement.

THE "BI-SQUARE" BROADSIDE ARRAY

Illustrated in figure 5-21 is a simple method of feeding a small broadside array first described by the author several years ago as a practical method of suspending an effective array from a single pole. As two arrays of this type can be supported at right angles from a single pole without interaction, it offers a solution to the problem of suspending two arrays in a restricted space with a minimum of erection work. The free space directivity gain is slightly less than that of a Lazy H, but is still worthwhile, being approximately 4 db over a half-wave horizontal dipole at the same average elevation.

When two Bi-Square arrays are suspended from a single pole, the Q sections should be well separated or else symmetrically arranged in the form of a square (the diagonal conductors forming one Q section) in order to minimize coupling between them. The same applies to the line if open construction is used instead of Twin-Lead, but if Twin-Lead is used the coupling can be made negligible simply by separating the two Twin-Lead lines by at least two inches and twisting one Twin-Lead so as to effect a transposition every foot or so.

When tuned feeders are employed, the Bi-Square array can be used on half frequency as an end-fire vertically-polarized array, giving a slight practical dx signal gain over a vertical half-wave dipole at the same height.

THE "SIX-SHOOTER" BROADSIDE ARRAY

As a good compromise between gain, directivity, compactness, mechanical simplicity, ease of adjustment, and band width the author recommends the array of figure 5-22 over the Lazy H for the 10 to 30 Mc. range when the additional array width and greater directivity are not objectionable. The free space directivity gain is approximately 7.5 db over one element, and the practical dx signal gain over one element at the same average elevation is of about the same magnitude when the array is sufficiently elevated. To show up to best advantage the array should be elevated suffi-

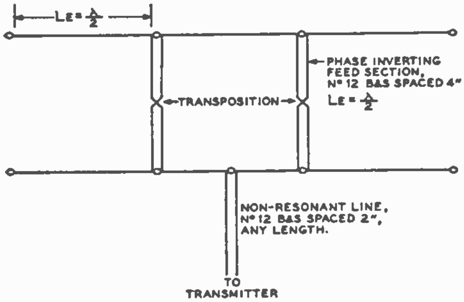


Figure 5-22.

THE "SIX-SHOOTER" BROADSIDE ARRAY.

This bidirectional curtain has a free space directivity gain of approximately 7.5 db, with a practical dx signal gain of the same order when it is well elevated and in the clear. A low SWR over a narrow band in the h-f range usually will be obtained if the two center elements are precut to 0.99 of a free space half wavelength and the outer elements and phasing sections are precut to 0.98 of a free space half wavelength, using no. 12 B&S gauge wire throughout. If the presence of surrounding objects causes an excessive SWR, the optimum lengths should be determined experimentally. The beam width at the half power points is approximately 40 degrees at low wave angles.

ciently to put the lower elements well in the clear and preferably at least 0.5 wavelength above the effective earth level.

THE "OCTAPUSH" 3-BAND AMATEUR ANTENNA

The antenna system illustrated in figure 5-23 is a "maximum utility" antenna designed for the amateur who is interested in the 7, 14, and 28 Mc. bands but is prevented by space limitations from erecting numerous fixed antennas and arrays each for a specific purpose. It is a highly effective, general coverage antenna for short and medium distance work on 7 Mc.; and as a bidirectional broadside curtain it provides good gain on 14 and 28 Mc.

On 14 and 28 Mc the SWR on the resonant line will be moderate, and the line losses therefore will be low. On 7 Mc. the SWR will be quite high, but the line loss will be tolerable for ordinary lengths of line. In view of the intended application on 7 Mc., a small amount of line loss on that band should not be considered objectionable.

The free-space directivity gain is slightly over 6 db on both 14 and 28 Mc. This may seem a bit low for an eight element broadside array on 28 Mc., and is explained by a wasted vertical lobe in the radiation pattern (due to one wavelength spacing between upper and lower elements), and by reduced current amplitude in the four outer elements and antenna effect on the vertical phasing sections (both due to unfavorable mutual couplings). However, it must be appreciated that compromises in performance almost invariably are part of the cost of increased utility.

To minimize the amplitude of the wasted vertical lobe on 28 Mc. and thus obtain the maximum possible practical dx signal gain, the *effective* elevation of the lower elements above ground should be an integral number of half wavelengths at 28 Mc. This can be accomplished by varying the elevation experimentally for maximum field strength, or by fixing the effective elevation by means of a ground screen. When either expedient is taken, the practical dx signal gain is appreciably greater on 28 Mc. than on 14 Mc.; otherwise it is nearly the same on both bands.

END-FIRE ARRAYS WITH PARALLEL COPLANAR ELEMENTS

The simplest form of high-frequency end-fire array consists of two parallel, non-staggered half-wave elements in the same horizontal plane. If they are fed 180 degrees out of phase, the bidirectional directivity is not greatest with half-wave spacing as might be supposed from a superficial consideration of time and space phase relationships, but rather with closer spacing. The directivity increases inversely with spacing quite noticeably down to a spacing of about 1/6 wavelength, then remains substantially constant for all spacings down to zero. However, the radiation resistance goes down as the spacing is reduced, and becomes very low for spacings less than about ⅛ wavelength, causing both the circulating

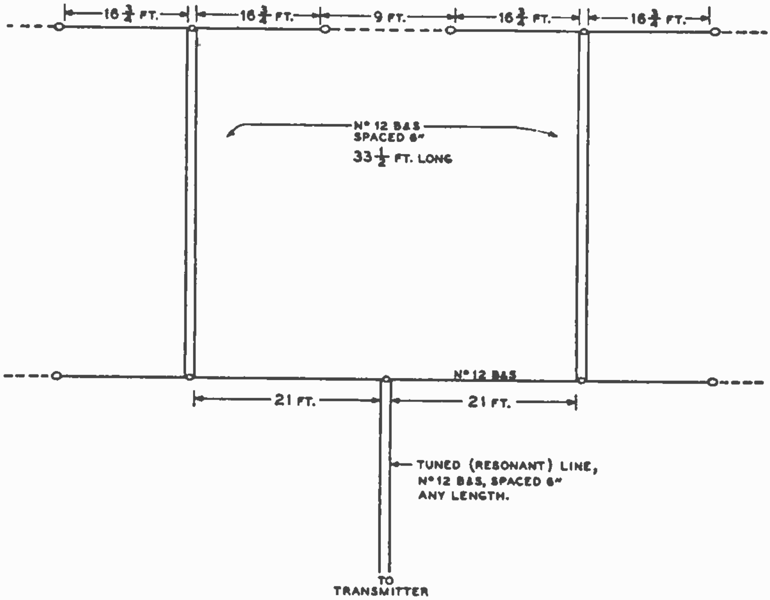


Figure 5-23.

THE "OCTAPUSH" AMATEUR ANTENNA ARRAY FOR THE 7, 14 AND 28 MC. BANDS.

Besides serving as a general coverage antenna for short and medium distance work on 7 Mc., the "Octapush" is a highly effective directional array on the 14 and 28 Mc. bands. The free space directivity gain on 14 Mc. and 28 Mc. is slightly over 6 db. Because the overall system is resonated at the transmitter, experimental adjustment of the element lengths is not necessary for best performance.

current and the peak voltages to reach very high values for a given power.

As a result of this situation, the signal gain begins to fall off rapidly as the spacing is decreased below a certain value which is determined by the loss resistance of the array. For a well-constructed two-element antenna using heavy conductors and very good insulation, this usually occurs at a spacing between 1/12 and ⅛ wavelength. Actually there is little difference in the signal gain with spacings between ⅛ and ¼ wavelength, and spacings within this range usually are employed in practical systems. It is desirable to use the greatest spacing which can be employed without significant reduction in signal gain, because increasing the spacing decreases the effective Q and thereby lessens the frequency sensitivity of the array.

End-fire arrangements of this type are preferable to stacked broadside arrays for long distance communication when the pole height is limited to ¾ wavelength or less, because the average elevation of the end-fire arrangement is the same as the pole height, while the average elevation of a two-tier stacked broadside array is ¼ wavelength less (assuming ½ wave spacing).

Another pertinent difference between stacked and end-fire arrangements is in the character of the directivity. For a given free space directivity gain, the beam width in a plane corresponding to azimuth directivity is *narrower* with the *end-fire* arrangement than with a stacked arrangement, because the end-fire arrangement contributes appreciable directivity in a plane containing the radiators as well as in a plane normal to the radiators. With a stacked arrangement practically all of the directivity

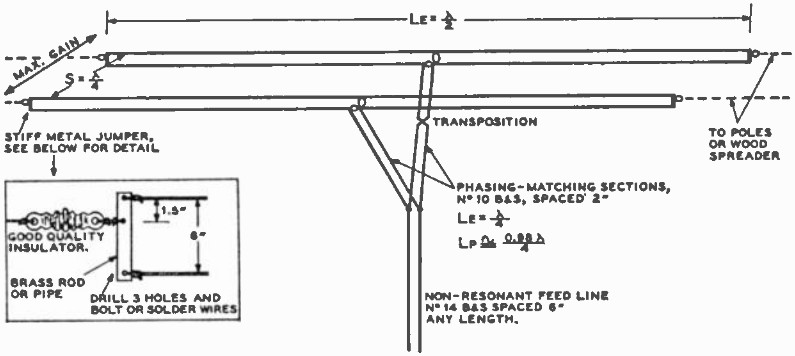


Figure 5-24.

SINGLE SECTION "FLAT TOP BEAM" END FIRE ARRAY.

This bidirectional h-f array has a free space directivity gain of about 4 db over a single element, a practical dx signal gain of the same order when well elevated, and somewhat more than 4 db when elevated between 1/4 and 1/2 wavelength above flat ground. The azimuthal beam width is about 40 degrees at the half power points. The phasing-matching sections can be precut to 0.98 of a free space quarter wavelength, but the element lengths preferably should be adjusted experimentally before being soldered. Usually the physical length of the elements will optimize at about 0.91 or 0.92 of a free space half wavelength. The element wire end-spacer shown makes precise rigging unnecessary, as the wires will be taut even though they differ slightly in length. Note that a 180 degree twist is taken in one phasing-maching section, but not in the other.

is in a plane normal to the radiators (except for the aspect directivity, which is the same in either case).

By spacing the elements an odd number of quarter wavelengths (usually ¼ or ¾) in an end-fire arrangement and driving them in *phase quadrature*, a unidirectional pattern is obtained. As this expedient is confined almost exclusively to large, commercially employed, multi-element arrays which combine both broadside and end-fire directivity, it will be discussed briefly in a later section on "combination" phased dipole arrays of the driven type.

At one time an end-fire configuration comprising many parallel half-wave elements in a horizontal row was popular, but the greater directivity obtainable with a broadside curtain has caused this type of array to fall into the discard except for a very few peculiar applications to which it happens to be particularly well suited. Using half-wave spacing, which simplifies the feeding of the elements in proper phase, an end-fire row of several half-wave dipoles has about 3 db less gain than the same number of parallel dipoles in a broadside row.

The gain of the two configurations will be approximately the same *for a given over-all row length* if the end-fire arrangement uses quarter-wave spacing, but this requires twice the number of dipoles and complicates the feed system, due to the necessary quadrature phase relationship between adjacent elements. Such an end-fire arrangement gives a undirectional pattern, which may or may not be an advantage, depending upon the particular application.

THE "FLAT TOP BEAM" END-FIRE ARRAY

Two horizontal, parallel, close-spaced, half-wave dipoles in a horizontal plane can be fed effectively in anti-phase by numerous means, and regardless of the feed method the combination usually is known as a "flat top beam" or "W8JK array" after the amateur who first popularized it. A recommended method of feed is shown in figure 5-24. This arrangement has practically the same signal gain as a closer spaced arrangement, and has the advantage of being much less frequency sensitive, due to the combination of comparatively wide spacing and folded dipole elements. Insulation requirements also are less stringent.

The free space directivity gain of the array shown in figure 5-24 is approximately 4 db, and the practical dx signal gain is of the same order when the array is elevated at least one half wavelength above the effective earth level. When the array is closer to ground, and the ground is comparatively flat, the practical dx signal gain will be somewhat higher (about 5 to 6 db over a single element for an elevation of ¼ wavelength). At low wave angles the beam width is approximately 40 degrees at the half power points.

A simple flat top beam antenna of the type shown in figure 5-24, using two radiating elements (a folded dipole is considered a single element), is referred to as a *single section* flat top. Additional sections may be added in colinear fashion to increase the azimuth directivity and power gain. However, because a single section already has considerable azimuthal directivity, two colinear sections with half-wave spacing between current loops produce but little additional directivity. The gain of two articulated sections over a single section is only about 1 db, which is hardly worthwhile.

To obtain a substantial gain over a single section it is necessary either to increase the current loop spacing between two sections to approximately one wavelength, or to employ three or more articulated sections. Two sections, each like that illustrated in figure 5-23, spaced with their centers one wavelength apart, are recommended in place of three articulated sections, because the use of two, separately fed sections makes available a choice of either a cloverleaf pattern or a figure 8 pattern simply by reversing the connections to one feed line. Two sections with a center spacing of one wavelength have a free space directivity gain of approximately 7 db, as compared to approximately 5 db for two contiguous sections.

It is possible to construct a flat top beam which will work with approximately equal effectiveness over a frequency range exceeding two to one, the directivity and gain remaining substantially the same throughout the range. Such an array is shown in figure 5-25. To minimize losses at the low frequency end of its useful range, the insulators supporting the elements should be the very best and have a long leakage path, because the single wire radiators and closer spacing result in higher impedances. Also, the resonant line should be of very low loss construction, because a rather high SWR will be present on the line. In view of the extensive useful frequency range, the element lengths are not especially critical.

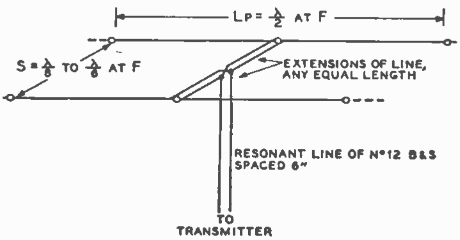


Figure 5-25.

TUNABLE, WIDE-RANGE FLAT TOP BEAM ANTENNA. THIS ARRAY IS IDEALLY SUITED WHERE SPACE LIMITATIONS MAKE IT NECESSARY TO USE ONE ANTENNA ON THE AMATEUR 14, 21, AND 28 MC. BANDS.

This version of the flat top beam antenna has a substantially uniform pattern and signal gain over a frequency range from F to 2(F), and gives good performance from 0.8(F) to 2.4(F). The overall system, antenna and feed line, must be resonated at the transmitter. To prevent excessive resistance loss, excellent insulation should be used throughout. If care is taken in construction, the practical dx signal gain over the useful frequency range is between approximately 4 and 5 db compared to a half-wave horizontal dipole at the same elevation, the gain being slightly higher near the upper frequency limit. As with the antenna of figure 5-24, the practical dx signal gain (but not the absolute field strength) is 1 or 2 db greater for elevations above flat ground on the order of 1/4 wavelength. The two extensions of the line should be made long enough that the half-wave radiating elements are not appreciably pulled away from a straight line by the weight of the feed system.

For maximum low-angle radiation a flat top beam antenna should be erected with the plane of the elements parallel to the ground. Tilting the array reduces the low angle radiation in both azimuth directions of maximum radiation. The only practical method of increasing the low angle radiation is to increase the effective elevation of the array.

PHASED DIPOLE ARRAYS WITH PARASITIC ELEMENTS

When a half-wave dipole is placed parallel to and closer than about one half wavelength to a driven element (non-staggered), the magnitude of the mutual impedance is sufficient to cause considerable current to flow in the free or "parasitic" element. Under these conditions the parasitic element will have appreciable effect upon the radiation pattern. Considerable variation in the phase angle of the currents can be obtained by detuning the parasitic element from resonance slightly, with only a small effect upon the relative magnitude of current flowing in the parasitic element. The relative magnitude of current flowing in the parasitic element can be controlled by choice of spacing, though from a practical standpoint the reduction in feed point resistance of the driven element with reduced spacing limits the extent to which the relative current in the parasitic element can be increased.

An array consisting of one driven dipole element and one or more parasitic dipole elements is commonly known as a *parasitic array* or *Yagi* array*. A parasitic array is ideally suited when a simple undirectional array is desired; in fact, nearly all parasitic arrays in practical use are of the undirectional type. A parasitic element is called a *reflector* when it is "behind" the driven element in a undirectional parasitic array, and a *director* when it is "ahead" of the driven element. More than one reflector seldom is employed in a simple parasitic array, but several directors (with or without a reflector) sometimes are employed one ahead of the other.

As a director a single parasitic element exhibits maximum directivity gain at about 0.1 wavelength spacing, the free-space directivity gain being about 5 db under optimum conditions. As a reflector the gain maximizes at a spacing of about 0.15 wavelength, the maximum free-space directivity gain being slightly less than that obtainable with a maximized director. However, it should be kept in mind that these spacings are optimum for maximum gain *only when a single parasitic element is employed.*

Maximum discrimination or "front-to-back ratio" is obtained with somewhat closer spacing than that which gives maximum gain. Therefore, when front-to-back discrimination is a prime consideration, a director spacing on the order of 0.05 wavelength can be used to advantage in receiving applications.

Parasitic elements ordinarily are "tuned" to give the desired amount and kind of reactance by control of their physical length. The parasitic element length which gives maximum gain ordinarily does not provide maximum front-to-back discrimination for that particular spacing. However, adjusting the parasitic element length for maximum discrimination usually results in only a small reduction in forward gain from the maximum gain obtainable at that spacing.

When the element spacing is made a small fraction of a wavelength, the effective "Q" of the overall system becomes quite high, and the antenna is quite frequency sensitive. Also, the "tuning" of the parasitic element(s) is more critical, further aggravating the condition. For this reason, when a parasitic array is to be used over a band of frequencies rather than on a single frequency, the element spacing sometimes is made greater than otherwise would be the case. Gain and front-to-back ratio at the mid frequency are sacrificed in the interest of improved performance at frequencies slightly removed from the "design center."

*"Yagi array" originally referred more specifically to an array employing parasitic elements in one of several particular configurations; it now is used loosely to refer to any parasitic array employing one or more reflectors and one or more directors, regardless of the element spacing.

In no case can a conventional parasitic array qualify as a "broad-band" antenna, because even though extremely large diameter elements may be used, a moderate change in frequency seriously upsets the reactance relationships of the various elements, which in turn affect the phase relationships.

While useful gain and discrimination can be obtained with a single parasitic element, the addition of another parasitic element (making one director and one reflector) provides a worthwhile increase in gain and discrimination. Because of the improved performance and the fact that the addition of another parasitic element complicates the mechanical construction only slightly, the combination of a single driven element and single parasitic element will not be treated in detail.

The remarks in the preceding paragraph do not necessarily apply in the case of an array using several driven elements, such as a broadside curtain. In the case of a driven curtain the usual practice is to use either a parasitic reflector curtain *or* a parasitic director curtain, as but little additional gain is to be realized by using both. While the use of "fore and aft" parasitic curtains does help the discrimination, effectively the same result can be obtained by using a total of only two curtains and feeding both curtains with equal currents (with an appropriate phase angle). Thus, when one parasitic curtain does not provide the desired discrimination, it is better to employ a two curtain driven array than add another parasitic curtain.

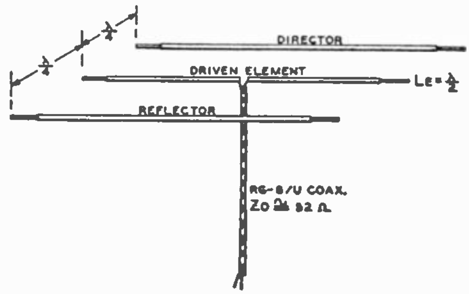


Figure 5-26.

THREE-ELEMENT PARASITIC ARRAY

This compact array has high gain with good discrimination, a low line SWR, and is ideally suited to mechanical rotation at the higher frequencies. For best results the element lengths should be adjusted experimentally, and they therefore should be made adjustable by incorporation of telescopic end sections or other means. The free space directivity gain is approximately 7 db, and the beam width slightly over 50 degrees at the half power points.

THE THREE-ELEMENT PARASITIC ARRAY

One of the most popular parasitic arrays employs a horizontal half-wave dipole as a driven element, with one parasitic reflector and one parasitic director, all parallel and in the same horizontal plane. This combination is hard to beat from the standpoint of size and simplicity vs. gain and discrimination, and its popularity is well deserved. Its compact construction makes mechanical rotation a simple matter at the higher frequencies, and when so employed the array usually comprises self-supporting tubular elements and often is referred to as a "three-element rotary array."

There is considerable latitude as regards element spacing. The optimum spacing depends upon the relative importance of gain and discrimination, but in any case the spacing is not very critical, and it is possible to obtain both a good degree of discrimination and a high order of gain with compromise design and adjustment.

A three-element parasitic array developed during the war for military applications and designed for high forward gain with simplicity of the feed system used quarter-wave spacing and RG-8/U 50-ohm flexible coaxial cable feeding directly into the center of the driven dipole, without benefit of a matching device. When the element lengths are correctly adjusted, both SWR and antenna effect on the line are quite tolerable with this arrangement. The schematic is shown in figure 5-26.

The front-to-back ratio of this arrangement when the element lengths are optimized for maximum forward gain is between 7 and 10 db. The beam width is slightly more than 50 degrees at the half power points. The free space directivity gain is approximately 7 db. While somewhat greater gains have been attributed to this type array, they have been based on experimentally derived data. The apparent discrepancy is explained by the fact that under certain conditions the practical signal gain of an array can be greater than the free space directivity gain, as discussed earlier in this chapter. The front-to-back ratio can be increased to approximately 15 db. without appreciably reducing the forward gain, simply by adjusting the elements for this compromise condition.

When a higher front-to-back ratio is desired, closer spacing (approximately 0.15 wavelength) is desirable. This lowers the feed point impedance of the driven element considerably, requiring the use of an impedance transformer, regardless of the type of line used, if a substantially flat line is desired. If low loss construction is employed and the feed line is matched to the feed point resistance of the driven element, so as to operate under flat line conditions, the practical signal gain will be about the same with 0.15 wavelength spacing as with 0.25 wavelength spacing, but the antenna will be more frequency sensitive.

The frequency sensitivity of a parasitic array can be reduced somewhat for a given element spacing by using large diameter elements, but the improvement thus obtainable is considerably less than that which can be obtained by applying this expedient to a simple half-wave antenna.

The use of a folded dipole as the driven element in conjunction with a line having a suitable characteristic impedance does not appreciably reduce the frequency sensitivity of the overall system, due to the fact that much of the frequency sensitivity is due to the reactance slope of the parasitic elements. A folded dipole or T-matched dipole often is used as the driven element, but not so much to reduce the frequency sensitivity of the array as to provide a more suitable feed point impedance for 300-ohm Twin-Lead or an open wire line, or to avoid some of the mechanical complications entailed in splitting and insulating the driven element, or to obviate the necessity for experimental adjustment of the length of the driven element.

The best method of determining the correct parasitic element lengths for a given frequency is to adjust them experimentally until maximum gain or maximum discrimination (or a desired compromise) is obtained. Fortunately the optimum element lengths are not significantly affected by height above ground once the array is at least a quarter wavelength above ground and *well clear of surrounding objects.* This makes it possible to optimize the element lengths at a height which is more readily accessible for experimental adjustment.

Experimental adjustment of the element lengths readily is accomplished by means of a remote reading field strength meter picking up at a point several wavelengths from the radiator and indicating at the antenna position. When front-to-back ratio is important, a second field strength meter is useful, as it makes it possible to determine the ratio at a glance without rotating the array through 180 degrees.

When the array is to be used over a band of frequencies rather than on one fixed frequency, the reflector should be lengthened until a perceptible decrease in field strength is noted, and the director shortened until a perceptible decrease in field strength is noted, after first adjusting the elements for maximum gain at the mid frequency. When adjusted in this manner the total frequency range over which the array may be used without a serious reduction in performance usually will be between 0.5% and 2% of the design center, depending upon the spacing and length/diameter ratio of the elements, upon the matching ratio, and upon other factors. This represents a deviation from the mid frequency of from about 0.25% to 1% plus or minus.

The rather laborious process of adjusting the element lengths experimentally can be avoided by either duplicating *exactly* the dimensions *and other mechanical construction* of an array known to give good performance on the desired frequency, or by purchasing one of the "precut" or "calibrated" arrays manufactured commercially. The precut or pretuned type is intended to give good performance over a specified narrow band of frequencies, such as an amateur band, when erected well in the clear. It is not readily adjustable. The calibrated

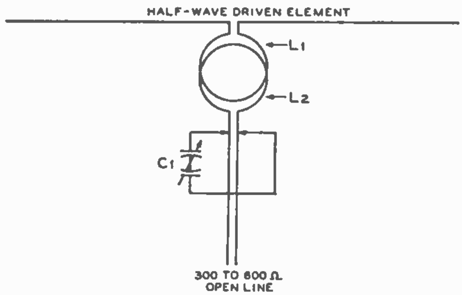


Figure 5-27.

COAXIAL LINK METHOD OF FEEDING THE DRIVEN ELEMENT OF A ROTATABLE PARASITIC ARRAY.

This inductively coupled arrangement permits continuous rotation without the use of slip rings or a "rotary joint". The loops L_1 and L_2 are identical physically and surround the shaft which supports the array. The capacity of C_1 and the point of attachment to the line are adjusted until "flat line" operation is obtained. In some cases the coupling (spacing) between L_1 and L_2 will have to be adjusted in order to achieve this condition. The physical length of the driven element need not be shortened to allow for the insertion of L_1 as the reactance of L_1 is effectively "tuned out" when C_1 is optimized. A shorted stub slightly less than one half wavelength long (the proper adjustment of the shorting bar being determined by experiment) may be substituted for C_1 if desired.

type is equipped with telescoping elements which are calibrated in frequency (on the assumption that the array will be installed according to instructions).

Rotating mechanisms and remote position indicators designed expressly for the purpose are availably commercially at reasonable prices, or they may be adapted or improvised from items originally intended for other purposes. Amateurs and others who for reasons of economy, pride, or enjoyment wish to design and construct their own rotary array will find a wide range of readily adaptable items at the junk yards and salvage stores in the larger cities. War surplus or salvaged ordnance and aircraft parts and equipment will be found an especially fruitful source of such items. For instance, an airplane propeller pitch changer of a widely used type makes an ideal rotator for a good sized array, with only slight modification.

A power driven reversible rotator with an electrical indicator of position is an operating convenience in any case, and is the only really practical arrangement when the array is located at a considerable distance from the operating position. However, the reversible motor, gear reduction, and electrical indicator often are not absolutely essential, and can be dispensed with by the use of manual drive, either direct (where the array is directly above the operating position) or by means of rope and pulleys.

There are various methods of transferring radio frequency power from the fixed supporting tower or pole to the rotatable driven element. The simplest is to use a flexible line, such as RG-8/U coaxial cable, and either coil it around the rotatable support shaft a few times in a helix, or to run it down through the center of the supporting shaft if the latter is hollow. Either of these arrangements requires either a "twist indicator" or an automatic stop (or automatic reverse) when power drive is used, in order to avoid subjecting the line to excessive strain as a result of continued rotation in one direction.

When unrestricted continuous rotation with power drive is desired, either sliding contacts or coaxial inductive loops are employed. The inductive loops, one stationary and one attached to the rotating assembly, are placed concentric with the supporting shaft and usually are tuned by means of lumped shunt capacity to accomplish the desired coupling characteristics. (See figure 5-27.) Sliding contacts take the form of slip rings or a rotary coaxial joint. The former are best adapted for use with a symmetrical ("balanced") line, while the latter is best adapted for use with a coaxial (asymmetrical) line.

Many of the commercially available arrays employ either "rotary link" inductive coupling or sliding contacts in order to achieve unrestricted continuous rotation. However, a home-constructed array is more easily built using a flexible feed line of appropriate construction and an automatic stop or automatic reversing switch.

Because the proper element lengths of a home constructed array are best determined experimentally when the array is not

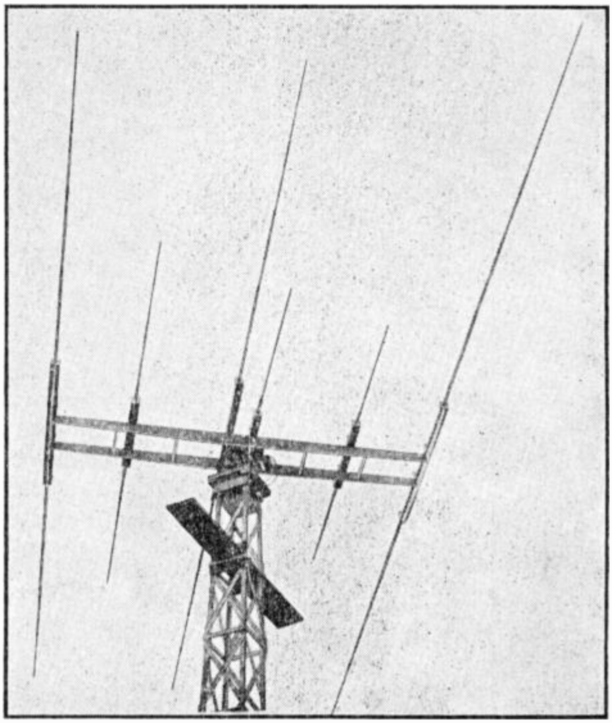


Figure 5-28.

DUAL BAND ROTARY PARASITIC ARRAY OF COMMERCIAL MANUFACTURE (MIMS-AMPHENOL "SIGNAL SQUIRTER") UTILIZING TOP OF TOWER ROTATOR AND COAXIAL LINK FEED IN CONJUNCTION WITH AN OPEN WIRE LINE. THE ROTATING LINK FEED PERMITS CONTINUOUS ROTATION.

an exact duplicate of one already optimized for maximum performance, and because the optimum lengths are affected considerably by the spacing and length/diameter ratio of the elements, no frequency vs. length conversion data will be given. However, in general the driven element will be about the same length as though no parasitic elements were used, the reflector will be from zero to about 6 per cent longer than a single half-wave element (not a folded dipole) in free space, and the director(s) will be from zero to about 6 per cent shorter than a single half-wave element in free space. If the driven element is not a folded dipole, all elements should be made symmetrically adjustable over these ranges by means of telescoping end sections or by means of an adjustable center clamp. If the driven element is a folded dipole, good results usually will be obtained by cutting the driven element to 0.92 or 0.93 of a half wavelength at the mid frequency and making only the parasitic elements adjustable.

Generally speaking, slight detuning or improper adjustment of the length of the driven element affects primarily the line SWR; it does not appreciably affect the directivity pattern or the directivity gain.

The tubular elements should be rigid enough to prevent excessive sag and whip. A moderate amount of either has no deleterious effect, especially when the element spacing exceeds 0.2 wavelength. Duralumin tubing of appropriate hardness and wall thickness makes excellent element material, though thin walled steel tubing also is good if it is copper or zinc plated to increase the r-f conductivity and protect the steel from effects of weather.

"All metal" construction of the array of figure 5-26 requires only one insulator (to support the half of the driven element connected to the inner conductor of the line). A metal supporting arm will have no significant effect upon the currents flowing in the elements, and vice versa, when connected only to points of zero r-f voltage and

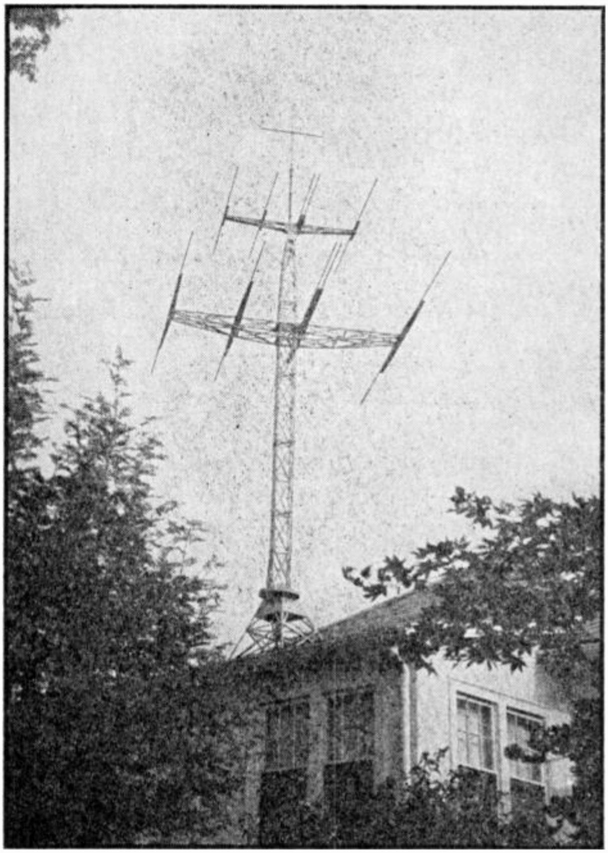


Figure 5-29.

HOME CONSTRUCTED ROTATING TOWER USED BY AMATEUR STATION W3EDD TO SUPPORT PARASITIC ARRAYS FOR THE 28, 50 AND 144 MC. BANDS, THE LATTER ORIENTED FOR VERTICAL POLARIZATION. THE WOOD MAST IS LIGHT BUT VERY STRONG, DUE TO THE EMPLOYMENT OF "AIRPLANE TYPE" CONSTRUCTION.

oriented so as to be normal to the elements it bisects.

A slight increase in forward gain can be obtained by adding a second director, though the frequency sensitivity of the array will be increased by the additional element. The maximum discrimination that is obtainable with four elements (one reflector and two directors) is not significantly greater than that obtainable with three elements (one director and one reflector). The addition of a third director does little to increase either gain or discrimination, and is hardly worthwhile unless the frequency is so high that the elements are quite small physically.

The accompanying photographs illustrate various rotary parasitic array components and details.

The construction aspects of practicable rotatable arrays are treated in somewhat greater detail in Chapter 30 of the *Radio Handbook.*

COMBINATION END-FIRE AND BROADSIDE ARRAYS

Any of the end-fire arrays previously described may be stacked one above the other or placed end to end (side by side) to give greater directivity gain while maintaining a bidirectional characteristic. However, it must be kept in mind that to realize a worthwhile increase in directivity and gain while maintaining a bidirectional pattern the individual arrays must be spaced sufficiently to reduce the mutual impedances to a negligible value.

When two flat top beams, for instance, are placed one above the other or end to end, a center spacing on the order of one wavelength is required in order to achieve a worthwhile increase in gain, or approximately 3 db.

Thus it is seen that, while maximum gain occurs with two stacked dipoles at a spacing of about 0.7 wavelength and the space directivity gain is approximately 5 db over one element under these conditions, the case of two flat top or parasitic arrays stacked one above the other is another story. Maximum gain will occur at a greater spacing, and the gain over one array will not appreciably exceed 3 db.

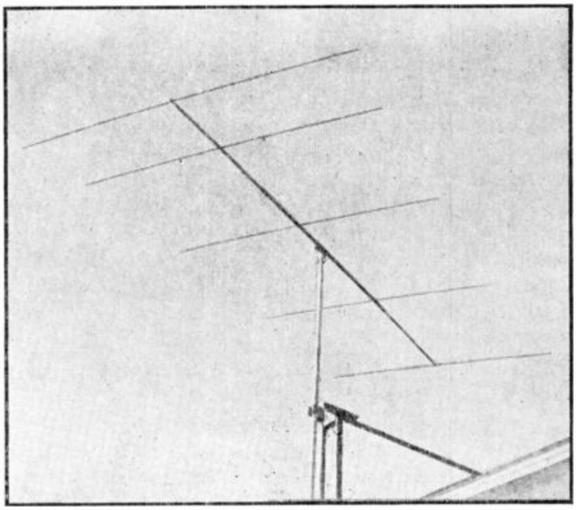
When two broadside curtains are placed one ahead of the other in end-fire relationship, the aggregate mutual impedance between the two curtains is such that considerable spacing is required in order to realize a gain approaching 3 db (the required spacing being a function of the size of the curtains). While it is true that a space directivity gain of approximately 4 db can be obtained by placing one, half-wave dipole an eighth wavelength ahead of another and feeding them 180 degrees out of phase, a gain of less than 1 db is obtained when the same procedure is applied to two large broadside curtains. To obtain a gain of approximately 3 db and retain a bidirectional pattern, a spacing of many wavelengths is required between two large curtains placed one ahead of the other.

A different situation exists, however, when one driven curtain is placed ahead of an identical one and the two are phased so as to give a unidirectional pattern. When a unidirectional pattern is obtained, the gain over one curtain will be approximately 3 db regardless of the spacing. For instance, two large curtains placed one a quarter wavelength ahead of the other may have a space directivity gain of only 0.5 db over one curtain when the two are driven 180 degrees out of phase to give a bidirectional pattern (the type of pattern obtained with a single curtain). However, if they are driven in phase quadrature (and with equal currents) the gain is approximately 3 db.

The directivity gain of a composite array also can be explained upon the basis of the directivity patterns of the component arrays employed alone, but it entails a rather complicated picture. It is sufficient for the purpose of this discussion to generalize and simplify by saying that the greater the directivity of an end-fire array, the farther an identical array must be spaced from it in broadside relationship to obtain optimum performance; and the greater the directivity of a broadside array, the farther an identi-

Figure 5-30.

FIVE-ELEMENT 28-MC. AMATEUR ARRAY OF THE PARASITIC TYPE, UTILIZING A ROTATING PIPE MAST (U-H-F RESONATOR CO.). THE MAST, BOOM, AND ELEMENTS ALL ARE OF ALUMINUM ALLOY TUBING, MAKING A VERY LIGHT ASSEMBLY HAVING GREAT MECHANICAL STRENGTH. IT ALSO IS LESS CONSPICUOUS THAN OTHER TYPES, AN IMPORTANT CONSIDERATION IN SOME RESIDENTIAL LOCATIONS.



cal array must be spaced from it in end-fire relationship to obtain optimum performance and retain the bidirectional characteristic.

It is important to note that while a bidirectional end-fire pattern is obtained with two driven dipoles when spaced anything under a half wavelength, and while the proper phase relationship is 180 degrees regardless of the spacing for all spacings not exceeding one half wavelength, the situation is different in the case of two curtains placed in end-fire relationship to give a bidirectional pattern. For maximum gain at zero wave angle, the curtains should be spaced an odd multiple of one half wavelength and driven so as to be 180 degrees out of phase, or spaced an even multiple of one half wavelength and driven in the same phase. The optimum spacing and phase relationship will depend upon the directivity pattern of the individual curtains used alone, and as previously noted the optimum spacing increases with the size and directivity of the component arrays.

A concrete example of a combination broadside and end-fire array is two Lazy H arrays spaced along the direction of maximum radiation by a distance of four wavelengths and fed in phase. The space directivity gain of such an arrangement is slightly less than 9 db. However, approximately the same gain can be obtained by juxtaposing the two arrays side by side or one over the other in the same plane, so that the two combine to produce, in effect, one broadside curtain of twice the area. It is obvious that in most cases it will be more expedient to increase the area of a broadside array than to resort to a combination of end-fire and broadside directivity. One exception, of course, is where two curtains are fed in phase quadrature to obtain a unidirectional pattern and a space directivity gain of approximately 3 db with a spacing between curtains as small as one quarter wavelength. Another exception is where very low angle radiation is desired and the maximum pole height is strictly limited. The two aforementioned Lazy H arrays when placed in end-fire relationship will have a considerably lower radiation angle than when placed side by side if the array elevation is low, and therefore may *under some conditions* exhibit appreciably more practical signal gain.

"BEAM WARPING" (ELECTRICAL ROTATION) AND PATTERN SWITCHING

One disadvantage of a conventional h-f broadside array having considerable azimuthal directivity is that the area which can be covered effectively by the beam goes down as the horizontal directivity and gain are increased.

When the size of the array becomes too great to permit economical utilization of mechanical rotation, it is more practical to achieve pattern control *electrically* rather than mechanically. This is done by controlling the phase relationships in the various bays of a broadside array.

The complexity of an electrically rotated array and its associated adjustable phasing network increases with the degree of arc over which the main lobe(s) must be shifted or "slued." In any case the usable arc is somewhat less than 180 degrees (for each lobe of a bidirectional "figure 8") when horizontally oriented dipoles are used in the array.

THE "ELECTROTATOR" ARRAY

The basic principles of electrical beam rotation are illustrated in the elementary version of an electrically rotated array shown in figure 5-31. This "Electrotator" bidirectional array has a space directivity gain of approximately 10 db when the beam is "aimed" due broadside, and a beam width of approximately 28 degrees at the half power points. The "nose" of the beam can be deflected to the right or left a maximum of 20 degrees in steps of approximately 5 degrees, thus permitting a 40 degree total sweep of each lobe of the bidirectional pattern. This means that each lobe is useful over approximately 68 degrees, and can be put "dead on" over an arc of approximately 40 degrees. Because of the bidirectional pattern, the array actually covers twice this amount of arc, and ordinarily two of these arrays can be oriented to cover all important compass directions from a given location.

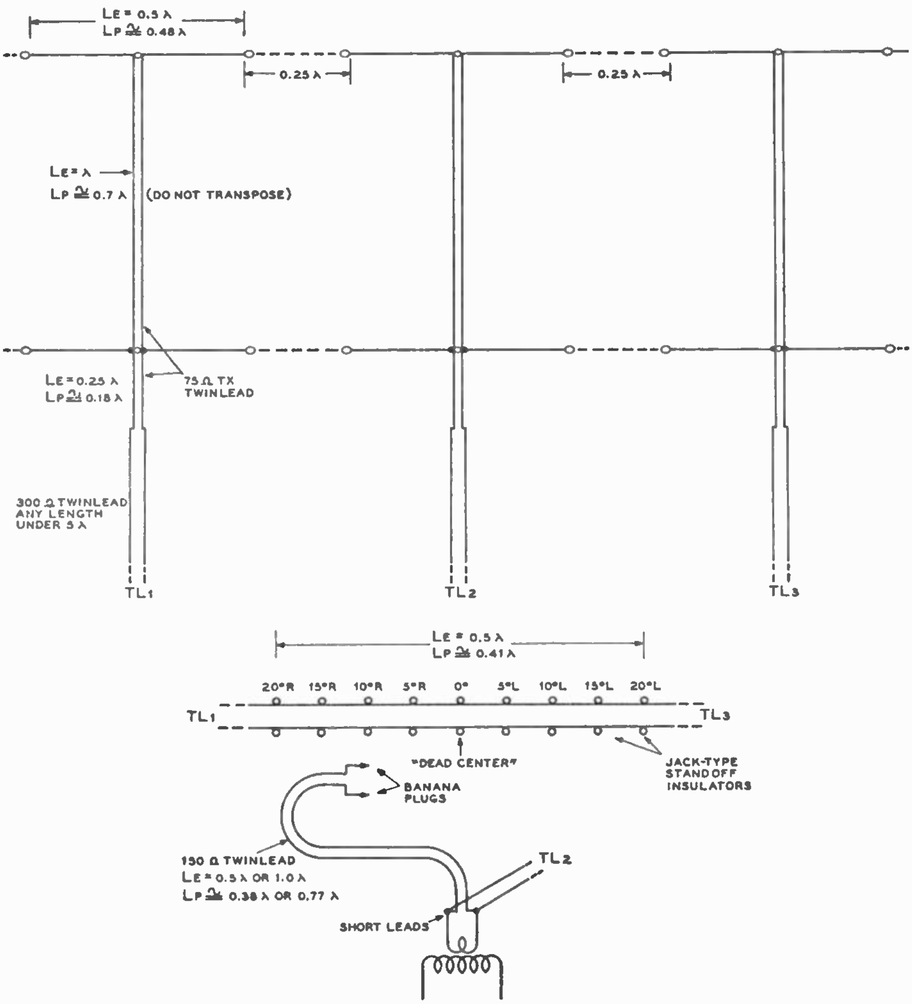


Figure 5-31.

"ELECTROTATOR" BIDIRECTIONAL ARRAY EMPLOYING ELECTRICAL BEAM ROTATION.

Each lobe of this bidirectional array can be swept through an arc of approximately 40 degrees in steps of approximately 5 degrees. The space directivity gain with the beam due broadside is approximately 10 db and the beam width about 28 degrees at the half power points. The gain and sharpness of the beam deteriorate slightly as the beam is shifted to the right or left. For details of the lobe shifter, refer to the accompanying text.

As the beam is deflected to the right or left from "due broadside," the beam width increases slightly, the free space directivity gain decreases slightly, and the amplitude of the minor lobes increases slightly with the amount of deflection. But the variations are so small that even under conditions of maximum deflection the performance of the array is not appreciably deteriorated.

It will be noted that each of the three identical bays of the array consists of two stacked dipoles, the spacing being that which gives maximum gain. The velocity factor of the 75-ohm transmitting type Twin-Lead used as a combined feed line and phasing section to the upper element is such that a section having an *electrical* length of one wavelength automatically gives optimum spacing between upper and lower elements.

75-ohm Twin-Lead also is used as a quarter wave "Q" section to match the feed point resistance of each bay to its individual 300-ohm transmission line. This arrangement results in a low SWR when the array is well elevated and in the clear.

The three feed lines, TL_1, TL_2, and TL_3, should be cut so that with the phase shifter set at "dead center" the *electrical* length of TL_1 is exactly equal to that of TL_3, and the *electrical* length of TL_2 is equal to that of TL_1 (including half the phase shifter) *plus* that of the 150-ohm flexible "pigtail" section. The lines should be poled so that the currents in all elements flow in the same direction at a particular instant. Usually these requirements can be met satisfactorily simply by allowing for the difference in propagation time on the basis of published "approximate" or "average" values of velocity constant for the various types of Twin-Lead, and being careful to observe polarity when splicing the various sections of Twin-Lead.

To minimize the effect of rain on the phase relationships (and therefore the direction of the beam), the three 300-ohm lines should be made as short as practicable, treated with Simoniz wax or Silicone waterproofing compound, and arranged so that approximately the same amount of each line is exposed to rain.

The step-by-step phase shifter is constructed as follows. An electrical half wavelength of 300-ohm Twin-Lead is divided into eighths, giving the nine plug-in contact points shown in figure 5-31. At each point the line is straddled by a pair of small, jack-type standoff insulators with solder lugs turned inward and soldered to the Twin-Lead at that point. The 150-ohm Twin-Lead "pigtail" shown in the illustration is terminated with two banana plugs mounted on an insulated spacing bar and spaced to mate with the jacks.

The overall system as shown will handle safely a transmitter output of approximately 400 w. amplitude modulated. When the array is constructed properly, the weak link or limiting component is the 150-ohm Twin-Lead pigtail, which must handle two-thirds of the transmitter output. If this pigtail is constructed of two lengths of RG-11/U flexible coaxial cable lashed together to form a symmetrical line section of approximately 150 ohms surge impedance, the power handling capability of the overall system is increased to approximately 1500 watts transmitter output, amplitude modulated. The lower velocity constant of this type of line (approximately 0.66) should be taken into consideration when cutting it to length. The outer conductors of the two parallel sections of coaxial line should be connected together at each end, and the inner conductors employed in the same fashion as the 150-ohm Twin-Lead.

Unless properly made, each junction between 300-ohm line and a 75-ohm Q section will be vulnerable to flashover and thus greatly restrict the power handling capability of the array. After a section of 75-ohm transmitting type Twin-Lead is solder-spliced to 300-ohm Twin-Lead, some scraps of polyethylene should be flowed in around the close-spaced bare wires and fused to the Twin-Lead dielectric with the aid of a hot iron.

The slight deterioration in performance in wet weather due to vulnerability of the 300-ohm Twin-Lead to moisture can be avoided by the use of an open wire line of approximately 500 ohms surge impedance and making the following substitutions. For each quarter-wave Q matching section, substitute a Q section made up of two, parallel lengths of RG-8/U coaxial cable lashed together, with the outer conductors jumpered at both ends, due allowance being made for the velocity factor (0.66). This gives a symmetrical transformer with a surge impedance of approximately 100 ohms. For the half-wavelength "phase shifter" section use the same construction as for the open wire lines, increasing the physical length of the section to 0.97 of a free-space half wavelength to allow for the greater velocity factor. The phase shifter pigtail should be constructed of 300-ohm Twin-Lead in this case, due allowance being made for the greater velocity factor (0.82). Because this section of 300-ohm Twin-Lead is indoors, it will be unaffected by weather. However, the safe current capacity of this pigtail will limit the transmitter carrier power which may be handled by the antenna to approximately 600 watts, amplitude modulated. (The pigtail carries two thirds of the total power.)

Strictly speaking, an array of the Electrotator type is a true broadside array only when the lobe shifter is in the "dead center" position and the lobe is normal to the line of the array. However, it is common practice to consider the array as falling in the "broadside" category under lobe shifted conditions so long as the phase difference between the outer bays does not exceed 180 degrees.

To ensure peak performance the element lengths should be adjusted experimentally, but when the array is well elevated and in the clear, and is to be used only over a very narrow frequency range, the radiators may be cut to 0.48 of a half wavelength with the assurance that good results will be obtained. This assumes the use of no. 14, 12 or 10 B&S guage wire at frequencies between 6 and 30 Mc., and the use of insusulators providing a gap of at least 6 inches at the end of each radiator.

When possible the array should be elevated sufficiently to put the lower elements at least one half wavelength above the effective earth level, though good results will be obtained with the lower elements as low as one quarter wavelength above effective ground.

If an "Electrotator" array is to be used for both transmitting and receiving, it is important that the send-receive relay be inserted right at the antenna coil of the transmitter (at the point marked "short leads" in figure 5-31). The line to the receiver should be made an exact multiple of an electrical half wavelength, the fewer the better (in order to minimize frequency sensitivity).

It is important also that the input impedance of the receiver accurately match the feed system, or otherwise deflection of the beam by the phase shifter will not always be the same on receive as on transmit, causing the azimuthal directivity to shift when switching from "transmit" to "receive." With some receivers it will be necessary to provide an impedance matching device at the receiver terminals in order to obtain acceptable performance.

Another solution is to use a separate bidirectional receiving antenna having sufficiently broad horizontal directivity to cover the entire usable azimuth range of the "Electrotator." This antenna may consist of two stacked dipoles similar to one bay of the "Electrotator," removed as far as practicable to one side in order to avoid undesirable couplings. Or, by suitable relay connections, just the center bay of the "Electrotator" may be used alone as the receiving antenna. In this case the relay may be inserted anywhere along TL_2 and the length of line from the relay to the receiver is not critical.

THE "X-CURTAIN"

A somewhat simpler system of pattern control, ideally suited to amateur applications, is incorporated in the 2 by 2 expanded curtain or "X-Curtain" of figure 5-32. The application of the phase/antiphase switch is similar to that illustrated in figure 5-14B.

The free space directivity gain of the array with the two bays in phase is approximately 8 db, or about 2 db better than the popular Lazy H, and the azimuthal beam width is about 50 degrees at the half power points.

With the two bays in anti-phase, a four-lobe "cloverleaf" pattern is obtained which effectively straddles the in-phase "figure 8"

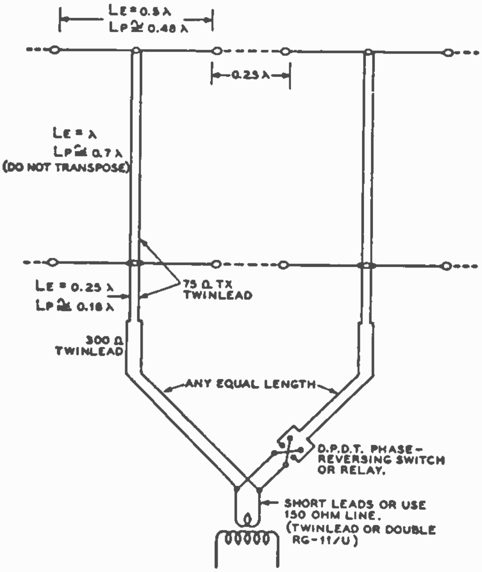


Figure 5-32.

THE "X-CURTAIN" TWO PATTERN ARRAY.

Using Twin-Lead line throughout, this expanded 2 by 2 curtain is but little more difficult of construction than a Lazy H, and has about 2 db greater broadside gain, or a space directivity gain of approximately 8 db. It has the additional advantage of pattern control, whereby the bidirectional figure 8 pattern can be changed to a cloverleaf which straddles the figure 8, thus approximately doubling the useful coverage of the array. If desired, a 450-550 ohm open wire line can be substituted for the 300-ohm Twin Lead by utilizing a different "Q" section as described in the accompanying text.

coverage, thus greatly increasing the useful coverage of the array. With the anti-phase arrangement, the space directivity gain is reduced slightly, and is comparable to the gain of a Lazy H. The character of the mutual impedance between the two bays is such that only a slight change in transmitter loading will occur when the phase switch is reversed.

The array will handle safely a transmitter power output of 1 kilowatt, amplitude modulated, if care is taken in construction.

Reference is made to those portions of the text on the "Electrotator" array previously described which deal with array elevation, feed line and element length, and with splicing and waterproofing of the Twin-Lead. It will be noted that the two arrays employ bays of identical construction.

If it is felt that the improvement in wet weather performance justifies the extra installation work involved, 450-550 ohm open-wire line may be used insted of 300 ohm Twin-Lead by substituting for each of the 75-ohm quarter-wave Q sections a 100-ohm Q section made up by lashing together two lengths of RG-8/U coaxial cable as previously described under the "Electrotator" array.

VERTICAL TRIAD ARRAY WITH PATTERN SWITCHING

While in general a horizontally polarized array is to be preferred in the h-f range, there are certain vertically polarized arrays which have important advantages to offer under particular circumstances. For instance, certain characteristics of the array of 5-33 could not be achieved with any combination of horizontal elements.

This "Vertical Triad" array offers a simple and compact means of obtaining directional radiation and response at low vertical angles at the lower end of the high freqency range, in any desired direction, without resorting to unhandy pole heights. It is highly recommended for the amateur who is interested in a full coverage "dx" antenna system for the 7-Mc. amateur band, yet is confined to a city lot.

Any two elements of the antenna constitute a single-section, close-spaced end-fire dipole array similar to a "Flat Top Beam" antenna turned on end. Such an array makes an effective low-angle radiator with moderate horizontal directivity, and not only provides worthwhile gain over a single vertical dipole, but also results in a symmetrical arrangement which is balanced to ground.

By spacing three such elements equally around a circle and providing for individual feed of each element, it is possible to activate any two elements from within the station by means of a two-pole triple-throw switch. The floating element always will be in a neutral plane with respect to the two driven elements and therefore will have

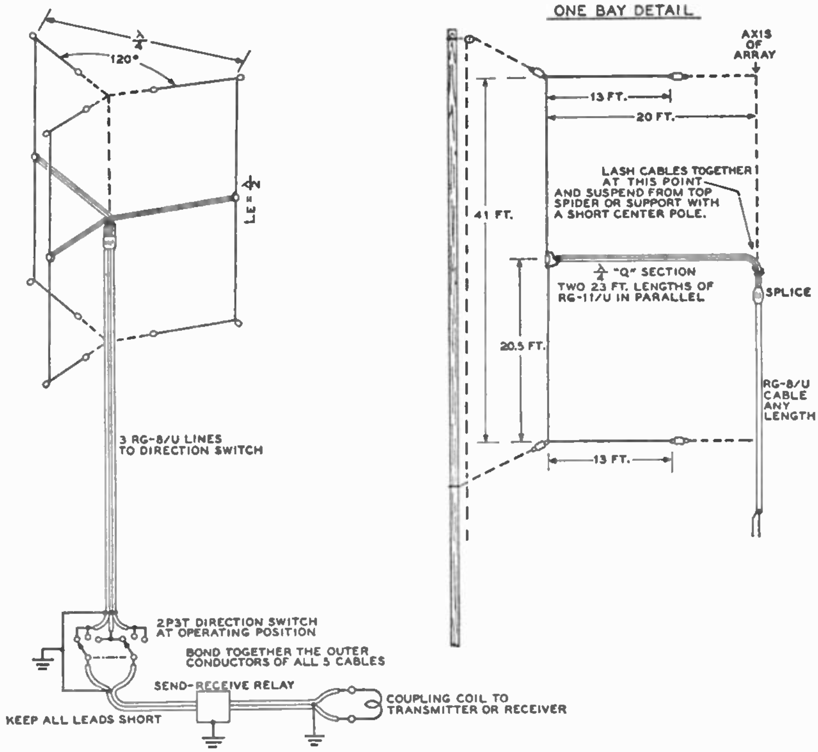


Figure 5-33.

ELECTRICALLLY ROTATED VERTICAL TRIAD ARRAY FOR 7-MC. AMATEUR BAND.

This compact array is a low angle radiator, and the "dx" performance when suspended from 50-foot poles is in all compass directions at least as good (when surrounding obstructions are not excessively high) as that of a half-wave horizontal dipole in its broadside direction when elevated approximately 100 feet. The bidirectional pattern may be switched to three positions which effectively cover all azimuthal bearings.

no effect upon the operation. Both the horizontal and vertical portions of all three feed lines which are close enough to the dipoles to give trouble from undesirable couplings also are in neutral planes, regardless of which two dipoles are being fed, and therefore will have no significant effect upon the operation of the dipoles and will not be susceptible to "antenna effect."

The direction switch makes available any one of three overlapping, equally spaced, bidirectional "figure 8" patterns. The horizontal directivity is broad enough that in a direction midway between two adjacent directions of maximum gain the gain will be only very slightly down from the maximum value.

The excellent "dx" performance of this array is due to the fact that most of the radiation is confined to vertical angles below 30 degrees regardless of the antenna elevation and type of soil. However, the performance will be slightly better over ground of high conductivity.

Although it is desirable to elevate the array sufficiently to put it well above surrounding obstructions so as to be "in the clear," good results will be obtained in an area of one-story buildings and low trees when the array is hung from three 50-foot poles.

Bending the ends of the half-wave dipoles towards the axis of the array shortens the vertical dimensions of the array and minimizes the required pole height without significantly affecting the gain of the array. However, radiation from and pickup by the horizontal portions of the array prevent a complete "null" in a direction broadside to the line of the two elements being employed. If maximum possible discrimination is desired, the half-wave elements should not be bent, but instead extended vertically to 67 feet.

The incorporation of the 36-ohm "Q" sections (two length of RG-11/U in parallel) results in a low SWR, making the length of the line uncritical. However, if so desired the Q sections can be dispensed with by accepting a moderate SWR and cutting the lines to an exact resonant length.

If sufficiently husky supporting poles are used, the wires forming the top "spider" can be pulled taut enough to support the lines as shown in figure 5-33, without excessive sag. If light poles are employed, it is recommended that the converging lines be supported by a short center pole rather than suspended as shown in the illustration. It is important that the converging lines from the three dipoles make approximately a right angle with the dipoles and then, upon reaching the axis of the array, drop vertically to a point at least ten feet below the bottom of the array or to the ground before taking another direction. If these precautions are observed, the antenna effect will be negligible.

While the optimum overall radiator length will depend somewhat upon the type of end insulators used, upon the nature and proximity of the supporting poles, upon the presence of surrounding objects, and upon other factors, satisfactory operation over the range 7000 to 7300 kc. usually will be obtained with the dimensions specified in figure 5-33 in spite of the fact that the dimensions are somewhat critical for best performance (due to the rather high effective Q). The response can be broadened out somewhat by substituting very large aluminum wire of the type sold for clothesline use and shortening the overall length of each half-wave element by 1 foot.

Except for the center insulator at the feed point of each element, all insulators are at very high impedance points, and long-gap insulators of good quality should be used. To avoid deterioration of the performance in wet weather the following precautions should be observed: The poles should be spaced at least 5 feet from the radiating elements and treated to minimize moisture absorption. The halyards should be of either Nylon rope or oil treated hemp. All insulators should have good wet weather characteristics.

The poles may be oil treated or given several coats of spar varnish. Aluminum paint or other paint with metallic pigment should be avoided because of the strong field in which the poles are located. Wire halyards should be avoided as they will detune the radiators.

It should be kept in mind that a vertically polarized antenna of this type delivers a strong surface wave which may cause interference to nearby broadcast receivers unless the transmitter power is low.

THE "BOBTAIL" BIDIRECTIONAL BROADSIDE CURTAIN

Another application of vertical orientation of the radiating elements of an array in order to obtain low-angle radiation at the lower end of the h-f range with low pole heights is illustrated in figure 5-34. When precut to the specified dimensions this single pattern array will perform well over the 7-Mc. amateur band or the 4-Mc. amateur phone band. For the 4-Mc. band the required two poles need be only 70 feet high, and the array will provide a practical signal gain averaging from 7 to 10 db over a horizontal half-wave dipole utilizing the same pole height when the path length exceeds 2500 miles.

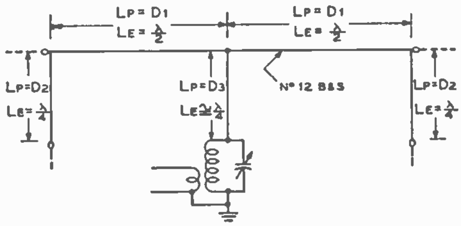
The horizontal directivity is only moderate, the beam width at the half power points being slightly greater than that obtained from three cophased vertical radiators fed with equal currents. This is explained by the fact that the current in each of the two outer radiators of this array carries only about half as much current as the center, driven element. While this "binomial" current distribution suppresses the end-fire lobe that occurs when an odd number of parallel radiators with half-wave spacing are fed equal currents, the array still exhibits some high-angle radiation and response off the ends as a result of imperfect cancellation in the flat top portion. This is not sufficient to affect the power gain appreciably, but does degrade the discrimination somewhat.

A moderate amount of sag can be tolerated at the center of the flat top, where it connects to the driven vertical element. The poles and antenna tank should be so located with respect to each other that the driven vertical element drops approximately straight down from the flat top.

Normally the antenna tank will be located in the same room as the transmitter, to facilitate adjustment when changing freqency. In this case it is recommended that the link coupled tank be located across the room from the transmitter if much power is used, in order to minimize r-f feedback difficulties which might occur as a result of the asymmetrical high impedance feed. If tuning of the antenna tank from the transmitter position is desired, flexible shafting can be run from the antenna tank condenser to a control knob at the transmitter.

The lower end of the driven element is quite "hot" if much power is used, and the lead-in insulator should be chosen with this in mind. The ground connection need not have very low resistance, as the current flowing in the ground connection is comparatively small. A stake or pipe driven a few feet in the ground will suffice. However, the ground lead should be of heavy wire and preferably the length should not exceed about 10 feet at 7 Mc. or about 20 feet at 4 Mc. in order to minimize reactive effects due to its inductance. If it is impossible to obtain this short a ground lead, a piece of screen or metal sheet about four feet square may be placed parallel to the earth in a convenient location and used as an artificial ground. A fairly high C/L ratio ordinarily will be required in the antenna tank in order to obtain adequate coupling and loading.

As with all vertically polarized h-f antenna systems, a strong surface wave is radiated that tends to aggravate broadcast interference, which may be troublesome if the transmitter power is high and broadcast receivers are located nearby.



	7150 KC.	3900 KC.
D1	68.5'	126'
D2	33'	60'
D3	30' TO 36'	54' TO 66'

Figure 5-34.

"BOBTAIL" BIDIRECTIONAL BROADSIDE CURTAIN FOR THE 7-MC. OR THE 4.0-MC. AMATEUR BANDS.

This simple vertically polarized array provides low angle radiation and response with comparatively low pole heights, and is very effective for dx work on the 7-Mc. band or the 4.0-Mc. phone band. Because of the phase relationships, radiation from the horizontal portion of the antenna is effectively supressed. Very little current flows in the ground lead to the coupling tank; so an elaborate ground system is not required, and the length of the ground lead is not critical so long as it uses heavy wire and is reasonably short.

LONG WIRE ANTENNAS

If all the dipoles in a row of driven, parallel (non-staggered), end-fire dipoles with half wavelength spacing are rotated simultaneously on their centers through 90 degrees in a plane containing the dipoles, so that the dipoles form a straight line, several pertinent effects are produced. First it will be noted that the "plus" end of one dipole is adjacent to the "plus" end of the next dipole, so that the dipoles may all be connected together. Individual feed of each dipole no longer is necessary in order to maintain the end-fire phase relationship; the articulated configuration permits feed of the system at any current loop or at one end of the string. We now have a simple "harmonic radiator," or "long wire antenna."

Next it will be noted that the aspect directivity and interference directivity no longer aid, but instead oppose each other, so that in the direction in which the interference directivity gain is maximum, the aspect directivity gain is zero, and vice versa. As the overall free space directivity pattern of the array is determined by the product of the two, maximum radiation occurs neither broadside to the wire nor in line with the wire, but at some oblique angle of revolution. It is for this reason that long wire radiators, used either alone or as elements of an array, often are called *tilted* long wires.

As the length of the wire in wavelengths is increased, the interference directivity is increased, but the aspect directivity remains the same regardless of the length of the wire. This means that as the length of radiator is increased, the interference directivity becomes more dominant, and maximum radiation occurs at a smaller angle with respect to the line of the radiator.

Because of the axial and coplanar symmetry of an unterminated straight conductor in free space, the lobes which constitute the directivity pattern are figures of revolution about the axis of the wire when the loop current magnitude is uniform, and a cross section of these lobes taken in a plane containing the wire also will be symmetrical about the mid point of the wire. This results in the familiar major lobe "cloverleaf" when the cross sectional pattern is represented by a conventional polar diagram.

A single long wire which is unterminated always has two identical major lobes, each forming a figure of revolution about the wire, and both lobes occur at the same angle with respect to the wire. If lines are drawn through all directions of maximum radiation, they will form two concentric cones, apex to apex. In actual practice the loop current magnitude will not be uniform, but rather will "taper off" with distance from the feed point. However, to simplify the present discussion of the basic properties of the long wire radiator, uniform loop current is assumed. The effect of non-uniform loop current upon the directivity pattern will be treated in a later section.

Just as in the case of a row of parallel dipoles, minor lobes are present, and the number increases with the length of the wire. The total number of lobes, major and minor, is equal to the number of half waves on the wire. The relative average amplitude of the minor lobes with respect to the main lobes decreases as the length of the radiator (and number of minor lobes) is increased. In the case of a long radiator having numerous minor lobes, the amplitude of the various minor lobes increases with the angle the minor lobe makes with a plane normal to the radiator.

As the length of the radiator is increased and the angle which the main lobes make with the wire decreases, the two main lobes become sharper, thus providing increased power gain. However, the directivity gain does not become pronounced until the wire is made several wavelengths long.

The relative amplitude of the minor lobes

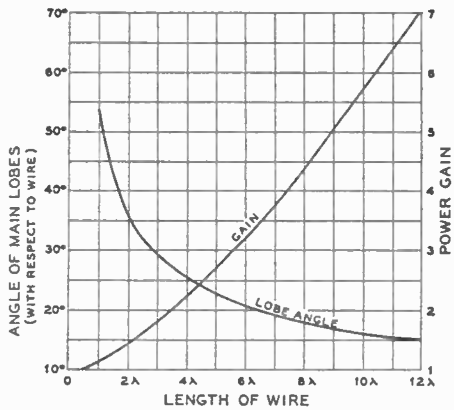


Figure 5-35.

FREE SPACE DIRECTIVITY AND GAIN CHARACTERISTICS OF A HYPOTHETICAL LONG WIRE RADIATOR HAVING UNIFORM LOOP CURRENT.

The lobe angle refers to the angle each of the two main lobes makes with the axis of the wire, and the gain refers to the free space directivity gain over a half-wave dipole. In the case of an actual antenna fed at one end or at a current loop (so as to achieve the desired phase relationships), the curves are only approximately correct. This is explained by the slightly tapered loop current distribution which results from articulated feed of the elements. (All interior half-wave sections act at the same time as radiator and feed line.)

is greater for a given end-fire array length in the case of a single long wire than it is in the case of parallel coplanar dipoles, because in the case of the long wire the aspect directivity tends to favor the minor lobes.

Figure 5-35 gives the pertinent directivity and gain characteristics of a long wire radiator in free space. For the longer lengths the curves apply only when the wire is pulled very taut. A sag at the center exceeding a small fraction of a wavelength will appreciably reduce the gain and the sharpness of the main lobes of a radiator 10 to 12 wavelengths long.

THE SINGLE LONG WIRE RADIATOR

A single long wire radiator is the simplest form of directional antenna, if we define "directional" as "having considerably more directivity than a half-wave dipole." However, the performance of a single long wire does not begin to compare with the performance of certain arrays to be described which utilize a combination of long wires, and the single long wire ordinarily is employed only for reasons of available space or convenience.

The radiation pattern and practical dx signal gain of a long wire radiator over typical soil bear only slight resemblance to the free space directivity pattern and gain. Furthermore, they usually are considerably different from the pattern and signal gain that would be obtained with the radiator over hypothetical perfect ground. The azimuthal directivity varies widely with the vertical angle involved, and calculation of the pattern obtained with the antenna over typical ground is complicated by the fact that directly off the ends the radiation is vertically polarized, broadside it is horizontally polarized, and at intermediate angles it has both a vertical and horizontal component.

In actual practice it will be noted that at frequencies below about 10 Mc. and with a radiator between two and three wavelengths long, fairly good coverage will be obtained in nearly all compass directions. This is explained by the fact that the minor lobes have appreciable amplitude, and by the fact that there will be appreciable radiation at *some vertical angle* in nearly all compass directions when the reflecting properties of typical earth are considered.

When the radiator is more than four wavelengths long the practical directivity begins to be pronounced, especially at the higher frequencies where only low-angle radiation is effective. While the conventional "free-space" polar diagram of such an antenna shows a four main-lobe cloverleaf with no radiation directly off the ends, in actual practice the antenna will be found broadly directional off the ends, with the azimuthal direction(s) of maximum effectiveness varying considerably with the wave angle, the elevation of the antenna, and the ground constants.

A long wire makes an effective means of obtaining low angle radiation at the lower end of the high-frequency range (3 to 6 Mc.) without resorting to high poles. With a radiator having a length of three or more wavelengths, the performance of the antenna in directions approximately in line with the radiator will be very good at low wave angles even with antenna elevations on the order of 1/8 to 1/4 wavelength (particularly over poor soil), and the 2500 mile performance of a low antenna three wavelengths long will be comparable to that of a broadside horizontal half-wave dipole elevated one wavelength above earth.

Unlike a string of cophased colinear dipoles, the radiation resistance of a long wire goes up rather slowly with overall length, and increases but little as the length is increased from four to twelve wavelengths.

An end-fed long wire can be made unidirectional by terminating it in its characteristic impedence by means of a non-inductive resistor connected from the far end to ground. At one time such an antenna was widely employed at medium and low frequencies for ground wave work, particularly as a receiving antenna. The antenna is quite effective under these conditions when used over poor soil, as a result of the forward tilt of a medium or low frequency surface wave when travelling over poor soil. (See Chapter 1.) An antenna of this type is called a *Beverage* or *wave* antenna.

Effect of Feed Point Location Upon the Pattern of a Long Wire Radiator

It is important that the reader appreciate the effect of the feed point location upon the radiation pattern of an unterminated long wire. To operate as a true long wire radiator, the antenna must be fed *either at one end or at a current loop* when conventional methods of feed are employed. And the pattern will be *roughly* the same regardless of which current loop is utilized as the feed point.

That the pattern does not remain exactly the same regardless of which current loop is fed is explained by the articulated feed inherent in a long wire system, whereby one element not only acts as a radiator but also serves as a feeder to the adjacent radiator. This results in a tapered current distribution (the current at each current loop becoming slightly less with increasing distance from the feed point). As the feed point on a long wire radiator is moved toward one end, the pattern is increasingly distorted, until with end (voltage) feed, or with current feed of the end section, the lobe directions are shifted noticeably from theoretical. The radiation off the unfed end is from 2 to 4 db. stronger than that off the fed end, the former being slightly greater than theoretical and the latter slightly less than theoretical. ("Theoretical" refers to the direction and amplitude the lobes would have if the current were the same at all current loops.)

If the antenna is opened at a *voltage* loop (an integral number of electrical half wavelengths from one end) to permit use of a symmetrical feed line, the currents in that portion of the antenna to one side of the feed point will be reversed with respect to those in the portion to the other side of the feed point, as compared to phase relationships existing under conditions of "free oscillation," current feed, or end feed.

This will affect the directivity pattern of the antenna, the effect being greatest when the high impedance feed point is at or near the center of the radiator rather than near one end. A wire four wavelengths long fed at the voltage loop in the exact center will be in effect, two, two-wavelength antennas fed in phase, and the free-space directivity pattern will be characterized by the main lobes occuring at a greater angle with respect to the wire.

If a long wire is voltage fed at a voltage loop which is off center, not only will the phase relationships be altered from the natural or "free oscillation" condition, but also there will be dissymmetry of the pattern due to the asymmetrical feed, as previously noted under the discussion of current feed. The latter does not seriously shift the direction of the major lobes, but does cause the lobe towards the "long side" of the radiator to be slightly stronger.

When the antenna is not an exact integral number of electrical half wavelengths long and is fed in the center, the feed point will be neither a current loop nor a voltage loop, but intermediate between the two. Likewise the directivity pattern will be intermediate between those obtained with the wire fed exactly at a current loop and with the wire fed exactly at a voltage loop.

When a center-fed long wire radiator is on the order of 10 or 12 wavelengths long, only a moderate frequency change (about 5 per cent) is required to change the feed point from a current loop to a voltage loop or vice versa. Thus it is seen that while the use of center feed with a resonant feed line permits operation of a long wire radiator over a very wide frequency range, the directivity pattern will be altered considerably for a frequency change of only a few per cent when the radiator is several wavelengths long.

A quantitative analysis of the effect of the feed point location upon the pattern and gain of a series-fed, *voltage*-fed long wire radiator of fixed length, and the effect of variations in length or frequency upon the pattern and gain of a center fed long wire radiator is quite complex and not within the scope of this book. However, the foregoing simplified qualitative discussion should enable the reader to appreciate the considerations involved.

While it is possible to "voltage feed" a long wire antenna at an intermediate voltage loop without breaking the wire, and thus avoid upsetting the natural phase relations, this method of feed has many disadvantages and therefore has not been considered in this discussion as a practical method of feeding a long wire.

Tilting a long wire radiator *in a vertical* *plane* has a marked effect upon the radiation pattern above typical soil, but the modification of the pattern as a result of the tilt does not lend itself to simple treament. The effect upon the pattern depends upon the angle of inclination, the average elevation of the radiator, the frequency, the length of the radiator in wavelengths, the ground characteristics, and other factors.

When a long wire radiator has a length exceeding 3 wavelengths, it may be end fed simply by bringing one end of the radiator down to the transmitter or antenna tuning unit, as the relatively short length of the downlead as compared to the overall length of the wire will result in but little distortion of the pattern and negligible reduction in gain, and the antenna may be used over a continuous, wide frequency range without upsetting the natural phase relationships on the wire. Certain disadvantages are inherent in such a simple arrangement, however, Pickup by the downlead will deteriorate the discrimination and noise reducing properties of the antenna when it is used for receiving, and the lack of symmetry may cause considerable trouble with "r-f to a-f feedback" in a high powered radiophone transmitter at the upper end of the h-f range if a low level microphone is employed.

There are many configurations of tilted long wire radiators used in combination which will provide better gain and directivity characteristics than a single long wire. One arrangement utilizes two staggered, parallel long wires excited out of phase, with the stagger and spacing so adjusted as to reinforce the radiation along a bidirectional line and more or less cancel in other directions. Such combinations are referred to as "staggered long wires" or "echelon antennas." When the tilted parallel wires lie in a vertical plane, at an angle with the earth, the array is known specifically as the RCA model B array or "projector." When the tilted parallel wires lie in a horizontal plane, at an oblique angle to the desired direction of maximum radiation, the array is known specifically as the RCA model C array or projector.

Long wire arrays of the echelon type will not be treated in detail because their performance is inferior to that of Model C projector, or "V Beam," to be described immediately following, and because echelon type arrays no longer are widely used.

THE "V" BEAM ANTENNA (RCA MODEL D ARRAY)

Two harmonic wires arranged in the form of a "V" will, when the apex angle is optimum, provide good directivity and gain characteristics. The two wires are fed at the apex of the V by means of a balanced line or transformer, so that equal currents of opposite phase flow in corresponding parts of the two wires. The apex angle is so chosen that the main lobes of the two wires reinforce along the bisector of the apex angle, and tend to more or less cancel in other directions.

When the legs are several wavelengths long, the optimum apex angle is approximately twice the angle the main lobes of one leg make with that leg when operated as a long wire radiator. For unterminated wires this angle is given in figure 5-36. Thus it is seen that for an unterminated V with legs 5 wavelengths long, the main lobe angle of each leg is 22.5 degrees; so the apex angle of the V should be made 45 degrees.

As the lobe angle changes rather slowly with length for a harmonic wire longer than approximately 5 wavelengths, and because the optimum apex angle of a "V" array is not especially critical percentage wise, a "V" array can be used with good results over a rather wide frequency range when the leg lengths are long, simply by resorting to tuned (resonant) feeders. However, the feed point impedance varies widely with frequency, and a highly flexible antenna coupling unit is required for wide-band operation (due to the high average SWR).

A "V" array will provide worthwhile gain and directivity when the leg lengths are as short as two wavelengths, and some gain when the legs are only one wavelength long. However, the optimum apex angle for such an array is slightly less than twice the lobe angle of the individual legs. Also, besides having less gain than a "V" array having longer legs, an array having comparatively short legs cannot be used with good results over as wide a frequency range.

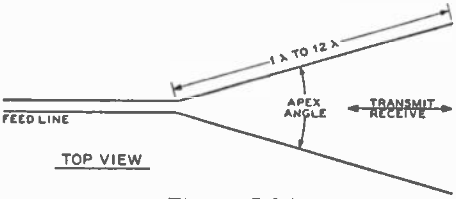


Figure 5-36.

THE BASIC "V" ARRAY (UNTERMINATED, BIDIRECTIONAL).

This array, used singly or in combination with similar units, is the most popular type of unterminated long wire antenna. When a tuned feed line is employed, the array may be used over a wide frequency range without serious deterioration of the performance. For leg lengths exceeding 3 wavelengths, the optimum apex angle is twice the main lobe angle of one leg used alone, as determined from figure 5-35. While the array is considered bidirectional, the lobe off the open end of the "V" has a slightly greater amplitude than the lobe off the apex end.

Best results will be obtained from a "V" array one wavelength on a leg when the apex angle is made approximately 90 degrees, rather than 108 degrees as obtained by doubling the lobe angle obtained from figure 5-35. For legs two wavelength long the apex angle can be made 65 degrees. For legs three wavelengths long (or longer) the apex angle may be made equal to twice the lobe angle.

A simple form of "V" array is illustrated schematically in figure 5-36. It is also known as the RCA model D "basic unit," and may be used either alone or in combination with similar units. A flat line can be achieved by incorporating one or more matching stubs, but this complicates matters when the array is to be used on more than one frequency.

The free space directivity gain of a properly designed "V" array of the type shown in figure 5-36 is approximately 6 db greater than the gain of a single leg operated alone when the leg length exceeds two wavelengths. Reinforcement of the overlapping main lobes along the apex bisector is responsible for approximately 3 db gain over one leg, while suppression of radiation in other directions raises it approximately 3 db more. The increased gain over a single wire twice as long as the one leg may also be explained on the basis of mutual impedances between the two legs of the "V." Thus it is seen from figure 5-35 that the free space directivity gain of a simple "V" array 8 wavelengths on a leg is approximately 12 db over a half-wave dipole, or approximately 3 db over a single long wire having a length of 16 wavelengths.

In actual practice the radiation off the open end of a "V" antenna is slightly stronger than that off of the apex end, for reasons previously explained under the discussion of the basic long wire radiator. However, as the difference in amplitude is not great, especially when the leg length does not exceed three wavelengths, the "V" is classed as a "bidirectional" array.

For the highest practical dx signal gain in the h-f range, a "V" array should be located over flat, level ground, and elevated at least one wavelength. However, at the lower end of the h-f range (below approximately 10 Mc.), good results will be obtained at elevations as low as 0.4 wavelength.

The radiation from a "V" array directly along the line of the apex bisector is horizontally polarized, because the vertical component from one wire cancels that of the other at all vertical angles. In other azimuthal directions, however, the radiation has both horizontal and vertical components.

A simple "V" array of the type shown in figure 5-36 has considerably more horizontal directivity than vertical directivity, which is fortunate. It permits the use of high azimuthal directivity without incurring excessive vertical directivity. When for some reason it is desirable to broaden the vertical pattern without increasing the azimuthal beam width, it can be done by inclining the plane of the array a few degrees so that the apex of the "V" is lower or higher than the open ends of the legs.

Variations within the range of typical ground constants have little effect upon the vertical pattern along the line of the apex bisector at vertical angles below 20 degrees. However, they do effect the amount of heat loss incurred as a result of ground currents.

Under certain conditions a substantial portion of the power fed to a "V" array with long legs may be dissipated in the ground below the array. This is explained by the fact that the radiation resistance of a "V" array is comparable to that of a half-wave dipole, which means that for a given total power delivered to the antenna, about as much current flows at each of the many current loops in a "V" array as flows in a single halfwave dipole. The appreciable ground loss is characteristic of all long wire arrays in which the overall wire length totals many wavelengths, and it causes the practical signal gain to be slightly less than would be indicated on the basis of directivity gain alone.

Basic "V" units of the type illustrated in figure 5-36 usually are employed commercially in various combinations of two or more which provide better discrimination for a given gain and directivity (by reducing the relative amplitude of the minor lobes). The use of two or more basic units also makes it possible to increase the vertical directivity without increasing the horizontal directivity, and vice versa. Common configurations are one "V" over an idential one (forming a "stack"), and two "V's" or stacks placed side by side in the form of a W.

The practice of spacing two identical bidirectional arrays one quarter wavelength along the line of maximum radiation and providing quadrature excitation to obtain a undirectional pattern (and additional gain) is applicable to the "V" array, but limits the frequency range over which it may be used, particularly if a parasitically excited reflector is used. By driving both units with separate lines, the direction of the beam may be shifted 180 degrees simply by reversing the polarity of one feed line. This arrangement requires stub matching at each array so as to provide two non-resonant lines to the transmitter, one of which is made an electrical quarter wavelength longer than the other to obtain a quadrature relationship at the antenna feed points.

It is also possible to obtain a undirectional characteristic by terminating the far end of each leg with a noninductive resistor to ground or to a ground screen. However, it is difficult to obtain a good termination in this manner using economical construction, and therefore such an arrangement seldom is employed. The rhombic antenna to be described immediately following makes a much more satisfactory undirectional antenna of the terminated type.

It is important that the wires of a "V" array be pulled very taut. This predicates the use of copper clad or galvanized steel wire for long spans. When a resonant line is used in order to permit operation of the array over a wide frequency range, a line constructed of no. 12 or no. 10 B&S gauge wire spaced 6 inches is satisfactory for transmitter outputs up to 1 kw., amplitude modulated. Higher powers call for greater spacing, thus limiting the transmitter power that can be handled by the tuned line at the upper end of the h-f range (due to excessive line radiation with wide spaced feeders at these frequencies).

Generally speaking, the discrimination (rejection of noise and undesired signals) for a given power gain is not so good with a single unit "V" array as with a broadside curtain. The same applies to interference caused to stations well off the direction of the "beam." When these considerations are of major importance, it is the usual practice to use *combinations* of "V" units which give comparable main lobe directivity and gain but lower amplitude of the minor lobes.

THE RHOMBIC ANTENNA

One of the most widely used h-f directional antennas in military and commercial service for point to point work consists of four, long-wire radiator elements disposed in the shape of a rhombus, from which the antenna gets its name. This antenna is simple to construct, has high gain, is unidirectional, and, most important, provides excellent performance over a broad, continuous frequency range. In fact, it is the latter characteristic, its ability to provide excellent performance over a frequency range of 2 to 1 and good performance over a range as great as 4 to 1, that is responsible for the wide popularity of the rhombic antenna in spite of the fact that it takes up much more room than, and provides less

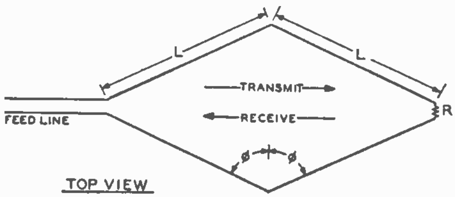


Figure 5-37.

SCHEMATIC OF THE BASIC HIGH-FREQUENCY RHOMBIC ANTENNA.

The pertinent design parameters are the side length, L; the tilt angle, ϕ; the elevation; and the value of the terminating resistor, R. When these are optimized the antenna is unidirectional, with high gain and good discrimination.

discrimination than, a broadside curtain and reflector combination having comparable gain. A further advantage of the rhombic antenna is the fact that it is not critical as to dimensions or adjustment, this simply being a corollary expression of its lack of frequency sensitivity.

The basic rhombic antenna is shown schematically in figure 5-37. The important dimensions are the side length, L, and the tilt angle, ϕ. The latter is equal to one half the included side angle, and represents the number of degrees by which the legs are tilted from a *broadside* direction.

The function of the terminating resistor, R, is to make the antenna unidirectional and aperiodic. While a rhombic array may be used either vertically or horizontally, for practical reasons it always is oriented in a horizontal plane when used for high frequency work, and figure 5-37 therefore is a top view. With horizontal orientation the radiation along the line of maximum output is horizontally polarized. The vertical components tend to cancel and in a simplified treatment may justifiably be ignored.

Comparison With "V" Antenna

The rhombic antenna sometimes is described, on the basis of geometrical appearance, as two articulated "V" antennas connected back to back and terminated to give a unidirectional pattern. While the rhombic does superficially resemble a pair of articulated "V" antennas, the considerations involved in the functioning of the two types of array are such that the rhombic requires distinctive treatment.

For example, in a basic "V" unit in free space the tilt angle (the complement of half the apex angle) is determined only by alignment of the lobes of the individual legs along the line of the apex angle bisector. However, in the case of a rhombic antenna, the time and space phase relationships between the forward elements and the rear elements are involved, as well as the alignment of the lobes of the individual elements along a common direction. And the tilt angle which aligns the individual lobes on the apex angle bisector (for a given leg length) is not the tilt angle which produces the most favorable phase relationship bewteen the forward pair and the rear pair of elements for maximum gain. Therefore, the optimum tilt angle for maximum forward gain from a rhombic is a compromise between the angle producing exact alignment of the individual lobes, and the angle producing optimum phase addition of the fore and aft pairs along the line of the apex angle bisector.

Another notable difference between the "V" and rhombic is the difference which the presence of the ground has upon the directivity of the two types of array, and therefore upon the optimum tilt angle. Another distinction is the difference in the lobe angle of the individual legs for a given leg length. The latter is due to the presence of sinusoidal current on a "V" leg as contrasted with uniform distribution on a rhombic leg. However, the difference is not appreciable except for short leg lengths.

Rhombic Antenna Design Parameters

The presence of the earth has considerable effect upon the directivity pattern of a rhombic array, and the optimum side length, tilt angle and elevation of a high frequency rhombic are determined by the optimum wave angle (vertical angle of wave arrival and departure). At low angles useful for long distance high-frequency communication, the effect of imperfect earth conductivity upon the directivity pattern is insignificant and may be ignored. Computations based upon the assumption of a perfectly

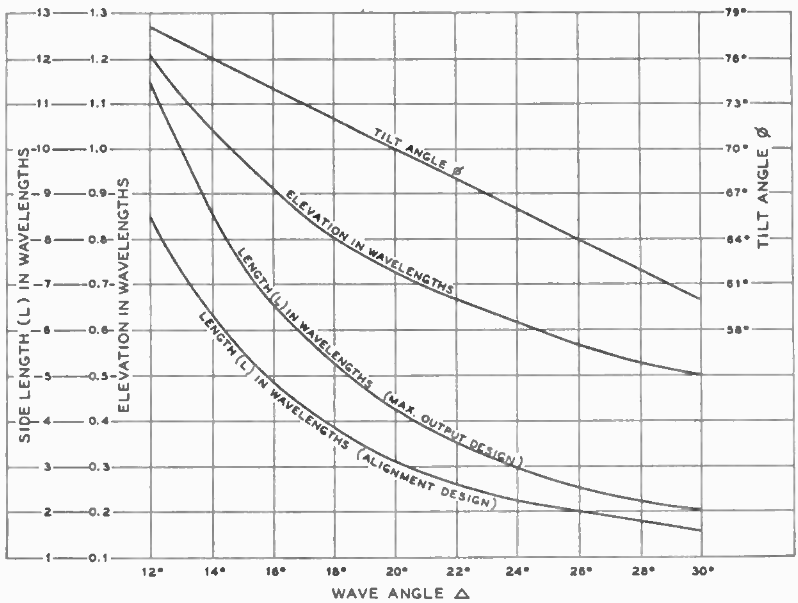


Figure 5-38.

RHOMBIC ANTENNA DESIGN CHART.

This chart gives the correct elevation, tilt angle, and side length as a function of wave angle for both a "maximum output" rhombic and an "alignment design" rhombic. The "alignment" type requires less side length but has slightly less gain. It is recommended when the available ground space is limited. It will be noted that the optimum elevation and tilt angle are the same for both types. The maximum output design method gives the maximum gain that can be obtained at a given wave angle from any rhombic antenna.

conducting earth have been found to check very closely with experimentally derived data on the pertinent portions of the directivity pattern. While the earth losses due to imperfect ground may cause the signal gain of a rhombic antenna to be 1 or 2 db less than that which would be obtained over perfectly conducting earth, the *shape* of the directivity pattern is not seriously altered at the lower angles of elevation by the finite conductivity of typical earth.

The high degree of vertical directivity obtained with a rhombic antenna makes it necessary to design the antenna *for a specific wave angle.* And for a particular wave angle and frequency there is an optimum combination of elevation, side length, and tilt angle which will give the maximum possible amount of radiation and response. It is important that the reader appreciate the fact that for a particular wave angle there is an optimum side length; *increasing the side length beyond the optimum length will result in reduced radiation and response at that wave angle,* regardless of what alterations may be made to the elevation and tilt angle.

Maximum Output Design

A rhombic antenna designed to give the maximum possible radiation and response at a particular wave angle and frequency is called a *maximum output* rhombic. The required tilt angle and the required elevation and side length in wavelengths for a maximum output rhombic are given in figure 5-38 as a function of wave angle. As an example of how the chart is to be used, suppose we wish to determine the dimensions of a maximum output rhombic for an 18 degree wave angle. It is done as follows:

Read up from 18 degrees to the intersection with the length curve marked "Max. Output Design," then left to the scale marked "Side Length in Wavelengths." This will give the proper side length (5.25 wavelengths). Then read up from 18 degrees to the intersection with the curve marked "Elevation," then left to the corresponding scale. This will give the proper elevation (0.8 wavelength). Next read up from 18 degrees to the intersection with the curve marked "Tilt Angle," and right to the corresponding scale. This will give the proper tilt angle (72 degrees).

An interesting fact about a maximum output rhombic is that the vertical angle of maximum transmission and response (or "nose" of the vertical lobe) is *slightly below the wave angle used in the design* computations. A rhombic designed by this method for a wave angle of 18 degrees will exhibit a vertical lobe which maximizes at about 16 degrees. Thus, while this rhombic gives the greatest possible output and response obtainable from any rhombic at a wave angle of 18 degrees, it actually exhibits slightly greater output and response at 16 degrees. Any attempt to beat the game by designing the rhombic for a wave angle slightly higher than that actually desired is futile, because the gain of a maximum output rhombic falls off as the design wave angle is increased, and at such a rate that there is no net gain by "fudging" the design angle.

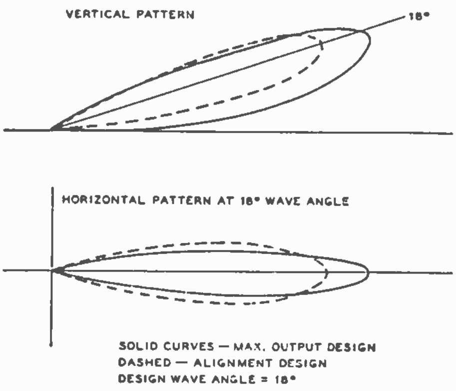


Figure 5-39.

COMPARISON OF DIRECTIVITY OF A "MAXIMUM OUTPUT" RHOMBIC WITH THAT OF AN "ALIGNMENT METHOD" RHOMBIC, BOTH DESIGNED FOR THE SAME WAVE ANGLE OF 18 DEGREES.

The two antennas have the same elevation and tilt angle, but the side length of the alignment rhombic is only 74 per cent of the side length of the maximum output rhombic. It is interesting to note that while maximum gain from a maximum output rhombic occurs at a vertical angle slightly below the design wave angle, nothing can be done to increase the gain at the design wave angle, because the gain of a maximum output rhombic falls off as the design wave angle is increased.

The elevation of a maximum output rhombic above ground is such that if an isotropic radiator were substituted for the rhombic, interference due to ground reflection would put the nose of the lowest lobe right on the design wave angle. However, the free space directivity of a maximum output rhombic is such that if the ground were perfectly absorbing, the vertical angle of the lobe nose would be somewhat below the design wave angle. The result in the case of a maximum output rhombic over reflecting ground is a composite directivity pattern which puts the nose of the lobe between these two angles.

From the foregoing discussion it is apparent that the gain of a maximum output rhombic increases as the design wave angle is decreased. However, there is a practical lower limit to the design wave angle, as the vertical and horizontal directivity both become excessive for practical applications when too much gain is attempted. The vertical and horizontal directivity (in terms of lobe width) of a maximum output rhombic are comparable, though the horizontal directivity is symmetrical while the vertical directivity is not. (Refer to the solid curves of figure 5-39.)

The magnitude of the variations in wave angle and of the deviation from a great circle path ordinarily encountered depends somewhat upon the particular geographical location involved. For instance, much greater azimuthal deviation will be observed when the path passes near the North magnetic pole than on an east-west circuit near the equator. Because of these factors, present commercial practice is to observe a maximum leg length of from 5 to 8 wavelengths for a high frequency "maximum output" rhombic, depending upon the geographical location and bearing.

Alignment Design

When the side length of a maximum output rhombic is reduced to approximately 74 per cent of the value giving maximum output, only a slight reduction in gain occurs (usually 1 to 2 db). This arrangement aligns both the space directivity of the rhombic and the ground reflection factor upon the wave angle, giving a somewhat sharper vertical pattern and broader horizontal pattern. A comparison of the patterns obtained using the two design methods is given in figure 5-39 for a wave angle of 18 degrees. The alignment method of design is the logical one to use when sufficient pole height is available for optimum height but the available ground space is not quite sufficient for a maximum output rhombic.

Compromise Designs

Considerable flexibility exists with regard to the design parameters. For instance, when the available pole height is not sufficient for optimum elevation, the reduction in elevation can be largely compensated for by reducing the tilt angle or increasing the side length. Reducing the tilt angle is not as effective as increasing the side length, but is a useful expedient when both the pole height and available land are limited. When the elevation is at least half the optimum value for the desired wave angle, a gain closely approaching that of a maximum output rhombic can be obtained by increasing the side length. The computation of an optimum compromise design under conditions of limited pole height is rather complicated and not within the scope of this book, but fortunately the design parameters are not critical and good results can be obtained by means of rough approximations based upon the foregoing data.

Wide-Band Operation

When a rhombic is to be used over a wide band of frequencies, the optimum dimensions depend upon such factors as whether the ground space or pole height or both are limited, and whether the application justifies a sacrifice in high-frequency performance in order to obtain improved performance at the low-frequency end of the range. To obtain a slight improvement at the low frequency end, considerable reduction in performance must be accepted at the high end.

The computations involved in designing an "optimum" rhombic for use over a specified wide frequency range become quite complex when all factors subject to compromise are considered, and are beyond the scope of this book. Usually the performance will be acceptable if the array simply is designed for optimum performance at a single frequency midway between the highest and lowest frequencies to be employed. This design frequency should be the arithmetical mean, rather than the geometrical mean, for best all around performance. Such an array will work very well over a frequency range of 2 to 1, and will provide substantial gain over a range as great as 4 to 1. However, to obtain worthwhile gain at the lowest frequency the side length must be at least two wavelengths long at that frequency.

As the operating frequency is lowered, the vertical angle of maximum output from a given rhombic slowly increases. However, as somewhat higher angles are effective at the lower frequencies, this vertical angle variation in the directivity of a rhombic is not detrimental and in some cases actually is advantageous. The usual practice is to design a wide band rhombic for a wave angle which is optimum for either the upper end of the frequency range or the arithmetical mean frequency, and in either case to let the vertical pattern take care of itself throughout the rest of the range.

The Rhombic Termination

As the unidirectional and aperiodic features of a rhombic antenna are determined by its termination, it is obvious that a rhombic must be correctly terminated in order to provide good performance. When the termination is correct, the front-to-back ratio of the antenna will be high and the input impedance will be substantially constant

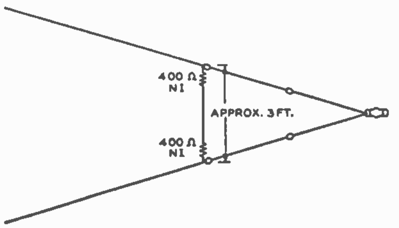


Figure 5-40.

RECOMMENDED METHOD OF TERMINATING A RHOMBIC WHEN LUMPED RESISTANCE IS USED.

A rhombic antenna terminated with lumped resistance behaves as though the terminated end were shunted with a small capacitive reactance. The above arrangement provides a more favorable termination, not only by minimizing the shunt capacitive reactance, but by neutralizing part of this shunt reactance with the inductive reactance of the jumper between the two terminating resistors.

over a wide frequency range. Usually the correct termination resistance for maximum uniformity of the feed point impedance will be on the order of 800 ohms, and when front-to-back ratio is not too important, it is permissible to design the antenna with an 800-ohm termination. When front-to-back ratio is of prime importance, it is desirable to determine the optimum value of terminating resistance experimentally, as the value is not necessarily the same as that giving the best impedance-frequency characteristic and is much more critical.

While it is possible to obtain almost complete suppression in the back direction by adjustment of the terminating resistance, the adjustment will not hold as the frequency is varied. Because the rhombic usually is employed over a substantial frequency range, the usual procedure therefore is to adjust the terminating resistor for maximum *mean* front-to-back ratio when front-to-back ratio is a major consideration.

The terminating resistor for a transmitting rhombic usually dissipates between 30 and 45 per cent of the power fed to the antenna, making necessary a husky resistor when much transmitter power is used. Commercially available noninductive resistors of the plaque or carbon rod type are suitable for powers up to one kilowatt or so. A worthwhile improvement can be obtained by splitting and rigging the resistance as shown in figure 5-40. This minimizes the stray capacitive reactance at the end of the antenna and the residual reactance is more or less cancelled by the inductance of the jumper connecting the two resistors.

Another widely used type of termination consists of a lossy open-wire line of stainless steel or similar resistance wire, proportioned to have a characteristic impedance of from 600 to 700 ohms (depending upon the desired terminating impedance). The line is either shorted or open circuited at the far end, or else terminated in a low wattage noninductive resistor. This arrangement is especially suited to powers of 1 kilowatt and higher. However, the line must be several wavelengths long at the lowest operating frequency and have suitable attenuation characteristics in order to provide a termination having substantially uniform impedance, with negligible reactive component, over the required frequency range.

In some commercial installations a dissipative line of limited attenuation is brought into the station. By means of a variable termination which permits control over the reactive and resistive components of the termination it is possible to improve the front-to-back ratio, or to effect a slight change in the direction of an objectionable minor lobe, over the entire frequency range by optimizing the adjustment each time the operating frequency is changed.

It is possible to avoid wasting the power dissipated in the terminating resistor (representing approximately 2 db) by using a reentrant circuit in which the rhombic is made to terminate itself. The power that ordinarily would be dissipated in the terminating resistor is fed back into the input end. Because the circuit for accomplishing this is rather complicated (as compared to a simple rhombic) and restricts operation to a single frequency, it is not widely used and therefore will not be described in detail.

Multiple-Wire (Space Tapered) Elements

The feed point impedance of a rhombic antenna can be lowered to a more conven-

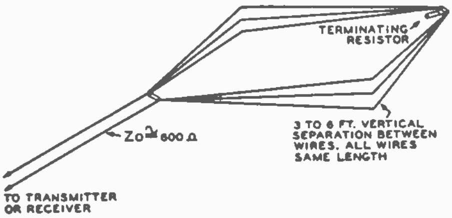


Figure 5-41.

ILLUSTRATING CONFIGURATION OF A MULTIPLE-WIRE RHOMBIC.

The use of two or three space-tapered wires arranged as above improves the impedance-frequency characteristic and lowers the impedance to a value more suitable for direct connection of a two-wire open line. Also, when the leg lengths are long, the gain is slightly greater than that of a single wire rhombic.

ient magnitude and made more uniform over a wide frequency range by the use of space-tapered multiple-wire elements as shown in the perspective schematic of figure 5-41. The fanning of the wires tends to compensate for the increased spacing between the two sides of the rhombus. It also has a beneficial effect upon the amplitude and phase relationships of the current, and oftentimes results in a signal gain of 1 to 2 db over a "single-wire" rhombic of the same dimensions when the leg lengths are long.

While three wires seem to be about optimum, almost as much improvement can be realized by the use of two space-tapered wires. Two are considerably simpler to rig than three, as a spreader can be used at each apex and a halyard connected to the mid point of the spreader.

Usually a good match can be obtained with two space-tapered wires by direct connection of an open wire line consisting of no. 14 B&S spaced 6 inches or no. 12 B&S spaced 9 inches. With three space-tapered wires a good match usually can be obtained by direct connection of an open wire line consisting of no. 12 B&S spaced 6 inches, no. 10 B&S spaced 8 inches, no. 8 B&S spaced 10 inches, or no. 6 B&S spaced 12 inches. The heavier line construction is necessary only when high power is employed, as the line loss will be acceptable even with lines of light construction. This is explained by the fact that the line will have a very low SWR.

The use of multiple wire construction substantially reduces the intensity of precipitation static when a rhombic is used for receiving, an important consideration in certain geographical locations subject to frequent, heavy precipitation static.

Rhombic Feed Systems

Customary commercial practice is to use an open-wire line of 600-650 ohms surge impedance for feeding a transmitting rhombic, and in locations where local noise is not excessive, to use the same type of line for receiving. While a flat line can be obtained with a single-wire rhombic at a particular frequency by the use of a matching stub, the VSWR is tolerable when no such stub is employed, and single-wire rhombics often are used without an impedance transformer in order to simplify multifrequency operation. When a multiple-wire rhombic is employed, direct connection of a 600 ohm line results in a very low SWR over the useful range of the antenna.

In locations of high ambient noise, good engineering practice calls for the use of a coaxial line or shielded pair. Highly engineered impedance matching devices which also serve as line balance converters have been developed which are so proportioned that they provide an excellent match to a coaxial line over a wide frequency range even when a single-wire rhombic is used, the coupling network being designed to compensate for the change in impedance with frequency which is characteristic of a single-wire rhombic antenna.

Rhombic Antenna Gain

With the usual types of phased dipole arrays, such as a broadside curtain having only moderate vertical directivity, the practical dx signal gain (referred to a half-wave dipole at the same elevation) is comparable to the free space directivity gain. However, in the case of a rhombic antenna such is not always the case. This is explained by the high order of vertical directivity realized with the longer leg lengths, by the waste

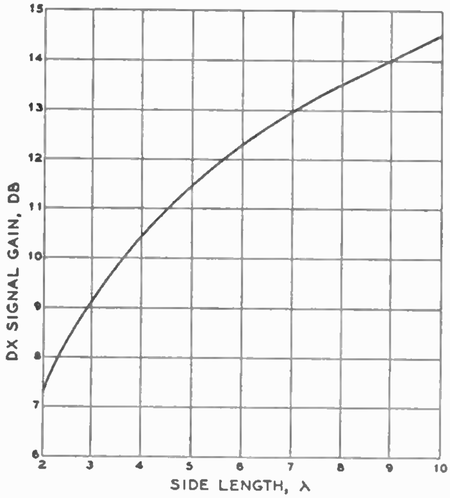


Figure 5-42.

PRACTICAL DX SIGNAL GAIN (APPROXIMATE) AS A FUNCTION OF SIDE LENGTH OF A MAXIMUM OUTPUT OR ALIGNMENT METHOD MULTIPLE-WIRE RHOMBIC DESIGNED IN ACCORDANCE WITH FIGURE 5-38.

The gain is referred to a half-wave dipole at the same elevation. It is assumed that the rhombic is favorably sited and well constructed, working at its design frequency, and that the most effective wave angle is the design wave angle. Because the most effective angle varies with ionosphere conditions, the gain curve represents a mean or average, and is only approximate. For the longer side lengths, the gain of a simple single-wire rhombic usually is from 0.5 to 1.5 db less than the gain of a similar multiple-wire rhombic.

of power dissipated in the terminating resistor, and by the appreciable earth losses under and around a rhombic antenna.

As the directivity gain is of practical interest only as it relates to practical dx signal gain, the accompanying gain chart (figure 5-42) shows the practical dx signal gain over a half-wave dipole at the same elevation. However, it must be appreciated that the curve applies only if the wave angle of the antenna is the most effective vertical angle, and that it is an average and at best only an approximation. The curve can be applied to a rhombic designed to either of the length curves given in figure 5-38 (either a maximum output rhombic or an alignment method rhombic). Note that the gain is greater with a maximum output rhombic for a given wave angle because the side length is greater, but that the difference is only about 1.5 db. The curve of figure 5-42 applies only to the frequency for which the antenna was designed.

It will be found that in actual practice the signal gain of a rhombic with long legs and a very sharp vertical directivity pattern will vary from well above to well below the gain figure given by the chart as the ionosphere conditions change, but that if the rhombic is designed for the optimum wave angle the *average* gain will be close to the indicated figure.

The gain curve is based upon the assumption that the mechanical construction of the rhombic is in accordance with good engineering practice and that the array is favorably sited.

Siting and Construction Considerations

While a makeshift rhombic antenna working under the handicap of an unfavorable site often will give creditable performance, careful construction and observance of certain basic siting rules are necessary in order to realize the maximum performance of which the antenna is capable.

The ideal location for a high frequency rhombic is over flat, level ground which extends for considerable distance along the azimuthal direction of maximum output and response. Obstructions under or surrounding the rhombic will have a detrimental effect upon the gain, though one or two small trees or buildings within the area will not appreciably deteriorate the performance. The angular elevation of the horizon in front of the rhombic should not approach the design wave angle too closely, even though the hills or mountains may be several miles distant. The conductivity of the soil is of minor importance.

When the ground is flat but not level, and slopes at the same angle for considerable distance in front of the rhombic, the performance will be as good as that obtainable with a level site *provided* the antenna is slanted so that the plane of the array is parallel to the ground and the slope of the ground is taken into consideration in picking the design wave angle. Thus, if the rhombic slopes downward towards its front at an angle of 4 degrees, the design wave angle should be 4 degrees higher than would otherwise be the case.

In constructing a rhombic antenna it is desirable that stray reactances be kept to a minimum. This calls for a minimum of hardware, the use of insulators with low shunt capacity, and a spacing of at least several feet between each apex of the rhombus and its supporting pole.

Design Data for a 7 to 30 Mc. Amateur Rhombic

A simple, easily erected compromise rhombic intended for amateur use over the frequency range 7 to 30 Mc. is illustrated in figure 5-43. The termination preferably should be in accordance with figure 5-40, using two 400-ohm noninductive resistors capable of dissipating safely about half the power delivered to the antenna.

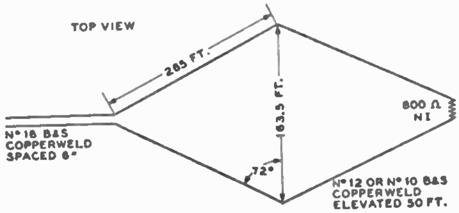


Figure 5-43.

SIMPLE, EASILY ERECTED COMPROMISE RHOMBIC FOR USE OVER THE RANGE 7 TO 30 MC.

The vertical directivity, horizontal directivity, and gain all increase with frequency, but nevertheless are quite good at the lower end of the range. The vertical angle of maximum radiation gradually increases as the frequency is decreased. However, as higher angles are effective at the lower frequencies, this is not detrimental. At no frequency within the range will the line VSWR exceed approximately 1.4 if the array is properly terminated. For details regarding the termination, refer to the accompanying text and figure 5-40.

CHAPTER SIX

V-H-F and U-H-F Antenna Systems

GENERAL CONSIDERATIONS; COMPARISON WITH H-F REQUIREMENTS AND PRACTICE

Many of the basic antenna types employed in the high frequency range are suitable for use in the v-h-f/u-h-f portion of the spectrum. However, the requirements for an antenna in the v-h-f/u-h-f range are somewhat different than for h-f use, and modifications often are desirable. Also, certain effects which are present to a minor degree in h-f systems become objectionable in the v-h-f/u-h-f range. For these reasons, antennas for the v-h-f/u-h-f range are treated separately.

As normal v-h-f/u-h-f propagation is via the space wave component of the ground wave, virtually unlimited vertical directivity can be employed to advantage in ground-to-ground work when the antenna is located on a stable platform. However, for vehicular, marine, and aircraft applications the amount of vertical directivity which can be used to advantage is limited.

When general coverage is desired, the horizontal directivity requirements are no different than for h-f work. However, for point-to-point work a much higher degree of horizontal directivity can be employed, because azimuthal deviations of the propagation path due to variations in the medium with time are insignificant.

Coupled with the fact that much greater vertical directivity, horizontal directivity or both usually can be employed to advantage in v-h-f/u-h-f applications is the fact that from a practical standpoint a high degree of directivity is much more easily obtained, due to the fortunate circumstance that the absolute dimensions of an antenna array of given configuration are inversely proportional to the design frequency.

Another pertinent difference between h-f and v-h-f/u-h-f requirements is the difference in the optimum practical antenna elevation under conditions of flat terrain. While antenna elevations exceeding one or two wavelengths seldom are necessary or even desirable for h-f sky-wave applications, v-h-f/u-h-f performance continues to improve with antenna elevation, up to the point where the transmission line becomes so long that the signal gain due to a further increase in antenna elevation is offset by the additional transmission line loss.

As the signal gain (in db) per foot of additional elevation goes down with increasing elevation, and as the transmission line loss (in db) per foot of nonresonant line is independent of line length, the antenna elevation (with respect to the transmitter or receiver) which gives maximum signal strength goes up as the transmission line attenuation goes down. However, the optimum elevation usually is limited by economical considerations, because in general the per-foot cost of line goes up as the attenuation goes down, and the per-foot pole cost goes up rapidly with increasing pole height.

Because the optimum wave angle for high-frequency sky-wave propagation is at some vertical angle which depends upon the frequency, path distance, and condition of the ionosphere, h-f directional antennas having considerable vertical directivity have a practical dx signal gain which may be considerably different from the free space directivity gain, even though the conductor losses and ground losses may be insignificant. The difference is greatest at certain critical heights above ground.

However, in the case of v-h-f/u-h-f antennas, the effective vertical angle for long distance communication always is approximately zero (with the possible exception of a very high flying aircraft at one end of the circuit). Also, well-designed v-h-f/u-h-f antennas always are highly efficient radiators, and always are elevated sufficiently to reduce earth losses and the effect of earth upon the antenna impedance to an insignificant value. Therefore, the practical dx signal gain of a v-h-f/u-h-f antenna over a reference antenna always is very nearly equal to the free space directivity gain.

While the best site for h-f use is a flat, level area which is free from obstructions for considerable distance (even when elevated sites are available), it is desirable in areas of hilly terrain to take advantage of terrain elevation when siting a v-h-f/u-h-f installation. When a well-elevated site is available, the pole height need be only sufficient to put the antenna well in the clear with regard to trees, buildings, and other surrounding obstructions, because a great increase in pole height will be required in order to effect a significant improvement in signal strength under these conditions.

Open-wire lines seldom are used in the v-h-f/u-h-f range, even though they are one of the most desirable types at the lower frequencies. Above approximately 50 Mc., line radiation becomes excessive when the usual spacings are employed. The usual procedure for v-h-f/u-h-f work is to use some type of low loss coaxial line, except for a few of the less exacting applications such as FM and television broadcast reception. For these applications 150 or 300-ohm Twin-Lead generally is employed for reasons of economy.

While vertical polarization is not often employed for high-frequency sky-wave work, it is standard for several applications in the v-h-f/u-h-f range because of the ease with which circular coverage can be obtained with a simple antenna and because the nulls in the vertical pattern are less pronounced. The first is an important consideration in mobile work, and the second is important in communicating with high flying aircraft. For point-to-point work or general broadcasting to stationary ground receivers, however, horizontal polarization has become standard practice for most services.

At frequencies above 50 Mc., capacitance effects at high impedance points on the antenna system become objectionable when insulators are used to support the antenna at these points. For this reason it is standard practice in the more exacting applications to support all radiating elements at low impedance points only. The small size of half-wave dipole elements at these frequencies makes such "self-supporting" construction quite feasible.

The short physical length of a dipole element at these frequencies also makes it practicable to obtain "broad-band" characteristics by the simple expedient of increasing the diameter of the antenna elements until the diameter is an appreciable fraction of a wavelength. The convenient wavelength dimensions also make practical the use of parabolic, flat-sheet, and "corner" type reflectors in the u-h-f and upper v-h-f range. These are simple yet highly effective directional systems which are not suited to h-f use because of their excessive physical size at such frequencies.

In the following discussion, v-h-f/u-h-f antennas will be grouped according to polarization because usually the polarization for a given application or service is predetermined and because the two types of polarization require different construction techniques. For the same reasons they will be further broken down within each polarization group into directional and nondirectional types. However, for applications requiring horizontal directivity, horizontal polarization usually is standard. This is fortunate, because horizontal orientation of the dipoles in a phased array simplifies the problem of obtaining the proper phasing and feeding of the elements without encountering difficulties from antenna effect. So, in view of the foregoing considerations, only limited space will be devoted to ver-

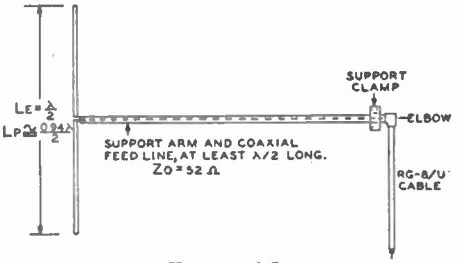


Figure 6-1.

VERTICAL HALF-WAVE DIPOLE WITH COAXIAL FEED.

When a moderate amount of eccentricity of the radiation pattern and a small amount of antenna effect on the feed line can be tolerated, the above arrangement offers a simple and effective v-h-f/u-h-f antenna system which is well suited to many types of installation. The optimum radiator length will depend upon the length/diameter ratio of the radiator. Large diameter elements will increase the useful band width.

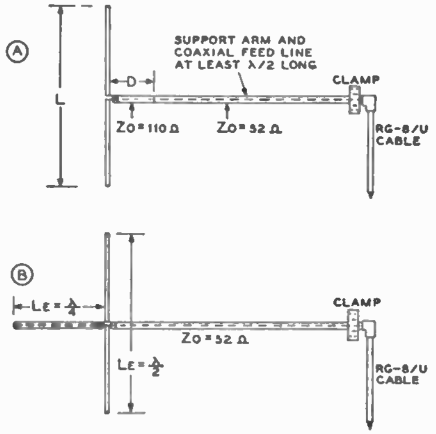


Figure 6-2.

BROAD BAND V-H-F/U-H-F ANTENNAS EMPLOYING REACTANCE CANCELLING SYSTEMS.

The arrangement shown at A limits the SWR to a tolerable value over a frequency range of about 1.8 to 1 when the dimension L is made 0.5 wavelength and the dimension D is made 0.1 wavelength near the low frequency end of the desired range. The arrangement shown at B gives a very low VSWR over a range of approximately 1.2 or 1.3 to 1.0 when the surge impedance of the quarter-wave reactance-cancelling section is optimized for the particular length/diameter ratio of the radiating section. It is assumed that fairly large diameter radiating elements are used with either arrangement.

tically polarized arrays which exhibit horizontal directivity.

VERTICALLY POLARIZED CIRCULAR RADIATORS

Circular radiators (omnidirectional in azimuth) for vertically polarized transmission and reception consist of either some sort of "Marconi" (resonant unipole worked against ground), some type of resonant half-wave dipole, or some vertically stacked arrangement of colinear half-wave dipoles. Often the resonance is very broad due to inherent characteristics of the type of radiator used. Regardless of the inherent broadness of the radiator it sometimes is enhanced by the incorporation of a reactance cancelling device or the proportioning of the system parameters to give a reactance cancelling characteristic.

THE VERTICAL HALF-WAVE DIPOLE

When properly mounted, a vertical half-wave dipole with conventional feed will exhibit a sufficiently low antenna effect on the line for many applications, though the antenna is inferior from this standpoint to various other types giving substantially the same radiation pattern. The important consideration with regard to mounting is to bring the feed line away from the radiator at a right angle for at least a half wavelength and preferably one wavelength, in order to avoid excessive distortion of the pattern and to prevent excessive antenna effect.

The most satisfactory arrangement, illustrated in figure 6-1, consists of a split radiator fed by a 50-ohm coaxial cable. The directly grounded half of the radiator should be mounted upwards, as the orientation has no effect upon the operation and this arrangement affords increased lightning protection.

The antenna can be broad banded by using large diameter elements. The useful frequency range can be further increased by either of the methods illustrated in figure 6-2. At figure 6-2A the last 0.1 wavelength of the coaxial line is proportioned to have a surge impedance of approximately 110 ohms. This considerably extends the useful range on the *high* side of the design frequency by about 70 per cent, the antenna covering almost a 2 to 1 frequency range with a tolerable VSWR when the dimensions are optimized experimentally. Note that the design frequency should be near the low end of the desired frequency range, rather than at the center of the range.

The arrangement at figure 6-2B makes use of the fact that a resonant half-wave dipole acts like a series tuned circuit, while a shorted quarter wavelength of line "looks" like a parallel resonant circuit. The correct surge impedance for the reactance-cancelling quarter-wave stub depends upon the length/diameter ratio of the radiator and upon the bandwidth to be covered. Usually the optimum surge impedance will be quite low. In actual practice the stub would be laid along the combined feed line section and support arm, rather than projected beyond the junction with the radiator as shown here in order to facilitate illustration of the electrical circuit.

A parallel resonant L/C circuit using lumped L and C can be substituted for the shorted quarter-wave section of line if desired. The capacitor should use polyethylene, ceramic, or other low loss dielectric material suitable for the operating frequency. The optimum values for a given frequency will depend upon the length/diameter ratio of the radiating section and the bandwidth to be covered. In typical installations the optimum capacity usually will run about 25 $\mu\mu$fd. per meter of wavelength as determined from the design frequency. The inductance can be a small loop or coil of wire, the optimum length of which is determined experimentally.

One disadvantage of the antennas shown in figure 6-2 is the fact that while they have good "broad-band" characterisitcs as regards feed point impedance, antenna effect on the line becomes rather high at frequencies considerably removed from that at which the elements resonate. The antenna effect can be greatly reduced at off-resonant frequencies by the use of a broad-band detuning sleeve when the application requires a low order of antenna effect over the entire working range of the antenna.

THE VERTICAL UNIPOLE (MARCONI)

In certain applications a large mass of metal is available to serve as an effective ground for a quarter wavelength vertical radiator. When sufficient elevation can be obtained by working a quarter-wave radiator against the metal body or frame of a vehicle, boat, etc., a simple and highly effective antenna is realized. Such an antenna is illustrated in figure 6-3. The feed point impedance will depend upon the length/diameter ratio of the radiator and upon the configuration of the ground body, and in some cases the surge impedance of the Q matching section will have to be

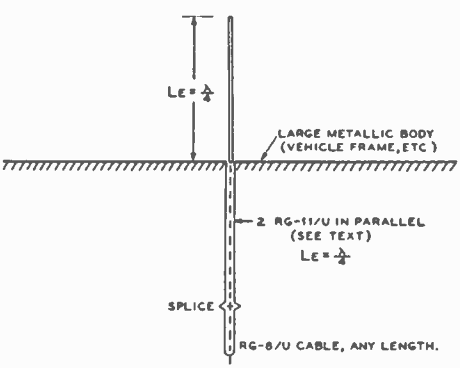


Figure 6-3.

RESONANT UNIPOLE OR "MARCONI" ANTENNA.

Below approximately 50 Mc. this antenna usually is somewhat flexible and is called a "whip" radiator, while above approximately 100 Mc. it usually is rigid and is called a "stub" radiator. The "ground" should preferably consist of a substantially flat plane extending at least a half wavelength in all directions from the radiator, and the radiator should preferably have an unobstructed "view" in all azimuth directions. However, the antenna often will work well when these conditions cannot be met, as when a 30 and 40 Mc. "whip" is mounted on the rear bumper of a vehicle in order to keep the height within acceptable limits. In the latter type of installation the pattern will not be truly circular, and will be affected considerably by the horizontal spacing between the radiator and car body. Also, the feed point resistance will be lowered, necessitating a lower value of surge impedance for the Q section.

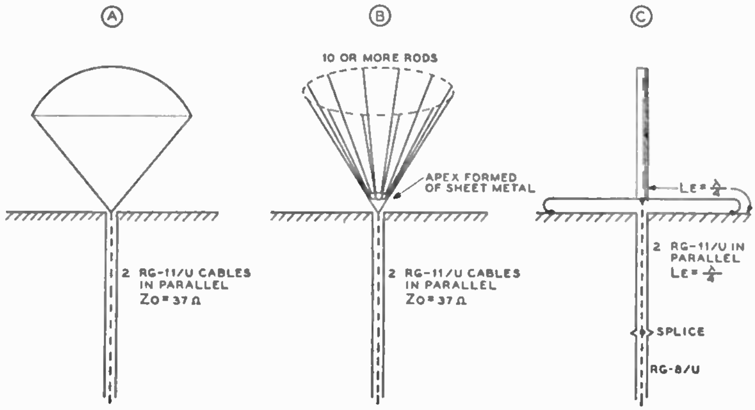


Figure 6-4.

BROAD BAND ASYMMETRICAL VERTICAL RADIATORS.

These radiators are especially well suited to broad band use at frequencies above 60 or 70 Mc. The arrangements shown at A and B will work well over a frequency range of 3 to 1; that of C over a range of approximately 1.3 to 1. For best results at least 10 spines should be used in the skeleton cone arrangement (B). They may be terminated on a metal hoop for additional stiffness if desired. A substantially flat or slightly domed ground plate is necessary for proper operation of the A and B arrangements; with the C arrangement this is not so important, though still desirable.

altered in order to obtain an accurate match.

A surge impedance of 26 ohms can be obtained by paralleling RG-8/U cable, 30 ohms by paralleling RG-8/U cable with an equal length of RG-11/U cable, and 37 ohms by paralleling RG-11/U cable. By proper choice of one of these three impedances for the "Q" section and the option of either 52-ohm cable (RG-8/U) or 75-ohm cable (RG-11/U) for the feed line, it is possible to obtain a very low SWR in almost any practical installation.

The optimum physical length of the quarter-wave radiator (called a "whip" or "stub" radiator) is determined for a given frequency not only by the length/diameter ratio of the radiator, but also by the configuration of the ground body.

It should be readily apparent that the optimum radiator length and Q section surge impedance will not be the same for a radiator located in the center of a car top as for a radiator mounted on the car's rear bumper. However, neither the radiator length nor Q section surge impedance is extremely critical for good performance, and usually a satisfactory combination can be arrived at experimentally with but little effort.

In the u-h-f range it is important that the shunt capacity provided by the supporting insulator be kept to the lowest possible value so that the feed point impedance will be substantially resistive. Excessive capacitive reactance shunted across the feed point may make it impossible to obtain a low line SWR, regardless of adjustment of the radiator length. Excessive shunt capacity can be rendered harmless by placing across the feed point a small inductance which resonates the capacitance to the operating frequency. This will also broaden out the useful frequency range of the antenna as already explained.

When the antenna is to be used over an appreciable frequency range it can be broad banded by any of the various schemes described for the vertical dipole in the preceding section. There also are other schemes which are especially well suited to an asymmetrical or "unbalanced" radiator working against ground when the frequency is not so low that the physical size offers a construction problem. Some of these are shown schematically in figure 6-4.

By proper dimensioning of the domed conical radiator shown at figure 6-4A, it is possible to obtain a low SWR over a frequency range considerably exceeding 3 to 1. However, when covering a range much greater than 2 to 1, the cone begins to act like a horn radiator at the upper end of the range, causing the vertical pattern at some frequencies to maximize at some angle above zero elevation. Thus, while it is possible to make the conical radiator virtually aperiodic above a certain "cutoff" frequency, with a frequency characteristic resembling that of a high pass filter, distortion of the radiation pattern occurs in a vertical plane as the frequency is increased.

The cutoff frequency (the frequency below which the SWR rises rapidly) is slightly below that at which the electrical "length" of the cone is one quarter wavelength. This occurs at a physical slant height which is a little less than one quarter wavelength.

The skeleton cone arrangement shown at figure 6-4B approximates electrically that shown at A. The stiff flared rods are lighter and present less wind resistance, but do not permit quite as flat an impedance-frequency characterisitc as the arrangement at A. However, if a sufficient number of rods are employed, the impedance-frequency characteristic still is superior to that of a thick cylindrical radiator. The spines can be terminated on a metal hoop to stiffen up the structure if desired.

The arrangement shown at figure 6-4C is simply a thick cylindrical radiator in conjunction with a combined "metallic insulator" and reactance cancelling section. This avoids the inherent physical weakness and shunt capacity of a dielectric supporting insulator. The reactance cancelling section is, in effect, two quarter-wave stubs in parallel across the series feed point. The effective surge impedance of the two combined reactance cancelling and physical support stubs is determined by the width of the two straps which form the stubs, and by the spacing between the straps and the ground plane. The wider the straps and the closer the spacing, the lower the surge impedance. The latter should be optimized for the lowest maximum SWR for the particular radiator employed over the bandwidth desired. Because of the low matching ratio of the Q section, it does not appreciably increase the frequency sensitivity of the system, and a low SWR over a frequency range of 1.3 to 1 is possible with this arrangement.

To avoid compromising the inherently good impedance-frequency characteristic of the A and B antennas of figure 6-4, the apex angle of the cone (twice the angle of revolution) is chosen so that the average feed point impedance approximates the surge impedance of the line. For the lowest possible maximum SWR, the apex angle should be optimized for the particular line impedance employed and frequency range to be covered. For RG-8/U cable the optimum angle of revolution will be roughly 45 degrees, and for RG-11/U cable it will be roughly 30 degrees.

A substantially flat or slightly domed ground plate which extends for at least one half wavelength in all directions is necessary for best operation of the A and B arrangements. With the C arrangement this is not so important, though still desirable.

In the case of all three antennas shown in figure 6-4 the feed line is almost completely decoupled or isolated from the radiating portion by the ground plane, provided the latter is extensive in terms of wavelength. Antenna effect on the line therefore will be insignificant at all frequencies.

THE BROWN "GROUND PLANE" ANTENNA

A resonant unipole can be used in fixed station "flag pole" installations by simulating a ground plane with four, equally spaced horizontal "radials" each an electrical quarter wavelength long as shown in figure 6-5. The radials effectively isolate the vertical radiator from the coaxial feed line, and as the opposite radials induce currents of opposite polarity in the feed line (causing complete cancellation), antenna effect on the line is negligible. The functioning of the radials differs from an infinite flat sheet ground plane in one respect, however. The feed point impedance, instead of being half that of a comparable dipole, is somewhat less. With typical length-to-

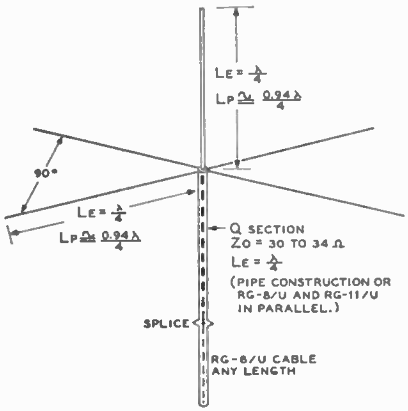


Figure 6-5.

THE BASIC GROUND PLANE ANTENNA.

Four resonant horizontal "radials" simulate a ground plane for this vertical unipole radiator. Various other methods may be used for matching the feed line to the feed point resistance of approximately 21 to 24 ohms. While the element lengths are best determined experimentally, they are not extremely critical and usually the above dimensions will give good results with thick elements of typical diameters.

diameter ratios of the vertical radiating element and radials, the feed point resistance will run between 20 and 25 ohms.

This feed point impedance can be transformed to an impedance comparable to the surge impedance of a coaxial line in a number of ways, the Q matching section shown in figure 6-5 simply being one of the more popular. For instance, a folded element can be used for the unipole radiator, so that the vertical radiator resembles half a folded dipole. The feed point impedance will be on the order of 80 ohms, permitting use of RG-11/U type 75-ohm coaxial cable without a matching transformer for applications not requiring an absolutely flat line. The use of a folded element also increases the useful frequency range (bandwidth) of the antenna.

"UNIPOLE"

THE "DROOPING GROUND PLANE" ANTENNA

It is possible to dispense with the quarter-wave Q matching transformer in figure 6-5 by arranging the radials so that they lie in an inverted cone. At a certain angle of revolution the feed point resistance will be equal to the surge impedance of the RG-8/U cable, permitting direct connection as in figure 6-7. If the radials were completely depressed, so they were vertical and parallel to the contained feed line, the feed point resistance would be comparable to that of a half-wave dipole. As the latter arrangement is essentially a skeletonized "sleeve" or coaxial antenna of the type shown in figure 6-8 and described in a later section, it is seen that the "Drooping Ground Plane" antenna is a hybrid which is more or less a compromise between the two.

When the radials are depressed as in figure 6-7, the vertical radiating element is not as completely isolated from the feed line as is the case for horizontal radials. Also, the drooping radials have a vertical component of current. The result is that while the antenna of figure 6-7 is somewhat simpler than that of figure 6-5, the feed line exhibits considerably more antenna effect.

While four drooping radials are shown in

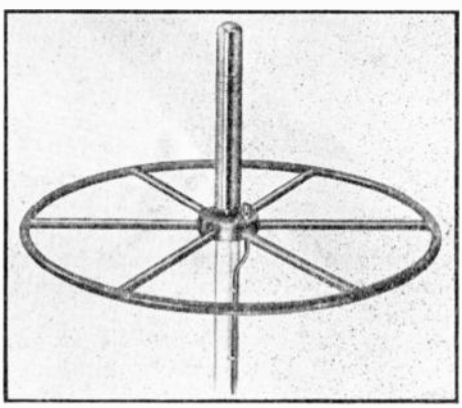


Figure 6-6.

COMMERCIALLY MANUFACTURED GROUND PLANE ANTENNA (AMPHENOL) FOR USE OVER THE 152-162 MC. BAND.

The grounded inner coaxial support rod serves not only as a metallic insulator but also as an impedance matching and reactance cancelling device. The "spoke wheel" ground plane is readily removable to permit mounting the radiator on a natural ground plane such as the metallic top of a railroad engine cab roof.

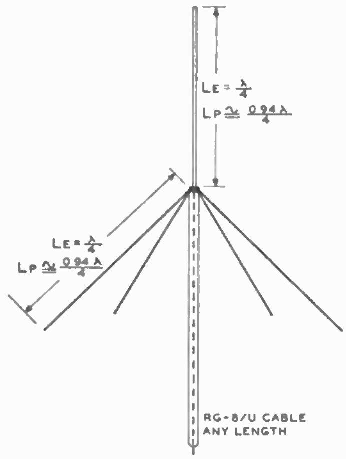


Figure 6-7.

THE "DROOPING GROUND PLANE" ANTENNA.

By optimizing the angle of revolution of the skeletonized conical skirt, a feed point impedance equal to the surge impedance of RG-8/U coaxial line (52 ohms) is obtained, thus permitting direct connection. Antenna effect on the line is slightly greater than with the ground plane antenna of figure 6-5.

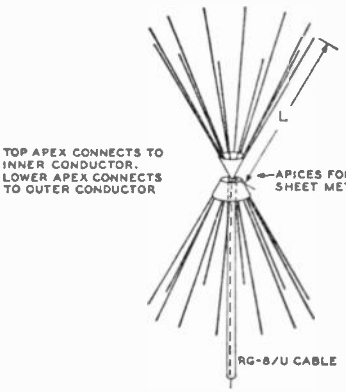


Figure 6-8.

THE DOUBLE SKELETON CONE ANTENNA.

A skeleton cone has been substituted for the single element radiator of figure 6-7. This greatly increases the bandwidth. If at least 10 elements are used for each skeleton cone and the angle of revolution and element length are optimized, a low SWR can be obtained over a frequency range of at least two octaves. To obtain this order of bandwidth, the element length L should be approximately 0.2 wavelength at the lower frequency end of the band, and the angle of revolution optimized for the lowest maximum VSWR within the frequency range to be covered. A greater improvement in the impedance-frequency characteristic can be achieved by adding elements than by increasing the diameter of the elements. With only 3 elements per "cone" and a much smaller angle of revolution a low SWR can be obtained over a frequency range of approximately 1.3 to 1.0 when the element lengths are optimized.

figure 6-7, three or five sometimes are employed (with little difference in performance).

DOUBLE SKELETON CONE ANTENNA

The bandwidth of the antenna of figure 6-7 can be increased considerably by substituting several space-tapered rods for the single radiating element, so that the "radiator" and skirt are similar. If a sufficient number of rods are used in the skeleton cones and the angle of revolution is optimized for the particular type of feed line used, this antenna exhibits a very low SWR over a 2 to 1 frequency range. Such an arrangement is illustrated schematically in figure 6-8. The maximum antenna effect on the line, while still low, is greater than that of a single skeleton cone worked against a horizontal ground plane (figure 6-4B). The antenna effect will vary considerably with frequency for a given installation.

Each set of spines may be terminated on a metal hoop for mechanical stiffness if the individual spines are not sufficiently rigid to prevent whipping in a strong wind.

THE COAXIAL OR "SLEEVE" ANTENNA

One of the most widely used v-h-f/u-h-f vertical radiators is the coaxial or "sleeve" radiator. The most common or series-fed arrangement is shown schematically in figure 6-9. It is essentially a verical dipole with the bottom radiating element or "skirt" made hollow to permit center feed of the radiator by running the coaxial feed line up through the center of the radiator. This effectively decouples that portion of the feed line within the bottom element from both elements.

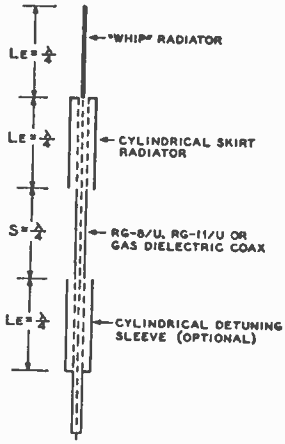


Figure 6-9.

THE COAXIAL OR "SLEEVE" ANTENNA (SERIES FED).

This antenna is widely used for fixed frequency v-h-f/u-h-f applications, both for mobile installations and fixed station "flag pole" installations. In the latter type installation a line detuning sleeve often is incorporated, as shown above, to minimize antenna effect. The skirt length of this antenna is very critical, making the antenna quite frequency sensitive. To simplify the illustration the hollow support staff is not shown. It goes outside the coaxial line and inside the sleeve elements, and electrically has no more effect upon the general operation of the antenna than would the outside conductor of the coaxial line if the support staff or mast were not present.

However, appreciable coupling still exists between the radiator (both elements) and that portion of the feed line immediately below the skirt, causing a moderate amount of antenna effect on the line or the hollow support shaft containing the line. When antenna effect must be kept to the lowest possible value, it can be minimized by incorporation of a quarter-wave detuning sleeve placed with its open end toward the bottom end of the skirt radiator and spaced one-quarter wavelength from it. The usual practice is to dispense with the detuning sleeve in mobile installations, and ordinarily it is incorporated only in fixed "flag pole" installations.

The chief disadvantage of the coaxial antenna is its frequency sensitivity. When the dimensions are optimized for a given frequency, considerable reduction in performance is observed at frequencies deviating from the design frequency by as little as 1 per cent. For this reason the antenna finds its greatest use in fixed frequency applications.

The high degree of frequency sensitivity is due to the combination of skirt and the contained feed line together with supporting staff acting as a high Q quarter-wave resonator or anti-resonant tank circuit. The length at which the skirt resonates as a linear tank is affected by the contained dielectric insulation used for spacing and support of the skirt. However, the length at which it resonates as a *radiator* is not affected by the dielectric material inside the skirt. The result is that the length which resonates the skirt as a high Q quarter-wave section of line is not quite the same as the length at which it resonates as the bottom half of the radiating system. However, as the length at which the skirt resonates as part of the radiating system is not critical, and deviations from optimum can be compensated for to some extent by alteration of the length of the top radiating element or "whip," the skirt length is adjusted for high impedance between the skirt rim and the contained line and support staff without regard to its resonant length as a radiator. As this adjustment is not affected by the presence of surrounding objects, it can be made before the antenna is installed, and the radiating portion of the antenna optimized after installation by adjustment of the whip length.

The optimum skirt length usually runs from about 0.98 to 1.04 of a free-space quarter wave, the length being defined as the length of the contained air column. The optimum whip length usually runs between 0.94 and 0.96 of a quarter wave. Commercial versions of the coaxial antenna usually are provided with a telescoping, calibrated skirt, and with a telescoping whip which can be adjusted for best performance under actual operating conditions (to allow for the effect of surrounding objects).

When the antenna is installed well in the clear, as in a "flag pole" installation, the feed point impedance is such that a good match is obtained with either 50-ohm line or 70-ohm line. When the feed point impedance is lowered by installation of the antenna on a car bumper, for instance, the better match is obtained with 50-ohm cable. A slight mismatch can be partially compensated for by adjustment of the whip and skirt length, so that at a single frequency a nearly flat line can be obtained even though the radiation resistance of the antenna differs somewhat from the surge impedance of the line.

The usual procedure is to run the coaxial feed line (ordinarily RG-8/U or RG-11/U) through a hollow staff which in turn runs up through the skirt to support the antenna at its center point. In such an arrangement any "antenna current" that ordinarily would flow on the outside of the outer conductor of the line flows instead on the outside of the support staff.

It is possible to shunt feed the antenna by tapping the center conductor of the feed line to a suitable point on the skirt, in which case the bottom of the whip is connected directly to the top of the skirt. However, this arrangement offers no worthwhile advantages for feeding a single coaxial radiator, and is somewhat more difficult to adjust. It does offer certain advantages, however, in a stacked array of colinear sleeve radiators, and will be described under vertically polarized v-h-f/u-h-f directional systems.

In fixed installations requiring a very long run of feed line, particularly when the frequency is over 100 Mc., gas dielectric coaxial line sometimes is employed in order to minimize line losses. However, in the average low power fixed installation using this type of antenna (such as various emergency services) the additional expense is not economically justified for the usual line lengths.

THE FLEXIBLE DIPOLE OR "HALYARD" ANTENNA

It is possible to make the outer conductor of a coaxial line serve as the skirt of a sleeve antenna at the same time it performs its duties as part of the transmission line. It is only necessary to detune the outer conductor of the line at a point one quarter wavelength from its end. That such an arrangement is feasible is explained by the fact that the outside of the outer conductor of a coaxial line "doesn't pay any attention to what is going on inside the line," and vice versa.

Exploiting this fortunate fact, it is possible to construct a very simple yet effective makeshift antenna of the "sleeve" family for use below 150 Mc. using nothing but RG-8/U or RG-11/U cable, some waxed or oiled cord, and a Boy Scout knife. With only a few additional materials a lasting, workmanlike job can be done, and the antenna will then make a highly desirable permanent acessory for portable equipment or as a stowed emergency "standby" antenna for use in the event of a failure of a regular fixed station antenna. The performance of the antenna compares favorably with that of a regular sleeve antenna with separate skirt at frequencies up to about 80 or 100 Mc., and the performance is still within 3 db of a well-constructed, conventional sleeve antenna at 150 Mc. The basic flexible dipole or halyard antenna (also sometimes known as a "pendant" or "limp" antenna) is shown schematically in figure 6-10.

Neither the whip length nor the skirt length is extremely critical, because slight variations in these lengths can be compensated for by adjustment of the detuning coil. The adjustment of the detuning coil is quite critical for maximum performance of the antenna, and the coil should be adjusted experimentally to the optimum fractional turn. Because this adjustment is critical, the antenna is quite frequency sensitive, just as is the case with a regular sleeve type antenna utilizing a separate skirt.

For a makeshift installation the detuning coil can be scramble wound "on air" and lashed with string. The inductance can be optimized by inserting a 2.5 volt 0.5 amp. dial light bulb in series with the bottom of the whip or touching a neon bulb to the top end of the whip and altering the coil turns until maximum indication is obtained, maintaining the transmitter input at a suitable level.

For a permanent or semi-permanent in-

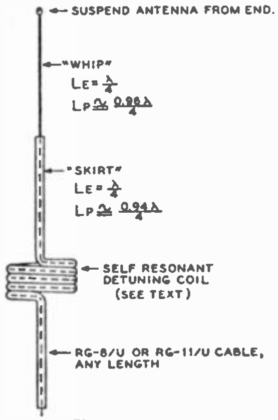


Figure 6-10.

THE FLEXIBLE DIPOLE OR "HALYARD" ANTENNA.

This antenna is easily improvised when RG-8/U or RG-11/U coaxial cable is available, and makes an excellent temporary or emergency antenna. The performance up to about 100 Mc. compares favorably with that of a regular sleeve antenna utilizing a separate skirt, and the efficiency is still fairly good up to about 150 Mc. The whip and skirt lengths are not critical, but the line detuning coil must be adjusted accurately to a fractional turn if maximum performance is to be obtained. The coil adjustment is best made with the help of a field strength meter.

stallation a workmanlike, waterproof connection should be made at the junction of the whip and skirt. Regular coaxial cable fittings can be altered to do the job. The detuning coil should be close wound on a low loss form, either of the X type or cylindrical, and firmly anchored in place after the correct adjustment is found. The detuning coil inductance should be adjusted for maximum field strength while the transmitter input power is held constant. As this type of antenna is primarily a fixed frequency antenna, it is permissible to adjust the coil for maximum field strength (with a given transmitter power) without regard to the line SWR.

To give some idea of the number of turns that will be required, the correct adjustment usually will be obtained with approximately 4 turns 5 inches in diameter for 30 Mc., 4.5 turns 3 inches in diameter for 50 Mc., and 2 turns 3 inches in diameter for 100 Mc.

As the feed point impedance with the antenna well in the clear usually is in the vicinity of 60 to 65 ohms, good results will be obtained with either RG-8/U or RG-11/U cable.

VERTICALLY POLARIZED CIRCULAR RADIATORS WITH SHARP VERTICAL DIRECTIVITY

For v-h-f/u-h-f ground-to-ground applications the only radiation that is effective is that which is directed parallel to the earth's surface. Energy radiated at an angle of departure exceeding one or two degrees serves no useful purpose except in unusual terrain situations or for commnication with aircraft. Therefore, if the radiation can be confined to very low vertical angles, an equivalent power gain is obtained without destroying the circular pattern in the horizontal plane.

The Colinear Coaxial Array

The most practical vertically polarized array for accomplishing this type of directivity control is a colinear stack of vertical "sleeve" antennas. The basic element is similar to the sleeve antenna of figure 6-8 except that shunt skirt feed is employed instead of series feed, because shunt feed permits a continuous length of pipe to be used for the "whip" portion of each radiator at the same time that it serves as the support staff or mast for the whole array. The basic system is illustrated by the schematic of the colinear coaxial array shown in figure 6-11. While only two sections or elements are shown, any number up to 10 or more can be used with a corresponding increase in power gain, the arrangement of the additional elements being the same as for the two illustrated. However, economic considerations usually limit the stack to from four to six sections for applications to which this type of array is suited. The practical signal gain is very nearly equal to the free space directivity gain, and can be determined from the colinear array gain curve of figure 5-15.

The point at which each feed line taps

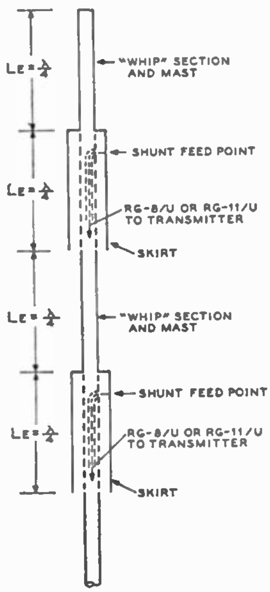


Figure 6-11.

VERTICALLY POLARIZED COLINEAR COAXIAL ARRAY.

This array gives an effective power gain while maintaining a circular horizontal pattern. While only two elements are shown, any number up to 10 or more may be employed, though more than four to six seldom are justified from an economical standpoint. To obtain the proper phase relationships, all feed lines are made exactly the same length and connected together at the transmitter coupling device. The shunt feed arrangement allows the use of a continuous tubular center member which serves both as radiator and supporting staff. If desired, a detuning sleeve or skirt can be added to minimize antenna effect, as illustrated in figure 6-9. Because the skirt lengths and feed point positions are quite critical, this type of array usually is purchased in pretuned form as a commercial "package."

on to the corresponding skirt is adjusted experimentally until negligible standing waves are observed. The length of the skirt of each section is quite critical, and when adjusted experimentally for best performance it will be found that this adjustment and the feed point adjustment interact, thus complicating the procedure considerably. Pretuned arrays of this type are offered commercially for various v-h-f frequency assignments. Experimental adjustment of such an array is not required.

VERTICALLY POLARIZED DIRECTIONAL ARRAYS

While stacked arrays which exhibit a circular pattern (such as the colinear coaxial array just described) are really directional arrays, the expression "directional array" by common usage refers to an array exhibiting pronounced horizontal directivity. For most v-h-f/u-h-f services which can use a high degree of horizontal directivity to advantage the use of horizontal polarization is standard practice. However, there are some applications where horizontal directivity is desirable in conjunction with vertical polarization. Therefore, some of the more practical arrays for this type of service will be described briefly.

PARASITIC ARRAY WITH SLEEVE TYPE DRIVEN ELEMENT

The horizontally polarized, three-element parasitic array described in the preceding chapter for high-frequency work easily can be modified for vertical polarization and scaled down for v-h-f/u-h-f work. Coaxial feed of the driven element offers a simple method of feeding this element when vertically oriented. The array is illustrated in figure 6-12. The use of quarter-wave spacing permits use of a regular sleeve antenna as the driven element without greatly affecting the feed point impedance. The practical signal gain of this antenna is between 6 and 7 db when adjusted for maximum forward gain.. For further details on the parasitic elements, refer to the discourse in Chapter 5 on the three-element parasitic array.

When the application requires that the azimuthal direction of the radiated beam be made adjustable, either the whole array may be constructed to turn, or the driven radiator may be made fixed and the parasitic elements arranged to rotate around the driven element.

Because a sleeve antennna by itself already is quite frequency sensitive due to the high "Q" of the skirt, the addition of parasitic elements does not appreciably increase

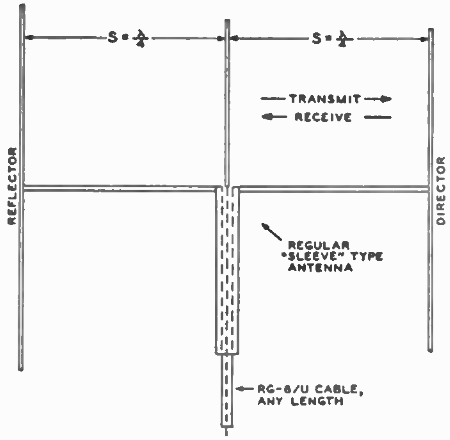


Figure 6-12

SLEEVE TYPE VERTICAL RADIATOR WITH PARASITIC REFLECTOR AND DIRECTOR.

This unidirectional array gives a practical signal gain of 6 to 7 db over the sleeve radiator alone. The driven element is a standard sleeve radiator of the type described in a previous section, and the operation and adjustment of the parasitic elements are covered under high-frequency arrays in Chapter 5.

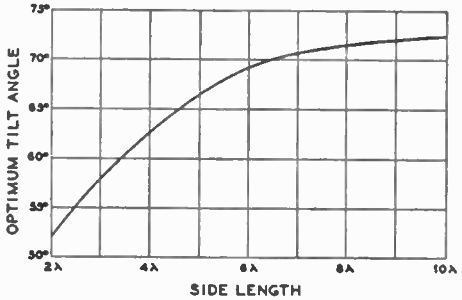


Figure 6-13.

OPTIMUM TILT ANGLE FOR A V-H-F RHOMBIC ANTENNA.

The curve applies to either a vertical or horizontal rhombic regardless of elevation above ground, so long as the elevation of the lowest portion of the antenna exceeds two wavelengths.

the frequency sensitivity of the antenna.

The use of quarter-wave spacing for the parasitic elements has two advantages. First, it does not greatly alter the feed point impedance, thus making a matching section unnecessary. Secondly, it simplifies the adjustment procedure by making it possible to adjust the sleeve radiator for optimum performance with the parasitic elements removed, then add the parasitic elements and adjust their lengths for any desired compromise between forward gain and front-to-back ratio, without making further adjustment of the sleeve radiator. This simplified precedure is not practicable with closer spacing.

THE VERTICAL RHOMBIC

One of the most effective vertically polarized directional arrays for the v-h-f range is a rhombic antenna oriented so that the plane of the array is vertical. Such an array requires a comparitively large space for erection, but in many applications this is no disadvantage. With a rhombic oriented in this manner a high signal gain can be obtained with only moderate horizontal directivity. This is due to the high degree of *vertical* directivity contributed by a rhombic antenna oriented in this manner.

In designing a rhombic array for v-h-f ground work the design considerations are somewhat different from those of a high-frequency rhombic designed for sky-wave propagation (with which the reader is presumed familiar). In the case of a v-h-f rhombic the optimum tilt angle is substantially independent of the height of the array above ground and the orientation of the array (vertical or horizontal). Also, from a practical standpoint there is no optimum elevation above ground if feeder losses and antenna cost are ignored; a v-h-f rhombic always should be elevated above ground as much as possible in level country, and erected on the highest available sight under conditions of irregular terrain. Unlike the high-frequency rhombic, a v-h-f rhombic need not be located over flat ground for best performance, and should not be inclined when located over sloping terrain.

The optimum tilt angle as a function of leg length for a v-h-f rhombic (either vertical or horizontal) is given in figure 6-13. For information regarding proper termination and other details common to all rhombics refer to the section on high-frequency rhombic antennas in the preceding chapter.

In constructing a rhombic for the upper portion of the v-h-f range, particular care must be taken to avoid undesirable stray lumped reactances and to avoid excessive net reactance in the termination. This becomes increasingly difficult as the frequency is increased, and limits the usefulness of the rhombic to a frequency of approximately 300 Mc. without resorting to "laboratory" techniques which are none too practical from the standpoint of obtaining a rugged, reliable, economical installation. Other arrays are better suited than the rhombic to u-h-f applications.

SHEET OR "SURFACE" REFLECTOR ANTENNAS

The radiation from a single dipole element can be concentrated into a restricted solid angle by the use of a sheet metal reflector, producing the same "beaming" effect as an array of dipoles. The sheet reflector can take any of numerous shapes, the most desirable being determined by the particular directivity pattern desired and by mechanical considerations. Sheet reflectors are not practicable in the h-f range because of their prohibitive size, but in the v-h-f/u-h-f range the required structures are easily and economically constructed.

The simplest "surface" reflector is a flat sheet reflector (figure 6-14A) but the gain obtainable with such a reflector is limited to about 7 db when a single driven dipole element is employed. When a broadside "curtain" consisting of a large number of driven dipoles is backed up by a flat sheet reflector the reflector simply converts the bidirectional pattern into a unidirectional pattern by "turning around" the radiation off the back side, and the gain due to the reflector is approximately 3 db. Under these conditions the reflector does not significantly affect the shape of the pattern off the front side, and hardly can be considered as a device for "beaming" the radiation into a smaller solid angle.

When a single driven dipole is employed in front of a large, flat-sheet reflector, the directivity and gain can be increased by forming the sheet into a cylindrical parabolic reflector in such a manner that the radiator is at the focus of the parabolic curve (figure 6-14B). Assuming that the width of the mouth is several wavelengths and that the dipole is vertical, most of the directivity will be in a horizontal plane. Increasing the distance across the aperature will result in a corresponding increase in horizontal directivity and gain. There is little point in increasing the height of the sheet very much beyond a wavelength, because nothing significant is to be gained thereby. In no case will the reflector provide much increase in vertical directivity.

While high gains can be obtained with cylindrical parabolic reflector systems, they will not be described in detail because they have nothing to offer over a *corner reflector* of the type shown at figure 6-14C and described in the following section, and the corner reflector is a simpler and more practical structure.

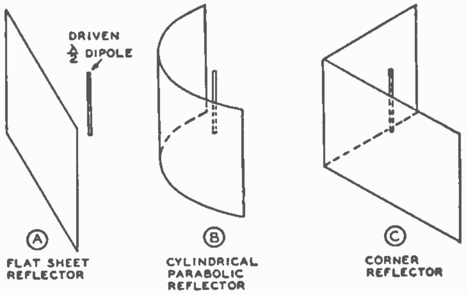


Figure 6-14.

BASIC "SHEET" OR "SURFACE" REFLECTING SYSTEMS FOR A DRIVEN DIPOLE WHICH PROVIDE DIRECTIVITY MOSTLY IN ONE PLANE.

At "A" is a plane or "flat sheet" reflector; at "B" it is rolled into a parabolic curve with the driven dipole at the focus; and at "C" the flat sheet is bent into the shape of a V with the driven dipole bisecting the included angle. With the above orientation the radiation is vertically polarized and most of the directivity provided by the reflecting surface is in the horizontal plane; but if the whole array is rotated through 90 degrees so that the dipole is horizontal, the radiation is horizontally polarized, and most of the directivity provided by the reflecting surface is in the vertical plane.

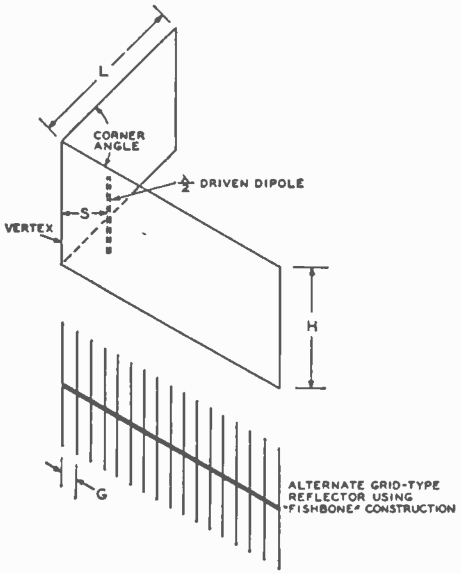


Figure 6-15.

ILLUSTRATING THE PERTINENT DIMENSIONS OF A CORNER REFLECTOR SYSTEM.

The manner in which the above dimensions affect the gain and radiation resistance is illustrated in figure 6-16. The half-wave driven dipole lies in a plane bisecting the corner angle and is parallel to the intersection of the reflecting surfaces.

THE CORNER-REFLECTOR ANTENNA

The corner-reflector antenna system is probably the most practical of all surface-reflector systems exhibiting a preponderance of directivity in one plane (a plane perpendicular to the driven dipole). In practical form it consists of two flat-sheet reflectors (or the equivalent) making an angle of from 45 to 90 degrees with each other, with the dipole radiator located parallel to the intersection of the sheets and lying in the plane bisecting the corner angle, as illustrated in figures 6-14C and 6-15. The practical signal gain of a well-constructed corner-reflector antenna runs between about 9.5 and 13.5 db for corner spacings within these limits.

The radiation resistance, directivity, and gain are all affected by the spacing S between the reflector apex and the dipole element. The variation in radiation resistance and gain with the dipole-to-corner spacing, S, for three common corner angles is shown in figure 6-16. The radiation resistance curves assume typical length-to-diameter ratios for the dipole radiator (rather than a hypothetical, infinitely thin dipole); the gain curves assume typical resistance losses in the antenna; and all curves assume that the side length S is great enough that the performance approaches that obtained with infinitely long sides.

By choosing an appropriate dipole-to-corner spacing for a given corner angle it is possible to realize a radiation resistance close to 52 or 75 ohms, thus permitting the use of a regular sleeve or "coaxial" radiator using RG-8/U or RG-11/U cable for a vertically polarized corner-reflector array.

The dipole also can be center fed, by bringing the feed line away from the dipole at a right angle along the line of the corner bisector so that the feed line is brought out the back of the reflector at the vertex. With this arrangement a quarter-wave Q section may be used to match the feed line to the feed point resistance of the dipole when the two are not approximately the same, thus permitting the use of the most desirable dipole-to-corner spacing without regard to the surge impedance of the feed line.

The height H (figure 6-15) should be equal to at least one half wavelength plus S, and the side length L should be at least 3S and preferably 4S if the maximum theoretical directivity and gain are to be approached for a given corner angle.

It is not necessary that the reflector comprise a solid surface. A screen of moderately small mesh may be used with substantially the same results. Or a grid arrangement consisting of close-spaced wire or rod conductors parallel to the dipole may be used with negligible loss in performance. The spacing between grid wires should not exceed 0.1 wavelength and preferably should not exceed .05 wavelength. The grid wires do not act as approximately resonant parasitic elements, but rather produce a combined effect simulating a solid sheet. Therefore their length is not critical, but preferably should be equal to at least 0.5 wavelength plus S. The grid wires need not be joined together electrically for proper operation, but they may be joined by metallic ribs for mechanical support if desired. The use of a grid structure instead of solid sheets produces a lighter reflector which has less wind resistance.

The dipole may be supported by one or more metallic support members attached to its center, without detrimental effects, provided that they run perpendicular to the dipole.

It will be noted from figure 6-16 that the gain increases as the corner angle is reduced. Theoretically it should be possible to obtain almost any desired degree of directivity and gain by reducing the corner angle, but 45 degrees is about the practical limit. This is explained by the fact that as the corner angle is reduced the spacing S must be increased in order to obtain enough radiation resistance to keep the resistance losses to a tolerable value. And as S is increased, the side length must be increased. Thus while a corner angle of, say, 10 degrees could be used to provide much higher gain, the required side length would become objectionable.

The corner reflector antenna system may be used with horizontal orientation to provide horizontal polarization if desired. In this case most of the directivity contributed by the reflector is in the vertical plane.

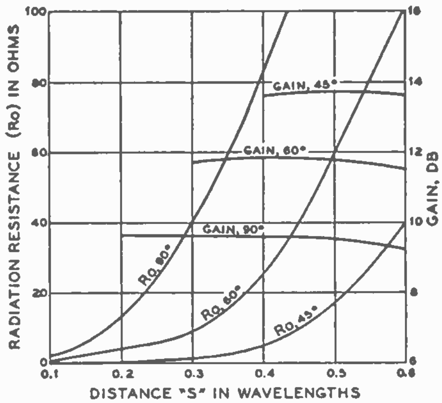


Figure 6-16.

ILLUSTRATING EFFECT OF DIPOLE-TO-CORNER SPACING, S, ON THE PRACTICAL SIGNAL GAIN AND RADIATION RESISTANCE OF A WELL-CONSTRUCTED CORNER REFLECTOR ANTENNA FOR THREE COMMON CORNER ANGLES.

The curves are based upon the following assumptions: The dimension H is equal to one half wavelength plus S; the length L is equal to 4S; the distance G does not exceed .05 wavelength; the dipole has a typical length/diameter ratio; and normal loss resistance is present (exclusive of feed line losses, which are assumed to be the same for the reference antenna). Because of the variability of the latter two factors (dipole shape factor and antenna loss resistance), the curves are approximate, but are close enough for practical purposes. The gain is referred to the dipole radiator used alone.

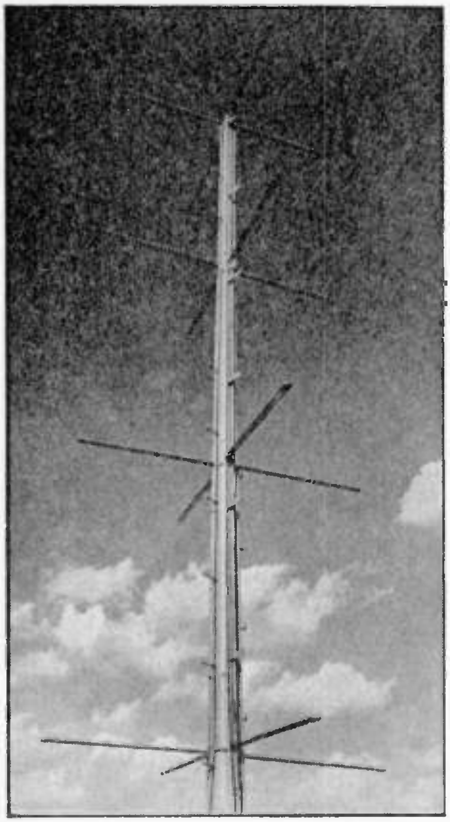
THE ROTATIONAL PARABOLA OR "DISH"

If a driven dipole (either vertical or horizontal) is placed at the focus of a reflector formed to the shape of a rotational parabola, considerable directivity will be provided in both the vertical and horizontal planes. This permits a much higher gain than that obtainable from a cylindrical parabolic reflector of comparable aperture size, because the latter reflector concentrates the energy in one plane only.

If the size of the reflecting paraboloid of revolution is very large in terms of wavelength, and the driven dipole is spaced at least several wavelength from the reflector, there will be little mutual coupling between dipole and reflector, and a spherical wave striking the reflector will be converted to a substantially plane wave upon reflection.

Unfortunately a dipole radiator does not provide a point source of "illumination." Also, the inherent aspect directivity of a dipole makes it impossible to provide precisely uniform illumination of the "dish." For these reasons the beam is slightly elliptical and somewhat broader than might be expected.

For a given dish diameter, the characteristics of the reflected wave are best when a rather shallow dish is employed, so that the focus is well spaced from the reflector. However, this decreases the solid angle represented by the dish; therefore, when an



"TURNSTILE" TYPE FM TRANSMITTING ANTENNA USING FOLDED DIPOLE ELEMENTS (WINCHARGER). THE TWO PANELS ARE FED 90 DEGREES OUT OF PHASE TO GIVE A SUBSTANTIALLY CIRCULAR PATTERN.

ordinary dipole is used, the dish is less intensely illuminated than would be the case with a shorter focal length. This waste of radiation can be minimized by using either a parasitic or small, flat-sheet reflector close to the dipole, in such a manner as to concentrate most of the radiation from the dipole into the solid angle represented by the dish.

If considerably more directivity is desired in the vertical plane than in the horizontal plane, or vice versa, the circular parabolic reflector can be truncated or "lopped off" so that the aperture measures considerably more in the horizontal plane than the vertical plane, or vice versa.

A rigorous treatment of the rotational paraboloid reflector and its various modifications is quite complex. For instance, the optimum position of the driven dipole actually may be very slightly removed from the exact focus, or the optimum dish shape may depart slightly from a true rotational parabola. (This is explained by the fact that the dipole does not serve as a point source of radiation.) Also, computation of the optimum dimensions for a rotational parabolic reflector system for a particular application is rather involved, and once determined the dimensions must be followed very precisely, thus making fabrication difficult unless proper facilities are available.

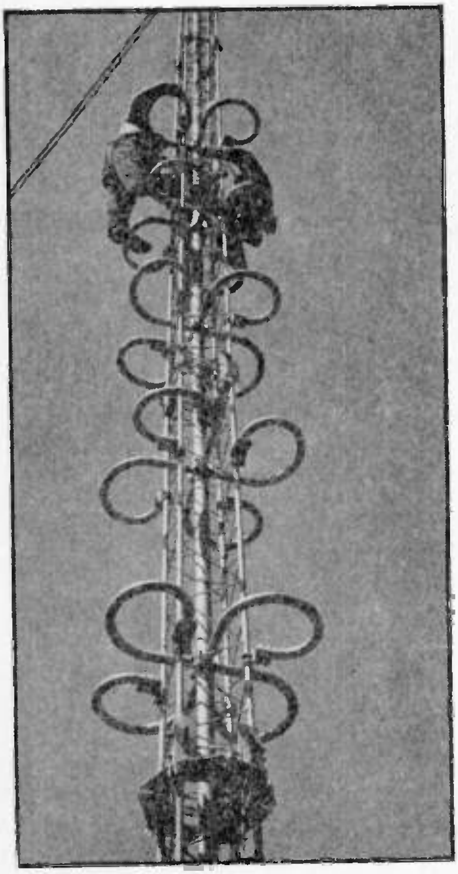
TWO ELEMENT "SQUARE LOOP" FM TRANSMITTING ANTENNA AS MANUFACTURED BY FEDERAL TELEPHONE AND RADIO CORP.

Because of these factors, plus the fact that the rotational parabolic reflector finds its greatest application above 1000 Mc. (where ordinarily it is used with wave guide feed), "dish" antenna systems will not be dealt with in detail.

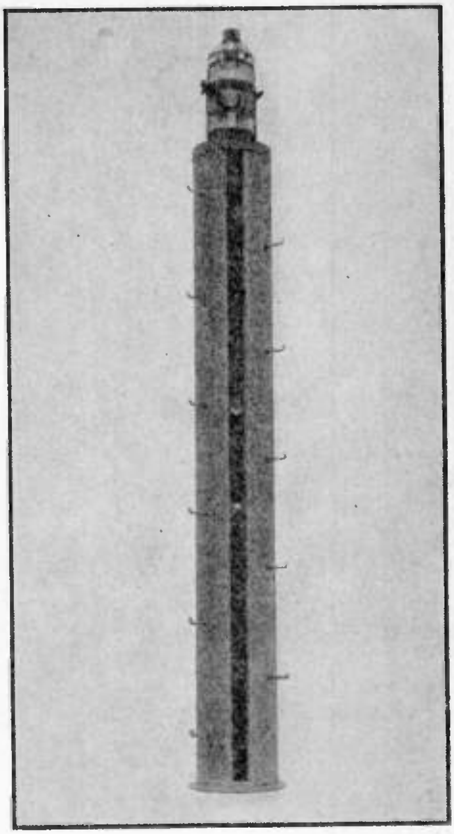
HORIZONTALLY POLARIZED CIRCULAR RADIATORS

The most important application of horizontally polarized circular radiation is in the field of FM broadcasting. Antennas for this application usually are of a "stacked" type in which two or more identical "bays" are stacked one over the other at a center spacing of from ½ to 1 wavelength and excited in phase.

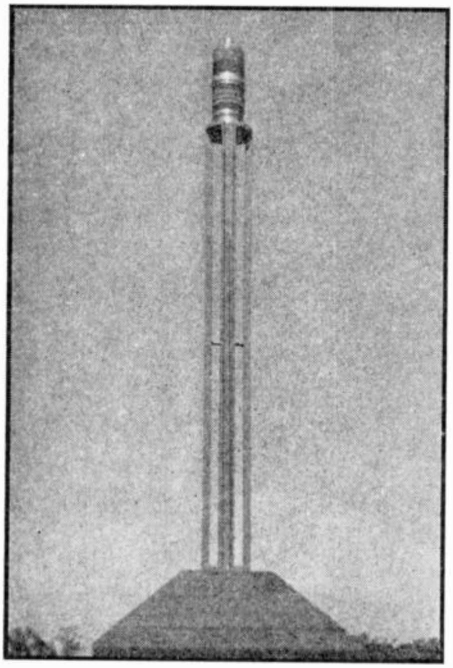
The basic bay takes various commercial forms, each manufacturer having developed a proprietary basic radiating element for



FIVE BAY WESTERN ELECTRIC "CLOVERLEAF" FM TRANSMITTING ANTENNA. THE TOWER STRUCTURE SERVES AS THE OUTER CONDUCTOR OF A COAXIAL FEED LINE. ALL RADIATING ELEMENTS ARE FED BY THE SINGLE CENTER CONDUCTOR, THE PROPER PHASE RELATIONSHIPS BEING OBTAINED BY REVERSING THE "TWIST" OF ALTERNATE BAYS.



BASIC SECTION OF RCA "PYLON" ANTENNA FOR FM BROADCAST USE, SURMOUNTED BY STANDARD BEACON. ALSO KNOWN AS A "SLOT" RADIATOR, THIS ANTENNA PRODUCES A SLIGHTLY ECCENTRIC HORIZONTAL PATTERN, WHICH, IN MOST CASES, CAN BE EXPLOITED TO PROVIDE BETTER COVERAGE. AS MANY AS EIGHT OF THE BASIC UNITS CAN BE STACKED. THE APPENDAGES ARE CLIMBING SPIKES; RADIATION IS FROM THE CYLINDER ITSELF



"TOWER" ANTENNA FOR FM BROADCASTING. THIS ANTENNA BY THE WORKSHOP ASSOCIATES, INC. ACTS AS A SECTION OF "LEAKY WAVEGUIDE". THE CHAMBER IS EXCITED BY MEANS OF TWO, CROSSED, HORIZONTAL DIPOLES AT THE CENTER OF THE SECTION. SEVERAL SECTIONS MAY BE STACKED FOR GREATER GAIN.

which he usually claims important advantages. However, they all meet the pertinent requirements imposed upon FM broadcast radiators, and generally speaking there is little to choose between them when all of the advantages and disadvantages are considered. Most of the commercial antennas have been given descriptive designations, such as "Square Loop," "Cloverleaf," "Turnstile," "Pylon," and so on. They all combine wide bandwidth or adjustable elements, low-angle radiation, and good electrical efficiency with rugged, economical construction.

Some of the antennas have a sufficiently broad response to permit use of a single size for the entire 88 to 108 Mc. band, while others have a calibrated tuning adjustment for each bay. Still others are purchased "precut" for the assigned frequency. Many of the designs have a sufficiently wide bandwidth to permit use for television broadcasting when the dimensions are suitably altered, and some are available with dimensions suitable for the TV bands.

The most commonly used center spacings between bays are approximately 0.5 wavelength and approximately 1.0 wavelength, as such spacing simplifies the phasing and feeding problems incident to stacking most of the basic radiating units. The gain due to vertical directivity is an approximate function of the overall height of the stack for all bay spacings (center to center) from 0.5 to 1.0 wavelength. For stack heights of 1 wavelength or more, the number of bays is important only as it determines the overall height of the stack. The approximate gain over one bay as a function of overall stack height is given in figure 6-17.

The most economical number of bays (or stack height) depends upon such things as the transmitter power and the elevation of the antenna site. Beyond a certain number of bays it is more economical to increase the antenna elevation or the transmitter power.

None of the various arrays is capable of an exactly circular pattern, but the distortion in every case is so slight as to be unimportant from a practical standpoint. While theoretically a perfectly circular pattern is obtainable from a circular loop carrying uniform current, or from crossed infinitesimal dipoles fed in quadrature, practical loops exhibit slightly nonuniform current distribution and a discontinuity at the feed point, and each of a pair of crossed half-wave dipoles exhibits interference directivity due to its finite length, thus distorting the pattern from a true cosine "figure eight" and making it impossible to obtain a circular resultant pattern by combining two such patterns in phase quadrature.

When an elongated service area makes some deliberate and controllable distortion of the horizontal pattern desirable, certain of the commercially manufactured FM antennas are better suited than others to modification of the normally circular pattern to one which more effectively covers the elong-

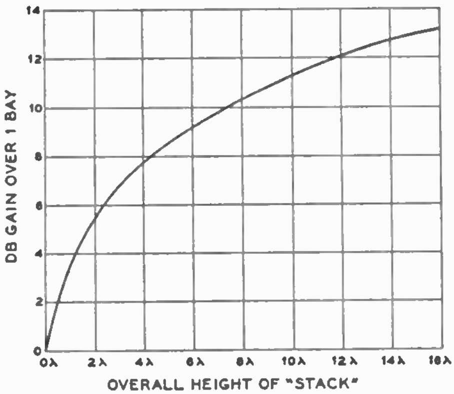


Figure 6-17.

APPROXIMATE GAIN (REFERRED TO ONE BAY) OF STACKED LOOPS (ROUND OR SQUARE OR CLOVERLEAF) FOR FM OR TV SERVICE AS A FUNCTION OF THE OVERALL STACK HEIGHT WHEN THE BAY SPACING IS BETWEEN 0.5 AND 1.0 WAVELENGTH.

The curve is not directly applicable to "crossed panel" antennas such as the Turnstile, or to "continuous bay" antennas such as the "Pylon" or "Slot" antenna. However, for stack heights exceeding one wavelength, the gain of a Turnstile or a Pylon referred to a half-wave dipole (not to one bay) is very nearly the same as for a stack of loops having the same overall height.

ated service area. For instance, antennas using circular loops are not well suited to this application, whereas a turnstile utilizing separate feed systems for each "panel" is ideally suited. By controlling the relative phase and amplitude of the current fed to the two panels, a wide variety of pattern choice is possible.

Because the fabrication of an FM broadcast antenna (exclusive of the tower) from basic materials is a formidable and time-consuming job, and at least as expensive as the purchase of a commercially manufactured "package" antenna, design details will not be discussed. Descriptive literature on these antennas is available from the various manufacturers.

HALF-WAVE HORIZONTAL DIPOLES FOR THE V-H-F/U-H-F RANGE

While broadly directional in azimuth, a horizontal half-wave dipole is not classed as a "directional" antenna. Horizontal dipole antennas are widely used for reception of FM broadcast and television signals, and for both reception and transmission by amateurs in the v-h-f/u-h-f bands. At frequencies below 150 Mc. a simple, inexpensive, and reasonably efficient arrangement is a folded dipole which uses 300-ohm Twin-Lead for both folded radiator and feed line.

A regular half-wave dipole fed with RG-8/U cable is to be preferred at higher frequencies, and can also be used in the lower portion of the v-h-f band if desired. When the application makes it desirable to do so, the frequency sensitivity of the latter antenna can be reduced by using large diameter elements, and the small amount of antenna effect can be reduced

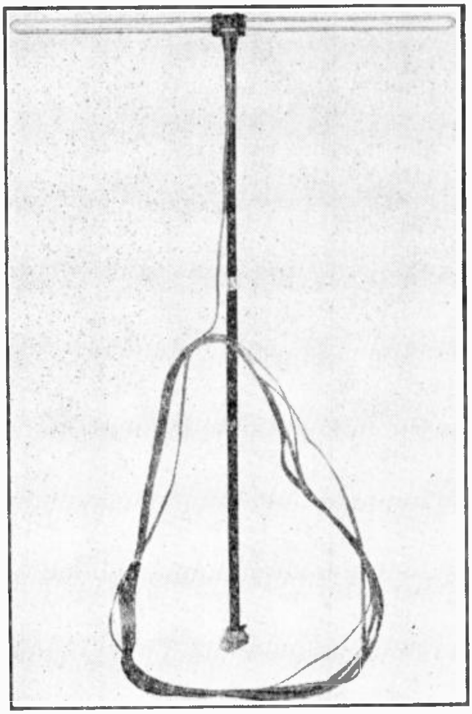


Figure 6-18.

COMMERCIALLY MANUFACTURED FOLDED DIPOLE (AMPHENOL) FOR 88 TO 108 MC. FM RECEPTION.

The use of tubular radiating elements permits "self-supporting" construction. The feed line is 300-ohm Twin-Lead.

FOLDED DIPOLE DESIGN DATA for FM and TV bands

The length dimension applies to the overall length of the unsplit element from the mid point of one end jumper to the mid point of the other end jumper. It is not the separation between these points, but rather the length of the conductor between these two points. A conventional folded dipole of typical commercial manufacture using air spaced rod or tubing elements is assumed.

FM BAND (88 TO 108 MC.)

Mid band (compromise for whole band)	57 in.
To favor high end	54 in.
To favor low end	60 in.

LOWER TV BAND (44 TO 88 MC.)*

Mid band (compromise for whole band)	88 in.
Channel 1 (44 to 50 Mc.)	118 in.
Channel 2 (54 to 60 Mc.)	98 in.
Channel 3 (60 to 66 Mc.)	88 in.
Channel 4 (66 to 72 Mc.)	80 in.
Channel 5 (76 to 82 Mc.)	70 in.
Channel 6 (82 to 88 Mc.)	65 in.

UPPER TV BAND (174 TO 216 MC.)

Mid band (compromise for whole band)	28.5 in.
Channel 7 (174 to 180 Mc.)	31.5 in.
Channel 8 (180 to 186 Mc.)	30.5 in.
Channel 9 (186 to 192 Mc.)	29.5 in.
Channel 10 (192 to 198 Mc.)	29.0 in.
Channel 11 (198 to 204 Mc.)	27.5 in.
Channel 12 (204 to 210 Mc.)	27.0 in.
Channel 13 (210 to 216 Mc.)	26.0 in.

*As this is written it appears probable that television channel 1 will be reassigned to another service. If this change is effected the optimum compromise length for the "low" band (channels 2 through 6) will be 80 inches.

To favor one or more certain TV channels, take the geometrical mean frequency or length between the two "outside" channels of those to be favored. Variation in performance over the band is the most noticeable on the lower TV band.

A convenient and easily remembered length formula for a conventional folded dipole (the length being defined as above) is: Divide the frequency in Mc. into 5555. The answer is in inches.

For basic information on the folded dipole antenna, refer to Chapter 5.

to insignificance by the incorporation of a detuning sleeve.

The reader is referred to that portion of Chapter 5 which treats high frequency dipole antennas, for basic information on these antennas. In addition, special attention should be paid to any insulators placed at voltage loops, as the losses and reactive effects increase with frequency. Preferred construction in the u-h-f range calls for support of the dipole only at or near the voltage node, as illustrated in figure 6-18. Some commercially manufactured antennas for FM and television reception utilize special designs which provide increased bandwidth or a more or less circular pattern, or in some cases both features. These antennas are considerably more complex electrically and mechanically than a simple dipole, and are much more difficult to fabricate from basic materials with ordinary tools. Therefore, it is recommended that such antennas be purchased ready constructed, as they are available at reasonable prices.

HORIZONTALLY POLARIZED DIRECTIONAL ARRAYS

Because of the slightly superior performance of horizontal polarization for most point-to-point ground applications, and because a horizontally polarized array is inherently balanced to ground and does not ordinarily require the use of detuning stubs or other special techniques in order to avoid excessive antenna effect, horizontal polarization is more or less standard for most point-to-point circuits. The only important exception involves two stations separated entirely by a sea water path and operating in the lower portion of the v-h-f band.

Most of the horizontally polarized high frequency arrays described in the preceding chapter will perform well in the lower portion of the v-h-f band, and several will provide acceptable performance well up into the u-h-f region. However, the only two that are recommended for such use without considerably modifying or expanding the construction are the 3-element parasitic array and the horizontal rhombic.

Because of the small size of half-wave dipole elements at these frequencies, it is the usual practice to increase the size of a broadside "curtain" considerably over those shown in Chapter 5, and to add a flat sheet (or grid type) reflector screen. "Billboard" or "mattress" arrays consisting of from 8 to 36 coplaner dipoles in a broadside curtain, backed up by a reflector screen, find numerous commercial and military applications, but are difficult to construct without special facilities.

THE HORIZONTALLY POLARIZED THREE ELEMENT PARASITIC ARRAY FOR V-H-F/U-H-F

The horizontally polarized parasitic array described in the preceding chapter and illustrated in figure 5-26 requires no modification for v-h-f/u-h-f use and will provide excellent performance up to 500 or 600 Mc. for fixed frequency work or over a narrow frequency band. It is particularly well adapted to rotation where requirements dictate that the direction of the beam be capable of rotation throughout 360 degrees.

THE HORIZONTAL RHOMBIC FOR V-H-F

The pertinent considerations involved in the electrical design and mechanical construction of a horizontal rhombic for the v-h-f range are the same as for a vertical rhombic and the reader is referred to the section in this chapter dealing with the v-h-f vertical rhombic, and to figure 6-13.

Rhombic Antenna for FM and Television DX Reception

Illustrated in figure 6-19 is a rhombic antenna designed for the purpose of increasing the range at which FM broadcast stations in the 88 to 108 Mc. band will provide good reception. If an ordinary dipole antenna broadside to the distant station(s) delivers a signal which at all times can be heard, even though it does not actuate the limiter and seldom if ever provides "program quality," the substitution of this rhombic antenna at the same elevation will boost the signal sufficiently to provide program quality reception all or substantially all of the time. However, if the signal at times fades clear out, due to tropospheric effects, then boosting of the signal by substitution of the rhombic may improve the quality of the reception but still will not provide consistent program quality.

The antenna is particularly well adapted to rural reception of 250-watt stations at distances at which the signal becomes too weak for good reception but does not fade badly. High powered stations usually fade badly before they become weak. The physical size of the antenna (including the four poles required) does not lend itself too well to installation in a residential district, but ordinarily will not be found objectionable for rural installations.

This antenna provides a practical signal gain throughout the FM band of considerably more than 10 db over a dipole which is working at its optimum frequency, provided that the rhombic is constructed carefully according to directions. These are given in considerable detail for the benefit of those whose technical knowledge is limited, yet can use such an array to advantage and wish to construct one themselves, without benefit of expert help.

The array does have one feature which in some cases will be a disadvantage. Top performance is obtained only over an azimuthal arc of about 10 degrees (five degrees to either side of the direction in which the array is pointed). When reception from only one station is involved, or when all of the desired *distant* stations are so closely grouped that they lie within a 10 degree arc, this high degree of horizontal directivity is not a disadvantage. But when the desired distant stations are widely separated, it will be necessary to effect a compromise when orienting the array, and possibly forego reception of some of the stations in order to hear certain others. Fortunately it is a common practice to locate the FM transmitters for a given city or town more or less in the same area.

While the discrimination of this array is good (meaning that the pick up in directions other that that in which the array is pointed is *comparatively* poor), the "off the beam" pickup still is adequate to receive nearby stations with good strength. Thus if one is located within 5 or 10 miles of one station and slightly beyond the "dipole

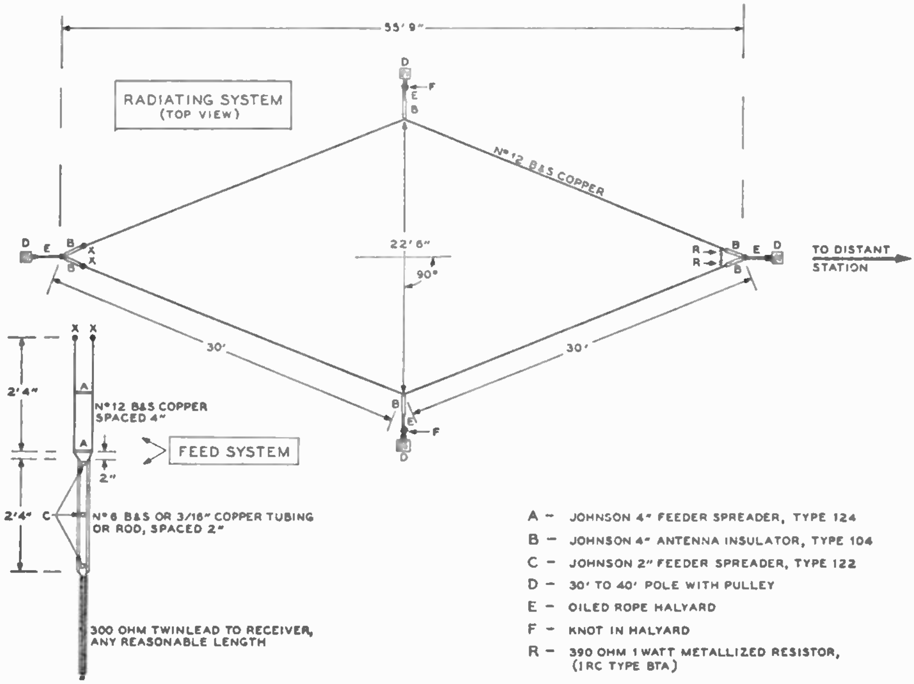


Figure 6-19.

RHOMBIC ANTENNA FOR DX FM RECEPTION AND GHOST-FREE DX TELEVISION RECEPTION.

The specified dimensions are for the 88 to 108 Mc. FM band. For the television bands the dimensions must be altered as described in the accompanying text. Any FM signal which does not fade completely out at any time when received on a half-wave dipole usually will be received with solid "program quality" on this rhombic receiving antenna elevated the same number of feet above ground. The practical signal gain is well over 10 db throughout the 88 to 108 Mc. band (compared to a horizontal half-wave resonant dipole). For television reception the antenna not only offers high, substantially uniform gain throughout either the "low" or "high" band, but also is much more free from "ghosts" (due to spurious reflections) than is an ordinary dipole or simple dipole array.

range" of another station, usually it will be possible to receive both stations satisfactorily simply by aiming the rhombic array at the distant station.

When possible the array should be elevated at least 10 feet above all trees, buildings, and other objects under or immediately surrounding the antenna, and particularly any which may be located between the front apex of the antenna and the desired stations. In any case the antenna should be elevated at least 30 feet above the actual earth, and preferably at least 40 feet. Regardless of the slope of the terrain, the array should be installed with the wires perfectly level, rather than parallel with the sloping ground.

Because of the sharp horizontal directivity, the array must be oriented very carefully with the aid of a suitable map and a compass capable of 1 degree of accuracy. The magnetic declination for the location must be known in order to determine true north. Alternatively, if the path is line of sight the pole locations may be lined up with the aid of a telescope. If neither compass nor telescope happens to be readily available, it may be possible to determine the "line of shoot" to a distant station by erecting a folded dipole well in the clear and rotating it until the center of the "null" off one end of the dipole is determined. However, this method is not practicable when the signal is fading or is very weak.

In positioning the poles, at least four feet should be allowed between each rhombic apex and its support pole. Then, to assure correct positioning of the rhombic with respect to the poles, a knot is tied in the halyard of each of the two *side* poles at such a point that when the two knots are pulled tight against their respective pulleys and the fore and aft halyards are then pulled fairly taut, the distance between the side corners is exactly 22 feet 6 inches. Each halyard should be connected to its insulator with the shortest practicable length of wire.

Any guy wires closer than 20 feet to the radiating portion of the antenna should be broken up with insulators spaced exactly 8 feet apart, in order to avoid possible resonance effects. The first guy insulator should be spaced exactly 4 feet from the pole (assuming that a set of guys is wrapped around the pole and that all guys of the set make contact with each other).

The terminating resistor is comprised of two 390-ohm 1-watt metalized resistors in series. They should be connected by means of wire having approximately the same size as the wire pigtails on the resistors. It is important that the termination be made exactly in this manner.

The illustrated construction will withstand high winds and perform well during rain, but is not suitable in locations subject to heavy icing.

In so far as practicable the Twin-Lead should be kept at least two inches away from buildings, etc. and should make the fewest possible sharp bends. The line should be made as short as possible in order to minimize line loss, but lengths up to 200 feet can be used when it permits a higher or otherwise more desirable location for the antenna.

It is strongly recommended that all joints throughout the system be soldered in order to assure quiet operation and permanence of performance.

This antenna is not suitable for locations which are subject to low field strength as a result of being in deep shadow (as for instance a receiving location immediately behind a range of high hills). The transmission path under such circumstances oftentimes deviates considerably from a direct line, and in some such cases is not constant, thus making a highly directional fixed array impracticable.

The "design center" of this array is 97 Mc. For other applications (such as television) it is necessary to use a different frequency for the design center. This can be done by scaling up or scaling down all dimensions given in figure 6-19 with the exception of the spacing between the conductors of the two matching sections. The antenna is particularly well suited to dx reception of television signals because of its wide band characteristics, and because of its discrimination against "ghost" signals caused by spurious reflections of the signal.

For the television bands the dimensions (with the exception of feed line and matching section spacings) should be altered as follows: For the "low" band (channels 1 through 6) multiply the dimensions by 1.5.* For the "high" band (channels 7 through 13) multiply the dimensions by 0.5.

SIMPLE 14-ELEMENT V-H-F BIDIRECTIONAL CURTAIN

The simple broadside curtain of figure 6-20 has almost as much gain at its design frequency as the rhombic array shown in figure 6-19, at frequencies below approximately 200 Mc. However, it is fairly frequency sensitive and therefore suitable only for fixed frequency work or use over a band of frequencies not exceeding plus or minus 2 or 3 per cent of the design frequency. Also, the length of the phasing stubs is quite critical, and they must be adjusted experimentally for best results.

The free space directivity gain is slightly over 10 db, and the practical signal gain is approximately 10 db at frequencies within

*As this is written it appears probable that television channel number 1 will be reassigned to another service. If this change is effected, multiply by 1.4 instead of 1.5.

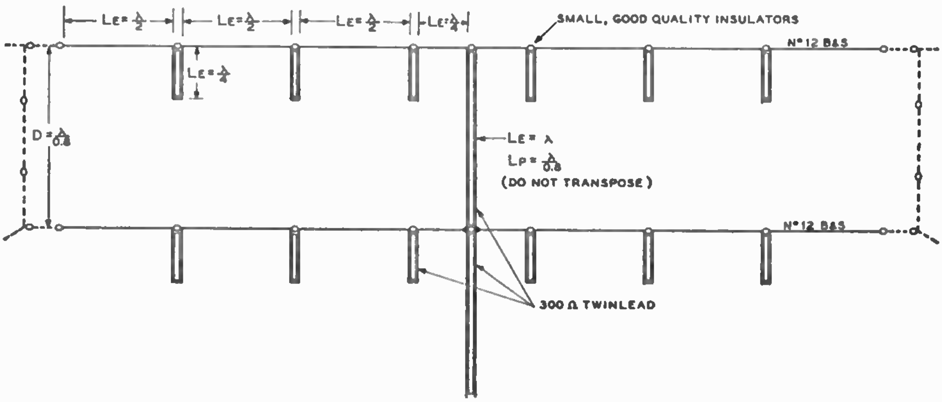


Figure 6-20.

SIMPLE, HIGH GAIN BIDIRECTIONAL BROADSIDE CURTAIN FOR V-H-F USE.

The use of Twin-Lead for transmission line and phasing stubs greatly facilitates construction of this 14 element array. It has a practical signal gain of approximately 10 db over a horizontal half-wave dipole. It is fairly frequency sensitive and therefore well suited only for fixed frequency use or for use over a narrow band of frequencies. For best results the physical length of the quarter-wave phasing stubs must be adjusted experimentally as explained in the accompanying text. The horizontal pattern of this array is quite sharp; the beam width at the half power points is approximately 19 degrees.

1 per cent of the design frequency. The bidirectional "figure 8" pattern is quite sharp when the array is properly adjusted, and the strongest minor lobes are somewhat weaker (compared to the main lobe) than those of the aforementioned rhombic.

The physical length of the half-wave radiating elements is not extremely critical, as moderate deviations from an electrical half wavelength can be compensated for by adjustment of the phasing stubs without deteriorating the performance of the array. The four, end radiating sections can be cut to 0.97 of a free-space half wavelength and the interior radiating sections to a full, free-space half wavelength.

The quarter-wave phasing stubs should be cut to 0.2 of a free-space wavelength, temporarily shorted at the bottom with a "pigtail," and then individually adjusted for optimum performance in the following manner.

Either a grid dip meter or a transmitter is required. If the transmitter power is 50 watts or more the problem is simplified. The array is erected temporarily in a clear spot, with the bottom row of elements at least a wavelength above ground. A stepladder or scaffold then is used to permit reaching the phasing stubs.

The four outside stubs are adjusted first, then the four adjacent to the outer four, and last the four inner stubs. If a moderate amount of transmitter power is available, this is done by adjusting the length of a shorting pigtail (or trimming off more of the Twin-Lead stub) until a small neon bulb glows with equal brightness when touched to the two points at which the stub under adjustment is attached to the antenna, and does not glow at all when touched to the center of the pigtail jumper at the bottom of the stub.

When the transmitter power is not sufficient to permit use of the neon bulb method, or if a grid dip meter is used, the feed line should be coupled to the transmitter or grid dip meter just tightly enough to give a pronounced change in grid or plate current when the feed line is connected. The indicating milliammeter should be placed where it can be read while making a stub adjustment and connected to the transmitter or grid dip meter by a pair of wires. Neither meter nor wires should be placed closer than one half wavelength to any radiating element. Each stub then is adjusted in the aforementioned order to a length which is arrived at as follows.

Mark the exact center of all the half-wave radiating elements except the four outer ones. Then adjust the length of the pigtail jumper on each of the outer shorting stubs until the following condition obtains on the radiating section connected thereto *on the side nearest the center of the array.* The radiator is pinched with a pair of pliers or diagonals (tightly enough to bite through the enamel should the wire be enameled) farther and farther off center until a barely perceptible indication on the meter is observed. The same procedure is repeated the other side of center.. The distance off center at which an indication is first observed should be the same both sides of center. If not, the mid-point of the two points at which an indication is observed can be moved towards the stub by lengthening the pigtail jumper, and can be moved away from the stub by shortening the pigtail jumper. After the outer stubs are adjusted, the four adjacent stubs are given the same treatment. The four inner stubs are simply made exact physical duplicates of the four intermediate stubs after the latter have been adjusted experimentally.

After all stub adjustments have been made, all connections should be soldered and the array hoisted to its regular position.

To avoid excessive deterioration of the performance of the array in wet weather, the Twin-Lead line and especially the phasing stubs should be given a light coat of Silicone waterproofing compound of the type manufactured for this purpose (such as Amphenol 307).

Once an array has been constructed and experimentally adjusted for a certain design frequency, the array can be duplicated physically with the assurance that good results will be obtained without experimental adjustment of the stubs. However, all physical details (such as the type of insulators used) should be copied exactly.

SURFACE REFLECTOR SYSTEMS USING HORIZONTAL POLARIZATION

Surface reflector systems such as the cylindrical parabolic, flat-sheet, and corner types (discussed earlier in this chapter in conjunction with vertically polarized arrays) can be used just as well with horizontal orientation and polarization. With the exciting dipole and reflector surface rotated so that the dipole is horizontal and the reflector surface maintains its relationship to the dipole, most of the directivity contributed by the reflector will be in a vertical plane.

When a paraboloid of revolution (or a truncated section thereof) is used as the reflector, the relationship between vertical and horizontal directivity is determined primarily by the aspect ratio of the reflector aperture, and not so much by the orientation of the dipole. Thus, with a horizontal dipole illuminating the "dish," it is still possible to have greater horizontal directivity than vertical directivity simply by cutting off the top and bottom of the dish so that the aperture is much wider than it is high.

CHAPTER SEVEN

Receiving Antenna Considerations

It was stated earlier that, generally speaking, a good transmitting antenna system makes a good receiving antenna system. However, it is desirable that the reader appreciate the considerations involved, because in some respects the requirements differ somewhat.

In the transmitting case, the primary consideration usually is maximum practicable signal gain. This is a function of line efficiency, radiation efficiency, and antenna directivity (including ground effects).

In the receiving case the primary consideration may be to obtain the highest possible signal-to-noise ratio, or it may be to obtain the greatest possible discrimination against undesired signals (even though it may mean a reduction in signal-to-noise ratio), depending upon the particular application.

With a reciver having very low inherent noise, or a low *noise factor*, only a small amount of signal (or atmospheric noise or other external noise) is required in order to override the inherent set noise, which, in a receiver with a very low noise factor, is mostly thermal agitation noise generated in the first tuned circuit.

Once an antenna provides sufficient receiver input that, of the total noise present in the receiver output, only an insignificant amount is represented by inherent receiver noise, then additional input (as would be obtained from a lower loss transmission line) will not significantly improve the signal-to-noise ratio. Substituting an anntena with greater directivity gain (assuming comparable line efficiency) will, under these conditions, provide an improvement in signal-to-noise ratio, but *not* because of the greater input to the receiver. The improvement is due to the fact that the more directional antenna picks up a greater amount of signal in proportion to the atmospheric noise picked up. Even if sufficient loss resistance were inserted in the feed line to reduce the amplitude of the signal to that obtained with the less directional antenna (with its smaller "capture area"), substantially the same improvement in signal-to-noise ratio would be realized.

In view of the foregoing, when an antenna is to be used only for receiving it usually can be constructed more cheaply than a transmitting antenna for the same type of service, *provided* the frequency is low enough and the receiver sensitivity high enough that *external* noise determines the receiver output noise even when a makeshift antenna is used. The saving is not so pronounced, however, in the case of a highly directional antenna, because a highly directional pattern can be obtained only with an array of considerable area, and any saving effected by economizing on the quality of transmission line, etc. may not amount to a substantial saving when the total cost (poles, land rent, etc.) is considered.

Concrete examples of the foregoing are the following two extreme cases. At 30 kc. a 70 foot vertical wire will provide about as good reception as will a large, non-directional transmitting antenna costing many thousands of dollars, and two 70 foot spaced wires used in an "Adock" arrangement for reception actually may provide better performance than a large, nondirectional transmitting type antenna utilizing an enormous "flat top" suspended between 500-foot towers.

For 1000 Mc. point-to-point work, however, a highly effective receiving antenna will resemble a transmitting antenna designed for the same application. Any attempt to economize on the construction when used for reception will result in a deterioration of the signal-to-noise ratio (and consequently the range in miles over which the circuit is reliable).

When a directional antenna is to be used for reception, more attention should be paid to the amplitude of minor lobes and to "antenna effect" than in the transmitting case, as minor lobe responses and pickup by the feed line due to "antenna currents" may seriously degrade the signal-to-noise ratio even though the power gain as a transmitting antenna is not significantly affected.

Ordinarily, less elaborate insulation may be used for a receiving antenna, because the only considerations are insulator loss and tensile strength, while in the case of a high-powered transmitter, flashover also must be considered. Another obvious saving occurs in the case of a rhombic antenna. A much less expensive terminating resistor may be used for reception than is required for a transmitter delivering several hundred watts of power. On the other hand, in geographical locations subject to precipitation static, it is desirable to employ dielectric covered wire for a receiving antenna, thus increasing the cost.

COMMON TRANSMIT-RECEIVE ANTENNA SYSTEMS

In many two-way communication systems the same antenna is used for both transmission and reception by incorporating an "antenna changeover relay," a single antenna thusly providing a high degree of performance for both purposes. However, there are practical limitations to the system as regards transmitter power and various other factors.

For "push to talk" radiotelephone applications, the use of a common antenna in conjunction with an antenna changeover relay is practicable for transmitter powers up to approximately 1 kilowatt. Suitable relays for higher power are not commonly available, partly because complications arise and partly because an antenna power exceeding 1 kilowatt seldom is employed for two-way "push to talk" communications circuits anyhow.

Unless the transmitter power is low, the relay system (both power and antenna circuits) should be arranged so that the antenna (or feed line) contacts close first and open last. This sequence is necessary in order to minimize sparking at the contacts. Even when this condition is realized, the tendency of some conventional relays to "bounce" at least once may produce a situation where the transmitter power is on and the output stage is, for an instant, not working into its normal load.

This condition may or may not cause trouble. For instance, if a class B modulator is used, the unloaded modulated amplifier may draw so little plate current as to endanger the class B modulation transformer unless conservatively rated, and also the plate tank condenser of the modulated stage may arc over if the spacing is not considerably greater than required for normal operation. However, the effect produced by a bouncing antenna relay or a relay which does not close the antenna contacts first and open them last *varies with the length of connecting line between the transmitter tank and the antenna changeover relay*.

If the electrical length of this connecting line is close to an integral number of half wavelengths, a bouncing antenna relay or a relay which does not close the antenna contacts first and open them last will permit the final amplifier to operate momentarily with practically no load, thus subjecting the modulation transformer and final tank condenser to abnormally high voltages. If the electrical length of the connecting line is close to an odd number of quarter wavelengths, a bouncing antenna relay or a relay which does not close the antenna contacts first and open them last will permit very high instantaneous voltages to exist across the line, possibly causing line or relay flashover. On the other hand if the electrical length of the connecting line does not approximate a resonant length yet is an appreciable fraction of a wavelength long, then it will act as a reactive load on the final amplifier when the relay is open, detuning the tank somewhat. Generally this is the desirable condition, as it minimizes the various breakdown possibilities previously enumerated.

For reliable operation the entire system should be so designed and so provided with sufficient safety factor that the antenna or feed line contacts on the antenna changeover relay can be opened and closed with the transmitter power on, with no detrimental effects other than excessive sparking at the relay contacts, and possibly slight overheating of the final amplifier tubes when the contacts are open.

Obtaining the proper contact closing and opening sequence is facilitated when the antenna relay is provided with a set of auxiliary contacts. These contacts can be employed to actuate the transmitter control circuits via other relays, and so adjusted that the control relays are not actuated until after the antenna or feed line contacts are closed.

Below 30 Mc., most any of the common antenna changeover relays may be used at impedance levels between 50 and 600 ohms without serious reflection due to line discontinuity. When coaxial line is employed, the outer conductor may be made common and only the inner conductor switched, in which case the relay contacts may be paralleled if the relay is of the usual d.p.d.t. type. At higher frequencies the reflection coefficient of the relay becomes more of a problem, requiring special relays for good performance above 50 Mc. or so. For u-h-f applications, special coaxial relays are available which, in effect, act very much like a "smooth" continuation of the line.

When choosing an antenna changeover relay, the maximum frequency, the power, and the impedance level should be considered. The lower the impedance level, the heavier the contacts should be for a given power, but the insulation requirements become less stringent. Low loss bakelite insulation is satisfactory for low impedance (50 to 75 ohm) circuits below 50 Mc., or for medium impedance circuits (up to 600 ohms) below about 30 Mc. For higher frequencies, or where the impedance level is much higher than 600 ohms, Mycalex, ceramic, or polystryene insulation is recommended.

In some cases electrical noise will reach the receiver via the relay control wires, particularly if the relay operates on 115 v. a.c. and is not of symmetrical construction. This pick-up can be reduced to an insignificant value by shielding all antenna relay wires carrying a.c. or d.c., by-passing these wires to the shield at the relay end, and grounding the shield. This expedient also will prevent r.f. from getting into the control or power circuits when the relay is in "transmit" position.

For ordinary c-w telegraphy applications in which "break-in" is not employed, the considerations involved in the use of a common antenna system for both transmission and reception impose somewhat less stringent requirements on the relays and transmitter equipment, because the antenna changeover relay ordinarily will be thrown only when the key is open and the final amplifier is not operative.

Break-in c-w telegraphy operation ordinarily requires the use of a separate receiving antenna unless the transmitter power is low. In some portable applications of certain low powered military equipment employing "break-in," a common antenna is utilized by means of a keyed antenna changeover relay. However, such an arrangement produces key clicks in nearby receivers over a wide frequency range, and is not recommended for general use.

TELEVISION RECEIVING ANTENNA CONSIDERATIONS

Because of the uniquely stringent requirements, antenna systems for television reception will be dealt with separately. Many complicating factors arise in connection with TV antennas which are not encountered in any other type of service. The following brief discussion of the more pertinent problems should assist the reader in understanding and solving the major difficulties which are encountered.

Susceptibility to Noise

Television requires a wide bandwidth, as a large amount of intelligence must be transmitted per unit of time. Also, no really effective receiver "limiting" is practicable with the standard amplitude modulation system employed for the video signal. This combination of wide bandwidth and no limiting means that much more transmitter power is required in order to produce a satisfactory signal than is the case for voice transmission, or conversely, the transmitter range is much less for a given transmitter power. In effect the transmitter power is "spread thin" over a wide band and therefore is not so effective in overriding background noise. The situation is aggravated by the fact that the carrier frequencies currently employed for television are susceptable to ignition interference, a particularly vexing type of interference because nothing can be done to eliminate it at the source.

"Ghosts" Due to Spurious Signal Paths (Reflections)

Cliffs, large buildings, water tanks, and other structures having large vertical surfaces often will produce a substantial signal at the receiving antenna by reflection. If the path of the "direct" signal is not a "clear shot," and the receiving antenna is in shadow, the strength of the reflected signal may be comparable to or perhaps even stronger than the direct signal. (The "direct" path is considered to be direct even when the signal must be *diffracted* over or around a hill or building.)

Indirect signals (those due to *reflection*) travel a somewhat greater distance in reaching the receiving antenna, and the time delay produces a "visual echo" which shows up on the screen as a secondary image or "ghost" which is displaced (usually in a lateral direction) from the primary image. If there are several strong reflections, there will be several secondary images produced with various amounts of displacement.

The picture quality under such conditions is intolerable, and to avoid this condition the usual procedure is to exploit the directional characteristics of the receiving antenna, enhancing them by a more elaborate antenna with greater directivity if necessary, so as to provide adequate rejection of the undesired signals. The antenna orientation is adjusted experimentally so as to alter the horizontal directivity pattern until maximum freedom from multiple images is accomplished. This is called "probing."

If one of the reflected signals produces the strongest signal, it may be desirable to "beam" the antenna so as to favor that wave path rather than the direct path, in which case the dominant reflection may be considered as a "substitute" primary signal and the image produced thereby as a substitute primary image.

"Ghosts" Due to Line Mismatch

When considerable mismatch exists at both the antenna end and the receiver end of the transmission line, one or more "ghosts" may be produced by the energy which is reflected from the receiver, back to the antenna, and back to the receiver. Usually only one secondary image will be formed in this manner, and unless the mismatch is very bad at both ends of the line, the secondary image will be considerably weaker than the primary image. But even a weak "ghost" image degrades the picture quality and is objectionable.

The lateral displacement of the image is proportional to the product of the physical length of the line and the reciprocal of its velocity factor. Even with a long line the time delay represented by the "round trip" taken by the reflected energy usually is much shorter than that resulting from indirect wave paths. Therefore a "ghost" caused by line mismatch usually occurs quite close to the main image, and has only lateral displacement, while a "ghost" caused by a reflected wave path may have most any lateral displacement, and may even be displaced vertically by a small amount if the location of the reflecting object is such as to produce a comparatively long time delay.

If the line is sufficiently short, the time delay being correspondingly short, the lateral displacement of the "ghost" will be so slight that instead of showing up as a distinct secondary image it will be manifest as a reduction in the definition or resolution of the main image. If there is a sufficiently strong double reflection, the picture quality is degraded in this manner if the product of the line length (in feet) and the reciprocal of its velocity factor exceeds approximately 20. A double image is apparent to the eye if the product exceeds approximately 70 (assuming the present standard of 30 frames and 525 lines, and an otherwise perfectly operating system).

If it is impossible to determine with certainty from the relative position of the images as to whether a "ghost" is due to multiple wave paths or to line mismatch, the question can be resolved by rotating the antenna slightly. If the ghost is due to line mismatch, the intensity will follow that of the main image, while if it is due to a reflected wave path the two images will not maximize and minimize at the same antenna orientations.

It will be noted that reference was made to considerable reflection at *both ends* of the line. This condition is necessary in order for the energy to make the round trip. If the line is perfectly matched at *either* end there will be no reflection at that end, and no energy can make a complete round trip regardless of how high the reflection coefficient may be at the other end. The "ghost" energy is dependent upon the line attenuation and the *product* of the reflection coefficients, and no difficulty will be encountered if the impedance match is either fairly good at both ends of the line or else very good at one end. Also, the greater the line attenuation, the greater the tolerable reflection coefficient product, because dissipating the reflected energy in the line is just as effective in reducing the "ghost" intensity as is cutting down on the amount of reflection. However, this should not be construed to mean that a line having high attenuation is desirable. The line attenuation should be as low as economic considerations permit, so as to deliver the strongest possible signal to the receiver, and any "line reflection ghosts" should be eliminated by improving the impedance match at one or both ends of the line.

Home television receivers of commercial manufacture are designed so as to have an input impedance which is quite constant over the entire frequency range covered by the receiver. This impedance usually is substantially resistive and corresponds closely to the surge impedance of a standard transmission line (ordinary 300 and/or 75 ohms). Therefore no difficulty should be experienced from "line mismatch ghosts" *when the recommended transmission line is employed,* regardless of the amount of mismatch at the antenna end of the line. Excessive mismatch at the antenna end of the line nevertheless is undesirable, however, because it reduces the amount of signal delivered to the receiver.

Necessity for a "Low Q" Overall System

To avoid distortion of the picture, the overall system including antenna, feed line, and receiver input circuit must have insignificant frequency discrimination over the pass band required for the video signal. This amounts to about 5 megacycles with present standards, and is easily met for a single channel. However, when the entire "low" or the entire "high" television band is to be covered with a single system, the antenna presents somewhat more of a problem. The Q of the line can be kept low by providing a close match at the receiver termination (a necessity anyway if the antenna is not closely matched to the line, in order to avoid "ghosts" from line reflections). But if the reactance of the *antenna* changes *too rapidly with frequency* over the range covered by *any one channel,* the picture quality will be impaired on that channel even with a perfect match between the line and the receiver.

This means that an antenna which is satisfactory for operation on a single channel under optimum conditions may or may not be satisfactory on other channels. The signal may not only be weaker on other channels due to mismatch, but distortion may result from an excessive reactance slope. It is for this reason that certain high gain, highly directional, close spaced arrays using parasitic elements are well suited for use on only one channel, or perhaps two or three adjacent channels.

The receiver input circuit is not so much of a problem, as it not only can be broadened out to have a bandwidth which will accommodate the video sidebands without excessive discrimination, but also can be tuned to the channel in use along with the oscillator and other front end stages. The only undesirable effect of deliberately limiting the Q of the input circuit is to limit the voltage step up between the line and the grid of the input tube. Fortunately at the higher frequencies (between 174 and 216 Mc.) where it is highly desirable to deliver the greatest possible signal to the grid of the first tube, the bandwidth of a television channel is a smaller percentage of the carrier frequency, and higher Q can be tolerated in the first tuned circuit.

Installation Considerations

The number and magnitude of the installation difficulties encountered vary with the number of stations to be received, the receiving location, and the distance from the transmitter. A suburban residence, in flat country, with no tall buildings nearby, set well back from the street, located not more than 15 or 20 miles from the only television transmitter in the area and blessed with a clear "line of sight" shot to the transmitting antenna offers but little more of an installation problem than does that of a good antenna for FM broadcast reception, but such a combination of favorable factors is comparatively rare.

When such a combination of favorable conditions does exist, a simple folded dipole mounted 15 or 20 feet above the roof top will do an excellent job if the transmission line impedance is comparable to the input impedance of the receiver.

The problem becomes somewhat more difficult when there are several stations in the area. The dipole should be cut or adjusted to the geometrical mean frequency, and if two stations differ in compass direction by roughly 90 degrees, careful experimental orientation of the dipole (and its horizontal directivity pattern) may be required in order to obtain satisfactory reception on all channels. In some cases two separate dipoles and feed lines in conjunction with a selector switch may be required.

A typical metropolitan location almost invariably presents difficulties which require care, patience, and at least an elementary knowledge of the considerations involved, particularly when there are numerous surrounding buildings capable of producing strong reflections and one or more of the desired stations are far enough away that the field strength is not adequate to override the ignition and other electrical noise common to such locations. The usual procedure in the more difficult locations is to employ some sort of directional array and to "probe" for the best antenna position and orientation while an assistant appraises the relative picture quality and relays the information via sound powered or battery telephone.

If it is impossible to find a compromise position and orientation which offers satisfactory performance on all of the desired channels, or the array has too much "Q" to permit operation on all of the desired channels, it may be necessary to employ two or more antennas and feed lines in conjunction with an antenna selector switch.

If the receiving location is in a large, high building having a flat roof, and the desired stations all lie in approximately the same direction, high gain and good directivity can be obtained over the entire low band or the entire high band with a single antenna by resorting to a v h-f rhombic array such as described in Chapter 6.

Such an elaborate installation is justified for demonstration purposes in the case of a store in a metropolitan area, as it is only with an especially good antenna installation in such a location that picture quality comparable to that which is obtained in most residential or suburban areas with a simple folded dipole can be realized. However, a rhombic array of this type is just about as difficult an installation job as is a combination of three or four parasitic arrays, and while slightly superior when the desired stations all lie in approximately the same direction, the rhombic array is not practicable when they do not.

In locations of high ambient electrical "noise," special attention must be given to the feed line and to whether it is operating under balanced conditions. A special 300-ohm "ribbon" or moulded pair line having a third, neutral conductor midway between the two feed line conductors is available, and its use sometimes will result in a substantial reduction in the amount of noise picked up by the line and fed to the receiver. The center, neutral wire should be grounded to the receiver chassis. This has no adverse effect upon a correctly operating line, as the center wire is in a neutral plane.

The noise pickup of conventional two-wire 300-ohm ribbon or "Twin-Lead" often can be reduced considerably by twisting it so that it makes one complete twist every 10 or 12 inches for the entire length.

When the location is especially bad, a coaxial line or shielded pair offers the best solution. This type of line is considerably more expensive than ribbon type line, but the expense is justified if the ambient electrical noise field through which the line passes to reach the antenna is unusually high. Many receivers are provided with optional 300-ohm symmetrical ("balanced") and 75-ohm asymmetrical ("unbalanced") input. This permits the option of inexpensive 300-ohm ribbon line in locations of low ambient electrical "noise," or the use of RG-11/U coaxial line in difficult installations.

In extremely noisy locations it may be advisable to suppress the worst of the electrical interference at its source, provided the latter is accessible.

Very Wide Band TV and FM Antennas

Special or "trick" variations of the dipole antenna which give creditable performance over the entire 44 to 216 Mc. range are available commercially for use in locations where the performance of such an antenna is adequate. Generally speaking, such an antenna is of necessity a compromise affair, and in most cases the average performance over the frequency range is definitely inferior to that of an ordinary folded dipole operating on the channel for which it is resonant.

At one or more certain frequencies in its prescribed frequency range the performance of such an "extra wide band" antenna may be comparable to that of a *resonant* folded dipole, but at other frequencies in the range it will be decidedly inferior. One exception is the "double skeleton cone" antenna, which can be made to give performance comparable to a *resonant* folded dipole over a frequency range of two or more octaves. However, the directivity increases considerably as the frequency is increased. The latter characteristic may or may not be an advantage, depending upon individual circumstances.

This characteristic, a substantial change in directivity over such a wide frequency range, is characteristic of all the common "extra wide band" antennas for FM and TV service. In some cases (such as the above example) the figure 8 lobes merely sharpen up as the frequency is increased. In others the directivity pattern will take on an altogether different configuration, with more than two major lobes.

CHAPTER EIGHT

Coupling to the Antenna System

COUPLING THE TRANSMITTER TO THE ANTENNA SYSTEM

General Considerations

There are three primary considerations involved in coupling a radiator or transmission line to a transmitter: (1) The antenna or feed line must be made to "look like" a resistive load to the final amplifier if it is not already purely resistive or very nearly so. (2) The final amplifier should "look" into the desired magnitude of load resistance. (3) The overall antenna system (including the coupling network) should discriminate strongly against spurious frequencies present in the transmitter output, to the extent that they will not be radiated with sufficient strength to be capable of causing interference to other services.

The various lumped or distributed reactances which may be required in order to accomplish the above usually are referred to collectively as a "coupling network" or "antenna coupler." The functions oftentimes are overlapping. For instance, the same coil and capacitor may be used to cancel a reactive component, convert the resistive component to the desired value, and at the same time provide attenuation of harmonics.

The function of transforming the resistive component of the antenna or feed line input impedance to the desired value of load resistance and cancelling the reactive component sometimes is referred to as "matching the impedance." Actually the impedance is not "matched" to anything; it is simply made resistive and transformed to the particular value which causes the final amplifier to draw the desired plate current. However, the term *matched* might be construed to mean that the resistive component is "matched" to (made equal to) the desired value of the load resistance.

It should be kept in mind that when a transmission line is employed there is nothing that can be done at the *transmitter* end of the line that will have any effect upon the standing wave ratio on the line (assuming a fixed frequency). It also should be kept in mind that when a tuned line is employed which is many wavelengths long and exhibits a high SWR, both the reactive and resistive components at the sending end will vary widely with a moderate change in frequency, thus complicating the design of an antenna coupler intended to work over an appreciable frequency range with such an antenna system. The same situation may arise in the case of a long wire (harmonic) radiator even when a short tuned line is employed.

There is a rather common misconception that when a purely resistive load is coupled to the plate tank of a radio-frequency power amplifier the tank condenser setting will not be disturbed, that if retuning is required the load must be reactive, and that if retuning is not required the load is nonreactive. This is not so. Even if a pure resistance is connected directly across some of the turns of the plate tank coil the resonance setting will change slightly because the mutual coupling between turns of the coil is not unity. The effect is most noticeable when a very low value of load resistance is clipped across a very small portion of the tank, and least noticeable when the same power is delivered to a high value of load resistance clipped across practically the whole tank coil. Thus, for a given power delivered to the load, the detuning effect increases with the transformation ratio.

The use of an untuned coupling coil to couple the load resistor also will effect the resonance setting of the plate tank condenser slightly, because of the inductance of the coupling coil and because of the leakage reactance introduced as a result of incomplete coupling between the tank coil and the untuned coupling coil. However, in view of the fact that a small amount of reactance coupled into the plate tank can be "tuned out" and is not detrimental, the matter of slight detuning of the plate tank is not of practical importance. It should be kept in mind, however, that slight detuning of the plate tank does not necessarily indicate a reactive load, and that such a check should not be used as a method of determining whether standing waves exist upon a transmission line.

Regardless of the method of coupling to an antenna system, precautions should be taken to see that under no conditions can d-c plate voltage appear on the feed line or antenna. While r-f burns are painful and may be serious, they rarely are fatal. But d-c plate voltage of even a few hundred volts is potentially lethal. When the r-f coupling circuits do not provide inherent blocking of the d-c plate voltage from the antenna and series plate feed is used, air or mica dielectric blocking condensers should be inserted in series with the leads to the antenna or feed line. These should have sufficient capacity that they do not significantly affect the r-f operation of the circuit (except when the coupling network calls for a critical value of capacitive reactance in series with the feed line), and sufficient voltage rating to provide considerable safety factor. Further protection is afforded by grounding the antenna system (and feed line if employed) to d.c. This is desirable anyhow, as a d-c ground affords increased lightning protection. If it is not convenient to attach a d-c ground at a point of zero r-f potential, the antenna or feed line can be grounded through an r-f choke(s) effective over the frequency range involved.

DIRECTLY-FED ANTENNAS

At the lower frequencies (especially in the l-f and v-l-f ranges) it is common practice to bring the radiator right in to the transmitter rather than use a transmission line. This is especially true in the case of shipboard installations, or in any installation where a single Marconi-type antenna must be used over a wide range of frequencies because of lack of space to erect numerous resonant antennas.

From an electrical standpoint, the simplest method of feeding such an antenna consists of adding enough lumped series reactance to cancel that of the antenna (either inductive or capacitive, depending upon whether the antenna is too short or too long) and coupling the antenna inductively (electromagnetically) to the transmitter in a manner which permits the coupling coefficient (and therefore the load on the transmitter) to be adjusted. The latter may be accomplished either by varying the number of effective turns on the pickup coil or by varying its position with respect to the transmitter tank coil. Such an arrangement is illustrated in the figure 8-1. The antenna can be "resonated" to the operating

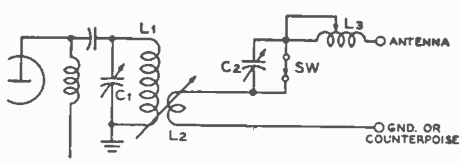


Figure 8-1.

SIMPLE METHOD OF COUPLING DIRECTLY TO AN ANTENNA SYSTEM WHICH MAY BE REPRESENTED BY A COMPLEX IMPEDANCE.

This arrangement may be used to couple to any complex impedance having a resistive component which does not exceed a few hundred ohms, and a reactive component which does not exceed a few thousand ohms. This means that in practice it may be used effectively for coupling to any antenna-against-ground system in which the electrical length of the radiator exceeds 1/20 wavelength and does not approximate an integral number of half wavelengths.

frequency by shorting out either L_3 or C_2 and varying the reactance of the other, depending upon whether the natural resonant frequency of the radiator is above or below the operating frequency. Adjustment is facilitated if a variometer or other continuously variable inductor is substituted for the tapped coil, L_3, though practically the same thing can be accomplished by utilizing somewhat more series inductance than the resonant value and then making the final adjustment with C_2, just as though the antenna were too long and L_3 were shorted out.

If the electrical length of the radiator (including ground lead) is equal to or very nearly equal to an odd number of quarter wavelengths, C_2 and L_2 may be dispensed with, the coupled reactance being such under these conditions that it can be "tuned out" by means of C_1. This arrangement usually requires that the physical length be optimized by "pruning," and limits the use of the antenna to a narrow range of frequencies.

While the arrangement of figure 8-1 is simple electrically, it possesses certain disadvantages. The movable pickup coil, L_2, introduces mechanical complications; little attenuation of certain harmonics is provided; and difficulty will be experienced in coupling to a radiator having an overall length approximating an integral number of half wavelengths.

The reactive component of the feed point impedance of a directly fed antenna also may be eliminated and the transmitter loaded to any desired plate current by the insertion of an adjustable reactive network of the "T," "Pi," or "L" type between the antenna feed terminals and the transmitter output stage (figure 8-2). This avoids the movable coupling coil L_2 of figure 8-1, thus permitting a simpler mechanical arrangement. Also, it allows proper feeding of a wider range of complex impedances. However, when the length of the radiator is but a very small fraction of a wavelength, there are practical limitations to the use of such an impedance transforming network. Under these conditions the series reactance presented by the antenna is many times the resistance, and a suitable coupling network

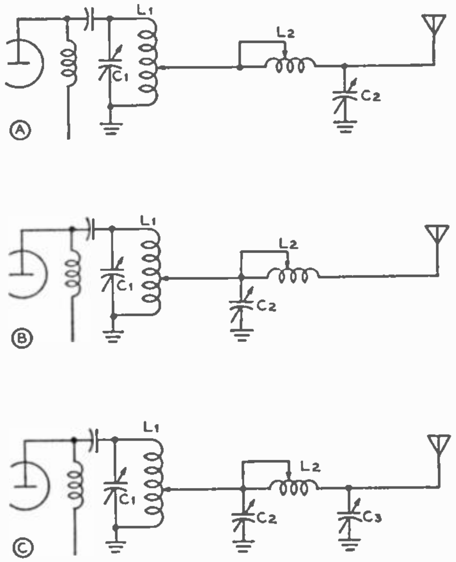


Figure 8-2.

REACTANCE NETWORKS USED AS ANTENNA COUPLERS.

The pi network at C will couple properly to a wider range of complex impedances than will the simpler L networks of A and B.

becomes awkward of construction because of the very high voltage and current encountered under such conditions. Also, the losses in the reactive elements (particularly the inductor) become excessively high. A point is reached where it becomes more economical to increase the length of the antenna than to provide a suitable coupling network (except of course in mobile and marine installations where the antenna length is definitely limited).

Usually there will be a variety of reactive network configurations which can be used to provide proper coupling of a transmitter to a certain load. However, simplicity and minimum response to harmonics usually dictate one of the "low pass" networks shown in figure 8-2, even though the harmonic attenuation of such a network often leaves much to be desired. The choice between A and B is determined by whether the resistive component of the antenna impedance is to be stepped up or down, and whether the reactive component is inductive or capacitive.

At certain critical antenna lengths difficulty may be experienced in obtaining the desired impedance transformation with an "L" network such as shown at A and B. In such cases the "pi" network of figure 8-2C will permit the proper coupling to be obtained. Correct adjustment of all three networks is indicated by a condition where attachment of the network and antenna to the final amplifier causes the amplifier to draw the desired plate current without *greatly* changing the setting of the amplifier tank condenser at which resonance (minimum plate current) occurs. It is not necessary to eliminate the reactive component of the antenna impedance *completely* with the coupling network, as a small amount of series reactance (which is equivalent to a high value of shunt reactance) can be "tuned out" by means of the plate tank condenser, C_1, with no undesirable effects.

The pi-section filter of figure 8-2C does not always provide effective harmonic attenuation, because the impedance of the antenna at a particular harmonic may be considerably different from the impedance at the fundamental, and the impedance at the harmonic may be such that the pi section filter is working under unfavorable conditions and offers little or no attenuation of the harmonic. This also applies to an L-section low-pass filter. Under certain conditions it is even possible for a particular harmonic to be accentuated, rather than attenuated. However, when used with an antenna which is essentially aperiodic, such as a terminated rhombic, the pi-section filter does a very good job of suppressing harmonics.

The reactive networks of figure 8-2 can be made to have high discrimination against a particular harmonic, regardless of the nature of the load impedance at that harmonic, by resonating the series arm L_2 to the harmonic by means of shunt capacity across L_2, the L/C combination being proportioned to give the desired inductive reactance at the fundamental frequency. The same thing can be accomplished by resonating the shunt arms (C_2, C_3) to a harmonic by means of series inductance, the L/C combination being proportioned to give the desired capacitive reactance at the fundamental frequency.

The shunt arm or arms may be resonated to trap the same harmonic as is rejected by the tuned series arm, so as to give additional attenuation of a particular harmonic, or they may be tuned to different harmonics, thus providing moderate attenuation of two different harmonics. In any case the correct initial adjustment of the network is rather tedious when the elements are tuned to provide harmonic attenuation, and therefore such an arrangement is practical only for an installation in which the adjustments can be logged (as for instance a transmitter used with a particular antenna on a few definite frequencies).

In low powered portable equipment the plate tank (C_1 L_1) of figure 8-2 C sometimes is omitted, the pi network serving at the same time as a tank "flywheel" and coupling network. However, for general use such an arrangement is not recommended. Under certain load conditions and with practical values of C_2 and C_3 the adjustments which give the desired loading will result in an effective "Q" which is not adequate for good "flywheel" effect and proper operation of the amplifier, and harmonic radiation will be excessive.

COUPLING TO RESONANT (TUNED) SYMMETRICAL FEED LINES

The considerations involved in coupling to a resonant two-wire open line *having a high SWR* are similar to those discussed under the directly fed antenna, except that the coupling network must be symmetrical. A symmetrical pi section network suitable for use with such a line is illustrated in figure 8-3. Symmetrical configurations of the L network will correspond in a similar fashion. If the "high Q" feeder is of such an electrical length as to produce a resistive component of impedance at the transmitting end not exceeding a few hundred ohms over the frequency range involved, a symmetrical version of figure 8-1 may be employed. This calls for a variable inductance and variable capacitor in each leg, unless the antenna is to be used only over a narrow frequency range and the feed line

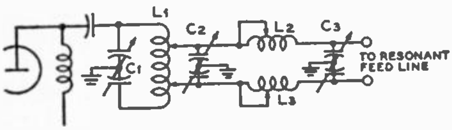


Figure 8-3.

SYMMETRICAL CONFIGURATION OF PI SECTION COUPLING NETWORK FOR USE WITH TWO-WIRE RESONANT FEED LINE.

The considerations involved in coupling to a resonant, two-wire open line are similar to those involved in coupling a directly-fed antenna-ground system (figure 8-2C) except that the network must be made symmetrical.

is cut to a resonant length which produces a current loop at the coupling coil, L_2, (in which case no inductor or capacitor is required in either leg). Regardless of the configuration, when using a symmetrical arrangement the elements in each leg should be adjusted so as to provide the same reactance in each leg.

When the line SWR is only moderate, the link coupled universal coupler described in the following section may be used, regardless of line length, and is recommended over the previously described arrangements. It also may be used when the SWR is high, if the line length is such that the overall system (radiator and feed line) is approximately resonant to the operating frequency.

THE LINK COUPLED UNIVERSAL ANTENNA COUPLER

The link coupled universal antenna coupler illustrated in figure 8-4 is suitable for use with symmetrical two-wire lines of the open type, shielded type, twisted pair type, or moulded pair type, regardless of electrical length, provided that the line is either flat or has only a moderate SWR. When the SWR is *high,* the coupler also can be used *if* the line is of such length that the series reactance is not excessive with respect to the resistive component, or in other words, if the line length is such as to make the antenna plus feed line combination approximately resonant at the operating frequency.

The chief advantage of this coupler is that when properly adjusted (and correct adjustment is a simple procedure), *all* harmonics and other possible spurious frequency emissions (such as subharmonics due to "leak through" from a multiplier string) are attenuated simultaneously and automatically so long as they are not too close to the desired transmitting frequency. In this respect it is far superior to the coupling arrangements of figures 8-1, 8-2, and 8-3.

The split stator variable condenser C_2 and the antenna tank coil L_4 should be similar to what would be appropriate for a split stator (balanced) plate tank circuit for the frequency and transmitter output power involved. In fact, if the final amplifier plate tank is symmetrical to ground (instead of as shown in figure 8-4), the plate tank and antenna tank components may be made identical, provided that the particular type coil employed is of such construction as to permit adjustment of feeder taps without having to "bite" through enamel insulation and without danger of shorting out turns with the clips.

The inductor L_5 and the condenser C_3 are not required when the line is substantialy flat, or when the line SWR is high and the length is such that the overall system (radiator plus feed line) is approximately resonant to the operating frequency. One or the other compensating reactance is clipped across the line only when the series reactance presented by the feed line is so great that the reflected shunt reactance cannot be "tuned out" by means of C_2 without the assistance of such compensating reactance (L_5 or C_3). The maximum inductance of L_5 need not be nearly as great as that of L_4, but the maximum capacitance of C_3 should be considerably larger than that of C_2. Smaller wire may be used for L_5 than for L_4, and C_3 need not have nearly as much spacing as C_2. The maximum capacitance of C_3 must be somewhat greater when a 70-ohm line (or thereabouts) is used than when a 600-ohm line (or thereabouts) is used, but less spacing is required.

Large values of compensating reactance (high values of shunt inductance or low values of shunt capacitance) are not required; instead the line reactance may be "tuned out" by means of C_2. The *minimum* value of shunt reactance (in ohms) required at C_3 to permit satisfactory operation regard-

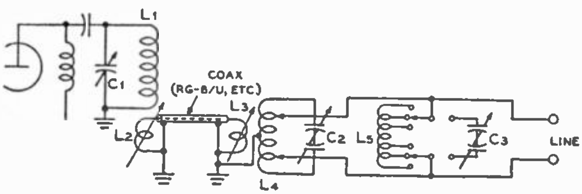


Figure 8-4.

LINK COUPLED UNIVERSAL ANTENNA COUPLER.

This coupler discriminates strongly against magnetic coupling of harmonics and other spurious frequencies by virtue of the selectivity of the resonant antenna tank circuit, and also minimizes stray electrostatic coupling of harmonics. The coupler as shown is suitable for coupling to two-wire lines having moderately high SWR, regardless of line length. When used with a flat ("nonresonant") line, the compensating reactances L_5 and C_3 may be omitted. Adjustment procedure is covered in the accompanying text.

less of line length is approximately equal to the characterisitic impedance of the transmission line divided by half the VSWR, the VSWR being expressed as an integer and not as a fraction. For instance, the minimum value of reactance required for a 600-ohm line having a VSWR of 6.0 is approximately 600 divided by 3, or 200 ohms. The amount of capacity which will exhibit this amount of reactance at a given frequency may be determined quickly by referring to one of the various reactance-frequency charts available in popular handbooks on radio and electronics.

The maximum capacity required at C_3 becomes quite high and therefore inconvenient with a combination of a low impedance line, a high SWR, and a low frequency. However, as the only line commonly operated at a high SWR is the open-wire line (having an impedance of approximately 600 ohms), and as this type of line with a high SWR is widely employed only in the h-f range, the required maximum capacitance at C_3 ordinarily is not inconveniently high. For example, assuming a frequency of 3 Mc., if the VSWR on a 600-ohm line is 6.0, as in the previous example, the maximum total shunt capacity of C_3 need be only about 300 $\mu\mu$fd. (600 $\mu\mu$fd. per section). A condenser of 500 $\mu\mu$fd. per section (a more common size) would be found satisfactory in most cases, as some reactance can be "tuned out" by means of C_2. At lower frequencies it is permissible to employ good grade mica padding condensers which may be switched in and out as required in conjunction with a variable condenser to cover the in-between values.

The use of a length of grounded low-impedance coaxial cable for link coupling between the final amplifier plate tank and the antenna tank virtually eliminates harmonic excitation of the antenna system via stray capacitive couplings. The link line may be made any reasonable length, but ordinarily the antenna coupler will be placed close to the final plate tank so as to permit reaching all controls and adjustments shown in figure 8-4 from one position, thus facilitating the tuning up process.

While the coupler will give pretty good results even when adjustments are not precisely correct, it is desirable that the coupler be operated correctly. No attempt should be made to "broad band" the coupler, as this is not compatible with maximum suppression of amplifier harmonics and other spurious frequencies. The correct initial tuning procedure is as follows.

Without the feed line attached to L_4, reduce the coupling between L_3 and L_4 to a very low value, and with some sort of resonance indicator (such as a neon bulb or an inductively coupled panel bulb or flashlight bulb attached to an insulating rod or stick), log the resonant settings of C_2 as a function of frequency or oscillator dial setting. If two resonance humps are observed, the coupling is too tight. If necessary, reduce the coupling between L_1 and L_2. With the antenna tank unloaded the Q is quite high, and only a small amount of coupling is required in order to exceed critical coupling and thereby obtain double resonance peaks.

If the transmitter power is low, it may be possible to observe resonance of the antenna tank by a flick of the amplifier plate current at the resonant setting of C_2. However, this system is not recommended when the transmitter power is high, as sufficient coupling to cause an indication of the plate meter oftentimes will cause the antenna tank condenser to arc when the antenna tank is unloaded.

Next, substitute for L_3 a noninductive dummy load having a resistance approximately equal to the surge impedance of the coaxial link line. Then determine how much coupling is required between L_1 and L_2 in order to load the amplifier to the desired plate current at the center of the frequency range covered by L_1. If the power handling capability of the dummy load resistor is inadequate, the adjustment can be made at reduced plate voltage (and correspondingly reduced plate current). The coupling between L_1 and L_2 then need not be changed unless a single coil is used at L_1 to cover a very wide frequency range by means of a large condenser at C_1. In the latter case it may be desirable to optimize the coupling between L_1 and L_2 at two frequencies, one towards the upper limit and one towards the lower limit. If there is much difference in the coupling, then one adjustment should be used for half the frequency range and the other adjustment for the other half. If there is little difference in the adjustments, it is permissible to split the difference and leave the coupling alone until the coil is changed.

The foregoing coupling adjustment is highly desirable when the coaxial line length is an appreciable fraction of a wavelength, but not so important when the length of the coupling line is an insignificant fraction of a wavelength. The idea is to prevent the link line from contributing undesirable reactive effects, either by operating it under substantially "flat line" conditions or by making it very short. The adjustment for "flat" operation of the link line is not especially critical, and once determined one simply substitutes the antenna coupler for the dummy resistor and makes all further loading adjustments by varying the coupling between L_3 and L_4, and by adjustment of the feeder taps on L_4. When the loading is adjusted in this manner so as to cause the amplifier to draw normal plate current, the coaxial line automatically is working under the desired conditions, (assuming, of course, that the reactive component of the main feed line impedance has been taken care of properly, as will be described in a later paragraph).

The foregoing presumes that the final amplifier always is run at the same plate input, or that if the power input is varied it is done by varying the plate voltage (as by a primary auto-transformer), because if the power input is varied by altering the coupling between L_3 and L_4 while maintaining the plate voltage constant, the link line no longer is terminated under the most favorable conditions. However, this is not too important a consideration if the section of coaxial link line is made very short in terms of wavelength; in this case the loading may be varied by manipulation of either or both links with no significant difference in results.

The antenna coupler is adjusted as follows. First tap the feeder wires on L_4, symmetrically with respect to the center tap, at points which embrace a small fraction of the total number of coil turns. Keeping C_1 always at resonance (tuned for minimum plate current), determine the resonant setting of C_2. If none can be found, clip either some inductance or capacitance across the feed line and check again. If no resonant point can be found with one kind of reactance, try the other. However, if the feed wires are clipped across a sufficiently small number of turns, resonance usually can be found at some setting of C_2 without adding shunt reactance across the line even when the line is highly reactive. Once a resonant point is found on C_2, the idea is to check the dial setting of C_2 against the "free" resonant setting for that frequency (as previously described), and then add enough compensating shunt reactance across the line to cause the two resonant settings to be very nearly the same.

If the loaded resonant setting represents more capacity at C_2 than does the unloaded resonant setting, the reactive component of the feed line impedance is inductive, and shunt capacitance in the proper amount must be employed to compensate for it. If the reverse condition prevails (less capacitance being required at C_2 for resonance when the feed line is attched to L_4), the reactive component of the feed line is capacitive, and shunt inductance in the proper amount must be employed to compensate for it. Once C_2 is made to resonate at the same setting regardless of whether the feed line is attached to L_4, the loading on the final amplifier may be adjusted to the desired value by varying the coupling between L_3 and L_4 and by varying the position of the feed line taps on L_4.

There will be numerous combinations of coupling and tap spread which give the desired loading. The most desirable combination is the one with the smallest tap spread which still permits the desired loading to be obtained, provided that this combination does not result in excessive heating of L_4 or arcing of C_2. If arcing or overheating occurs, a greater tap spread must be employed.

Decreasing the tap spread increases the loaded Q of the antenna tank circuit, making it more selective and increasing the discrimination against harmonics, but it also increases the circulating current in the coil and the peak voltage across the condenser. There is no point in carrying the idea of a high loaded Q to an extreme, as a loaded Q of 10 or more in both the final plate tank and the antenna tank will attenuate harmonics to the extent that interference to other services is very unlikely unless there is appreciable radiation from the power wires due to stray couplings, or considerable radiation directly from the tank circuits. The latter may occur with unshielded tank coils of large diameter.

If the antenna plus feed line combination is not highly frequency sensitive, the coupler can be used over a moderate frequency range without changing the adjustment of the compensating reactance, as a certain amount of reactance may be "tuned out" by means of C_2. If the antenna system is highly frequency sensitive (as in the case of a harmonic radiator fed with a long feed line having a high SWR), both the compensating shunt reactance (L_5 or C_3) and the antenna tank condenser (C_2) will have to be readjusted even for a frequency change of but a few per cent. The loading (coupling and possibly the tap spread) also will have to be readjusted, thus making the procedure for readjusting the coupler for a slightly different frequency rather a complicated procedure when the antenna system (radiator plus feed line) is highly frequency sensitive.

It is recommended that the antenna changeover relay (if used) be inserted in the coaxial link coupling line rather than in the regular feed line, especially if the feed line has a high SWR. The grounded outer conductor can be made common to both transmitter and receiver, and the relay contacts of a conventional d.p.d.t. antenna relay paralleled to carry the current of the inner conductor with less sparking and heating. The selectivity of the antenna tank affords improved image rejection.

COUPLING TO NONRESONANT (FLAT) LINES

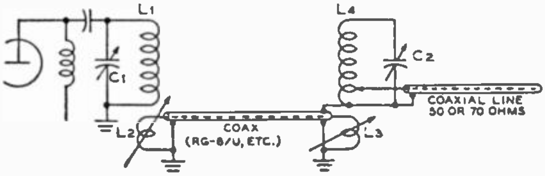
When a two-wire open line is operated under substantially "flat line" conditions (as indicated by a VSWR of approximately 1.5 or less), the compensating reactances L_5 and C_3 of figure 8-4 may be dispensed with and the coupler used as previously described.

Coaxial lines almost invariably are operated under flat line ("nonresonant") conditions. When a coaxial transmission line is employed, the harmonic reduction afforded by the selectivity of a resonant antenna coupler of the type shown in figure 8-4 (less the compensating reactances L_5 and C_3) may be realized by modifying the configuration of the coupler so as to make it "single ended" as shown in figure 8-5. Adjustment is the same as for the symmetrical coupler of figure 8-4, except that there is only one feeder tap, and the "tap spread" embraces that part of the coil between the tap and

Figure 8-5.

SINGLE-ENDED VERSION OF THE LINK COUPLED UNIVERSAL ANTENNA COUPLER, FOR USE WITH COAXIAL FEED LINE.

This coupler is similar to that of figure 8-4 except that it is not symmetrical with respect to ground. No shunt compensating reactances are provided, as it is assumed that the coaxial feed line always will be operating under substantially "flat line" conditions. A considerably higher C/L ratio should be employed in the antenna tank when it is single ended.



the ground end of the coil.

Because of the single-ended configuration and the fact that the impedance into which the antenna coupler "looks" always is low, the coupler of figure 8-5 should employ a much higher C/L ratio for the antenna tank than that of figure 8-4. While the antenna tank condenser preferably should have at least four times as much capacity as the effective series capacity of C_2 in figure 8-4, the condenser need be no larger physically for a given transmitter power, because the maximum voltage across the antenna tank coil will be much less (assuming a comparable order of loaded Q in either case).

Simple Inductive Coupling to Flat Lines

In certain low power applications, particularly in the v-h-f and u-h-f range, little or no discrimination against harmonics is required other than that provided by the plate tank circuit of the oscillator or amplifier stage feeding the antenna. In such cases the simple coupling arrangement of figure 8-6A may be employed for coupling to a line which is operating under substantially "flat line" conditions. Best results will be obtained with a comparatively small coupling coil tightly coupled, as an attempt to use a larger coupling coil loosely coupled usually is doomed to failure. Therefore the coupling coil should have the fewest number of turns which will provide the desired loading over the frequency range involved.

Somewhat improved performance will be obtained if the inductive reactance of the coupling coil is "tuned out" by means of a series or shunt condenser as shown at figure 8-6B and 8-6C. Oftentimes when the line is flat and the impedance transformation ratio is high, the desired loading cannot be obtained, regardless of the number of coupling turns, unless the coupling coil is resonated in this manner. This is especially true when the final plate tank is "low-C," but also may be true of a "high-C" plate tank if the physical arrangement of the coupling coil is such that tight coupling to the plate coil cannot be attained.

If the line is substantially flat, a single adjustment of C_2 will be found satisfactory over a fairly wide frequency range for usual values of C_2 and L_2. When the coupling coil is tuned in this manner it is possible to obtain the desired loading with looser coupling, thus simplifying the mechanical arrangement of the variable coupling coil for a given range of load impedances.

Because looser coupling to the coupling coil will suffice when the B or C arrangement is employed, some reduction of harmonic radiation is realized. However, the reduction is not nearly so great as is obtainable with a properly designed link coupled universal antenna coupler. The (slight) discrimination against harmonics obtained with the B and C arrangements also may be explained on the basis of the selectivity of the coupling circuit. However, as the effective Q is low with usual values of L_2 and C_2, the harmonic discrimination provided

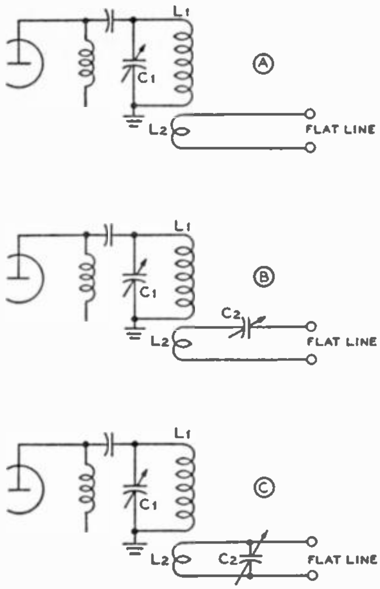


Figure 8-6.

SIMPLE INDUCTIVE COUPLING TO A FLAT (NONRESONANT) LINE.

The coupling method shown at A is considerably simpler than those previously illustrated, but is suitable only for low power applications in which a moderate amount of harmonic radiation can be tolerated. Either a coaxial or a symmetrical ("balanced") line may be used. The B and C arrangements permit tuning out the reactance of the coupling coil, and provide some discrimination against harmonics (other than that provided by the output tank of the transmitter). When the B arrangement is used with a two-wire (symmetrical) line, it is recommended that a condenser be put in series with each conductor of the line. When the C arrangement is used with a two-wire line, it is recommended that a split stator condenser be used to preserve symmetry. However, if the impedance of the two-wire line is low (say, 70 ohms or so) the arrangements illustrated may be used without an objectionable amount of unbalance, particularly at the lower frequencies. Because the coupling loop ordinarily will not possess much inductance, the condenser C_2 usually must be rather large for the frequency involved (as compared to values of capacity commonly used in tank circuits).

by the B or C arrangement ordinarily is not very great.

When a coaxial line is employed, one side of the coupling coil (and the outer conductor of the coaxial transmission line) may be grounded. When a symmetrical line is used, one side also may be grounded if the line impedance is low (say 70 ohms or thereabouts) without detrimental effects. When an open-wire line is employed, grounding one side of the coupling coil will produce excessive unbalance. Grounding the center of the coupling coil is recommended in this case, but this presents mechancial complications. Grounding the coupling coil (either the center or one end) is not particularly important when the coupling coil is placed at the "cold" end of a single-ended plate tank coil as illustrated in figure 8-6, but is recommended when a split stator (balanced) tank circuit is employed with no r-f bypass to ground at the center of the tank coil. In any case the coupling link should be placed at the r-f voltage node (usually the center or one end of the tank coil).

With a coaxial or low impedance two-wire line, the B and C arrangements may be used as illustrated. When a two-wire open line is employed, however, this may result in an objectionable amount of line unbalance, particularly at the higher frequencies. The balance may be preserved by inserting a series condenser in each conductor of the line in the B arrangement, or substituting a split stator condenser at C_2 in the C arrangement. Grounding the rotor of the split stator condenser will substitute for a grounded center tap on the coupling coil when a feed line ground is necessary in order to avoid capacitive coupling of harmonics.

Shortcomings of Commercial "Swinging Link" Tank Coils

One popular style of "swinging link" plug-in transmitter tank coil utilizes a single variable coupling link for a set of coils covering a frequency range which may exceed 10 to 1. The number of turns in the coupling coil or "link" obviously must be a compromise, because the optimum number of turns varies with the particular plug-in coil in use and with the particular load impedance employed. A variable tap on the link ordinarily involves mechanical complications. Therefore the following expedients are recommended.

When difficulty is experienced in coupling to a low impedance line or load at the higher frequencies because of an excessive number of link turns, a series condenser employed as shown at figure 8-6B usually will remedy the situation. When difficulty is experienced in obtaining sufficiently tight coupling to an open wire line at the lower frequencies because of insufficient link turns, the coupling can be increased by employing a shunt condenser to resonate the coupling link, as shown at figure 8-6C.

One line of swinging link tank coils employs a plug in link arrangement. This is a highly desirable feature as it permits the optimum number of link turns to be used with each different tank coil.

COUPLING THE ANTENNA SYSTEM TO THE RECEIVER

EFFECT OF COUPLING UPON RECEIVER SENSITIVITY AND OVERALL GAIN

Receiver input circuits designed to couple the antenna or transmission line to the grid of the first tube take a wide variety of forms, as the requirements of the input circuit depend upon many factors such as frequency, receiver bandwidth, equivalent noise resistance and input resistance of the first tube, receiver gain, versatility, image rejection requirements, simplicity, and economy. These factors are related in a complicated manner which is properly a phase of receiver design, and therefore a detailed discussion is not within the scope of this book. However, it is desirable that the reader have an appreciation of the factors involved, even though not concerned with the details of receiver design. With this in mind the following brief treatment is presented, but first it is important that there be no confusion between *sensitivity* (which is a function of signal to noise ratio) and *gain* (which is a matter of amplification).

Generally speaking, the design factors are less critical and the requirements less stringent at the lower frequencies, as the atmospheric noise level is sufficiently high that certain of the factors mentioned need not be optimized nor emphasized in order to obtain good performance.

In the v-h-f range and above, it is desirable to choose for the first tube a tube having high input resistance (this is largely a function of electron transit time) and a low equivalent noise resistance. The input circuit is designed to have low losses and usually is designed to match (approximately) the antenna resistance to the resultant resistance of the input circuit as shunted by the input resistance of the first tube. When the antenna is matched to the receiver input in this manner the *gain* will be maximum.

Under certain conditions a slight improvement in the signal-to-noise characteristics (sensitivity) of the receiver can be obtained at a sacrifice in gain by deliberately mismatching the impedances. The maximum possible improvement is determined by (1) the ratio of the first tube's input resistance to its equivalent noise resistance, (2) by the ratio of the electronic input resistance of the first tube to the impedance of the input circuit (with first tube connected but antenna disconnected), and (3) by the normal or "flat line" attenuation of the transmission line. If the latter is high, no improvement may be possible, as any increase in the sensitivity of the receiver may be completely offset by increased line loss due to increased SWR resulting from the deliberate mismatch.

Deliberate mismatching to improve the S/N ratio always is accomplished by increasing the coupling beyond that which gives matched impedances and maximum gain. In every case the loss in gain due to the mismatching is at least as great as the improvement in the S/N ratio, and in some cases will be much greater. However, the loss in gain is not serious, as it may be made up elsewhere. More important is the fact that overcoupling reduces the selectivity of the first tuned circuit and therefore the image rejection. And still more important are the undesirable effects resulting when a mismatched transmission line is used. For instance, the frequency sensitivity becomes excessive, a particularly objectionable characteristic in wide band applications such as television.

In view of the foregoing the usual practice is to match the impedances throughout, the required coupling adjustment being determined by observing what degree of coupling provides maximum *gain*. With an otherwise perfect receiver (one using tubes having infinite input resistance and zero equivalent noise resistance) the matched impedance condition will give a noise level 3 db above the antenna noise (a "noise factor" of 3 db), while adjusting the coupling for optimum signal-to-noise ratio will give a noise level substantially that of the antenna noise (a noise factor of approximately zero db) if a lossless line is assumed.

With practical tubes having finite input resistance and equivalent noise resistance, the maximum improvement obtainable by adjusting the coupling for maximum S/N may be somewhat more or somewhat less than 3 db, but in general the maximum improvement obtainable by deliberate mismatching will not exceed 2 or 3 db. The maximum improvement obtainable by adjusting the coupling for maximum S/N is greatest with tubes having a low equivalent noise resistance and at frequencies sufficiently high that the input resisistance is low compared to the impedance of the *unloaded* tuned grid circuit. In any case the improvement hardly can be considered worthwhile when the disadvantages inherent in deliberate mismatching are considered.

The foregoing discussion assumes that the gain of the first stage (whether an r-f amplifier or mixer) is great enough that the noise contributed by the following stages can be ignored. This is a legitimate assumption when good design practices are followed.

It also has been assumed that the transmission line, if used, has a characteristic impedance which approximates or is converted to the antenna impedance. This is the most desirable condition, but if it does *not* obtain, then it should be noted that a "matched impedance" condition at the receiver end is obtained when the receiver input circuit is made to "look like" a resistance having a value which is *not* equal to the surge impedance of the line, but rather a value which will cause the reflected impedance at the antenna end of the line to match the resistance of the antenna. In this case the line will exhibit standing waves when the receiver coupling is adjusted for maximum gain (representing a "matched" condition), just as it will exhibit standing waves when the antenna is used for transmission.

It should be kept in mind that "matched" as used here means that the *antenna* is matched to the load (taking into consideration the impedance transforming characteristics of the interconnecting feed line and any intervening impedance transformers). The load at the receiver end required to do this does *not* "match" (equal) the characteristic impedance of the line unless the latter approximates or is converted to the antenna impedance.

It is interesting to note that when the normal or "flat line" loss of the line is high (say, 10 db or so), the impedance transforming characteristics of the line are largely "washed out," and the optimum load at the receiver end for maximum gain will approximate the surge impedance of the line regardless of any mismatch between the line and antenna. If the reasons for this are not clear, refer to the discussion of the effect of line loss on SWR in Chapter 2.

It should be kept in mind that when a transmission line is connected to a receiver, the load is the *receiver*, rather than the antenna. The antenna in this case may be considered as a *generator* having an internal impedance equal to the impedance of the antenna at the point where it connects to the transmission line.

COUPLING METHODS

Because the principle of reciprocity applies to all reactive networks and transformers, any of the coupling methods illustrated in figures 8-1 to 8-6 inclusive may be used "in reverse" for coupling an antenna or transmission line to the grid of the first tube in a receiver. The plate element then becomes the grid of the first tube in the receiver, and the plate blocking condenser and r-f choke shown in these figures may be ignored.

The two receiver input circuits most commonly used in the h-f range and above, and also widely used at lower frequencies, are illustrated in figure 8-7. The similarity of figure 8-7A to that of figure 8-6A, as well as that of figure 8-7B to that of figure 8-2C (with L_1 and C_1 of the latter omitted) is apparent. The coupling coil L_2 of figure 8-7A sometimes is provided with an "antenna trimmer" consisting of a series condenser of a value appropriate to the frequency involved. If C_1 is ganged with other "front end" condensers, a small trimmer condenser, accessible from the front panel, sometimes is used in parallel in order to permit the operator to compensate for the frequency detuning effect of the antenna upon the first tuned circuit.

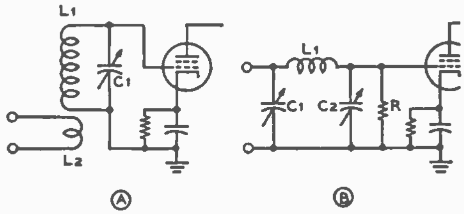


Figure 8-7.

COMMON H-F, V-H-F, AND U-H-F RECEIVER INPUT COUPLING CIRCUITS.

The arrangement at A is suitable for use with either symmetrical or asymmetrical feed lines or antennas, and commonly is employed in commercial receivers of the tunable type. When used with an asymmetrical antenna system (such as one using a coaxial line), one of the L_2 terminals is grounded. On receivers intended for use with antennas having a wide range of impedances, the turns ratio usually is made such that the input impedance is in the neighborhood of 300 ohms, this value being a good compromise value. The arrangement at B permits adjustment of the impedance transformation, thus allowing an accurate match regardless of the antenna or feed line impedance, but introduces an additional control and presents ganging difficulties. For this reason it is useful primarily for fixed frequency receivers or for receivers used only over a narrow frequency band. To avoid unbalance of the feed line when a symmetrical feed line is employed with the B arrangement, a half-wave section of coaxial line may be used as a "phase inverter" as described in Chapter 2.

The arrangement at A is suitable for use with either a symmetrical or asymmetrical line or antenna, while that at B is recommended for use only with an asymmetrical antenna or line (such as a Marconi antenna or a coaxial line). While a two-wire open line or other symmetrical feed system can be used with the pi-section arrangement at B, unbalance of the line will result, causing reduced discrimination of a directional antenna system and increased local noise pickup by the feed line.

The B arrangement can be used over a narrow frequency band without unbalancing a symmetrical feed line, however, by the use of a section of coaxial line having an electrical length of one half wavelength and connected as a "phase inverter" as described in Chapter 2.

The resistor R of figure 8-7B, except in broad-band applications, serves only as a d-c return for the grid and has a high value (one megohm or thereabouts). The B arrangement is most popular for fixed frequency receivers, while that shown at A is the one commonly used in "tunable" receivers.

Commercial receivers covering a wide band of frequencies often are provided with a "compromise" value of input impedance which may not provide a very good "match" with certain antenna systems. The arrangement of figure 8-7B may be employed to advantage to couple the line to the receiver under such conditions. If the line is symmetrical, the balance can be maintained by using a phase inverting section of "coax," as previously mentioned, or a symmetrical pi-section network can be employed, using two inductors and two split stator condensers (as in the transmitter antenna coupler of figure 8-3).

A simple modification of the conventional coupling arrangement of figure 8-7A is shown in figure 8-8. Because coaxial lines usually run from 50 to 75 ohms characteristic impedance, and two-wire symmetrical lines from 300 to 600 ohms impedance, it is impossible to obtain a good match with both when using a fixed number of turns as in figure 8-7A. With the figure 8-8 arrangement an impedance ratio of 4 is obtained by the use of the grounded center tap, making a good match possible with either a coaxial line or a 300 to 600 ohm symmetrical line. The grounded center tap also will minimize feeder pickup when a symmetrical line is employed, as capacitive coupling of in-phase line currents to the grid circuit is minimized when the antenna coil is grounded properly. Grounding one side of L_2 in figure 8-7A when a symmetrical line is employed produces line unbalance and is about as objectionable (in receiving applications) as is the capacitive coupling which is reduced thereby.

Faraday screens (electrostatic shields which are so designed as to have negligible effect upon the electromagnetic coupling) formerly were used rather extensively in order to avoid electrostatic coupling of undesired energy into a receiver (such as in-phase currents on a transmission line), but they no longer find much use. Coupling via means of a grounded low-impedance winding (such as in figure 8-8) accomplishes substantially the same result, is simpler mechanically, and does a sufficiently good job for most applications.

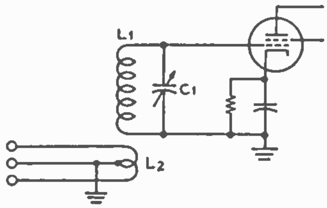


Figure 8-8.

RECOMMENDED "COMPROMISE" RECEIVER INPUT CIRCUIT FOR GENERAL USE.

This input circuit permits ganging with other "front end" circuits and a reasonably good match with either a 300-600 ohm open or Twin-Lead line or a 50-75 ohm coaxial line. When a coaxial line is employed, only half of L_2 is utilized, giving an impedance transformation of four as compared to the whole antenna coil, and therefore a better match. The grounded center tap also minimizes "antenna effect" on the line when a two-wire symmetrical line is employed, by minimizing capacitive coupling and at the same time maintaining line balance. The circuit is simply that of figure 8-7A with a grounded center tap provided on L_2.

INTERFERENCE REDUCTION

Interference to reception takes many forms. Several of these are tied up closely with the subject of coupling to the receiver, if the term "coupling" is considered in its broadest sense. One example is electrical interference picked up and coupled into a receiver as a result of "antenna effect" upon the feed line. Another is electrical interference coupled into a receiver via the power supply wires as a result of inadequate isolation or decoupling of these wires where they enter the receiver. Another example is interference from a powerful, nearby transmitter as a result of inadequate rejection by the antenna coupling circuits. This may take the form of "i-f leak through," image interference, or cross modulation. In many cases the most effective method (and sometimes the only effective method) of eliminating such interference is to provide additional selectivity *ahead of the first tube,* which means that the filter or resonant circuits incorporated for the purpose of such rejection become part of the antenna coupling or "input circuit."

Some of the more common types of interference and effective methods of avoiding or dealing with them will be discussed briefly.

Generally speaking, the effective measures boil down to first minimizing the interference at the source (as for instance reducing the sparking of a commutator by dressing it down), "bottling up" the interference at the source by shielding, and minimizing radiation and conduction of noise from the source by filtering all incoming and outgoing leads. Then, by means of shielding and filtering, pickup at the receiver is confined to the antenna proper, which is placed as far from the noise as possible (using a properly operating transmission line to connect the antenna to the receiver). The foregoing applies primarily to aperiodic electrical noise which is not *intentionally* radiated.

When the interference is from an intentionally radiated signal, nothing much can be done at the source (assuming that the transmitting equipment is operating properly and that the interference is not a spurious emission); but such signals almost invariably are confined to a comparatively narrow band of frequencies, and selective rejection circuits can be employed at the receiver to attenuate them. These consist of parallel resonant L/C circuits in series with the line or antenna, and series resonant L/C circuits in shunt, as shown at figure 8-9A.

When several interfering signals make the use of such rejection circuits or "traps" impracticable, a low pass filter oftentimes may be used to provide attenuation of all of the undesired signals, *provided* that the interfering signals all are somewhat higher than the highest desired frequency. (See figure 8-9B.) The filter may be designed with a cutoff frequency equal to the highest desired frequency in accordance with filter design data presented in various standard electronics handbooks. The filter is simply inserted between the antenna or feed line and the receiver input terminals, the number of sections being determined by the amount and sharpness of attenuation required.

The most desirable arrangement of all is the incorporation of sufficient signal frequency tuned circuits ahead of the first tube to prevent cross modulation, and a sufficient total number of signal frequency tuned circuits to prevent image troubles, with all of the signal frequency tuned circuits made an integral part of the receiver and tuned by a single condenser gang.

Packaged interference filters suitable for isolating power supply wires are available commercially at reasonable prices. Basically these consist simply of r-f chokes and bypass condensers having a current carrying capacity and voltage rating appropriate to the job, but there is more to obtaining good attenuation characteristics over a wide band of frequencies than just hooking together some condensers and chokes. Special bypass condensers are employed, and the dressing of the various leads is quite critical. Therefore, when the maximum possible attenuation is desired, it is recommended that commercial filters be employed. Descriptive literature which will help in choosing a suitable filter is available from several manufacturers.

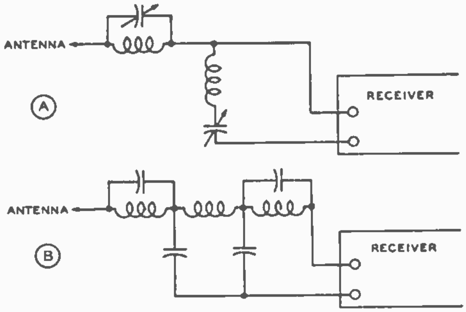


Figure 8-9.

EXAMPLES OF TWO COMMON METHODS OF INCREASING DISCRIMINATION AGAINST STRONG UNWANTED SIGNALS.

The "wave trap" arrangement at A utilizes resonant L/C circuits to provide a high series impedance and low shunt impedance to a particular frequency. Additional resonant L/C circuits may be added in cascade to increase the attenuation. By the use of traps tuned to different frequencies, two or three particular frequencies may be rejected. In general the L/C ratio of the series traps should be high and that of the shunt traps low. The low pass filter arrangement at B attenuates all frequencies above the cutoff frequency. Increasing the number of sections increases the amount and sharpness of attenuation. Design parameters depend upon not only the cutoff frequency but also the input impedance of the receiver. Design data will be found under a treatment of wave filters in most any standard handbook on electronics.

NOISE REDUCTION IN VEHICULAR INSTALLATIONS

Elimination of electrical noise in vehicular receiving installations presents special problems, and therefore will be treated separately.

The measures required in order to reduce "car noise" to an insignificant magnitude depend upon the frequency range involved and upon the particular vehicle. Even vehicles of the same make and model may require different treatment, and a vehicle that has been effectively "suppressed" may develop noise and require further treatment after it has been driven a few thousand miles. A certain amount of "cut and try" is unavoidable, but certain general rules can be followed to advantage. Also, certain measures can be taken without first checking on the necessity for them, because they invariably are required for a noise-free installation.

Military vehicles and aircraft, as well as the more expensive commercial aircraft, are factory equipped with "anti-noise" electrical systems and otherwise thoroughly suppressed when radio installations are anticipated, and ordinarily no corrective measures need be taken when radio equipment is installed. It is only necessary to comply with the installation instructions and follow up with the proper maintenance. Therefore, the following discussion applies only to ordinary gasoline powered passenger cars utilizing conventional ignition systems.

The various noises may be broken down into the following categories: (1) Ignition interference (originating primarily in the high tension system); (2) Wheel static (including tire static); (3) Generator "whine"; and (4) Voltage regulator "hash."

Noise from the High Tension System

The high tension ignition system is the most vicious offender. First make sure that all spark plugs are in first class condition and correctly adjusted, and that the distributor rotor gap is no greater than absolutely necessary. Then eliminate all unnecessary gaps at the usual "pinch fit" terminals on the high tension leads by soldering the ends of each high tension wire to the terminating lugs or jacks.

The best procedure is thorough shielding of all high tension wiring, and this is recommended in police and other installations where the cost can be justified. Special shielded high tension wire, shielded spark plugs, and shielded connectors for the purpose are available commercially.

However, a pretty good job can be done on the high tension system without resorting to complete shielding. First, replace all high tension wiring with new ignition wire of the type having a steel inner conductor. Old, "leaky" wire will give trouble regardless of the type. If high and low tension wires are run side by side through cabling channels, isolate the low tension wires from the high tension wires by running the low tension wires outside the tube or channel, and keep low tension wires separated at least six inches from high tension wires at all points where practicable unless one or the other is effectively shielded.

If the ignition coil is located behind the dash instead of on the motor side of the bulkhead, the high tension lead from the coil should be shielded as far as the bulkhead, and the shield grounded at both ends.

The main object is to keep the r-f energy generated at spark gaps in the high tension system confined to the motor compartment. The low tension leads will pick up some electrical noise from the high tension system by induction and distribute it throughout the car over the light wires, etc. if precautions are not taken to prevent it. The 6-volt wire from the ignition coil to the dash switch is about the worst offender.

Fortunately, induced high tension noise can be kept out of the low tension system by the same chokes and bypass condensers that are employed to isolate noise generated in the low tension system, as will later be described, if the systems are well isolated.

Spark plug suppressors and a distributor suppressor will aid in reducing high tension noise, particularly if the high tension system is not completely shielded, without noticeably affecting the car performance *if the suppressors are of the proper design.* They should have from about 7000 to 15,000 ohms resistance, and all spark plug suppressors should have approximately the same resistance when checked on an ohmmeter. On some cars (such as certain model Fords) a universal type distributor suppressor will not fit, but a special suppressor is availiable from authorized dealers handling the particular car.

RECEIVER AND ANTENNA CONSIDERATIONS

The receiving antenna preferably should be mounted at the rear of the car, to keep it as far from the various spark gaps in the ignition system as possible, and connected to the receiver via a coaxial feed line. If the receiver is to be used only on a fixed frequency, the receiver itself can be mounted in the trunk of the car, thus minimizing feed line loss and also reducing stray pickup.

If not already provided within the receiver, a shielded low pass filter should be inserted in the "hot" 6-volt lead to the receiver right where it enters the chassis (assuming that the vibrator or dynamotor power supply is integral with the receiver). If the power supply is separate, shielded low pass filters should be placed in both the hot heater lead and the B plus lead where they enter the receiver chassis. Suitable filters are available commercially as packaged units.

The amount of work involved in obtaining virtually complete elimination of interference on weak signals is minimized if the receiver is equipped with a "noise silencer" of the peak chopper type. Such noise silencers are available commercially as packaged units for external attachment to receivers not so equipped.

Quieting the Low Tension System

All low tension wires should be bypassed to the bulkhead at the point where they pass through the bulkhead. The bypass condenser leads should be as short as possible. The optimum value of bypass condenser to be used for the purpose depends upon frequency. Above approximately 30 Mc. it will be found that .01-μfd. or .006-μfd. mica condensers do a better job than paper condensers of larger capacity. In the AM broadcast band, paper tubular condensers of about 0.25-μfd. capacity will be found satisfactory. Metal cased "auto radio" condensers are ideal for applications requiring the large capacity of a paper condenser. At the higher frequencies the length of the condenser leads or "pigtails" should not be over an inch long for best results. The 6-volt wire from the ignition coil to the switch is especially troublesome, and should be bypassed again at the switch.

Eliminating "R-F Pipelines"

All control rods, metal tubing, the steering post, and anything else (other than electrical wiring) that passes through the bulkhead without making a positive ground to the bulkhead should be jumpered to the bulkhead by means of a flexible grounding "pigtail" of shield braid of the shortest practical length. The motor itself should be grounded to the frame in the same fashion at each rubber engine mount, assuming that the engine is floated on rubber mounts. The exaust pipe also should be grounded at two or three points to prevent conduction of noise to the rear of the car. If the exhaust pipe and muffler are not so grounded, they should be jumpered to the frame by means of short pieces of shield braid.

Quieting Generator Whine

Generator noise can be greatly reduced by bypassing the "hot" (armature) lead to ground right at the generator terminal in accordance with the bypassing data previously given. Above about 15 Mc. considerable improvement can be realized by inserting an r-f choke in series with the hot (armature) lead right at the generator. The choke should be wound with heavy wire (no. 10 B&S for a 30-ampere generator) in order to minimize voltage drop. When the receiver is used only on a single frequency or over a narrow band of frequencies, stubborn cases of generator hash can best be handled with a resonant r-f choke, the choke being resonated to the signal frequency by means of a small, compression type variable condenser (trimmer) placed across it. This also reduces the amount of inductance required in the choke, and therefore the voltage drop.

Voltage Regulator Hash

Voltage regulators utilizing vibrating points to control the amount of current flowing through the generator field (and therefore the charging rate) will produce a "hash" which becomes increasingly objectionable with the reception frequency.

A large condenser placed across the regulator will interfere with the operation of the regulator and also may cause short vibrator life. However, at the higher frequencies, where the hash is the strongest, considerable reduction in the amplitude of the hash will be obtained with a bypass condenser having a capacity as low as .001 μfd., and this amount of capacity ordinarily will not have an adverse effect upon the operation or life of the regulator. The condenser should be placed right at the regulator, from the generator field wire to ground. All other terminals on the regulator (assuming it is of the conventional type) may be bypassed to ground with whatever capacity is necessary in order to do the job.

If a .001-μfd. mica bypass on the regulator output does not keep r-f hash off the lead to the generator field, the bypass can be augmented by an r-f choke placed right at the regulator. Unlike the choke in the armature lead, this choke does not have to carry heavy current, and a compact choke wound from no. 18 B&S ordinarily will do the job satisfactorily. If reception is on but a single frequency or a narrow band of frequencies, the choke may be shunted by a compression trimmer and tuned to the signal frequency in order to increase its effectiveness.

Wheel Static

Wheel static is caused by "leaking" of static electricity generated by the rotating tires and brake drums. It is worst in a very dry climate while driving on a dry, paved road. A special conducting powder is available which, when injected into the inner tubes in the proper amount in accordance with accompanying instructions, virtually eliminates trouble from "tire static." The remaining wheel static results from the fact that the presence of wheel bearing grease precludes a constant, low resistance "ground" contact from the wheel to the rest of the car through the bearings.

To get around this condition, a "static collector" should be installed on each wheel. While so-called universal types are manufactured, it is best to purchase the ones made for the particular make car involved, from an authorized dealer. This is especially true of the rear wheel type. Not all car dealers carry static collectors in stock, and oftentimes they must be ordered from a zone warehouse. The same applies to tire powder, though of course in the latter case the powder need not be purchased from a dealer handling a particular make car.

Residual Noise

When all the foregoing corrective measures have been applied, a listening test may or may not show that a satisfactory reduction in car noise has been accomplished. If not, further bonding and bypassing is indicated, and it is largely a cut and try process. If the addition of a certain bypass or the grounding of a certain member results in any improvement at all, it should be left on, as it may be found later when all other noise sources have been eliminated that removing the bypass or ground makes a much greater difference than is the case when it is contributing only a small portion of the total noise.

It is especially important that the hood of the car be effectively grounded at several points, particularly if the high tension system is not shielded, because a poorly grounded hood will permit radiation from the high tension wiring to reach the antenna. An examination of the hood latches will show whether they must be augmented, worked on, or modified to ensure a good ground contact.

Sometimes it will be observed that when the car is in motion there is a crackling or grinding noise present, after all other noise has been substantially eliminated. This usually can be traced to one or more scraping contacts somewhere in the car body or frame. It is cured by tightening or bonding the offending joint or joints.

COUPLING TO TELEVISION RECEIVERS

As explained in Chapter 7, special considerations are involved in the coupling of a feed line to a television receiver. The line must be terminated in an impedance very nearly equal to the characteristic impedance of the line in order to avoid the necessity for an exact match at the antenna end of the line. Also, the avoidance of pickup by the transmission line is of prime importance, because of the fact that a television receiver is highly susceptible to electrical "noise."

The Q of the input circuit also is of major importance. If it is too high, the video sidebands will not be amplified faithfully, and distortion of the picture will result.

When a grounded grid triode is employed as the first r-f amplifier, tube noise is not as much of a problem, and the importance of delivering the maximum possible signal to the first tube is minimized. This permits use of an aperiodic input circuit using only a non-inductive resistor from the cathode (the "hot" input element) to ground. For balanced input a "dummy" resistor is connected to the other input terminal. A resistor-load input circuit suitable for use in the 44 to 216 Mc. range with either 300-ohm Twin-Lead or RG-11/U coaxial line is shown in figure 8-10.

300-ohm Twin-Lead is connected to terminals 1 and 2, giving a symmetrical 300-ohm resistive termination. The continuity of the folded doublet with which Twin-Lead ordinarily is used effectively puts R_1 and R_2 in parallel for d.c., which gives an effective cathode bias resistance of 75 ohms.

If three-wire neutral-ground type of 300-ohm ribbon line is employed, the connections are the same except that the neutral (center) wire is connected to chassis ground, terminal 3.

If RG-11/U coaxial line is used, terminals 1 and 2 are jumpered and connected to the inner conductor, and terminal 3 is connected to the outer conductor. This makes the line termination and cathode bias resistance both 75 ohms.

Thus it is seen that when 300-ohm ribbon line is employed, the line is terminated in a balanced, symmetrical load equal to its surge impedance. And when 73-ohm coaxial line is employed, the line still is properly terminated (well within acceptable tolerances). In either case the cathode bias resistance is 75 ohms, a suitable value for a 6J4.

The condenser C_1 is made approximately equal to the cathode-to-ground capacitance of the 6J4, and maintains circuit symmetry when a symmetrical feed line is employed. For the same reason R_1 and R_2 should be

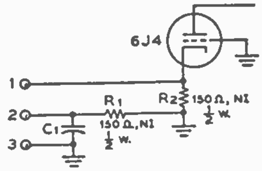


Figure 8-10

RESISTOR-LOAD INPUT CIRCUIT FOR A GROUNDED GRID R-F STAGE, SUITABLE FOR A TELEVISION RECEIVER

This circuit offers a simple solution to the problem of obtaining a substantially constant, resistive load for the transmission line over the wide range of frequencies employed for television broadcast service. Either two-wire or three-wire (neutral ground) 300-ohm ribbon may be employed by connecting the "hot" conductors to terminals 1 and 2, and the center ground wire (if used) to terminal 3. To use a 70-75 ohm coaxial line, terminals 1 and 2 are jumpered and connected to the inner conductor, and terminal 3 is connected to the outer conductor. Further details are covered in the accompanying text.

selected with an accurate ohmmeter so as to be within 2 or 3 per cent of each other and between 150 and 160 ohms. Even a small amount of unbalance will produce an objectionable amount of antenna effect, and these precautions are necessary in order to provide a well balanced, accurately matched termination. Ordinary metalized rod type composition resistors of good quality are suitable at R_1 and R_2.

Strictly speaking this arrangement does not provide a purely resistive load termination, because of the input capacity of the tube and the balancing capacity C_1. However, the shunt reactance is comparatively low even at 216 Mc., and the reflection coefficient is tolerable unless an antenna is used which exhibits an unusually high mismatch. The input capacity of the tube (and the required value at C_1) can be reduced slightly by "floating the heater of the 6J4 by means of a low resistance u-h-f type r-f choke in series with each heater lead right at the socket terminals.

The noise factor is degraded slightly by the effect of the d-c cathode current flowing through the terminating resistors, but with good quality resistors the increase in noise is small. It can be avoided if desired by keeping the d.c. out of the 150-ohm terminating resistors. This requires the addition of a blocking condenser in series with the cathode, and an r-f choke in series with a 75 to 100 ohm cathode resistor to ground.

Parallel-tuned L-C circuits also are used to couple the feed line to the first tube, the Q being limited to a value which will allow the necessary pass band. The problem of maintaining the Q at the changing optimum value over the entire frequency range is simplified when a turret type tuner is employed, because the input circuit for each channel can be designed for optimum Q. At the lower TV frequencies it is difficult to obtain a sufficiently low Q simply by choice of coil design, and the coil usually is shunted with a resistor of such value as to give the desired Q.

The usual practice is to employ no shunt capacity other than that of the wiring and interelectrode capacity of the tube, the circuit being tuned inductively either by means of a slug or sliding contacts. In the latter case, the "coil" often is a section of line, rather than a solenoid. This permits the greatest possible gain, as the gain obtainable for a given bandwidth or Q at a given center frequency increases with the L/C ratio.

A low-Q, low-C, tuned input circuit of this type feeding a grounded cathode 6AK5 or 6AG5 pentode r-f amplifier will provide somewhat more *gain* than a 6J4 in the arrangement of figure 8-10, but the sensitivity (noise figure) will be approximately the same.

CHAPTER NINE

Measuring Equipment and Techniques

There are many measurements to be made in the process of adjusting a complex antenna system or checking its operating characteristics. Many of the measuring devices and techniques are based on ordinary, standard procedures for measurement of r-f voltage, current, power, reactance, impedance and resistance. The following discussion assumes that the reader is in general familiar with such measuring equipment and conventional techniques for employing such equipment. Measuring equipment and techniques more or less unique to antenna and transmission line measurements will be discussed in somewhat greater detail.

MEASUREMENT OF POWER OUTPUT

Measurement of r-f power dissipated in a load resistor is not strictly an antenna problem, because the load resistor is only a "dummy" antenna; but because such measurement is so closely related it will be included in the discussion.

The amount of power dissipated in a load resistance may be measured either by determining the voltage across and current through the resistor, or by measuring the amount of heat developed in the resistor, by means of a calorimeter.

POWER DETERMINATION BY E AND I MEASUREMENTS

The former method (measurement of E and I) is simpler of execution, but is not conveniently applicable to measurements of very high power because of the difficulty in obtaining a suitable load resistance. Also, the method becomes increasingly subject to error as the frequency is raised, regardless of the power level, due to the following: (1) It becomes increasingly difficult to obtain a purely resistive load as the frequency is raised above the h-f range. (2) R-f ammeters and voltmeters tend to display increasingly large error as the frequency is increased. (3) Stray reactances and couplings become increasingly troublesome as the frequency is raised, giving rise to deceptive voltage and current indications as well as affecting the operation of the circuit to which the instruments are connected.

Load Resistor Considerations

With regard to (1), large carbon rod resistors or "noninductive" wire wound resistors can be used up to 30 Mc. or so without introducing serious error due to reactive effects and skin effect, and can be used to dissipate considerable power. At higher frequencies it becomes necessary to "tune out" the reactive component and to determine the actual, effective resistance of the load *at the frequency involved,* the latter procedure being outside the scope of this discussion.

Small, metalized resistors of the common 1-watt size in values not exceeding a few thousand ohms exhibit much less reactive effect and are less subject to skin effect than the large resistors previously mentioned. The reactance is insignificant (for all practical purposes) and the resistance substantially constant up to several hundred megacycles, and the r-f resistance can be assumed to be the same as the d-c resistance (within these limitations) when an especially high degree of accuracy in measurement is not required. The characteristics of ordinary carbon composition resistors are not nearly so good in these respects, due to greater skin effect (as a result of the greater diameter of the current carrying element) and due to the comparatively high capacitance between conducting granules.

Because the undesirable effects become increasingly troublesome as the resistance is increased, high values of v-h-f and u-h-f load resistance are best simulated, by means of a low value of resistance in conjunction with a quarter wavelength section of transmission line. As the characteristic impedance of a section of line can be determined accurately from its physical dimensions, and the impedance inverting characteristics of a quarter-wave "Q section" are well known, such a setup and procedure involves no particular difficulties.

Use of Attenuating Line

If a load resistance having desirable characteristics at v.h.f. or u.h.f. does not have sufficient power handling capability for a particular job, it still can be utilized for r-f power measurement by using it as the termination of a section of lossy transmission line having a known normal attenuation. The arrangement is analagous to that commonly employed to terminate rhombic transmitting antennas when high power is used.

This method of measurement is simplified if the load resistance is equal to the characteristic impedance of the lossy line, so that the line is working under "flat line" conditions, because then no conversion factor need be applied to the attenuation to take care of the effect of standing waves, and also no care need be taken to see that the line has a certain exact electrical length. In addition, the power handling capability of the attenuating line is maximum.

The load resistance can be converted to some other value, if desired, by connecting a low loss quarter-wave "Q" (linear) transformer between the lossy line and the terminals across which the desired load resistance is to be simulated. Using this method it is possible to employ a 1-watt metalized resistor of known resistance to measure accurately power outputs up to several hundred watts at impedance levels up to a thousand ohms or higher. The only limitations are the safe current handling capacity of the attenuating line and the requirement that the attenuation per wavelength of attenuating line not be too great.

Frequency Sensitivity of Measuring Instruments

With regard to the frequency sensitivity of measuring instruments, the following considerations are pertinent. Thermocouple ammeters and galvonometers are frequency sensitive, and above a certain frequency (depending upon the instrument) a correction factor must be applied. A correction curve (error vs. frequency) usually is supplied by or is available from the manufacturer of the instrument.

Because the error is primarily due to skin effect in the heater element, the error almost invariably causes the meter to read high, and is less (at a given frequency for a given style of instrument) in the lower current ratings. While a 120-ma. thermocouple instrument may exhibit an error of only a few per cent at 100 Mc., an instrument of the same style having a range of several amperes may exhibit an error of as much as 50 per cent at the same frequency.

Special thermocouple ammeters designed for minimum error and reactive effects at the higher frequencies are available for v-h-f and u-h-f measurements, and are recommended where the greatest possible accuracy must be obtained, because they not only are less frequency sensitive but also are designed to minimize stray and spurious reactances. Obviously a correction curve is not of much help when an instrument looks like a complex impedance having considerable reactance both in series and in shunt with the resistive component. These special high-frequency meters often employ an external thermocouple unit in order to permit short, heavy connections to the circuit under test.

Stray reactive effects which tend to show up at the higher frequencies will be minimized if the thermocouple is inserted at a point as near zero r-f potential as possible, rather than at a point having high impedance to ground, regardless of the type of meter employed. In the case of a symmetrical circuit, this will require that the load resistance be made up of two identical resistors, with the thermocouple inserted at their junction.

Vacuum tube voltmeters, used for measuring r-f voltage, also are frequency sensitive. However, the frequency above which the error becomes significant and a correction factor is required is on the order of 100 Mc. in most any well-designed "general purpose" instrument of recent manufacture, and considerably higher in instruments designed expressly for v-h-f and u-h-f work. The error is due to resonance effects in the probe circuit and to transit time effects in the tube, regardless of whether the instrument is of the diode type or the triode type. As the former effect tends to make the meter read high and the latter low, the resultant error may be either on the high side or the low side, depending upon instrument and frequency.

It should be kept in mind that while the voltage across the probe may be determined accurately by means of a correction curve, the probe has a few $\mu\mu$fd. of effective shunt capacity which may produce an undesirable effect upon the circuit under measurement, causing the voltage across the circuit under measurement to change when the probe is connected. Under these conditions, the reactance introduced by the probe must be "tuned out."

To determine the voltage across a symmetrical circuit element (such as a two-wire open line) when only one v.t.v.m. is available, it is necessary to measure the voltage from each side to "ground" separately and then add the two voltages. Usually it can be assumed that the two voltages are of the same phase. The "ground" can be a midpoint on the line termination. If two identical instruments are available, a more accurate measurement can be made by connecting the two probes across the symmetrical line or load "back to back" (probe shells strapped together and connected across the balanced line or load).

Stray Reactances and Couplings

With regard to detrimental stray reactances and couplings introduced as a result of connecting measuring instruments into a circuit at the higher frequencies, they can be minimized by taking the rather obvious precautions of making all leads as short as possible, and, where series reactance must be kept to an absolute minimum, by using heavy conductors. In some cases shielding will be necessary, and in any case it will help if the instruments themselves are designed for operation up to the highest frequency under consideration.

The residual stray reactances and couplings must be compensated for or allowed for if of objectionable magnitude, and to do this a good working knowledge of L, C, and R circuits in combination is essential. If the reader's knowledge of such circuits is inadequate to cope with such problems, he is referred to one of the several excellent general texts on electronics or radio treating such circuits in detail.

Regardless of how the measurements are made, it should be kept in mind that if the actual, effective value of load resistance is known accurately, it is necessary to measure only the current *or* the voltage in order to determine power, but that if both voltage and current measuring instruments are available the exact value of the load resistance need not be known, though of course it should be as nearly nonreactive as possible considering the frequency involved.

POWER DETERMINATION BY CALORIMETRIC METHODS

When the frequency is too high or the power too great for convenient and accurate determination of power dissipation by E, I, and R measurements, calorimetric methods oftentimes are employed. The usual arrangement utilizes a water cooled load with continuous flow and some method of measuring the rate of flow and temperature rise of the water. To minimize the error due to radiation and conduction of heat from the water, the rate of flow is made great enough that the temperature rise is small. When judiciously executed, this method is capable of very high accuracy, even in the u-h-f range.

MEASUREMENT OF ANTENNA POWER

Measurement of the actual power delivered to a radiator is in most cases made somewhat more complex by the fact that the feed point resistance of the antenna is not always easily determined, and by the fact that care must be taken to see that either the voltage and current are in phase at the point of measurement or that the power factor is taken into acount.

If the overall antenna system is resonant, the power at *certain points* in the system can be determined simply by measuring the voltage and current at that point. These points are points of unity power factor where the voltage and current are in phase. They occur at voltage loops (current nodes) and at current loops (voltage nodes). Thus, in the case of an antenna fed by a long resonant line having appreciable but unknown attenuation, it is a simple matter to determine the power at a nonreactive point near the sending end (by measuring E and I) and again at a point near the receiving (or antenna) end. The transmitter output then is known, the power reaching the antenna is known, and the line attenuation is known. For well-designed antennas having a physical length which is a substantial fraction of a wavelength, the radiated power may be taken as approximately equal to the antenna power, because the radiation efficiency will in most cases approach 100 per cent.

As the characteristic impedance of the line usually is known with good accuracy, the power also can be determined by measuring only E *or* I, by checking the VSWR at that point on the line under consideration and resorting to some simple calculations. Thus, if the VSWR on a 600-ohm line is 2.0 at a certain portion of the line, then the impedance at a voltage loop is 1200 ohms, the impedance at a voltage node 300 ohms, and the impedance is purely resistive at both points. Knowing that the line has a certain value of purely resistive impedance at a certain point, it is only necessary to measure the current or voltage at that point to determine the power flowing at that point.

If the transmission line is "flat," as determined by SWR measurements, the voltage and current may be measured at any point on the line at which it is desired to check the power flow, as the voltage and current will be in phase at all points on the line.

DIRECTIONAL COUPLERS

By means of various devices known as the "Reflectometer," "Micromatch," "Power and Impedance Monitor," etc. (all belonging to a family carrying the designation of "directional couplers"), it is possible to measure directly the power flowing at any point on a line, without regard to the position of nodes and regardless of SWR. This very handy ability of these devices is explained as follows.

Standing waves are due to the combination of a direct wave and a reflected wave, as explained in Chapter 2. Until energy has had time to reach the termination and be reflected back there can be no standing wave, regardless of the mismatch of the termination, and the line will (until some energy reaches the termination) act like a "flat" line. (A "flat" line is a line terminated in an impedance which dissipates all of the energy reaching it, thus exhibiting no reflection.) This means that, theoretically, *for an instant* the power flowing in a line could be measured at any point on the line by simply measuring E and I. Of course in actual practice the wave would reach the line and be reflected back, resulting in a standing wave, before a measurement could be made. And anyway, even if it could, the measurement would not give us the amount of power *transferred to the load* except in the case of a line terminated in Z_0.

Directional couplers have the interesting feature of being able to respond to the transmitted wave while ignoring the reflected wave, and by reversing the polarity of the coupler, to respond to the reflected wave while ignoring the transmitted wave. Thus we can measure the amount of energy going and the amount of energy returning, and the difference is the amount of power which is transferred to the load.

This algebraic addition of the "going" power and the "returning" power is done automatically when E and I are measured at a current or voltage loop, but the use of a directional coupler makes it possible to measure the power at *any point on the line,* and not just at specific points. Directional couplers also are useful for measuring standing wave ratios, and will be discussed later in further detail under that topic.

PHANTOM (DUMMY) ANTENNAS

The expressions *"phantom" antenna* and *"dummy" antenna* are used to indicate an *artificial,* nonradiating antenna which is used to check transmitter operation, to facilitate power output measurements, to determine what coupling network parameters or adjustments are required in order to couple to a certain type of antenna, or to "tune up" a transmitter without radiating a signal capable of causing interference.

In connection with receiver measurements, the term "dummy" antenna indicates an artificial antenna which is used (in conjunction with the signal generator) to simulate a real antenna having certain characteristics. The dummy antenna is connected between the signal generator and the receiver input.

RMA standards specify the characteristics of certain dummy antennas to be used for receiver testing, thus providing a common "yardstick" for gain and sensitivity measurements. These apply primarily to standard broadcast and "all-wave" broadcast receivers which usually are employed in conjunction with simple, single-wire antennas of indeterminate length, and they are intended to simulate the reactive and resistive characteristics of an average antenna of this type. They consist of specified values of L, C, and R in combination.

Communications receivers usually are designed with an input impedance which is substantially resistive, and if the specified input impedance is comparable to that of the internal impedance of the signal generator, no dummy antenna need be inserted between the generator and receiver, as the generator itself acts as a dummy antenna. If the impedance of the generator is lower than that of the rated or desired receiver input impedance, the dummy antenna consists simply of a non-inductive resistor in series with the output of the signal generator, the value being such that the total series resistance (including the internal impedance of the generator) is the desired value. In the v-h-f range and above, receiver input circuits and signal generators are both usually designed for 50 ohms impedance, thus simplifying the problem of measurements.

The type of phantom or dummy antenna appropriate for a particuar transmitter application depends upon whether it is desired only to load the transmitter to facilitate tuning up without radiation, whether it is desired to simulate the exact characterics of a certain antenna or the average characteristics of a certain class of antennas, or whether a load is desired for accurate power output measurements without attempting to simulate any particular antenna characteristics.

While some effort has been made to establish a distinction between the expressions "dummy" antenna and "phantom" antenna, the former to indicate a simulated load having specified characteristics and the latter to indicate a load designed only to take power from the transmitter without regard to the exact characteristics of the load, the two expressions long have been used interchangeably in a rather loose fashion. The term "dummy *load*" also is sometimes used to indicate a nonradiating transmitter load.

Ordinary Mazda lamps often are employed as a phantom antenna. While adequate for some purposes, they have certain limitations. These may be summarized as follows: (1) They are fragile. (2) At the higher frequencies they exhibit reactive effects. (3) The resistance varies widely with the amount of power dissipated in the lamp. (4) For a given amount of dissipation, the brilliancy varies considerably with frequency (due to skin effect). (5) They are not well suited to water cooling.

These shortcomings make Mazda lamps none too well suited to accurate measurement of transmitter power output, but nevertheless Mazda lamps make an inexpensive, simple, readily availiable load for moderate powers when the purpose is simply to provide a nonradiating load which will permit loading the transmitter to normal output. If it is not convenient to obtain the desired load resistance by a series-parallel combination of various wattage lamps, any lamp or combination of lamps capable of dissipating the transmitter output power can be used by coupling the lamp(s) to the transmitter through an adjustable network. The link coupled universal antenna coupler described in the previous chapter is well suited to this application.

Special noninductive resistors are utilized in phanton or dummy antennas when the characteristics of Mazda Lamps are not acceptable. For very high power applications these usually are water cooled.

STANDING WAVE MEASUREMENTS

Oftentimes it is desirable to know the position of a current loop or node, or voltage loop or node, on a radiator or transmission line, without regard to the standing wave ratio. Sometimes it is desirable to know the VSWR on a line, without regard to the position of the nulls and loops. In other cases it is desirable to know both the magnitude of the standing waves and the location of loops and nulls. In fact, when both the null position and VSWR can be measured accurately, it is possible to determine the resistive and reactive components of a complex load impedance with a high degree of accuracy, and this system is widely used for making such measurements in the v-h-f and u-h-f range.

LOCATING LOOPS AND NODES

To locate the nodes and loops on a radiating system or open-wire line, the only thing required is a voltage or current indicator which has adequate sensitivity and does not disturb the normal operation of the line or radiator.

Vacuum tube voltmeters, crystal rectifier voltmeters, and small neon bulbs are most generally employed as voltage indicators. Obviously only a relative measurement is required for location of the nodes and loops, and absolute accuracy is not important.

The voltage indicator is capacitively coupled to the radiator or to one wire of an open line, with the least amount of coupling which will give a usable indication (so as to disturb the line constants as little as possible). The more sensitive the indicator, the less coupling is required, making the use of a highly sensitive indicator desirable (particularly when the available transmitter power is low).

Because the voltage loops never are sharply defined, and because the voltage nodes are sharply defined only when the Q of the radiator is very high or the SWR on the line is very high, it usually is necessary to "fix" the exact location of a voltage loop or null by locating two points of equal voltage far enough removed from the loop or null that a slight displacement along the line gives a change in reading. The loop or null then is halfway between these two points.

When using this "equal amplitude" method of locating a null or loop it is important that exactly the same coupling between the indicator and line be maintained as the indicator is moved along the line. A polystyrene "guide spacer" can be employed to facilitate constant coupling by maintaining a fixed physical spacing of appropriate amount (and therefore a constant capacity) between the line and the pickup electrode or "probe."

When the indicating device is coupled to first one feeder wire and then the other with exactly the same amount of coupling, at the same position on the line, it should register the same amount of voltage at a given point on a symmetrical line (such as a two-wire open line). If the readings are not the same, the line is not balanced and "antenna effect" is present.

To check for the position of nodes and loops on a radiating system or a transmission line by means of a *current* indicator, the most commonly employed arrangements consist of a small pickup or "sampling" loop in series with either a small Mazda bulb, an r-f ammeter or milliammeter (usually of the thermocouple type), or a d-c milliammeter or microammeter in series with a crystal rectifier. In the latter case the sampling loop, the meter, and the crystal all are connected in series.

The sensitivity can be reduced, if it is inconvenient to lower the transmitter power to a more convenient value while the checking is being done, either by reducing the size of the sampling loop, by shunting the current indicator, or both.

The sampling loop is oriented with respect to a two-wire line in such a fashion that it lies in a plane which is parallel to that which contains the feeder wires, and in such a manner that the loop is symmetrical with respect to the two wires. A spacing jig made up of polystyrene or other suitable material should be used to provide constant spacing between the sampling loop and the line.

If the sampling loop is rotated 90 degrees, on any axis which is parallel to the two feed wires and equidistant from them, there should be no indication of current at any point on the line if the line is operating under balanced conditions, because under such conditions the current induced in the loop by one wire should cancel that induced by the other when the loop is oriented in this manner, regardless of the displacement of the loop axis with respect to the plane of the wires. An indication of current with the loop so oriented means that antenna effect is present.

When using the device to check the position of loops and nulls on a linear radiator by means of the "equal amplitude" method, the orientation of the sampling loop with respect to the radiating conductor is not important so long as both the orientation and spacing are maintained very exactly as the indicator is moved along the radiator.

Checking the node and loop locations on a "moulded pair" or "twisted pair" line (especially the latter) is somewhat more difficult than in the case of an open-wire line, because the field is much more confined. However, while not so convenient or accurately done, it can be accomplished. But in the case of coaxial cable, the inaccessibility of the inner conductor makes it impracticable to observe the current distribution simply by moving an indicator along the line.

THE TRANSMISSION LINE AS A WAVEMETER; LECHER WIRE MEASUREMENTS

When connected to a uniform line, the wavelength (and therefore the frequency) of an r-f source can be determined from the standing wave pattern simply by measuring the distance between loops or nodes and allowing for the velocity factor of the line, because the adjacent loops and adjacent nodes always are exactly an electrical half wavelength apart. If sufficient care is taken, it is possible to obtain good accuracy with this method, and because of its simplicity it is widely used for wavelength measurement in the v-h-f and u-h-f ranges.

The simplest arrangement of this type for measuring an unknown frequency consists of a pair of bare, parallel wires stretched taut, with the closest practical spacing for the length involved. They should not be supported except at the ends, so as to obtain a uniform impedance and a velocity factor approaching 1.0 as closely as possible. They are fed at one end, and an adjustable "shorting bar" is used to establish resonance. Such a device, often called "Lecher Wires," may be considered as a form of absorption wavemeter using linear circuit elements.

A Lecher Wire line should be at least one wavelength long at the lowest frequency to be measured. The shorting bar should have a knife edge to permit an accurate "fix," should be of low resistance material, and should be provided with a slender "handle" of low loss dielectric material, so as to minimize the amount of dielectric (including the operator's hand) in the field of the line. If the Lecher Wire line is stretched on a wood or other dielectric frame in such a manner that the wires are not separated from the frame by a distance which is large compared to the spacing of the wires, the velocity factor will be lowered sufficiently that it cannot safely be assumed as equal to 1.0. However, if some means is available for accurate measurement of the velocity factor, this is not important except that it makes it necessary to apply a convesion factor to all physical measurements.

If a scale is provided on the frame to facilitate wavelength measurements, the velocity factor can be taken into consideration in the calibration, so that the scale reads in free space wavelength even though the velocity factor is appreciably less than 1.0. Obviously the metric system should be used, so that

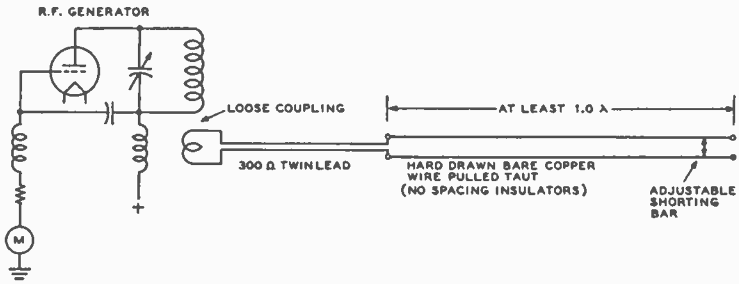


Figure 9-1.

BASIC ELEMENTS OF A LECHER WIRE WAVEMETER.

The Lecher Wire system is essentially an absorption wavemeter using linear circuit elements. The construction of a practical setup is described in the accompanying text.

readings can be interpreted directly in terms of meters, decimeters, or centimeters.

A typical Lecher Wire setup is shown schematically in figure 9-1. The 300-ohm Twin-Lead does not have to be any particular length, and does not require any special care, just so long as it is not excessively long and is not moved while measurements are being made. (Moving it with respect to adjacent objects will affect the point at which the shorting bar resonates at a given frequency).

The principle upon which the Lecher Wire wavemeter works is as follows: When the shorting bar is located so that the electrical length of the entire line is an integral number of electrical half wavelengths, the line will act as a high Q parallel-resonant tank circuit with a current loop at the coupling coil, and will produce the same effect upon another circuit (such as an oscillator tank circuit) as will a conventional absorption wavemeter which is tuned to the same frequency. Because of the high Q of the unloaded linear tank, very loose coupling will suffice, and a single turn coupling coil ordinarily will be adequate.

When the shorting bar is slid over the resonance point, or the shorting bar is left in one position and the frequency of the r-f oscillator is varied, there will be a sharp dip in grid current at resonance. The plate current will exhibit a "hump" at resonance, but when a self-excited triode oscillator is employed, the grid current indication is sharper and more sensitive. In the case of a superregenerative receiver using no r-f stage, resonance will be indicated by the receiver going out of superregeneration as the shorting bar is slid over the resonance point. The coupling should be adjusted so that the nonsuperregenerating condition exists only over a very short travel of the shorting bar, and the resonant point is taken as the mid-point of the "dead spot."

When the Lecher Wire wavemeter is coupled to an r-f amplifier, or to certain u-h-f oscillators not of the negative grid type (klystron, magnetron, etc.), the grid dip method is either impracticable or less effective. In such cases a sharp, sensitive indication of resonance can be obtained by means of a sensitive current indicator loosely coupled to the Lecher Wire system at the coupling loop. Resonance then is established by adjusting the shorting bar for maximum current indication. If the r-f current meter is inserted directly in series with the coupling coil, care must be taken to see that the Q of the Lecher Wire system is not lowered excessively thereby, because this may cause an objectionable amount of broadening of the resonance indication.

Measuring Procedure

A wavelength measurement is taken as follows: The Lecher Wires are coupled just closely enough to the r-f source to give a readily discernible resonance indication when the shorting bar is moved along the line. The bar then is adjusted carefully to the resonance position closest the *unfed* end of the Lecher Wires. This position is carefully noted, and the shorting bar then is slid towards the generator end until the adjacent resonant point is located. The distance between the two points is then carefully measured, preferably on a metric scale. This distance is equal to an electrical half wavelength. The distance is doubled in order to arrive at the electrical wavelength, and divided by the velocity factor to convert to free space wavelength As mentioned before, if the Lecher Wires are well isolated from surrounding objects the velocity factor may be assumed to be 1.0 (with negligible error), in which case no conversion is necessary. The conversion likewise is unnecessary if the Lecher Wires are supported on a frame which is provided with a permanent scale calibrated directly in terms of free space wavelength.

Sources of Error

The accuracy obtained with a Lecher Wire wavemeter depends upon the precision with which the system is constructed and upon the care taken in making a measurement, and usually runs from about 0.1 to 1.0 per cent. The two primary sources of error, other than poor execution of the technique and errors in the measuring scale, are listed below along with methods of minimizing them.

Because of the finite spacing of the Lecher Wires, the side of the line towards the unfed end is coupled to the rest of the system, producing variable reactive effects upon the "live" portion of the Lecher Wires as the shorting bar is moved along the line. This effect can be minimized by making the spacing between the two wires as close as practicable, and by terminating the unfed end of the line in a noninductive resistor which closely approximates the surge impedance of the Lecher Wires. This not only minimizes the coupling between the fore and aft sections of the Lecher Wires, but also effectively shunts the shorting bar with a substantially constant resistance rather than a widely varying reactance. This lowers the Q slightly, causing resonance to be less sharp, but a slightly broad resonant point which is correct is better than a sharper one which is subject to error.

The second potential source of considerable error is excessive coupling between the Lecher Wire wavemeter and a self-excited oscillator, particularly if the latter does not employ a very high Q tank circuit. This causes bad "pulling" of the oscillator frequency. The cure, obviously, is to employ looser coupling, and, when the greatest possible accuracy is required, to use a high Q oscillator.

Conversion Data

When using a regular steel tape or rule, the following conversion data are convenient:

Wavelength in meters is obtained by multiplying the distance between resonance points in inches by 0.0508, then dividing by the velocity factor.

Frequency in megacycles is obtained by dividing the distance between resonance points in inches by the velocity factor, then dividing the answer *into* 5907.

MEASURING SWR

The standing wave ratio at a given point on a transmission line can be measured by a voltage or current indicating device of the type described earlier in this chapter if it is capable of accurate measurement in absolute units. The current or voltage is measured at a loop and again at an adjacent null, and the standing wave ratio, expressed in terms of either voltage or current, is the ratio of the two readings. The current standing wave ratio at a given point on a line is equal to the voltage standing wave ratio; so either a current measuring device or a voltage measuring device can be used to check the "VSWR," which is the conventional method of expressing a standing wave ratio quantitatively.

THE SLOTTED LINE

When it is desired to determine the VSWR existing at a certain point on a *coaxial* line, it can be done by breaking the line at that point and inserting a section of *slotted line* having the same characteristic impedance. The slotted line is simply a

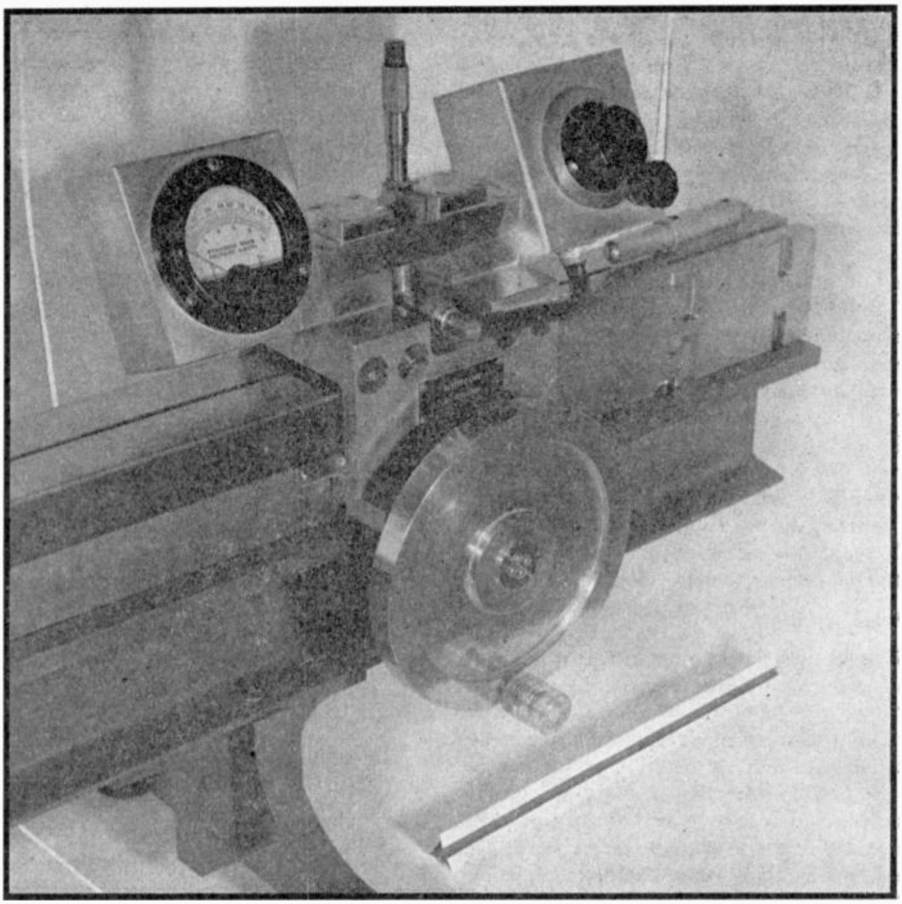


Figure 9-2.

ILLUSTRATING CONSTRUCTION OF A LABORATORY TYPE SLOTTED LINE USED FOR PRECISION MEASUREMENTS (COURTESY COLLINS RADIO COMPANY).

Visible in this view of the carriage are the indicating meter (left), the micrometer type probe assembly (top center) and the tuning system for the crystal detector (right). The slotted line itself, only part of which is visible in this illustration, is eight feet long.

section of air dielectric coaxial line constructed to close tolerances and provided with a milled slot along the outer conductor, parallel to the axis of the conductor, to permit insertion of an r-f voltmeter probe. No spacing insulators are used except at the ends of the slotted section, resulting in a uniform line having a velocity factor of substantially 1.0.

The voltmeter usually consists of a diode or crystal rectifier (detector) with tuned (resonant) input in conjunction with a sensitive d-c microammeter, the whole assembly being arranged to slide along the line on a grooved track which keeps the probe in an exact position with respect to the slot as the probe is moved along the line.

As the probe is not permitted to project into the line very far, its effect upon the operation of the line is insignificant. The slot itself has no significant effect upon the operation of the slotted section of line because currents on the inner wall of the outer conductor travel parallel to the axis of the line (assuming only the normal mode exists).

Oftentimes the "detector box" or "carriage" is provided with a rack and pinion arrangement which facilitates fine adjustment of the carriage position, and a scale is provided to permit reading of distances. A typical laboratory type slotted line is illustrated in figure 9-2.

An r-f generator suitable for slotted line testing up to 1000 Mc. may consist of a power type "Lighthouse" triode oscillator employing a butterfly tank, or a coaxial tank with an adjustable plunger. A section of attenuating line often is placed somewhere between the oscillator and the slotted line while making load adjustments, so as to minimize "pulling" of the oscillator frequency, though this is not always necessary with an oscillator employing a high Q tank.

The length of a slotted line measuring section should be at least one half wavelength and preferably one wavelength at the lowest frequency at which measurements are to be made. Slotted line instruments usually are made with a characteristic impedance in the 50 to 55 ohm range or in the 70 to 75 ohm range so as to equal the surge impedance of commonly used coaxial lines. For accurate measurements the surge impedance of the slotted section should equal that of the regular line within very close tolerances.

The slotted line section usually is of considerably larger diameter than that of many of the widely employed types of coaxial cable, such as RG-8/U, and to minimize reflection as a result of line discontinuity (even though the joining sections have the same surge impedance), a short, tapered section often is included at the ends of the slotted line in order to minimize the "bump" when the slotted section is joined to a smaller diameter line.

The approximately square law characteristic of a crystal detector (and of a diode at small signal levels) makes it difficult to measure accurately standing wave ratios above 10 or so. Also, because the characteristics of a crystal change with age, the calibration of the indicating meter must be checked and sometimes corrected from time to time. To get around these difficulties, a calibrated adjustable attenuator sometimes is employed between the probe and the detector. This usually consists of a calibrated depth adjustment on the probe, resembling a micrometer. Then the amount of attenuation required to maintain the same indicator reading when the probe is moved from a voltage "null" to a voltage "peak" is the standing wave ratio. When using this method the absolute accuracy of the indicating meter is unimportant, though of course the indicator should be consistent (stable) and capable of responding to a very slight variation in the amount of probe pickup.

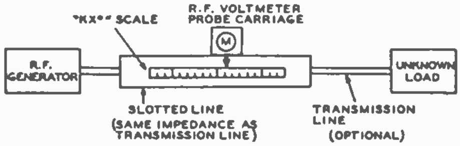


Figure 9-3.

SLOTTED LINE MEASURING SETUP FOR DETERMINING LOAD IMPEDANCE CHARACTERISTICS IN ACCORDANCE WITH THE ACCOMPANYING TEXT.

The slotted line preferably should have a length of at least 1 wavelength at the operating frequency. It is permissible to connect the load directly to the slotted line, but if a line is used it should have exactly the same surge impedance as the slotted line section. The surge impedance of the line connecting the slotted section to the r-f generator need not have the same surge impedance, but it preferably should be approximately the same. The "KX°" scale is shown in detail in figure 9-4.

LOAD MEASUREMENTS WITH A SLOTTED LINE

The VSWR may be determined at any point on a coaxial line much more readily by means of an appropriate *directional coupler* than by means of a slotted line, and of the two methods is the only practical one

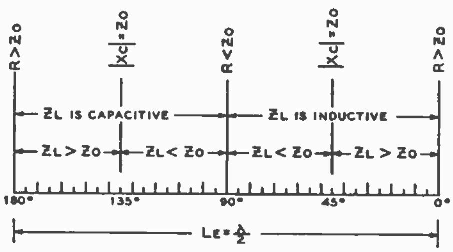


Figure 9-4.

SPECIMEN SLOTTED LINE MEASURING SCALE ("KX°" SCALE) FOR USE WITH FIGURES 9-5 AND 9-6.

The scale should be accurately drawn so that the overall physical length (180 degrees) corresponds to an electrical half wavelength at the operating frequency. For a slotted line section having a velocity factor substantially equal to 1.0, the length of the scale in inches should be equal to 5907 divided by the frequency in Mc.

for continuous monitoring of the VSWR on a line connected to a v-h-f or u-h-f transmitter of considerable power (such as an FM broadcast transmitter). However, the slotted line is better adapted to certain measurements required for the initial adjustments of a v-h-f or u-h-f antenna system. For instance, the resistive and reactive components of a load impedance (impedance and phase angle) can be determined from slotted line measurements and reference to the two accompanying charts*, provided that the slotted line has a velocity factor approaching 1.0 so closely that no significant error will be introduced by assuming that it is exactly 1.0. The method also is applicable to open-wire lines and symmetrical loads, provided that the measuring section has a velocity factor closely approaching 1.0. If little dielectric material

*These slotted line charts and the data pertaining to their use are taken from *Practical Analysis of Ultra High Frequency*, by J. R. Meagher and H. L. Markley, with the permission of the publishers, RCA Service Company, Inc. (a Radio Corporation of America subsidiary).

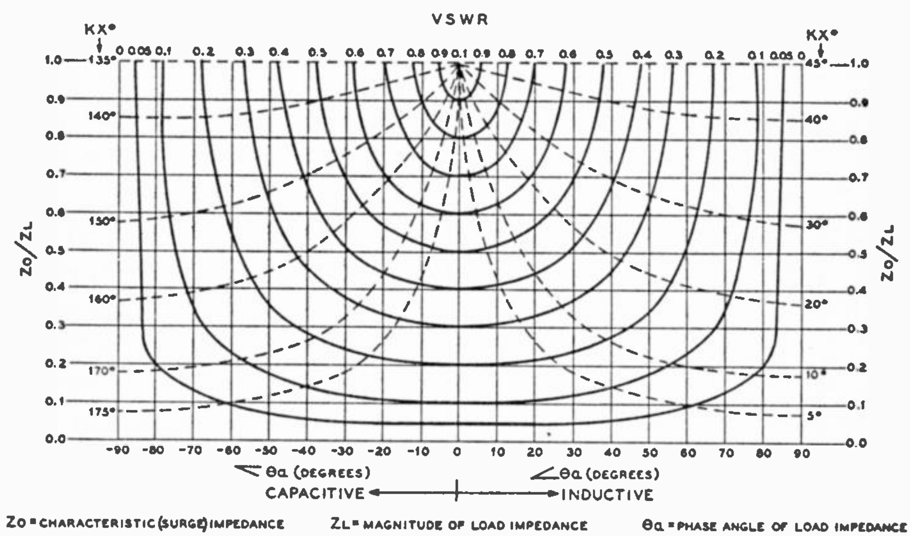


Figure 9-5.

INTERPRETATION CHART FOR SLOTTED LINE MEASUREMENTS, FOR VALUES OF KX° BETWEEN 0 AND 45 DEGREES AND BETWEEN 135 AND 180 DEGREES.

is used to support a two-wire line, the line itself may be used as the measuring section with only a slight error. However, in the following discussion it will be assumed that a slotted measuring section having a velocity factor of 1.0 is used in conjunction with a coaxial transmission line and asymmetrical load. The characteristics of a load impedance are determined in the following manner. (Refer to figure 9-3 for the equipment setup.)

1. *Make up a scale* for the particular operating frequency involved, as shown in figure 9-4.
2. *Establish a reference point* as follows:
   (a) Short circuit the load as effectively as possible.
   (b) Adjust the generator output to the desired frequency and an appropriate output level.
   (c) Locate accurately the voltage null closest the center of the line by moving the probe carriage along the line. This is the *reference point* which will be referred to in the following instructions.
3. *Attach the scale* (1) to the slotted line so that the 90 degree mark lines up with the reference point.
4. *Remove the short circuit* and check the oscillator frequency for shift, correcting it if necessary.
5. *Determine the VSWR* by noting the ratio of the voltage reading at a voltage "node" to the voltage reading at a voltage "loop." Convert to a decimal fraction so as to correspond to the charts of figures 9-5 and 9-6.
6. *Find "KX°."* To do this, accurately locate the voltage "node" closest to the scale reference point (90 degrees). The scale reading at this point is taken as "KX°."

After determining the VSWR and "KX°," refer to the interpretation charts, figures 9-5 and 9-6, to determine the impedance and phase angle of the load. If KX° is between 45 and 135 degrees, use figure 9-6; otherwise use figure 9-5. It should be noted that the magnitude of the load impedance is given in terms of the surge impedance of the slotted line; this permits a universal

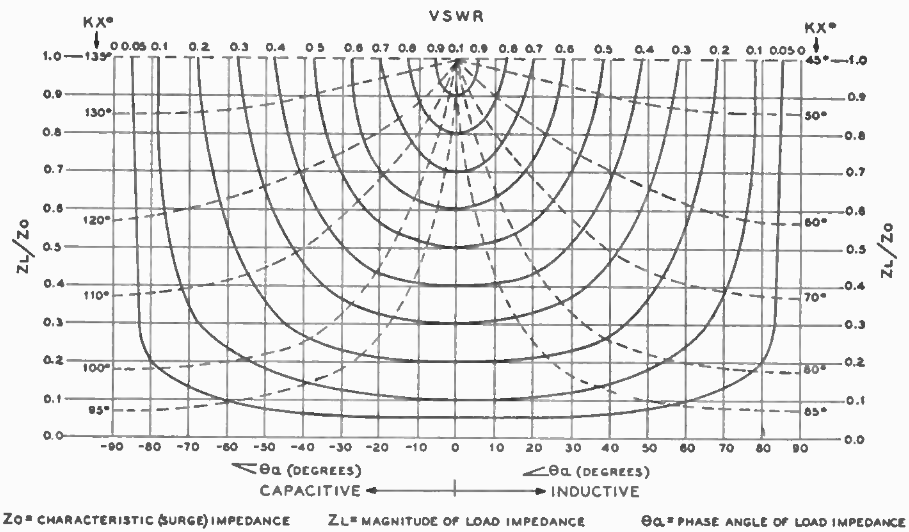


Figure 9-6.

INTERPRETATION CHART FOR SLOTTED LINE MEASUREMENTS, FOR VALUES OF KX° BETWEEN 45 AND 135 DEGREES.

chart which is applicable to any slotted line, regardless of surge impedance.

IMPEDANCE MEASUREMENTS AT THE LOWER FREQUENCIES

The reason for using a slotted line for antenna impedance measurements or receiver input impedance measurements in the v-h-f and u-h-f ranges is to minimize the detrimental effects of stray couplings and reactances. At lower frequencies these are not so troublesome, and in the m-f range and below, more conventional methods and instruments may be employed.

Some of the more common methods of measurement are the various resonance methods (substitution, resistance or conductance variation, reactance or susceptance variation, and "Q meter" measurements), and the various null methods (impedance bridge, Schering bridge, and "double-T"). As such impedance measurements are not strictly or primarily an antenna problem, and as they represent ordinary laboratory technique and practice, they will not be treated here in detail. Care must be taken, however, in applying certain of these methods to the measurement of symmetrical (balanced to ground) loads.

FIELD STRENGTH MEASUREMENTS

Measurement of the field intensity or field strength permits evaluation of antenna radiation efficiency, directivity characteristics, and signal "coverage." It also facilitates "tuning up" an antenna system.

For tuning up an antenna system or measuring the directivity characteristics of an array, the field strength measuring device, commonly called a *field strength meter,* need not be particularly sensitive and need only be calibrated in relative units (db from an arbitrary reference level).

To determine the "coverage pattern" by means of a field strength survey, the meter also must either read in absolue units or else be used in conjunction with a standard field generator. In addition, it must be quite sensitive, the required sensitivity being comparable to that of a good receiver. In fact, such a field strength meter actually is a special form of receiver.

Field strength measuring equipment suitable for making coverage surveys or checking the actual radiated power from an antenna is rather complicated, and a discussion is not within the scope of this book. It is available commercially, though rather expensive (due to the complexity and rather limited demand).

Field strength meters with comparatively low sensitivity and giving only a relative indication of field strength are much simpler of construction, and readily can be improvised. Essentially the simplest form of such a meter consists of a pickup antenna or a loop, a rectifier or detector, and a sensitive d-c meter. Usually the input circuit to the rectifier is tuned.

SIMPLE DIODE TYPE F-S METER

A relative reading field strength meter typical of the simple diode types is illustrated in figure 9-7. A germanium or silicon crystal serves as a rectifier, and a sensitive d-c microammeter as the indicator. To increase the usefulness of the instrument, it is constructed in two parts: the pickup and detector unit, and the remote indicator unit. This makes it possible to read, at the antenna, the field strength some distance from the antenna. This facilitates antenna adjustment by making it unnecessary to run back and forth between the meter and antenna or to enlist the aid of an assistant.

Care must be taken with this type of instrument to avoid excessive pickup, as the microammeter may be damaged thereby. Another disadvantage of this type of instrument is the restricted range over which readings can be taken. The approximately square-law characteristic of the crystal rectifier limits the useful range to not much over 10 db, which is inadequate for many purposes. The db range can be doubled by the use of a meter shunt and switch, but this is rather awkward, besides requiring two scale calibrations. The need for separate scale calibrations is explained by the fact that the crystal characteristic is neither strictly linear nor strictly square law, which means that a db increment on the low scale is not necessarily the same as a db increment on the high scale.

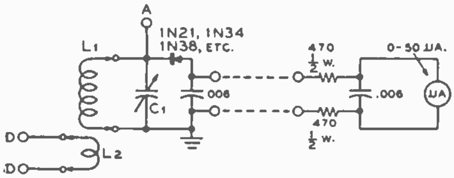


Figure 9-7.

SIMPLE DIODE-TYPE FIELD STRENGTH METER WITH REMOTE INDICATOR.

The useful single-scale range of this type of meter is about 12 or 14 db. The coil windings L_1 and L_2 are wound on a single plug-in coil form, the turns ratio being about 4 to 1. The combination L_1, C_1 should be capable of tuning to the desired frequency. A short whip is connected to terminal "A," or a short doublet is connected to the "D" terminals. For measurement of ratios, the microammeter should be calibrated in db. This type of field strength meter is not suitable for measuring absolute field strengths except when used in conjunction with a standard field generator.

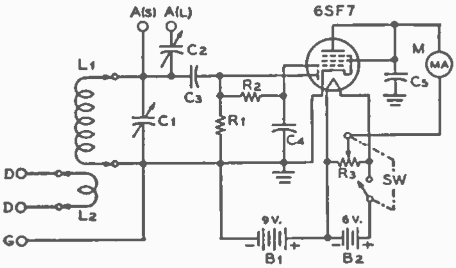


Figure 9-8.

RELATIVE READING FIELD STRENGTH METER USING DIODE RECTIFIER AND D-C AMPLIFIER.

The useful single-scale range of this instrument is approximately 30 db, making it more convenient to use than the meter of figure 9-7 when checking the directivity of an array. With a suitable r-f input circuit the above instrument can be used up to approximately 300 Mc. The constants are as follows.

L_1, L_2—Plug-in r-f coil, closely coupled windings, turns ratio approximately 4 to 1

C_1—Variable condenser, suitable for frequency range to be covered

C_2—3-30 $\mu\mu$fd. adjustable ("trimmer" type)

C_3, C_4, C_5—.001 μfd. mica

R_1, R_2—100,000 ohms, ½ watt metalized resistors

R_3—1000 ohm wire wound potentiometer with integral switch

MA—0-1 ma. d-c (invert meter unless zero right)

B_1—Two 4½ volt "C" batteries in series

B_2—6 volt "latern" battery

WIDE RANGE F-S METER

The useful single-scale range can be increased to nearly 20 db by substituting for the crystal rectifier a 1.5-volt vacuum tube (1G4, etc.) connected as a diode. However, if a tube and battery are to be employed, the rectifier-amplifier type meter described in the following paragraphs and illustrated in figure 9-8 is recommended. While containing a few more components, the cost is about the same, because of the difference in the cost between a 0-50 μa. microammeter and a 0-1 ma. milliammeter. The latter, incidentally, is much more rugged than a 0-50 or 0-100 μa. microammeter.

When calibrated in accordance with the procedure to be described later, the useful single-scale range of the instrument exceeds 30 db. Also, there is no danger of damaging the meter from excessive input, and the tube can be damaged only by r-f inputs considerably exceeding 100 volts peak (which is well above the safe limit for a crystal, as the inverse or "back" voltage is roughly twice the peak r-f input voltage). The sensitivity is comparable to that of the field strength meter of figure 9-7.

A 6SF7 is connected as a variable-μ triode and used as a diode rectifier and d-c amplifier. The variable-μ characteristic increases the useful range of the meter over that obtainable with a straight diode arrangement, by compressing that portion of the scale which would be excessively expanded (figured on a db or logarithmic basis).

By the use of short, heavy leads and an appropriate input tank circuit, the instrument can be used up to about 300 Mc. This limit can be increased considerably, if necessary, by substituting a "radar" type silicon crystal for the diode; but the useful scale range will be decreased slightly because the crystal is more nearly "square law" than is the diode.

The use of a tube with an indirectly

heated cathode permits use of the heater battery (B2) as part of the plate voltage supply. The remainder of the required plate voltage is obtained from two 4.5-volt "C" batteries in series. As the drain on the latter is less than 1 ma., their life under normal use will approach their "shelf life." The life of the 6-volt "lantern battery," B2, will be about 24 hours (aggregate) under conditions of intermittent use, representing a battery maintenance cost of about 2½ cents per hour. However, considering the number of hours a meter of this type ordinarily is employed, battery expense is not a major item.

Because of the manner in which the d-c amplifier is used, the indicating milliammeter (MA) reads backwards; that is, an increase in signal causes a decrease in plate current. This requires that either the meter be inverted or else a "zero right" meter be employed. The latter are manufactured for certain receiver "S meter" applications, and are available on special order.

The potentiometer R_3 is used to set the plate voltage on the 6SF7 to a value which causes exactly 1 ma. of plate current to flow with no signal input. This adjustment is made when the instrument is calibrated in db, and again each time the instrument is turned on to take a reading.

The potentiometer should be of the integral switch type, so that turning the knob all the way counter-clockwise turns off the heater supply. To avoid possible damage to the meter, it is important that the potentiometer be connected with the polarity such that *clockwise* rotation of the knob *increases* plate voltage.

Provision is made for either a short "whip" antenna, a long whip or wire, or a doublet. The short whip is connected to the A(S) post, the long antenna to the A(L) post, and a doublet to the "D" posts. When a long antenna is employed, a ground should be connected to the G post at frequencies below about 20 Mc., and regardless of frequency the adjustable coupling condenser C_2 should be optimized when maximum sensitivity is required. When pickup is excessive with any of the three antennas, it effectively can be reduced to any desired value by detuning C_1.

Calibration Procedure

The db scale either should be drawn directly on the meter face, or on a blank scale cemented over the regular scale. The instrument is calibrated with the aid of an oscillator which is provided with means for cutting the plate voltage exactly in half by throwing a switch. The full voltage should be the normal rated voltage for the tube, and the oscillator should have plenty of feedback and be lightly loaded. The frequency at which the instrument is calibrated is not important.

The coupling to the field strength meter and the setting of C_1 are adjusted so that with full plate voltage on the oscillator the field strength meter reads what would correspond to slightly less than 0.1 ma., and the needle position marked on the scale. The oscillator voltage then is cut in half and the needle position again marked on the scale. Full voltage is applied to the oscillator and the coupling reduced or C_1 detuned until the needle falls exactly on the last mark. The voltage then is cut in half, and the needle position marked. Full voltage then is applied again to the oscillator and the process repeated. This gives 6 db increments. The procedure is followed until the 6 db increments become too close together to be useful. The last usable increment is taken as the "zero" or reference point and all of the calibration marks are labelled accordingly: 0, 6, 12, 18, etc. db.

ANTENNA ADJUSTMENT BY THE "RECIPROCITY METHOD"

It is possible, and under some conditions even preferable, to make antenna and feed line adjustments by the "reciprocity method." The position of the r-f generator and field strength meter are simply reversed. An r-f generator drives an "exciter" antenna which is located some distance from the regular antenna or array, and the regular antenna feed line is terminated in a noninductive resistor having a value closely approximating the surge impedance of the line. The voltage developed across this resistor then is measured by means of a sensitive vacuum tube voltmeter suitable for the frequency under measurement, or, if the

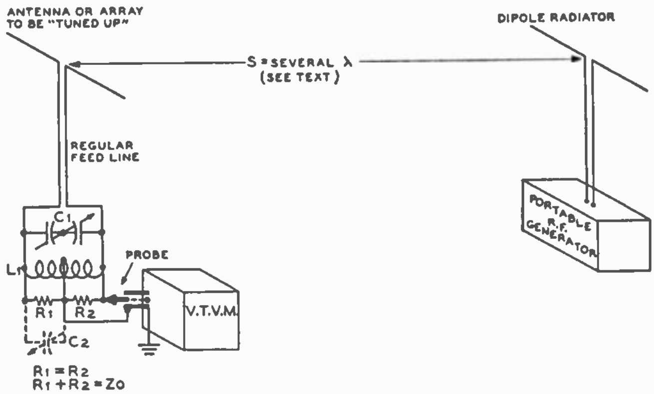


Figure 9-9.

ILLUSTRATING "RECIPROCITY METHOD" OF TUNING UP AN ANTENNA SYSTEM.

The anti-resonant tank circuit L_1, C_1 acts as a line balance converter, thus preventing in-phase currents (antenna effect) from registering on the v.t.v.m., and also preventing the v.t.v.m. from unbalancing the line. It also permits "tuning out" of the probe reactance. When a coaxial line is employed, a line balance converter is not required, though it still will be necessary to tune out the reactance of the probe if the reactance is appreciable at the frequency involved. Also, when a coaxial line is employed it is not necessary to split the load resistance as is shown above; instead a single noninductive resistor may be used. At the higher frequencies, short, heavy connecting leads should be used.

r-f generator has sufficient power, the current through the resistor is measured by means of an r-f milliammeter.

Because only relative accuracy rather than absolute accuracy is required, a v.t.v.m. may be used for this purpose up to its extreme frequency limit. The probe reactance should be tuned out, however, if it constitutes an appreciable portion of the load when the probe is connected across the terminating resistor. Used in this manner a v.t.v.m. such as a "Vomax" is suitable for use up to approximately 200 Mc., and a Hewlett-Packard model 410-A is suitable for use up to at least 1000 Mc.

When a symmetrical ("balanced") line is employed, a suitable "phase inverting" network is necessary in order to prevent the v.t.v.m. from unbalancing the line and to prevent in-phase currents (antenna effect) from affecting the v.t.v.m. reading. Such an arrangement is shown in figure 9-9. The coil L_1 should be close wound, so as to provide maximum coupling between the two halves, and the loaded Q of the L_1, C_1 combination should be at least 10, for the same reason. The balancing condenser C_2 is required only if the shunt capacitance contributed by the probe is an appreciable fraction of the capacitance of one section of C_1, in which case its value should be the same as the shunt capacitance of the probe.

The L_1, C_1 combination should *not* be tuned for maximum reading of the v.t.v.m. when set up to make antenna measurements. Instead it should be tuned to the operating frequency with the feed line disconnected. This can be done by very loosely coupling L_1 to the transmitter and tuning C_1 to resonance. The feeders then are connected and C_1 is left alone.

A frequencies below approximately 200 Mc. a much simpler arrangement for use with a symmetrical line consists of an r-f milliammeter in series with two noninductive resistors having a total resistance equal to the surge impedance of the line, the combination simply being connected across the end of the line as shown at figure 9-10A.

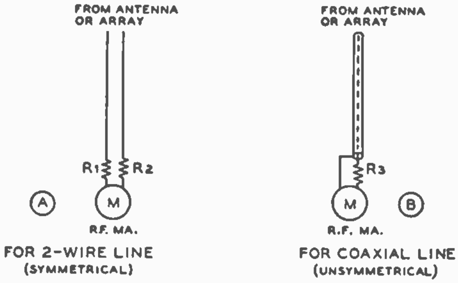


Figure 9-10.

ALTERNATIVE LINE TERMINATION AND MEANS OF MEASUREMENT WHEN USING THE "RECIPROCITY METHOD" OF ADJUSTING AN ANTENNA SYSTEM.

An r-f current meter of the themocouple type is used instead of an r-f voltmeter. While this system is simpler than that of figure 9-9 (particularly when a symmetrical line is employed), it is not as sensitive, and requires a more powerful r-f field generator.

Splitting the load resistance permits the r-f meter to be inserted at a point of zero r-f potential, a desirable practice above 20 Mc. or so. Below approximately 20 Mc. a single resistor may be employed, as the unbalance will not be sufficient at these frequencies to cause trouble.

The chief disadvantage of the figure 9-10 arrangements is their poor sensitivity as compared to the v.t.v.m. method of figure 9-9, thus requiring a much more powerful generator.

At frequencies up to several hundred megacycles, ordinary metalized resistors in the ½ watt size may be used as terminations (either in figure 9-9 or in figure 9-10), as they present little reactance, and the r-f resistance is substantially the same as the d-c resistance at such frequencies.

When a coaxial line is employed, the problem of line unbalance and antenna effect is not encountered, and a much simpler v.t.v.m. measuring setup can be employed. The probe is simply connected across a single terminating resistor if the reactance presented by the probe is not appreciable at the frequency involved. If the reactance is appreciable, it can be tuned out by connecting an inductance and variable capacitance in parallel directly across the resistor, and resonating the combination with the feed line disconnected as previously described.

If a thermocouple type meter is used in series with the load resistor as an indicator, it should be placed at the ground end of the resistor.

Advantages of the "Reciprocity Method"

When using a field strength meter to "tune up" an antenna (to optimize radiator length, stub position or length, etc.), the antenna adjustments will have a considerable effect upon the transmitter loading. It is therefore necessary to check and correct the loading each time an antenna adjustment is made, so as to keep the power fed to the antenna system constant. Obviously this can be a rather tedious procedure when a large number of adjustments are to be made.

The use of the reciprocity method obviates the necessity for frequent checking and altering of the loading while antenna or feed line adjustments are being made, as adjustment of the main antenna will have negligible effect upon the amplitude of the power radiated from the exciter antenna.

If especial care is taken to avoid spurious coupling the regular transmitter can be used as the r-f field generator, the transmitter being connected to the exciter antenna via a transmission line. The best type of line for such use when a dipole is employed as the exciter antenna is a shielded pair, though regular "coax" or even low impedance twisted or molded pair may be used if particular care is taken to avoid antenna effect on the line. It also is necessary that direct radiation from the transmitter proper be kept to a sufficiently low value that the direct radiation picked up by the regular antenna system (including the feed line, and the measuring equipment at the termination) is negligible.

The possiblity of spurious couplings can be avoided by using instead of the regular transmitter a shielded, portable oscillator with self-contained power pack to feed the exciter antenna, with the oscillator located right at the exciter antenna. The 115-volt a-c supply wires should be bypassed to the oscillator cabinet and so oriented or spaced with respect to the main antenna and feed line that no spurious excitation of the main antenna occurs as a result of supply line radiation.

OPTIMIZING ANTENNA ADJUSTMENTS WITH THE AID OF FIELD STRENGTH MEASURMENTS

When using a field strength meter, certain considerations must be kept in mind in order to avoid misleading indications or erroneous interpretations. Some of the more pertinent considerations are the following:

The first thing to keep in mind is the fact that surrounding objects, particularly telephone and power wires, can cause energy to be transferred to the field strength meter in such a manner that an *accurate* check on the horizontal directivity pattern of an array by means of nearby field strength measurements is impossible.

An important precaution to be observed is to locate the field strength meter (or at least the pickup portion of it if it is the remote indicating type) a sufficient number of wavelengths from the main antenna that there is negligible mutual impedance. In the case of a dipole, the separation need be only 1 or 2 wavelengths, but in the case of a highly directional array it may be necessary to employ a separation of 10 or more wavelengths in order to reduce the mutual impedance to a negligible value. This is especially important when the field strength meter uses a resonant dipole for an antenna.

The pickup antenna on the field strength meter always should be oriented so as to give the same polarization as that of the antenna under test.

If the antenna under test is a rotatable array, it usually is preferable when checking the horizontal pattern to locate the field strength meter in a fixed position and rotate the main antenna, rather than use a fixed orientation of the main antenna and take field strength readings on a circle of fixed radius. This is especially true when surrounding objects are likely to distort the horizontal pattern at ground level and the antenna is to be used for sky-wave communication.

When vertical polarization is employed and both the main antenna and field strength meter antenna are located less than one or two wavelengths above ground and are separated by only a few wavelengths, much if not most of the energy is propagated to the field strength meter via the surface wave even at frequencies well into the u-h-f range. It should be kept in mind that the strength of the surface wave is quite sensitive to both antenna height and the character of the soil.

CHAPTER TEN

Antennas for Navigational Aids

A few navigational aids employing electromagnetic radiation and reception such as "Loran," the British "Gee" system, etc. do not depend upon antenna directivity for their operation. Instead they give a "fix" by virtue of the time differential between pulses arriving from coordinated pulse transmitters spaced a considerable distance apart. However, the great majority of "radio" aids to navigation (meaning those utilizing electromagnetic waves as contrasted to ultrasonic and certain other means of propagating intelligence) all take advantage of the directional characteristics of an antenna system to give bearing (and sometimes angular elevation), even though a pulse system may be incorporated to give distance or range indication.

Because of the large number of different radio aids to navigation and the complexity of all but the most elementary types, a detailed discussion of all of the specific systems is beyond the scope of this book. However, a general idea of how the directional characteristics of an antenna are exploited in these applications does not entail extensive treatment, and hence such a discussion will be included so as to familiarize the reader with the use of directional antennas in navigational aid applications and to enable him to appreciate some of the more important considerations involved.

PROPAGATION AND SITING CONSIDERATIONS

"Locating" devices such as radar, radio beacons, GCA, radio direction finders, etc. take advantage of the fact that within line-of-sight a radio wave follows substantially a straight line, and beyond line-of-sight tends to follow a great circle route. The error introduced by assuming that the radio wave follows exactly a rectilinear or a great circle path ordinarily is quite small except in the case of h-f sky-wave propagation over very long distances, or in the case of m-f or l-f ground-wave propagation over medium or long distances when the effective conductivity of the earth is not fairly constant for considerable distance to either side of the great circle path, or in the case of v-h-f or u-h-f ground-wave propagation over considerable distance when the intervening terrain is hilly or mountainous.

By measuring the direction of arrival of a radio wave, it is evident that, to the extent of the foregoing exceptions and qualifications, the direction of the antenna which radiated the wave can be determined with respect to the receiving antenna.. The simplest equipment capable of determining the direction of wave arrival comprises a rotatable directional antenna in conjunction with a receiver which indicates relative field strength, and the combination constitutes an elementary "direction finder."

Not only can there be a significant or even a considerable error due to the transmission path deviating from a straight line or from a great circle path under certain conditions, but there also are numerous factors tending to introduce an error when measuring the direction of arrival of the wave. The most serious are due (1) to distortion of the field at the receiving point due to the presence of nearby objects, and (2) to a false indication obtained with certain directional antennas when the arriving wave is eliptically polarized. The former (1) are called "site errors," and the latter (2) are called "polarization errors."

Site Errors

It was shown previously that the ground and other objects in the vicinity of an antenna must be considered as being effectively part of the antenna system in so far as its directional characteristics are concerned. If the earth is flat for considerable distance in every direction and there are no trees, buildings, etc. to distort the pattern, the *horizontal* pattern will be the same as in free space. But otherwise the horizontal pattern will be distorted.

When the antenna is in a fixed location with regard to objects which affect the directivity pattern, it is possible to measure the deviation or error for all compass directions, and thus "calibrate" the direction finder so as to permit correction of the error. In the case of a direction finder used over a wide range of frequencies, the error characteristic must be determined at numerous frequencies over the band, because the error characteristic is frequency sensitive.

When the receiving antenna is not in a fixed location with respect to all objects having a significant effect upon the azimuthal directivity pattern, as is the case of an aircraft flying over mountainous terrain, compensation is not practicable and the bearing is unreliable. When an airborne direction finder exhibits this type of site error, it is called *mountain effect*. Fortunately it is quite obvious when mountain effect is present and the bearing cannot be relied upon. It is manifested by erratic fluctuations in the bearing indication. The erroneous bearing is the result of an interference pattern being set up between the direct and mountain-reflected waves. In the case of highly irregular terrain the amplitude of the mountain-reflected waves can be quite strong—comparable to the intensity of the direct wave—and they can combine in a complex manner which is highly sensitive to both frequency and position.

Polarization Errors

When antenna systems for navigational aids are taken up in following sections it should be noted that some of the commonly used types (for example the simple loop) have one directivity pattern for horizontally polarized waves and a quite different directivity pattern for vertically polarized waves. If the working frequency and distance are such that there is appreciable sky-wave component in the received signal, the received signal ordinarily will have both horizontally polarized and vertically polarized components, because of the usual random or eliptical polarization of down-coming sky waves regardless of their polarization upon leaving the radiator.

With such a combination of antenna and propagation conditions, either it will be impossible to obtain a bearing, or else the bearing will be in error. In the latter event the bearing will slowly shift, rather than fluctuate as in the case of "mountain effect," provided mountain effect is not present also. In fact, if there is sufficient time, a fairly accurate bearing can be obtained in spite of this condition by averaging the bearings obtained at uniform intervals over a period of time.

This type of error, called *polarization error*, is negligible when the ground wave dominates the sky wave. Therefore it becomes objectionable only beyond a certain distance, determined by the frequency and time of day. At the frequencies at which polarization error ordinarily is encountered (primarily within the m-f range), the sky wave, and therefore the bearing error, is much worse at night. It is for this reason that polarization error of this type also is known as *night effect*.

THE LOOP ANTENNA

One simple type of antenna widely used in conjunction with navigational devices employing radio waves is the *loop* antenna. Essentially the loop antenna is a coil of one or more turns having dimensions small compared to a wavelength, and close wound so that all turns of the coil occupy as nearly the same position in space as possible. The loop may be wound in the shape of a circle, square, triangle, or most any "open" configuration without significantly affecting its operation. The amount of voltage pickup at a given frequency is determined primarily by the area enclosed by the loop and the number of turns in the loop, provided that the number of turns is small enough to result in substantially uniform current in all turns.

The free space directivity pattern of a loop which is operating under balanced conditions resembles that of an infinitesimal dipole placed at the center of the loop and normal to the plane of the loop. However, the dipole will, under these conditions of orientation, have a polarization mutually perpendicular with that of the loop.

When the loop encloses an area which is small in terms of wavelength, which is the usual case when a loop is employed at frequencies below approximately 2000 kc., the radiation resistance is very low compared to the loss resistance. This results in poor efficiency when the loop is employed as a transmitting antenna, but with a highly sensitive receiver may permit reception of any signal above the atmospheric noise level at these frequencies. As explained in previous chapters, an antenna need not provide much energy pickup at the lower frequencies in order to make atmospheric noise the predominant noise present in the receiver output. It is for this reason that the loop antenna is widely used not only for navigational aids, but also for AM broadcast receivers in order to make an external antenna unnecessary.

A loop ordinarily is oriented so that the plane of the loop is vertical. The horizontal directivity pattern of such a loop, provided that it is operating under balanced conditions, is, *for vertically polarized waves*, as shown at figure 10-1A. It will be noted that while the maxima are broad, the two nulls are quite sharp, and the directivity characteristics therefore are suitable for application to certain navigational aids.

However, a loop oriented in this manner is responsive also to *horizontally* polarized waves having a wave angle greater than zero, the response increasing with the wave angle until it is maximum for a wave coming straight down. Unfortunately the horizontal directivity pattern for horizontally polarized waves is displaced 90 degrees in azimuth with respect to the horizontal pattern for vertically polarized waves, causing the nulls for horizontal polarization to coincide with the maxima for vertical polarization, and the maxima for horizontal polarization to coincide with the nulls for vertical polarization. This results in a polarization error when an appreciable proportion of the signal arrives via sky-wave propagation, as previously discussed under *polarization error*. While this effect greatly reduces the usefulness of a loop antenna for navigational aid applications, the loop nevertheless is widely employed for such purposes because of its simplicity and compactness. However, its limitations with respect to polarization error always are taken into consideration.

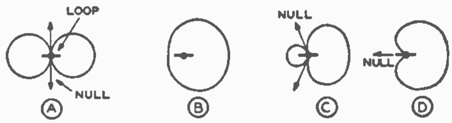


Figure 10-1.

DIRECTIONAL PATTERNS OF A "LOOP" AND OF LOOP PLUS SENSE ANTENNA.

The patterns apply only to vertical polarization, and only when the sense antenna is coupled so as to provide a voltage which is either in phase or 180 degrees out of phase with the resultant loop voltage. The pattern at C also is typical of a loop (alone) which is imperfectly balanced and exhibits "antenna effect" which is in phase with the resultant loop voltage. When the antenna effect is 90 degrees out of phase with the resultant loop voltage, the pattern resembles that at A except that the minima are "blurred" and do not go to zero. Thus it is seen that "antenna effect" can produce substantially the same result as the incorporation of a "sense" antenna, the latter simply being a short, vertical antenna placed near the loop.

The polarization error exhibited by loop antennas can be reduced considerably and under certain critical conditions of propagation eliminated altogether by incorporating a compensating winding. Or, by using two spaced loops one ahead of the other and connecting them in phase opposition, downcoming horizontally polarized waves will cancel when the loop combination is in the null position for vertically polarized waves. This position is the same as for a single loop. The use of two loops in opposition greatly reduces the energy pick up as compared to a single loop, especially when the spacing is small.

Distortion of Loop Pattern by Antenna Effect

Unless a loop antenna is electrostatically balanced to ground, "antenna effect" will be present and the pattern will be distorted. The manner in which it is distorted depends upon the relative amplitude and phase of the antenna effect, and resembles the distortion obtained with a "sense" antenna, which will be described in a later paragraph. Typical patterns obtained with a loop exhibiting antenna effect, or with a loop used in conjunction with a "sense" antenna, are illustrated in figure 10-1.

It is desirable, for reasons which will become apparent, to eliminate all antenna effect and to achieve pattern control by the use of a separate sense antenna. Antenna effect is minimized by using a symmetrical arrangement, and by electrostatic shielding of the loop; any residual antenna effect then is eliminated by a "compensator" or "balancer." The latter is adjusted until the two nulls are exactly 180 degrees apart. A detailed discussion of the numerous loop coupling circuits commonly employed for tuning, coupling, and balancing a loop is properly a matter of radio direction finder design and beyond the scope of this book*, but a typical circuit, minus the components used to couple in a sense antenna, is shown in figure 10-2.

The electrostatic shield is provided with a gap in the form of an insulated bushing, so as to prevent it from acting as a shorted turn. When fabricated of nonmagnetic metal and arranged in this manner, the shield has no significant effect upon the performance of the loop except to minimize

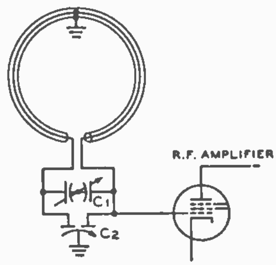


Figure 10-2.

SIMPLE, DIRECTLY-TUNED LOOP USING ELECTROSTATIC SHIELD.

Ordinarily the loop will consist of more than one turn, with a center tap. A single turn is shown in order to simplify the illustration. The loop is tuned to the operating frequency by means of C_1. C_2 is a special "balancing" condenser of the differential type (with opposed stators), and is used to eliminate any residual antenna effect. The metalic shield must be provided with a gap as shown in order to prevent it from acting as a shorted turn.

capacity unbalance (and thereby minimize antenna effect) and to reduce the effects of precipitation static, the latter being an important consideration in aircraft applications.

Use of a "Sense" Antenna to Resolve 180 Degree Ambiguity

It is apparent from figure 10-1A that a loop antenna operating without antenna effect exhibits a 180 degree ambiguity, because of the two nulls and the symmetry of the directional pattern.

By coupling in some input from a short vertical antenna located near the loop, the composite directivity pattern can be made to resemble figure 10-1B, C, D, or most anything in between, depending upon the relative amplitude and phase of the resultant voltage contributed by the sense antenna with respect to that contributed by the loop. By providing a switch to permit switching the sense antenna in or out, it is possible to use the loop alone for accurate bearing indication, and to switch in the sense antenna to give a pattern which makes it possible to resolve the 180 degree ambiguity.

*For an excellent and comprehensive treatment of loop antenna coupling circuits, as well as direction finders in general, see D. S. Bond, *Radio Direction Finders,* McGraw Hill Book Co., Inc., 1944.

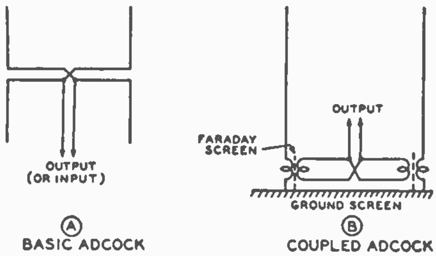


Figure 10-3.

BASIC ADCOCK ARRAY (A) AND COUPLED ADCOCK ARRAY (B)

The basic Adcock exhibits no polarization error when sufficiently removed from earth, but in practical installations the elements are sufficiently close to the ground to unbalance the system and make it responsive to horizontally polarized waves due to imperfect concellation of currents induced in the horizontal members. The coupled arrangement, when properly constructed, is not subject to this shortcoming. In certain navigational aid systems, Adcock arrays also are employed as transmitting antennas.

THE ADCOCK ANTENNA

The inherent polarization error of a loop antenna can be avoided by using two vertical (linear) elements in anti-phase, in conjunction with a transmission line system which is operating under balanced conditions and is free of antenna effect. The horizontal directivity pattern resembles a "cosine" figure 8 if the spacing is a small fraction of a wavelength, and therefore resembles the horizontal directivity pattern of a loop for *vertically polarized waves.*

Because a vertical wire can transmit and is responsive to vertically polarized waves only, it is apparent that two such wires used in combination (such as figure 10-3A) still will be responsive only to vertically polarized waves. However, care must be taken to avoid signal pickup by the horizontal phasing sections as a result of imperfect cancellation in the horizontal members, because this will make the overall system responsive to horizontally polarized waves.

Imperfect cancellation of currents flowing in the horizontal members can be considered a form of "antenna effect," and can occur as a result of the lower half of each vertical element having greater capacity to ground than the top half of each element. The polarization error introduced thereby is much less than that exhibited by an ordinary loop, but nevertheless is objectionable unless the antenna elements are well removed from the earth.

There are various methods of reducing the response of the horizontal phasing-feeding sections, one very effective method being the "coupled Adcock" shown at figure 10-3B.

The pickup of an Adcock array is comparable to that of a single turn loop having corresponding dimensions. Therefore, because many turns can be employed in a loop and because the voltage developed by a loop is proportional to the number of turns (within the limits previously discussed), a much larger structure is required for an Adcock array than for a loop having comparable pickup. At the lower frequencies the required structure is too large to permit convenient mechanical rotation. However, this difficulty can be resolved by using a pair of "crossed" Adcock arrays and rotating the figure eight pattern in azimuth *electrically* by means of a *goniometer,* the basic principles of which are covered in the following paragraph.

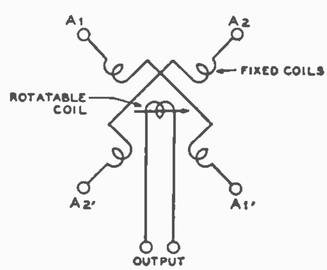


Figure 10-4.

SCHEMATIC REPRESENTATION OF GONIOMETER

The terminals A_1 connect to the feed line from one panel of a pair of crossed loops or Adcock arrays, and the terminals A_2 connect to the feed line from the other panel. This permits the figure eight space pattern to be rotated through 360 degrees. An electrostatic shield (not shown) ordinarily is used between the rotating coil and the fixed coils.

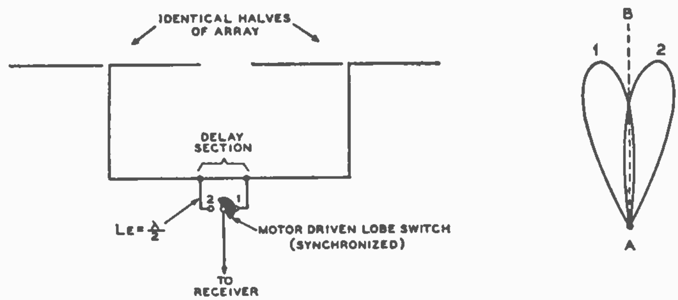


Figure 10-5.

ILLUSTRATING PRINCIPLE OF LOBE SWITCHING, USED TO GIVE EFFECT OF TWO SEPARATE ANTENNAS AND INCREASED POSITION ACCURACY

The skeleton schematic of the basic system shows only the inner conductor of a coaxial feed system, the omitted outer conductor of all sections of line being connected together and to chassis ground. It should be noted that in either position of the lobe switch the entire antenna is operative. The nose of either lobe alone is comparatively broad, and not capable of registering direction with a high degree of accuracy. However, by rapid commutation of the lobe back and forth between positions 1 and 2, direction can be ascertained with a high degree of accuracy by orientating the antenna for equal response from the two lobes, as explained in the accompanying text.

THE GONIOMETER

When because of size or other physical limitations it is impracticable to provide mechanical rotation of a loop or an Adcock array, *electrical* rotation of the figure eight pattern can be accomplished by locating two loops or Adcock arrays on a common axis, crossing them at 90 degrees, and combining the outputs with the aid of a device called a *goniometer*. Essentially the goniometer consists of an electrostatically shielded, rotatable coil which is symmetrically disposed with respect to two, crossed coils (or pairs of coils), each of which is connected via a transmission line to one "panel" of the crossed loops or Adcock arrays. Rotation of the movable coil produces effectively the same result as physical rotation of a single loop or Adcock array in space.

LOBE SWITCHING

The "nose" of the lobe of even a highly directional antenna array is not especially sharp, and therefore is not very sensitive to change in direction. So, to provide a high degree of accuracy, electromagnetic navigational aids (as well as radar fire control devices) often employ *lobe switching*, instead of simply maximizing the response of a single lobe having a fixed position with respect to the array. The lobe is switched back and forth *rapidly* by an amount which puts the direction of equal amplitude sufficiently down on the side of the lobes (usually about 85% of maximum) that a very slight variation in direction represents an appreciable change in response.

Referring to figure 10-5, the associated indicating mechanism is synchronized in such a manner that the response to the antenna with the lobe in position "1" is displayed separately from the response of the antenna with the lobe in position "2," or else the two responses are differentiated. The ability of the system to distinguish very small differences in direction can be explained as follows.

For a given orientation of the antenna, the response or field strength will have a constant amplitude in a direction along the line A-B during commutation of the lobe, but slightly off this direction the response or field strength provided by one lobe will increase while that of the other will decrease. It is also apparent that it is an easy matter to correlate the synchronized indicator or display apparatus to give "sense." That is, if the lobe is shifted back and forth in azimuth, it is possible not only to tell when the antenna orientation is "dead on," but also to tell whether it is to right or left when it is off. Or, if the lobe is shifted in angular elevation, it is possible to tell whether the antenna orientation is dead on, high, or low.

The lobe usually is deflected back and forth (or up and down) an equal amount from a line normal to the array in a manner similar to or related to that employed in the "Electrotator" array described in Chapter 5. For instance, if the fixed contacts of a synchronized d.p.d.t. switch were connected to "straddle" the dead center position on the phase shifting section of the Electrotator array, and the movable contacts connected to the pigtail, the lobe could be deflected equal amounts to either side of dead center by means of the switch.

Essentially the same thing is done in the arrangement illustrated in the skeleton schematic of figure 10-5, which represents a simple form of array consisting of two colinear half-wave dipoles fed with coaxial line. However, the schematic also applies to a large curtain or "mattress" divided into two identical sections, regardless of the number of dipoles employed.

The amount the beam is deflected from the normal depends upon the electrical length of the delay section. For instance, if both the schematic and polar diagrams of figure 10-5 are taken as top views, then with the lobe switch in position 1 the energy picked up by the right dipole does not have to travel as far to reach the receiver as does the energy picked up by the left dipole. However, the energy from the left and right dipoles will arrive at the receiver at the same time and thus provide maximum reenforcement if the wave front strikes the left dipole sufficiently ahead of the right dipole, as could occur in the case of a wave front arriving from a particular direction (determined by the length of the delay section) to the left of a line normal to the dipoles.

It is assumed, of course, that the electrical length of the delay section is small in terms of wavelength, because as the electrical length of the delay section is increased indefinitely the lobe rotates through 360 degrees in a counterclockwise direction, causing it to be deflected *toward* the normal over part of each cycle.

In actual practice the beam never is deflected more than about half the width of the lobe at the half power points; so aspect directivity of the dipoles, which was ignored in the preceding discussion of the effect of the delay section on lobe travel, need not be considered.

While the contacting type lobe switch shown in figure 10-5 can be used in a practical system, the contacts eventually tend to become noisy and also may give trouble when the same antenna is used for both transmission and reception of pulse signals, as in radar, if the peak power of the transmitter is high. This is avoided by the use of a *capacity* type lobe switch using a grounded rotating plate and two stationary electrodes. In this case the transmission line is connected directly to the mid-point of the line joining the two halves of the array, and the stationary plates of the capacity switch are connected via a half wavelength section of line to the points where the stationary switch contacts are connected in figure 10-5. This produces the effect of throwing lumped reactance first across one branch and then the other, thus altering the electrical length of the line first to one half of the array and then the other half.

Analysis of the principle upon which lobe switching systems operate will show that the lobe need not of itself be sharp in order to obtain good accuracy (though the obtainable accuracy does increase with the sharpness of the lobe). In fact, if the lobes are too sharp, difficulty will be obtained in "getting on" the target, beacon, etc. Also, because a moderate displacement of a lobe-shifted array exhibiting very sharp lobes will put the target, beacon, etc. completely outside the effective coverage of both lobes, the "tracking" of rapidly moving targets will be made more difficult.

The lobe switching system just described can be used to give either azimuth (left-right) or angular elevation (up-down). One such installation can be used to give one and then the other by first determining the azimuth and then rotating the antenna system through 90 degrees in order to determine angular elevation. However, physical rotation of the antenna through 90 degrees presents somewhat of a problem when a large array is used, and anyway it is desirable to be able to determine both azimuth and angular elevation simultaneously when information on both is desired.

It is possible to obtain both azimuth and elevation information simultaneously (or at least so it seems to the senses) by resorting to a more complicated switching arrangement. The array is divided into *four* equal segments (two over two), instead of just two, and the lobe is consecutively and very rapidly deflected up, left, down, right, and so on around, by means of a four segment lobe switch. The upper and lower lobes are paired off and fed to one indicator system, and the left and right lobes are paired off and fed to another indicator system.

CONICAL SCANNING

By going one step further, it is possible to simplify the system in the following manner. At frequencies above approximately 500 Mc. it is possible to rotate a dipole radiator physically without encountering mechanical difficulties. If the dipole is used to illuminate a parabolic reflector and is mounted slightly off center with respect to the axis of the paraboloid (displacing the dipole a short distance laterally with respect to the focal point), then rotating the dipole by means of a "scanning motor" which is aligned on the axis of the paraboloid *produces the same effect as rotating the entire array*. The lobe is deflected away from the axis of the reflector by a fixed angle, and as the dipole is rotated or "spun," the axis of the lobe describes a cone symmetrical with respect to the axis of the reflector. The scanning motor also drives a four segment commutator or switch which is synchronized with the indicating system in such a manner as to give up-down and left-right sampling in the same fashion as the four-sector lobe switching arrangement.

SPARK GAP T/R SWITCHES—DUPLEXING

Most pulse type radar apparatus employs the same antenna system for both transmitter and receiver. An ordinary mechanical switch cannot be made to operate fast enough, for obvious reasons, and therefore is unsuitable for switching the antenna back and forth between transmitter and receiver. The switching problem is solved by utilizing one or more spark gap switches, which simply are spark gaps so placed that when the transmitter fires it triggers off the gaps and causes a low impedance arc to appear across each gap, instead of an air gap having practically infinite impedance. While the impedance of an arc is not as low as that of an actual short circuit, it is sufficiently low that by means of sections of transmission line acting as linear transformers the antenna is effectively disconnected from the receiver and connected to the transmitter when the application of transmitter power causes the gap(s) to arc across.

This arrangement, called *duplexing*, provides almost instantaneous switching, and effectively prevents the loss of transmitter power and receiver sensitivity that would occur if switching were not employed, as well as protecting the components of the receiver input circuits from damage caused by excessive input voltage.

The basic principle of the spark gap T/R (transmit-receive) system is illustrated in figure 10-6. When the transmitter is inoperative, the line sections included between D and A constitute a ¾ wavelength shorted section, producing effectively a very low impedance at point C, where the transmitter is connected, and, because the transmitter does not significantly load the line when connected to the low impedance point at C, presenting a very high impedance at point D. Under these conditions practically all of the energy from the antenna, upon reaching the T junction at D, is delivered to the receiver input circuit. In effect, the receiver may be considered to be fed by a 300-ohm flat line which is shunted with a very lightly loaded antiresonant section of line. Because the antiresonant section presents an impedance of

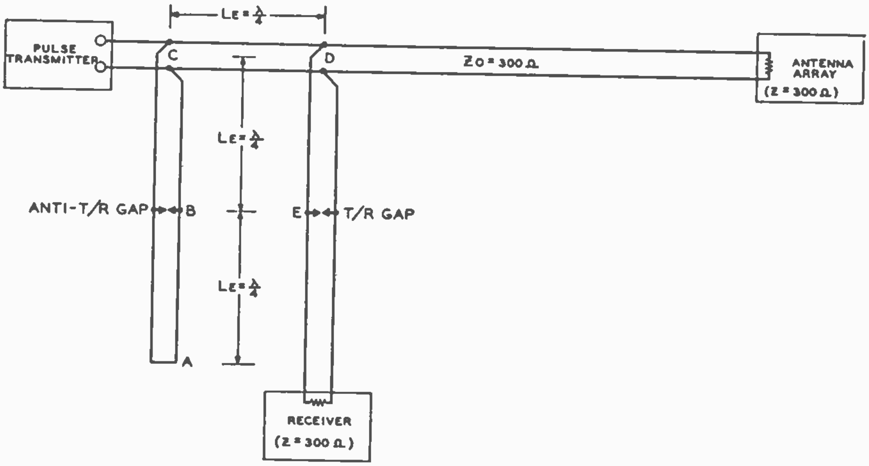


Figure 10-6.

ILLUSTRATING BASIC PRINCIPLE OF SPARK GAP T/R SYSTEM (DUPLEXING)

The antenna system is "self-switched" from "receive" to "transmit" by means of spark gap switches when the transmitter is keyed. The anti T/R gap at B is required only under certain conditions. The spark gaps together with their associated resonant line sections are known collectively as a DUPLEXER. The manner in which the spark gaps effectively switch the antenna from receiver to transmitter is described in the accompanying text.

several thousand ohms at point D, it produces negligible reduction in the strength of the signal reaching the receiver.

When the transmitter is "keyed" with a pulse, both spark gaps fire and produce very low impedance arcs at points B and E. The circuit then may be considered as consisting of a flat 300-ohm line from transmitter to antenna, shunted by two anti-resonant quarter-wave stubs. The impedance presented by the stubs at points C and D is not as high as would occur with actual shorts at points B and E, but because the impedance of the arcs at B and E is much lower than the line impedance, the reflected impedance at points C and D is much higher than the 300-ohm impedance of the line, and therefore very little power is wasted.

If the transmitter is so designed that when inoperative between pulses the only loading it produces on the feed line is that which is represented by the transmitter tank circuit losses, then the anti-T/R spark gap at B (and its associated stub A-C) may be eliminated with no detrimental effects. In such a case the length of line from transmitter tank to point D is made an integral number of electrical half wavelengths. The impedance which the transmitter tank circuit reflects at point D then is the same as the impedance across the line at the transmitter, which under the aforementioned conditions is quite high (several times the line impedance of 300 ohms). As in the previous case, most of the energy from the antenna will be delivered to the receiver, with only a very small percentage being dissipated by the transmitter tank circuit and tubes. When the transmitter is pulsed the conditions are the same as before except that there is only one anti-resonant stub shunted across the line instead of two.

Types of T/R Spark Gap

Various types of spark gap are used for T/R switching. The basic construction of the simplest type resembles an ordinary automobile spark plug. However, special features sometimes are incorporated to improve the operation in certain applications.

An ordinary open type gap exhibits a resistance during conduction of from about 25 to 50 ohms for typical values of current, and requires from about 9 to 12 microseconds for complete deionization. This resistance is low enough for switching a two-wire open line, but is not satisfactory for use with coaxial lines because of their low characteristic impedance. Also, radar and beacon equipment utilizing a very short pulse for good range resolution and a short minimum range requires a gap which will deionize more quickly.

The striking voltage, arc impedance, and the deionization time all can be reduced considerably by placing the gap in a glass envelope which contains a very small amount of water vapor at a pressure of less than 1 millimeter of mercury. A "keep alive" or "ignitor" electrode sometimes is added to the T/R tube to facilitate breakdown. It is connected to a constant source of power and maintains a glow discharge within the tube between pluses, but *not across the gap electrodes.*

INDEX

ERRATA AND ADDENDA

The *optical horizon* referred to on pages 22-23 and 63-65 is the *geometrical* horizon. It assumes rectilinear propagation of light waves. In actual practice light waves may be bent slightly by an atmosphere which is not homogenous, and there may be a slight difference between the visual horizon and the geometrical horizon.

* * *

The first line of paragraph 3 in column two, page 73, should read as follows: "Clouds do not produce *substantial* reflection of . . ."

* * *

Throughout the book reference is made to a very slight superiority of horizontal polarization (as compared to vertical polarization) for v-h-f and u-h-f communication. The reader is warned that as of the time this book goes to press the assertion is based upon experimental data not accepted by all investigators as adequate or conclusive, and that as more data become available for correlation it may be necessary to revise or qualify the statement. There is good reason to believe that the relative effectiveness may be determined by the receiving site and the nature of the intervening terrain to such an extent that no general statement as to equality or the superiority of one polarization over the other is justified.

A

B

C

T

U

V

W

X

Y