# ANNEX 1
# THE PROPAGATION MEDIA

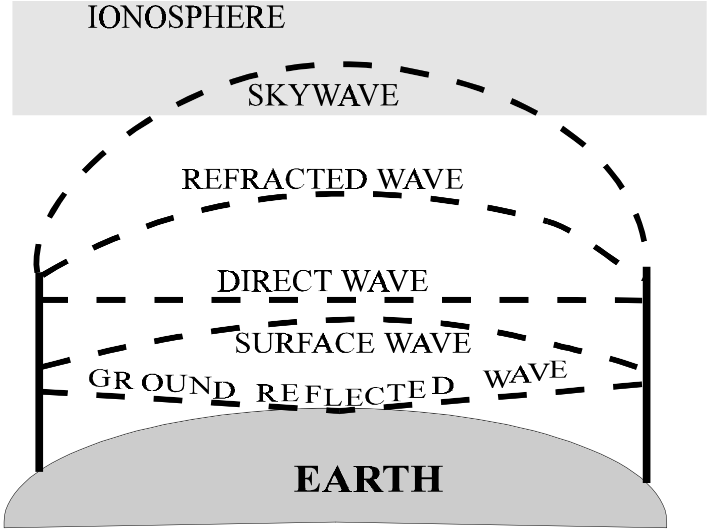
## Introduction

There is a high correlation between solar activity and HF skywave propagation. The reflecting and absorbing layers of the ionosphere are produced by and are largely controlled by the radiation from the Sun. Some of the variations in the ionospheric characteristics are more or less regular and can be predicted with reasonable accuracy. Other variations, resulting from abnormal behavior of the Sun, are irregular and unpredictable. The one selectable parameter in radio communication is the frequency to be used. Frequencies that allow the best propagation vary with time of day, day of the week, season of the year, and even with the 11-year solar cycle. This annex deals with the concepts of understanding the propagation media. It introduces the concepts associated with radio-wave propagation and illustrates how the Sun influences all radio communication beyond ground-wave or line-of-sight communication. But, primarily, this annex discusses the details of HF radio-wave communication, and (since this frequency band is by far the most sensitive to ionospheric effects) how the propagation effects can be handled.

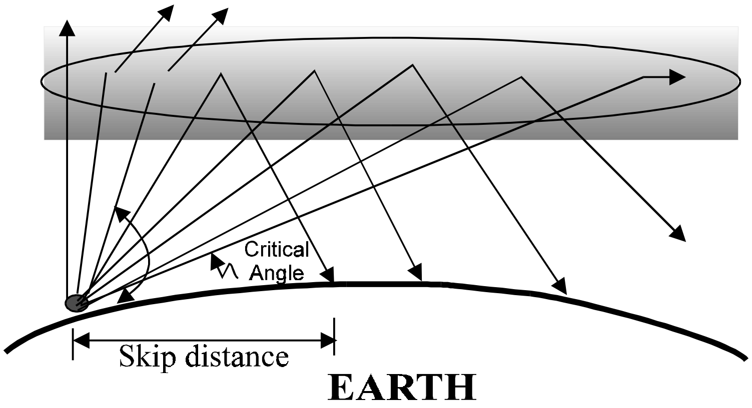
## A1-1 Radio propagation

The usable frequency range for radio waves extends from the highest frequencies of sound, about 20 kHz, to above 30,000 MHz. The frequency band from 3 to 30 MHz is designated as the high frequency (HF) band. Most of the newer HF radios can operate in a larger range of 1.6 to 30 MHz, or higher. Most long-haul communications in this band, however, generally take place between 4 and 18 MHz. Depending on ionospheric conditions and the time of day, the upper frequency range of about 18 to 30 MHz may also be available. The HF band, of all of the frequency bands, is by far the most sensitive to ionospheric effects. HF radio waves, in fact, experience some form of almost every known propagation mode. The sun influences all radio communication beyond ground-wave or line-of-sight ranges. Conditions vary with such obvious sun-related cycles as time of day and season of the year. Since these conditions differ for appreciable changes in latitude and longitude, and everything is constantly changing as the Earth rotates, almost every communications circuit has unique features with respect to the band of frequencies that are useful and the quality of signals in portions of that band.

The two basic modes of radio wave propagation at HF are ground wave and skywave. Figure A1-1 illustrates these two modes.



a. Groundwave



b. Skywave

FIGURE A1-1
**Ground wave and skywave**

---

### A1-1.1 Ground waves

A ground wave, as the name implies, travels along the surface of the earth, thus enabling short-range communications. Ground waves are those portions of the radio-wave radiation directly affected by the surface of the Earth. The principal components are

1) an Earth-guided surface wave,

2) a direct wave,
3) a ground-reflected wave,
4) a space wave, and sometimes
5) a tropospheric-reflected/refracted wave.

*Ground-wave* communication is more straightforward than skywave and it is generally assumed that the ground wave is merely an attenuated, delayed, but otherwise undistorted version of the transmitted signal. The received strength of transmitted radio signals in the ground-wave mode is dependent on such factors as: transmitter power, receiver sensitivity, ground conductivity and terrain roughness, antenna characteristics (such as height, polarization, directivity and gain), the radio frequency, and the type of path traveled. For a given complement of equipment, the range may extend out to as far as 400 km (250 mi) over a conductive, all-sea water path, but over arid, rocky, non-conductive terrain, however, the range may drop to less than 30 km (20 mi), even with the same equipment. Ground-wave propagation is almost always vertically polarized.

The *surface wave* is that component of the ground wave that is affected primarily by the conductivity and dielectric constant of the Earth and is able to follow the curvature of the Earth. When both transmitting and receiving antennas are on, or close to, the ground, the direct and ground-reflected components of the wave tend to cancel out, and the resulting field intensity at the receiving antenna is principally that of the surface wave. The surface-wave component is not confined to the Earth’s surface, however, but extends up to considerable heights, diminishing in field strength with increased height. Because part of its energy is absorbed by the ground, the electric intensity of the surface wave is attenuated at a much greater rate than inversely as the distance. This attenuation depends on the relative conductivity of the surface over which the wave travels. The best type of surface for surface-wave transmission is sea water. The electrical properties of the underlying terrain that determine the attenuation of the surface-wave field intensity vary little from time to time, and therefore, this type of transmission has relatively stable characteristics. The surface-wave component generally is transmitted as a vertically polarized wave, and it remains vertically polarized at appreciable distances from the antenna. This polarization is chosen because the Earth has a short-circuiting effect on the electric intensity of a horizontally polarized wave but offers resistance to this component of the vertical wave.

Absorption of the radio wave increases with frequency and limits useful surface-wave propagation to the lower HF range. At frequencies below about 5 MHz, the surface wave is favored because the ground behaves as a conductor for the electromagnetic energy. Above 10 MHz, however, the ground behaves as a dielectric. In the region below 10 MHz, conductivity of the surface is a primary factor in attenuation of the surface wave. As frequencies approach 30 MHz, losses suffered by the surface wave become excessive and ground-wave communication is possible only by means of direct waves.

*Direct waves*, also known as line-of-sight (LOS) waves, follow a direct path through the troposphere from the transmitting antenna to the receiving antenna. Propagation can extend to somewhat beyond the visible horizon due to normal refraction in the atmosphere causing the path to be somewhat bent or refracted. Because the electric field intensity of a direct wave varies inversely with the distance of transmission, the wave becomes weaker as distance increases, much like the light beam from a lantern or headlight. The direct wave is not affected by the ground or by the tropospheric air over the path but the transmitting and receiving antennas must be able to “see” each other for communications to take place, making antenna height a very critical factor in determining range. Almost all of the communications systems above 30 MHz use the direct (LOS) mode. This includes the commercial broadcast FM stations, VHF, UHF, microwave, cellular telephone systems, and satellite systems.

*Space waves* constitute the combination of all signal types which may reach a receiver when both the transmitting and the receiving antennas are within LOS. In addition to the direct signal, space waves include all of any earth-reflected signals of significance and, under specific conditions, would include undesirable strong secondary ionospheric modes as well. Space waves will support a relatively high signal bandwidth, as compared to ionospheric modes.

*Ground-reflected waves* result from a portion of the propagated wave being reflected from the surface of the earth at some point between the transmitting and receiving antenna. This causes a phase change in the transmitted signal and can result in a reduction or an enhancement of the combined received signal, depending on the time of arrival of the reflected signal relative to the other components.

*Tropospheric-reflected/refracted waves* are generated when abrupt differences in atmospheric density and refractive index exist between large air masses. This type of refraction, associated with weather fronts, is not normally significant at HF.

### A1-2.2 Skywaves

Skywaves are those main portions of the total radiation leaving the antenna at angles above the horizon. The term skywave describes the method of propagation by which signals originating from one terminal arrive at a second terminal by refraction from the ionosphere. The refracting (bending) qualities of the ionosphere enable global-range communications by “bouncing” the signals back to Earth and keeping them from being “beamed” into outer space. This is one of the primary characteristics of long-haul HF communication -- its dependence upon ionospheric refraction. Depending on frequency, time of day, and atmospheric conditions, a signal can bounce several times before reaching a receiver which may be thousands of kilometers away. Ionospheric skywave returns, however, in addition to experiencing a much greater variability of attenuation and delay, also suffer from fading, frequency (doppler) shifting and spreading, time dispersion and delay distortion.

Nearly all medium- and long-distance (beyond the range of ground wave) communications in the HF band is by means of skywaves. After leaving the transmitting antenna, the skywave travels from the antenna at an angle that would send it out into space if its path were not bent enough to bring it back to Earth.  The radio wave path is essentially a straight line as it travels through the neutral atmospheric region below the ionosphere.  As the radio wave travels outward from the Earth, the ionized particles in the ionosphere bend (refract) the radio waves.  Figure A1-2 is an idealized depiction of the refraction process.  Depending on the frequency and ionospheric ionization conditions, the continual refraction of the radio waves results in a curved propagation path.  This curved path can eventually exit the ionosphere downward towards Earth so that the radio waves return to the Earth at a point hundreds or thousands of kilometers from where they first entered the ionosphere.  In many cases, the radio waves reenter the ionosphere by “bouncing” from the Earth, and again be refracted back at a further distance.  This is known as multihop and, under the right conditions, will give global reach.  On a single HF link, many single-hop and multihop propagation paths are frequently possible.

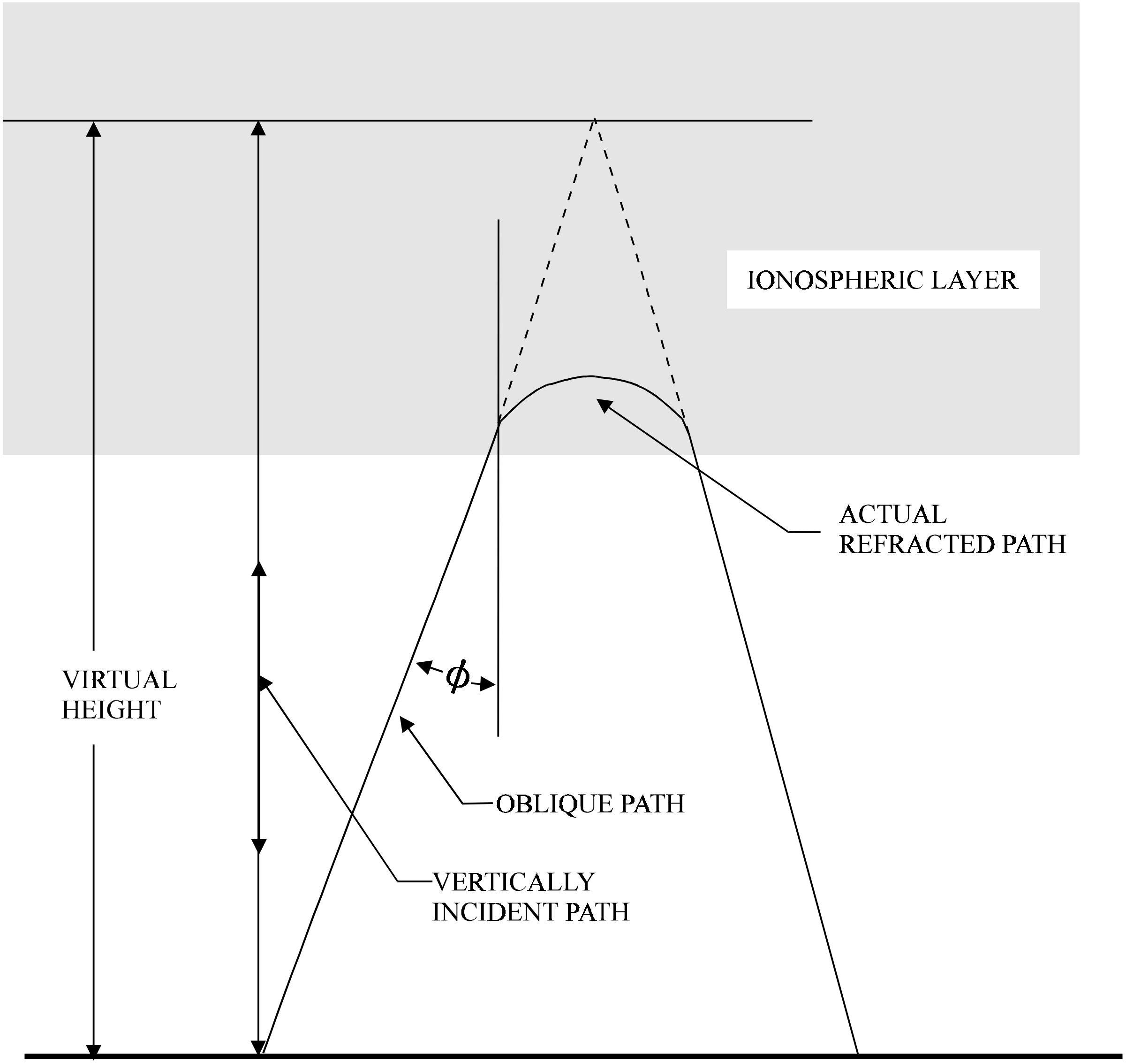


FIGURE A1-2
**Refraction of radio waves**

HF communication is widely regarded as the most challenging radio communication medium. HF systems that can adapt their operating frequencies to changing environmental conditions do much to offer reliable skywave communications. These systems must also accommodate received signal fading, frequency (Doppler) shift and spread, time delay shifts and spreads, and the results of multipath propagation (a radio signal that has been reflected from more than one ionospheric layer). The enormous variability of the ionosphere, however, makes it difficult to model the HF channel adequately.

The angle at which sky waves enter the ionosphere is known as the angle of incidence. This is determined by wavelength and the type of transmitting antenna. A radio wave refracts from the ionosphere at the same angle that it enters. Thus, the angle of incidence is an important factor in determining communications range. For long distances, the incident angle needs to be large, and conversely, for relatively short distances, the incident angle needs to be small. The frequency and the angle of incidence can often be changed to optimize link performance. The angle of incidence is critical, because if it is too nearly vertical, the radio wave will pass through the ionosphere without being refracted back to Earth. If the angle is too great, the waves will be absorbed by the lower ionospheric layers before reaching the more densely ionized upper layers. As shown on Figure A1-2, the virtual height is considerably greater than the actual layer height.

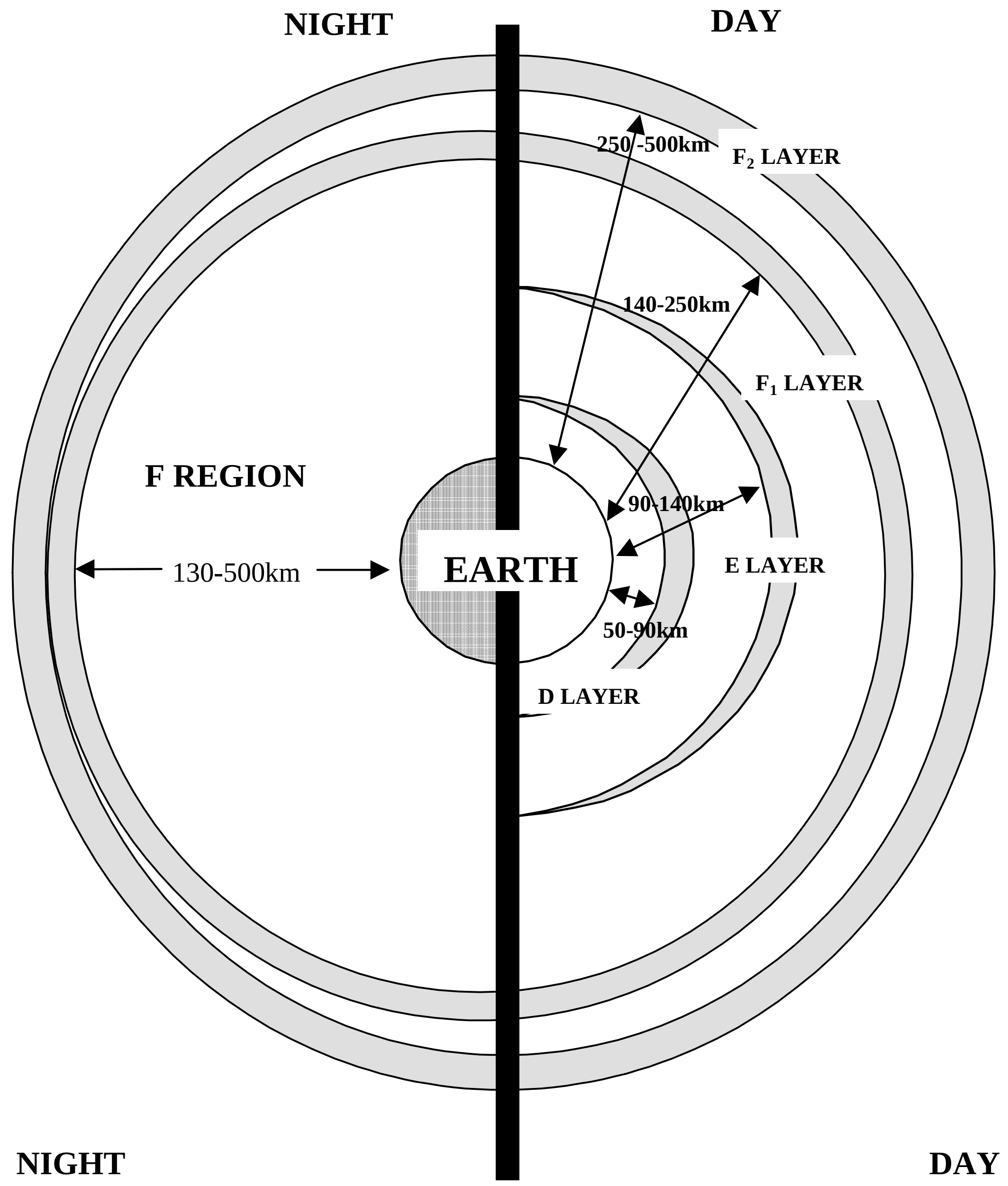
Virtual height is a convenient and an important quantity in measurements and applications involving ionospheric reflections. Because an ionospheric layer is a region of considerable depth, for practical purposes it is convenient to think of each layer as having a definite height. The height from which a simple reflection from the layer would give the same effects of the gradual bending that actually takes place is called the virtual height of that layer.

Among other qualities of a radio wave are strength and polarization. The strength (field strength or field intensity) of a radio wave decreases rapidly as the wave travels from the transmit antenna into free space because the original energy must be distributed over an increasingly larger surface area. Polarization generally matches the orientation of the electric (E) field, that is, if the E-field is parallel to the surface of the earth, the wave is said to be horizontally polarized. And, conversely, if the E-field is perpendicular to the Earth’s surface, the wave is said to be vertically polarized.

Ionization is caused primarily by solar radiation. As was mentioned earlier, there are two types of radiation from the Sun that influence propagation -- electromagnetic radiation and particle emissions. The electromagnetic radiation includes EUV, UV, and X-rays. Each type of radiation has an impact on a different part of the ionosphere. A long term periodic variation results from the 11-year sunspot cycle. Sunspot cycles average 10.7 years in length, but have been as short as 7.3 years and as long as 17.1 years. The highs and lows also vary greatly. Sunspots generate bursts of radiation that cause higher levels of ionization. The more sunspots, the greater the ionization. At solar maxima (high sunspot number), the radiation from all the active regions around the sunspots makes the ionosphere capable of returning higher-frequency radio signals to the Earth instead of allowing them to pass through. It has not been unusual to have worldwide propagation on frequencies above 30 MHz. Frequencies up to 40 MHz or higher are often usable during solar maxima for long-distance communication. During periods of minimum solar activity, however, the amount of radiation is reduced, and frequencies above 20 MHz become unreliable for long-distance communication on a day-to-day basis. This is because the E- and F-layers are too weakly ionized to reflect the higher frequency signals back to Earth. There are no absolutes at either solar maximum or solar minimum. Periods of extremely intense activity have been observed from a single sunspot region during solar minimum and large numbers of very low radiation sunspots have also been observed during a solar maximum. These conditions allow unexpected propagation conditions to exist at either solar minimum or maximum. Particle emissions come in two forms -- high-energy particles (high-energy protons and alpha particles) and low-energy particles (low-energy protons and electrons). The different types of radiation travel from the Sun to the Earth at different speeds and have varied effects on skywave propagation. EUV ionizes the ionospheric F-region and is always present at some level. During periods of increased solar activity, the increased number of active regions on the Sun provide for increased EUV, which in turn increases the F-region ionization. This makes the use of higher frequencies for long-distance communication possible. UV and X-ray emissions, however, ionize the D-region, which absorbs HF energy. During solar flares, UV and X-ray emissions generally increase considerably and cause increased signal loss on those HF circuits facing the Sun. Traveling at the speed of light, it takes about 8.3 minutes for electromagnetic emissions from the Sun to travel to the Earth. High-energy particle radiation travels more slowly than does light, and reaches the Earth in from 15 minutes to several hours after a large solar flare. The high-energy particle emissions cause much higher absorption in the Earth's polar regions. They also create a radiation hazard to satellite systems and personnel orbiting the Earth in spacecraft. The lower-energy particles travel even more slowly, typically taking 20 to 40 hours, and cause magnetic disturbances, auroras, sporadic-E layer ionization, and increased polar-region absorption.

## A1-2 The ionosphere

The ionosphere is a region of electrically charged gases and particles in the earth's atmosphere, which extends upward from approximately 50 km to 600 km (30 to 375 miles) above the earth's surface. See Figure A1-3. During daylight hours, the lower boundary of the ionosphere is normally about 65 to 75 km above the earth's surface, but can be as low as about 50 km. At night the absence of direct solar radiation causes the lower boundary to move upward to about 100 km.



**NOTE: NOT DRAWN TO SCALE**

FIGURE A1-3
**Earth with ionospheric layers (day and night)**

---

The ionosphere is made up of several ionized regions, which play a most important part in the propagation of radio waves. These regions have an influence on radio waves mainly because of the presence of free electrons, which are arranged in generally horizontal layers.

Ionization is the process of creating free electrically charged particles (ions and free electrons) in the atmosphere, thus establishing the ionosphere. The Sun is the primary “engine” of ionization. The earth’s atmosphere is composed of many different gases. Because the sun emits radiation in a broad spectrum of wavelengths, different wavelengths ionize the various atmospheric gas molecules at different altitudes. This results in the development of a number of ionized layers. Extreme ultraviolet (EUV) radiation from the sun is a primary force in the ionization process. The various types of gas molecules in the upper atmosphere have different susceptibilities to ionization, primarily based on the wavelengths of the ionizing radiation. The short-wavelength solar radiation, including EUV, is sufficiently intense during daylight hours to alter the electronic structure of the various gas molecules above altitudes of about 65 km. In general, the interactions between ions, free electrons, and background neutral molecules in the ionosphere involve chemical, electrodynamic, and kinetic forces. The existence of charged particles in the ionosphere allows electrical forces to affect the motions of the atmospheric gas.

The intensity of solar radiation and therefore ionization varies periodically allowing prediction of solar radiation intensity based on time of day and the season. Ionization is higher during spring and summer because the hours of daylight are longer and conversely lower during the fall and winter because the hours of daylight are shorter. This ionized ionospheric structure also varies widely over the Earth’s surface, since the strength of the sun’s radiation varies considerably with geographic latitude, time of day, season, sunspot activity, and whether or not the ionosphere is disturbed. The intensity of the solar radiation tends to track solar activity, especially the sun spot activity. In addition to ionizing a portion of the neutral gas, solar radiation also breaks down some of the neutral molecules, thereby changing the composition of the upper atmospheric gas. Although the principal source of ionization in the ionosphere is electromagnetic radiation from the sun, there are other important sources of ionization, such as solar particles and galactic cosmic rays. The ionization rate at various altitudes depends upon the intensity of the solar radiation and the ionization efficiency of the neutral atmospheric gases. Collisions in the atmosphere, however, usually result in the recombination of electrons and positive ions, and the reattachment of electrons to neutral gas atoms and molecules, thus decreasing the overall ionization density.

For the purpose of propagation prediction and ionospheric studies, it is frequently useful to separate the environment (especially the ionosphere) into two states, benign and disturbed. The benign ionosphere state is that which is undisturbed by solar flares, large geomagnetic storms, and known manmade (including nuclear) events. Even then, there is still a significant variability, partly due to the effects of such phenomena as traveling ionospheric disturbances (TIDs), sudden ionospheric disturbances (SIDs), sporadic-E, and spread-F, as examples. The disturbed ionosphere is a state that includes the effects of several disturbing influences which occur quite naturally. Solar flares, geomagnetic storms, and nuclear detonations will cause significant ionospheric changes. Disturbances may also be produced by the release of certain chemicals into the ionosphere. The magnitudes of the introduced effects vary widely. Certain regions of the ionosphere, such as the auroral zone and the equatorial region (in certain categories), are always in the disturbed state.

### A1-3 Ionospheric layering

Within the ionosphere, there are four layers of varying ionization that have notable effects on communications. As has been noted, solar radiation (EUV, UV, and X-rays) and, to a lesser extent cosmic rays, act on ionospheric gases and cause ionization. Since these ionization sources vary both in energy level and wavelength (frequency), they penetrate to different depths of the atmosphere and cause different ionization effects. The natural grouping of energy levels results in distinct layers being formed at different altitudes.

At altitudes below about 80 km, winds and weather patterns cause a turbulent mixing of the atmospheric gases present at these lower levels. This turbulent mixing diminishes as altitude increases and as the stratification (or layering) of the constituent gases becomes more pronounced. The density of ionized gases and particles increases with altitude to a maximum value, then decreases or remains constant up to the next layer. The higher layers of the ionosphere tend to be more densely ionized and contain the smaller particles, while the lower layers, which are somewhat protected by the higher ones, contain the larger particles and experience less ionization. The different ionospheric gases each have different ionizing wavelengths, recombination times, and collision cross sections, as well as several other characteristics. All of this results in the creation of the ionized atmospheric layers. The boundaries between the various ionospheric layers are not distinct, because of constant motion within the layers and the changeability of the ionizing forces.

The ionospheric layers that most influence HF communications are the D, E, $F_1$, and $F_2$ layers, and, when present, the sporadic-E layer. Of these, the D-layer acts as a large rf sponge that absorbs signals passing through it. Depending on frequency and time of day, the remaining four ionized layers are useful (necessary!) to the communicator and HF communications.

Due to the ionization effects of the solar zenith angle (height of the Sun in the sky), the altitudes of the various layers and their relative electron densities at any time depend on the latitude. For mid-latitudes, the following are typical layer (region) altitudes and extent:

D-region -- 70 to 90 km (a bottom level of 50 km is not too unusual)
E-region -- 90 to 140 km
Sporadic-E region -- typically 105 to 110 km
F-region -- from about 140 km to as high as 1000 km

$F_1$-region -- 140 to over 200 km (during daylight only)
$F_2$-region -- 200 to about 500 km

The hourly, daily, seasonal, and solar cycle variations in solar activity cause the altitudes of these layers to undergo continual shifting and further substratification.

### A1-3.1 D-layer

The D-layer, which normally extends from 70 to 90 km above the Earth, is strongest during daylight hours with its ionization being directly proportional to how high the sun is in the sky. This layer often extends down to about 50 km. The electron concentration and the corresponding ionization density is quite small at the lowest levels, but increases rapidly with altitude. The D-region electron density has a maximum value shortly after local solar noon and a very small value at night because it is ionized only during the day. The D-layer is the lowest region affecting HF radio waves. There is a pronounced seasonal variation in D-region electron densities with a maximum in summer. The relatively high density of the neutral atmosphere in the D-region causes the electron collision frequency to be correspondingly high. The main influence of the D-region on HF systems is absorption. In fact, this region is responsible for most of the absorption encountered by HF signals which use the skywave mode. Because absorption is inversely proportional to frequency, wave energy in the lower end of the HF band is almost completely absorbed by this layer during daylight hours. The rise and fall of the D-layer, and the corresponding amount of radio wave absorption, is the primary determinant of the lowest usable frequency (LUF) over a given path. Due to the greater penetration ability of higher radio frequencies, the D-layer has a smaller effect on frequencies above about 10 MHz. At lower frequencies, however, absorption by the D-layer is significant. Absorption losses of the higher-frequency waves depend on the D-region ionization density, the extent of the region, the incident angle, the radio frequency, and the number of hops, among other factors. (For every hop, the rf wave traverses the D-region twice, once on the way up, and once on the way down.)

### A1-3.2 E-layer

The lowest region of the ionosphere useful for returning radio signals to the Earth is the E-layer. Its altitude ranges from about 90 km to about 130 km and includes both the normal and the sporadic-E layers. The average altitude of the layer’s central region is at about 110 km. At this height, the atmosphere is dense enough so that ions and electrons set free by solar radiation do not have to travel far before they meet and recombine to form neutral particles. It is also dense enough to allow rapid de-ionization as solar energy ceases to reach it. Ionization of this layer begins near sunrise, reaches maximum ionization at noon, and ceases shortly after sundown. The layer can maintain its ability to bend radio waves only in the presence of sunlight. At night, only a small residual level of ionization remains in the E-region. The normal E-layer is important for daytime HF propagation at distances of up to about 2000 km. Irregular cloud-like layers of ionization often occur in the region of normal E-layer appearance and are known as sporadic-E ($E_S$). These areas are highly ionized and are sometimes capable of supporting the propagation of sky waves at the upper end of the HF band and into the lower VHF band.

### A1-3.3 Sporadic E

In addition to the relatively regular ionospheric layers (D, E, and F), layers of enhanced ionization often appear in the E ($E_S$)-region and the lower parts of the F-regions (sporadic F). The significant irregular reflective layer, from the point of view of HF propagation, is the $E_S$-layer since it occurs in the same altitude region as the regular E-layer. Despite what their name implies, these layers are quite common. A theory is that $E_S$ occurs as a result of ionization from high altitude wind shear in the presence of the magnetic field of the Earth, rather than from ionization by solar and cosmic radiation. Another theory is that $E_S$-layers are thin patches of long-lived ions (primarily metallic) that are believed to be rubbed off from meteors as they pass through the atmosphere, and then are formed into thin layers by the action of tidal wind systems. Layers of sodium ions produced by similar mechanisms commonly appear in the 90-km altitude range. Because the recombination rates of metallic ions are extremely low in the ionosphere, these thin layers can persist for many hours before being neutralized by recombination and dispersed by diffusion and are most commonly observed at night when the background densities are low.. Areas of $E_S$ generally last only a few hours, and move about rapidly under the influence of high altitude wind patterns. Different forms of $E_S$, having different characteristics and production mechanisms, are found in the auroral zones and, at an attitude of about 105 km, in the low and middle equatorial latitudes. They share the common characteristics that they are all E-layer phenomena, their occurrence is not predictable, and they all have an effect on HF radio communications. When $E_S$ occurs, it produces a marked effect on the geometry of radio propagation paths which normally involve the higher layers. Their peak densities can sometimes exceed that of the higher altitude F-region. When this occurs, these layers can reflect incident HF waves at much lower altitudes and prevent reflections from the F-layer, thereby greatly reducing the expected range of transmission. Although $E_S$ is difficult to predict, it can be used to advantage when its presence is known. It has been found that close to the equator, $E_S$ occurs primarily during the day and shows little seasonal variation. By contrast, in the auroral zone, $E_S$ is most prevalent during the night but also shows little seasonal variation. In middle latitudes however, $E_S$ occurrence is subject to both seasonal and diurnal variations and is more prevalent in local summer than in winter and during the day rather than at night.

### A1-3.4 F-layer

The F-layer is the highest and most heavily ionized of the ionized regions, and usually ranges in altitude from about 140 km to about 500 km. At these altitudes, the air is thin enough that the ions and electrons recombine very slowly, thus allowing the layer to retain its ionized properties even after sunset. The F-layer is the most important one for long-distance HF propagation. If sporadic ionospheric disturbances are ignored, the height and density of this region varies in a predictable manner diurnally, seasonally, and with the 11-year sunspot cycle. Under normal conditions it exists 24 hours a day. The F-layers ionize very rapidly at sunrise and reach peak electron density early in the afternoon at the middle of the propagation path. The ionization decays very slowly after sunset and reaches the minimum value just before sunrise. At night, the layer has a single density peak and is called the F-layer. During the day, the absorption of solar energy results in the formation of two distinct density peaks. The lower peak, the $F_1$- layer, ranges in height from about 130 km to about 300 km and seldom is predominant in supporting HF radio propagation. Occasionally, this layer is the reflecting region for HF transmission, but in general, obliquely-incident waves that penetrate the E-region also penetrate the $F_1$-layer and are reflected by the $F_2$-layer. The $F_1$-layer, however, does introduce additional absorption of the radio waves. After sunset, the $F_1$-layer quickly decays and is replaced by a broadened $F_2$-layer, which is known simply as the F-layer. The $F_2$-layer, the higher and more important of the two layers, ranges in height from about 200 km to about 500 km. This $F_2$-layer reaches maximum ionization at noon and remains charged at night, gradually decreasing to a minimum just before sunrise. In addition to being the layer with the maximum electron density, the $F_2$-layer is also strongly influenced by solar winds, diffusion, magnetospheric events, and other dynamic effects and exhibits considerable variability. Ionization does not completely depend on the solar zenith angle because with such low molecular collision rates, the region can store received solar energy for many hours. In the daytime, the $F_2$-layer is generally about 80 km thick, centered on about 300 km altitude. At night the $F_1$-layer merges with the $F_2$-layer resulting in a combined F-layer with a width of about 150 km, also centered on about 300 km altitude. Due to the Earth/ionospheric geometry, the maximum range of a single hop off of the $F_2$-region is about 4000 km (2500 miles). The absence of the $F_1$-layer, the sharp reduction in absorption of the E-region, and absence of the D-layer cause night-time field intensities and noise to be generally higher than during daylight. Near the equator, there are significant latitudinal gradients in the F-region ionization. In the polar regions (high latitudes), there is a region of strongly depressed electron density in the F-layer. These can have important effects upon long-distance radio wave propagation.

### A1-4 Ionospheric effects on radio wave propagation

As was noted earlier, the ionosphere is created primarily by solar radiation and affects the area from about 50 to 600 km above the Earth. When radio waves strike the ionized layers, depending on the frequency, some are completely absorbed, others are reflected/refracted so that they return to the earth, and still others pass through the ionosphere into outer space.

### A1-4.1 Absorption

Absorption is the power robber and tends to be greater at lower frequencies and of greater effect as the degree of ionization increases. Below about 65-km altitude, the neutral atmosphere generally has very little effect on HF propagation. The sizes of water vapor molecules and other absorptive gases in the medium are insignificant compared to the HF wavelength, so there is very little interaction between the propagating HF radio waves and the components of the medium. In the D-region, however, collision-related absorption is usually the largest ionospheric-caused loss factor encountered in HF propagation. The primary cause of ionospheric absorption is the ionization due to the atmospheric bombardment primarily by X-rays and UV radiation. Absorption also increases markedly during solar flares and auroral disturbances. The D-region organizes at dawn, reaches a maximum at local noon, and dissipates when the Sun sets. Even though only about one tenth of one percent of the layer molecules are ionized, enough free electrons are present in this region to cause significant absorption of propagating HF signal energy. Absorption generally reaches a maximum during local daylight hours when free electrons exist at relatively low altitudes. Since the ionization efficiency within this region is a strong function of the solar zenith angle, the number of electrons available to contribute to the loss process varies with time of day, season, and geographic location. Although the 11-year solar cycle produces changes in solar radiation output, these changes are sufficiently slow that they are associated with the benign ionosphere. By contrast, abrupt changes in solar radiation influence not only the intensity of the ionization, but the distribution in the affected area as well. These disturbances can significantly affect HF propagation modes and disrupt HF communication systems.

### A1-4.2 Attenuation

The principal measure of the effects caused by the disturbances is the attenuation of the signal. This attenuation is the result of increased absorption plus the fading and distortion resulting from the delay spread (time dispersion) and Doppler spread (frequency dispersion) which are caused by time-varying multipath propagation. Skywave signals propagated over mildly disturbed ionospheric paths exhibit minimal Doppler spread (about 1 Hz) and moderate delay spread (about 0.1 ms). When the signals are propagated over moderately to severely disturbed channels, the multipath plus scatter from moving irregularities will typically result in Doppler spreads of 5 to 20 Hz and delay spreads of 3 to about 10 ms. The larger spread values are associated with the most disturbed ionospheric conditions in which signal scattering is the primary propagation mechanism. The received signal level from such a scatter propagation path can be as much as 30 dB below the level of signals propagated by the skywave refraction.

### A1-4.3 Day/night terminator

Normal day-to-night (sundown) and night-to-day (sunrise) transitions create sharp density irregularities in the ionosphere and form the so-called “day/night terminator.” Ionization increases dramatically as the sun rises at ionospheric altitudes. Along with abrupt increases in absorption, the sunrise transition can create relatively sharp horizontal gradients that refract HF waves in both azimuth and elevation. Path lengths for propagation at a given frequency will often change dramatically as the ionization density and layer heights of the ionosphere are modified by sunrise. For about an hour around the sunrise terminator, propagation conditions can be very unstable and can adversely affect HF communications. As the Sun rises, however, the sharp gradients smooth and very soon have little impact on propagation. After sunset, the fact that some ionization exists is due primarily to the relatively slow recombination of some of the ions and electrons.

Although the effects of the day/night terminator on propagation can be briefly significant, the impact on well-engineered HF systems can be minimized through prediction and mitigation efforts.

## A1-5. Radio wave propagation in the ionosphere

The ionosphere is responsible for several conditions that affect HF radio-wave propagation, both for ground wave and for skywave modes.

### A1-5.1 Attenuation and absorption

The usability of an HF path is highly dependent on the amount of attenuation the signal undergoes. The greatest attenuation of the skywave-propagated radio wave is usually caused by free-space spreading loss. Attenuation, in addition to the free-space spreading loss, can also be due to ionospheric absorption as well as the absorption effects associated with imperfect ground, foliage, and atmospheric factors. Ionospheric absorption occurs predominately in the D-region where there is little refractive bending at HF. In this region, free electrons, in addition to being influenced by the ionizing and electromagnetic forces, are also influenced by the electromagnetic field of any incident HF wave. As the electrons move under the influence of the applied electric field, they frequently collide with the more abundant neutral particles. Since the mass of an electron is thousands of times less than the mass of a neutral molecule, these collisions result in a significant loss of electron momentum and energy. Because the neutral particles do not oscillate in the applied field, they do not re-radiate the energy gained during these collisions but transform the energy into heat.

The net result is that free electrons in the upper atmosphere effectively absorb energy from the wave and disperse it throughout the neutral gas. Such absorption is greater at lower frequencies because it is proportional to the inverse square of the frequency. It is this absorption that defines the LUF (lowest usable frequency) (see par. 5j, below) of a circuit. If too much energy is absorbed, the signal will be too weak to be heard at the receiving station. A radio signal loses strength every time it passes through the D-region. For this reason, a three-hop signal will have more path loss (or less signal at the destination), than a one-hop signal.

In order to minimize absorption, frequencies as near the MUF (maximum usable frequency) (see par. 5h, below) as possible should be used. To obtain the maximum range for a single-hop propagation mode, the highest possible frequency and the lowest possible elevation angle that will produce an F-region reflection should be used. Moving to higher frequencies is also generally beneficial for increasing the propagated signal strength and therefore the SNR of the received signal. An HF wave is not simply reflected once at a single critical altitude, it is continuously refracted as it traverses the ionosphere.

### A1-5.2 Diffraction, reflection, and refraction

Diffraction is a wave property which is associated with the re-radiation of a wave when it encounters a surface or an obstacle.  This is of major concern for VHF and higher frequency systems.  At HF, diffraction can cause a phase shift and possible destructive multipath.

The relationship between radio frequency, the path incidence angle, and the amount of ionization can lead to skywave effects that range from benign (small additional losses) to severe (no signal refracted back to earth).  In the benign ionosphere, the signal is refracted without spending much time in the lower ionosphere, where absorption is substantial.  In the severe case, if the frequency is too high or the path incidence angle is too small, the transmitted signal will pass completely through the ionosphere and continue into space.  Between these two extremes lies a very wide range of possibilities.

Refraction occurs when the electron concentration changes, so that waves are refracted back towards the Earth if their frequency is not too high.  Reflection and refraction are two words that often seem to be used interchangeably, even though they describe quite different phenomena.

Reflection occurs at any boundary between materials with different dielectric constants.  Depending on their wavelength (or frequency), radio waves may be reflected by buildings, trees, vehicles, the ground, water, ionized layers in the outer atmosphere, or at boundaries between air masses having different temperatures and moisture content. Some of the radio energy will be absorbed in the medium that the wave hits, and some of it may pass into (or through) the material.

Refraction is the bending of a wave as it passes from one medium into another.  This bending occurs because the wave travels at a different speed in the new material.  Radio waves bend when they pass from one material into another in the same way that light waves do.  The degree of bending depends on the difference in speeds at which the waves move through the two materials.  Refraction occurs only when the wave approaches the new medium in an oblique direction.  If the whole wave front arrives at the new medium at the same moment, it is slowed up uniformly and no bending occurs.  The amount of bending also increases at higher frequencies.  On frequencies below 30 MHz, long-distance communications is the result of refraction.   Depending on the frequency used and the time of day, the ionosphere can support communications from very short ranges of less than 90 km (called near vertical incidence skywave -- NVIS) to distances of greater than 4000 km on a single hop and up to global distances with multihop propagation.

### A1-5.3 Other HF propagation factors

The MF and HF bands use the F-region ($F_2$-layer in the daytime) for the primary mode of long-distance propagation. The MF band (300-3000 kHz with wavelengths of 1000m to 100m) suffers from extreme daytime D-layer absorption. In this band, only daytime signals which enter the ionosphere at very high angles will escape complete absorption and will be returned to Earth. This limits communications to distances of only up to about 100 km or so. At night, when the D-layer dissipates, lower-angle signals can be propagated for very long distances.

Another general factor affecting propagation in this band is noise, both atmospheric and man-made. Since noise-producing thunderstorms and man-made noise are usually greater in warm weather, this band is generally more effective in winter. The lower part of the HF band has similar characteristics.

Although the ionosphere is multilayered, the F-layer typically has the most electrons and is the most heavily ionized. This F-layer electron dominance, however, does not always correspond on a one-to-one basis with radio-wave propagation effects. Lower layers (D, E, $E_S$) may have a major impact on HF system operation. As we have seen, the normal D-region acts as a power robbing attenuator to the HF signal. This is especially strong during the daytime when the Sun-caused ionization is at its greatest. Additionally, patches of sporadic-E ionization may blanket the upper ionopheric layers and serve as the primary reflection layer. This can have both favorable and negative implications for HF coverage.

The phase speeds of the various propagating modes are different, and depend on the characteristics of the propagation path as well as on the wave characteristics themselves. As a result, at the receiver, the propagating modes may be out of phase with one another. Because of the dynamic behavior of the ionosphere, the relative phase relationships between the propagating modes are usually in constant change. The received signal is the vector sum of all of the signals received at each moment and, with many propagating modes, generally results in destructive interference. This produces received signal level fading and the limited-bandwidth that is a characteristic of HF circuits.

The radiation from an HF antenna generally covers a rather large area, even for highly directive antennas. In considering the effects of frequency and angle of incidence on propagation, it is usually more convenient to think of several individual rays instead of a large illuminated area. For a fixed frequency, the paths of rays leaving the transmitter are as shown in Figure 3-4. For low angles of elevation, path **A** is long and the range (surface distance) is large. As the elevation angle increases, the range decreases (**B**) until the skip distance is reached at **C**. Note that as the elevation angle increases, the angle of incidence (Figure 3-3) decreases. Skip distance is the distance from the transmitter to the closest point that can be reached by skywave. For still higher angles of elevation, the ranges can increase rapidly (**D** and **E**). The reflection height of a fixed-frequency HF wave increases as the incidence angle decreases until the point is reached where the electron density is not sufficient (for that frequency) for the ray to be totally reflected and the ray (escape ray) penetrates the layer and does not return (**F**). This point is the critical angle of incidence. The small group of rays between the skip ray and the escape ray is dispersed over a great range. These are high-angle rays. The distances, and even the capability of propagation support for these high-angle rays are problematical in that all of the ionospheric variables come into play. Although their signal strengths may be small, workable signals can be received over high-angle paths. The critical angle is a function both of frequency and of ionization density. For any given ionization distribution, the frequency at which the critical angle reaches zero is known as the critical frequency. This is the maximum frequency which can be reflected as vertical incidence. Radiation at angles more than the critical angle above the horizon will not be returned to Earth, while radiation at angles less than the critical angle will be returned, the distance depending on the elevation angle (Figure A1-4). As the elevation angle is lowered, less refraction in the ionosphere is required to bring the radio wave back to Earth, or to maintain useful signal level. This is the basis for the emphasis on low radiation angles for long-distance circuits, although absorption (and thereby attenuation) increases in proportion to the length of the path through the lower (absorption) ionospheric layers. The high angle waves (escape rays) are bent only slightly in the ionosphere, and so pass on through it.

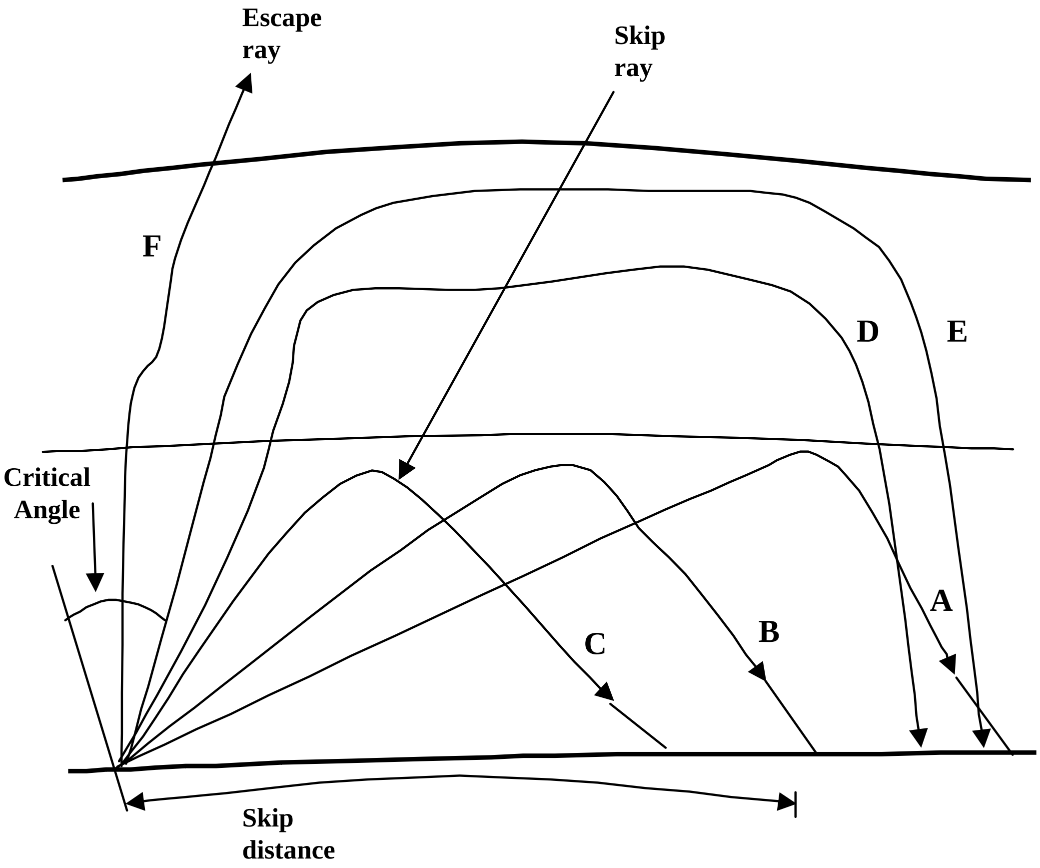


FIGURE A1-4
**Ray paths as a function of elevation angle for a fixed frequency**

### A1-5.4 Maximum usable frequency

The Maximum Usable Frequency (MUF) is the highest frequency at which radio waves are returned to Earth by ionospheric refraction and which can be used to transmit over a particular path under given ionospheric conditions at a specific time. In the theory, frequencies higher than the MUF penetrate the ionosphere and continue on into space, while frequencies lower than the MUF tend to refract in the ionosphere and return back to Earth. Because the ionization of the ionospheric layers is extremely variable, the MUF must be statistically defined based on the intensity of F-region ionization. The accepted working definition of MUF is that it is the highest frequency predicted to occur via a normal reflection from the $F_2$-layer (F-region at night) on 50 percent of the days of the month at a given time of day on a specified path. This definition depends only on the mode of propagation and the path geometry and is independent of HF radio system parameters and received noise. During a solar maximum, the MUF can rise above 30 MHz during daylight hours, but during solar minima, can be at or below the lower part of the HF band. In practice, the MUF is not a sharp limit and propagation is often possible on frequencies greater than the theoretical MUF.

From an operational perspective, the highest frequency that provides communications between two HF stations at a given time and under specific conditions is defined as the operational MUF. This operational MUF accounts for all HF system parameters such as transmitter power, antenna patterns, and receiver sensitivity and includes actual environmental conditions including the levels of man-made noise at each station. The maximum operational frequency may be appreciably higher than the statistically-predicted MUF. Neither the ground nor the ionosphere are smooth reflectors, as assumed in the simplified theory. Scattering from irregularities will frequently allow signals to propagate to distances beyond the limit of the refracted wave. The received signal characteristics are nearly always optimal for skywave communications when the link operating frequency is chosen at or just below the classical MUF.

Vertical-incidence ionosondes and vertical-incidence sounders are valuable tools for both real time and long-term pictures of the key MUF-determining parameters, such as peak densities and layer heights. The frequency at which the vertical-incidence signal is no longer reflected is the critical frequency at that time and at that location. HF users can interpret the vertical-incidence testers’ data (ionograms) from stations worldwide to help choose the appropriate frequencies and to modify the predicted MUFs with real and near-real time data.

The ionospheric layers, particularly the D- and E-layers, absorb radio-frequency (rf) energy as the radio waves pass through. This absorption is higher at the lower frequencies and is inversely proportional to the square of the frequency. Consequently, it is desirable to use the highest frequency possible -- the MUF. The MUF, as defined, is suitable for reliable communications only 50 percent of the time. The statistical MUF is also unreliable due to the ionospheric irregularities and turbulence as well as to the statistical deviation of day-to-day ionospheric characteristics from the predicted monthly median MUF value. Thus, a lower frequency is chosen to provide a margin for the daily variations in MUF. This frequency, the Frequency of Optimum Traffic (FOT), is usually calculated as 85 percent of the monthly median MUF. Statistically, the FOT so calculated lies below the actual daily MUF 90 percent of the time. It should be noted that the FOT bears this statistical relationship with the MUF and is not calculated as the frequency which will provide the maximum received signal power. Sophisticated propagation prediction programs do not use the 0.85 factor, but instead compute the FOT directly from statistical distributions.

### A1-5.5 Lowest useful frequency

The lowest useful frequency (LUF), as statistically calculated, is the lowest frequency at which the field intensity at the receiving antenna is sufficient to provide the required signal-to-noise ratio (SNR) on 90 percent of the undisturbed days of the month. Unlike the MUF and FOT, the LUF, in addition to the amount of absorption on the particular path, is also dependent on the link parameters such as transmitter power output, antenna efficiency, receiver sensitivity, required SNR for the service required, transmission mode, existence of multipath, and the noise levels at the receiving station. As frequency is lowered, the amount of absorption of the signal by the D-layer (and frequently the E-region) increases. This absorption is proportional to the inverse square of the frequency. Eventually, as the frequency is lowered further, the signal becomes completely absorbed by the ionosphere. This is the LUF. The LUF changes in direct correlation to the movement of the Sun over the radio path, and peaks at noon at the midpoint of the path. The “window” of usable frequencies, therefore, lies in the frequency range between the MUF and LUF.

When an ionospherically propagated wave returns to Earth, it can be reflected upward from the ground, travel again to the ionosphere, and again be refracted to Earth. This process, multihop propagation, can be repeated several times under ideal conditions, leading to very long distance communications. Propagation between any two points on the Earth’s surface is usually by the shortest direct route, which is a great-circle path between the two points. A great circle is an imaginary line drawn around the Earth, formed by a plane passing through the center of the Earth. The diameter of a great circle is equal to the diameter of the Earth (12,755 km or 7,926 mi at the equator). The circumference of the Earth -- and the “length” of any global great-circle path is about 40,000 km or 24,900 mi. Due to ionospheric absorption and ground-reflection losses, multiple-hop propagation usually yields lower signal levels and more distorted modulation than single hop signals. There are exceptions, however, and under ideal conditions, communications using long path signals may be possible. The long path is the other (long) route around the great-circle path. The same signal can be propagated over a given path via several different numbers of hops. The success of multihop propagation usually depends on the type of transmitter modulation and type of signal. For interrupted continuous wave (ICW) (also known as CW or Morse code) there is no real problem. On single-sideband (SSB), the signal is somewhat distorted due primarily to time spreading. For radio-teletypewriter (RTTY) or data service using frequency-shift keying (FSK), the distortion may be sufficient to degrade the signal to the point that it cannot be used. Every different point-to-point circuit will have its own mode structure as a function of time of day. This means that every HF propagation problem has a totally unique solution.

## A1-6. Anomalous ionospheric propagation and ionospheric disturbances

Deviations from normal ionospheric propagation occur as the result of certain irregularities and transient conditions in the ionosphere. There are also significant disturbances as a result of solar activity. Both the ionospheric disturbances and ionospheric propagation deviations can cause increased absorption, selective fading, signal dispersion and even complete loss of the signal. The most significant of the anomalous propagation mechanisms are sporadic E (previously discussed), spread-E or spread-F (depending upon the ionospheric layer affected), sudden ionospheric disturbances (SIDs), ionospheric storms, and polar cap absorption. Most of these anomalous propagation mechanisms are solar-related. Solar flares are the most frequent and best known solar events. Less common and less known are coronal mass ejection (CME) and significant proton events. These, especially if focused on Earth, can cause severe, even catastrophic, damage to Earth communications throughout the electromagnetic spectrum. Additionally, proton events can be a significant hazard to the safety of orbiting astronauts in addition to the strong threat of equipment damage and communications failures on orbiting communications satellites. It has happened in the past and will happen again in the future.

### A1-6.1 Solar flares

Solar flares are explosive releases of solar energy in the form of radiation at wavelengths from x-rays to visible light and mass ejecta ranging from alpha particles to very hot plasma. The radiation reaches the Earth in about 8.3 minutes, but high-energy alpha particles and protons can take from as short as 15 minutes to several hours to reach Earth depending on their energy. It takes about a day and a half for hot plasma to arrive. Although they can occur throughout the solar cycle, solar flares are most frequent during solar maximum. The solar flares, themselves, are relatively short-lived events, but the associated increase in ionizing x-ray and other high-energy radiation can lead to dramatic increases in the ionospheric electron content in the 65- to 90-km altitude range over the entire sunlit hemisphere. During a flare event, the electron density in the D-region may increase by as much as ten times, leading to a sudden and drastic increase of radio-wave absorption. Solar flares and the increased radiation fluxes associated with them also affect the E- and F-layers of the ionosphere, but to a lesser degree.

These solar flare events trigger many of the short duration ionospheric phenomena such as SIDs, and are closely related to other solar-related activity such as geomagnetic substorms, increased auroral activity, and ionospheric storms. Within the HF portion of the radio spectrum, the prediction and analysis of solar flares and specifically the various ionospheric disturbances produced by the flares present difficulties.

### A1-6.2 Ionospheric irregularities

There are also ionospheric irregularities which seem to be only indirectly related to the Sun. Such irregularities, including tilts, layer height variations, traveling ionospheric disturbances (TIDs), spread-F, and sporadic-E are difficult if not impossible to predict. There are also sudden frequency deviation (SFD) and short wave fades (SWF), both of which occur within minutes of the appearance of a solar flare. The effects, though, are only on the sunlit hemisphere. Less immediate but near-term phenomena associated with energetic solar protons are also encountered.

Irregularities in ionospheric surfaces scatter, or defocus, radio waves. When this occurs in the F-layer, the disturbance is called spread-F and in the E-layer, spread-E. The surface abnormalities occur as a result of random motion within the ionized layers and changes in ion density profiles. The ionospheric electron density normally varies rather smoothly with height. Spread-F occurs when there are irregularities of electron density embedded within the otherwise benign ionosphere. This causes a distortion of signals which are propagated through that area. It is said that the name spread-F comes from the appearance of the traces on a standard vertical ionogram. Instead of a sharply defined trace, the trace appears to be either spread in frequency, or in delay time. Multipath reflection is a probable cause of frequency spread, while delay spread usually seems to be caused by scattering of the radio wave due to irregular ionization in the signal path. Spread-F, in the middle latitudes, is primarily a sporadically occurring nighttime phenomenon whose production mechanism is not well understood. The resulting signal amplitude will depend on whether the received signals are multipath-reflected or scattered from the irregularities. It has been reported that delay spreads of up to 1 ms have been observed, but Doppler spreads are usually much less than 1 Hz.

Sudden ionospheric disturbances (SIDs) are another group of unpredictable phenomena which can affect HF communications. A SID can totally disrupt HF propagation and usually affects all radio links operating within, or even partially in, the daylight side of the earth. SIDs occur without warning and usually last from a few minutes to several hours, but can disrupt skywave communication for hours or days at a time.. Because the majority of SIDs are associated with solar flares, the frequency of their occurrence is related to the 11-year sunspot cycle. The primary cause of the propagation disruption (fadeout or blackout) is a sudden, abnormal increase in the ionization of the D-region. That region then absorbs all the lower range of HF frequencies and, depending on the intensity of the ionization, partially or completely absorbs the higher HF and sometimes the lower VHF frequencies. A SID can also enhance the field strength of LF and VLF waves.

TIDs are ionospheric disturbances whose signature is a change in the ionospheric density that appears to propagate from the magnetic polar regions to the equatorial regions. TIDs are sometimes associated with solar flares and magnetic disturbances, but at other times they seem independent of such events. In any case, the induced fluctuations in the ionization density and/or layer heights can have noticeable effects on the effective range and received signal strength of HF signals passing through the affected portion of the ionosphere.

### A1-6.3 Ionospheric storms

Ionospheric storms are the result of sudden large increases in solar activity, primarily solar flares and CME, which produce streams of charged particles, sometimes called "solar wind". When these particles arrive in the Earth's atmosphere, they tend to be deflected by the Earth's magnetic field towards the auroral zones. In addition to the particle stream, there is an increase in electron density in the D-region and an expansion and diffusion of the $F_2$-layer. This increased activity can produce large variations in critical frequencies, layer heights, and absorption and may last from several hours to several days. The intensity of the storms varies and the effects usually extend over the entire Earth. The storms typically follow solar-flare-initiated SIDs by anywhere from 15 minutes to 72 hours. During periods of very high sunspot activity, ionospheric storms can occur without initial SIDs. Charged particles from the storms have a scattering effect on the F-layer, temporarily neutralizing its reflective properties. During the storms, ionospheric propagation is characterized by low received signal strengths and flutter-fading, a form of fading that especially affects voice communications. In the first few hours of a severe ionospheric storm, the ionosphere is in a state of turbulence, layer-forming stratification is destroyed, and propagation is, consequently, erratic. In the later stages of severe storms, and throughout more moderate storms, the upper part of the ionosphere is expanded and diffused. As a result, the virtual heights of the layers are much greater and the MUFs much lower. Absorption of radio waves in the D-layer is also increased. These storms can be devastating in the HF band because of the limitation of ionospheric support in the normally-propagating higher frequencies causing nonabsorptive blackout of HF trunks in the affected area. The effects of ionospheric storms are most severe at solar maximum but seem to have more effects on communications at solar minimum.

Ionospheric storms are accompanied by magnetic storms and magnetospheric substorms with subsequent auroral effects. Magnetic storms affect primarily the equatorial and mid-latitude ionospheres, whereas magnetospheric substorms are confined to the high-latitude (auroral and polar) regions.

Magnetic storms often follow the eruption of solar flares within 20 to 40 hours. and are disturbances of the Earth’s magnetic field. The earth’s magnetic field is an important feature since it generally prevents a direct encounter between the ionosphere and energetic particles of solar origin, and especially solar wind streams. Magnetic storms may last from a few hours to several days and can cause changes in the movement of charged particles in the polar cap, strengthening of electric currents in the ionosphere, equator-ward expansion of the auroral region, and increased ionospheric density in all layers. While the event is occurring, the various components of the Earth’s geomagnetic field fluctuate over much wider limits than they normally do.

The geomagnetic field is composed of an internal main field, which very roughly resembles the field resulting from an embedded bar magnet, and an external field generated by currents which flow within the ionosphere and magnetosphere. The geomagnetic field and the solar wind interact like a blunt object in a supersonic flow field. The earth’s field is compressed on the Sun-ward side and distended on the opposite (dark) side, giving rise to a characteristic shape resembling a comet. Within the magnetosphere, solar wind particles are generally excluded, due to deflection by the severely distorted geomagnetic field. Although there is a general correlation between long-term averages in magnetic activity and solar activity, the correlation is not precise and is sometimes rather low.

### A1-7. Equatorial, mid-latitude, auroral, and polar regions

Skywave communication in equatorial, auroral, and polar regions is generally more of a problem than in mid-latitude regions. Ionization density irregularities that significantly impact HF radio performance are an important, and sometimes dominant, feature of the ionosphere in these regions. In particular, skywave signals received over these paths tend to show large delay and Doppler spreads that result in significant distortion of received signals. This distortion is due primarily to the size, motion, and spatial distribution of electron density irregularities. Even relatively weak irregularities within an otherwise benign ionosphere can cause small but noticeable signal distortion.

#### A1-7.1 Equatorial region

In the equatorial region, equatorial spread-F (ESF) is a significant recurring irregularity. ESF seems to be primarily caused by an uprising of low electron density globules, or depletion regions, moving upwards from the bottom to the top of the F-region ionosphere. Because of the presence of the geomagnetic field and the conductivity of the lower ionosphere, ESF does not occur until shortly after sunset at low (equatorial) latitudes, where the magnetic field is nearly parallel to the earth’s surface. ESF, also known as equatorial clutter, is particularly damaging to HF communications because the typical dimensions of the irregularities of the upwelling globules range from centimeters to hundreds of meters. It has been reported that the largest globules can measure as much as 1000 km from north to south and 100 km from east to west. The upwelling of these globules always follows the sunset and moves in an easterly direction at speeds of several hundred (300-700)km/hr. At HF, ESF produces multipath from both the globules and from normal ionospheric reflection, Doppler shift from the moving irregularities, and dispersion caused by refraction of the incident wave from the globules and the background ionization.

HF propagation at mid-latitudes, as has been shown, can be adversely affected by a number of factors, including the day/night terminator, SIDs, TIDs, sporadic-E, spread-F, solar flares, and magnetic storms.

The auroral ionosphere is the region of greatest particle precipitation from both the earth's outer radiation belt (the magnetosphere) and the interplanetary environment.

### A1-7.2 Auroral region

Auroral effects are classified as discrete or diffuse, based on their space and time characteristics. Both the auroral oval (asymmetrically distributed around the magnetic pole) and the diffuse aurora are regions in which particle precipitation (primarily electrons) plays a significant role in ionization production. The various phenomena associated with the auroral region are inherently irregular in both space and time. Based on observations, discrete auroral effects occur most frequently at night, while diffuse effects, although fairly widespread, are primarily a morning phenomenon. The discrete aurora is generally structured, dynamic, and the visual effects are quite bright (the Northern Lights, for example). The diffuse aurora, on the other hand, is much more diffused, slow to change, and the visual effects are only faintly seen. The diffuse aurora is present during both quiet and disturbed magnetic conditions, while the discrete aurora is primarily caused by disturbances. Both types of aurora intensify during magnetospheric substorms. During severe substorms, auroral absorption can be strong enough to cause extended periods of radio blackout.

As in other areas, sunlight fosters an intense ionospheric density in both the auroral and polar areas that tends to hide many possible irregularities. In the absence of nighttime UV ionization in the fall, winter, and early spring, the available ionization is almost wholly dependent on particle precipitation and ion movement from the sunlit ionosphere to the dark (nighttime) areas. During these periods, any HF radio wave passing through or terminating or originating in the auroral and/or polar regions is subject to many forms of density irregularities and frequently shows the effects from the great range of disturbances. (Note: When it is winter in the northern hemisphere, it is summer in the southern hemisphere and vice versa. Spring and fall are also reversed between the two hemispheres.) The effects of ionization and the electron precipitation are the visible aurora, increased electron density at D-, E-, and F-layer heights, heightened auroral radio-wave absorption, and enhanced geomagnetic field perturbations affecting the E-layer.

D-layer ionization, and the accompanying radio-wave absorption, is much like that of the other sunlit latitudes. During an intense solar flare, however, the D-layer ionization region expands sharply upward with the net effect of a sudden and severe increase in radio-wave absorption. This frequently results in an HF radio blackout that can last for the duration of the flare event. This increase in absorption also happens during significant electron precipitation events. During major magnetic storms, communications links that pass through the D-layer in the diffuse auroral absorption zone may be blacked out for the duration of the storm.

Auroral ionization in both the E- and F-layers is a combination of the ionization from direct solar radiation in the sunlit sector and collision-produced ionization from electrons streaming earthward along the geomagnetic field lines. Additionally, the nighttime F-layer is also populated by ionized "blobs" (as contrasted with depleted globules) which are large regions of enhanced ionization that enter the auroral F-region from the midnight sector of the polar cap. Collision-produced ionization and ionized blobs are the dominant factors in the winter nighttime ionosphere and give it its unique character. The amount of electron precipitation and the entry of energetic electrons into the earth's atmosphere is dependent upon a complex interaction between the solar wind magnetic effects and the Earth's magnetosphere. The intensity of this electron precipitation increases with magnetospheric substorm activity.

Collision-generated E-layer ionization is not dependent on solar radiation and is sporadic by nature with sporadic effects on radio wave propagation. During late autumn, winter, and early spring, collision-generated ionization is the dominant production mechanism. At these times, E-layer ionization is quite variable and unpredictable, and sporadic E-layer formation becomes a common phenomenon. Auroral sporadic-E ($E_S$) differs from mid-latitude $E_S$ in that the irregularities tend to be aligned with the magnetic field. During magnetic storms, the electron precipitation increases greatly. This results in a sharp increase in E-layer electron density and the development of very intense $E_S$ layers. These intense $E_S$ layers usually have a critical reflection frequency that far exceeds that of any higher layers. Once these $E_S$-layers are formed, they blanket the F-layers, preventing ground-launched HF radio waves from reaching and refracting from the F-layer regions. Any communication link into or out of the auroral ionosphere must then propagate via the E-layer, at least on the initial or final hop of the path. Generally speaking, the E-layer is so irregular and variable during the existence of strong $E_S$-layers that signals passing through these layers will be subject to significant delay and Doppler spread effects. During magnetically stable daytime conditions, radio signals have about the same characteristics as those in the mid-latitudes. As was reported in earlier reports, during disturbed conditions, multipath signals are observed that are comparable in amplitude to those of the quiet daytime channel, but with considerable more spread in delay time and Doppler frequency. Scatter-type signals, which are about 30-dB less in amplitude than the specular-multipath signals but with somewhat more spread can also be observed.

### A1-7.3 Polar region

Radio-wave propagation in the polar region and polar cap absorption can pose significant problems for any radio link originating, terminating, or passing through the polar area.

An important disturbance feature of the polar ionosphere is a solar proton event (SPE). During these events, solar protons are ejected from the sun and arrive in the vicinity of the earth in about 30 minutes. These protons are deflected toward the polar regions by the geomagnetic field where they enter the Earth's atmosphere. The sharply increased D-layer ionization associated with an SPE attenuates, at least to some degree, all radio waves passing through the D-region of the polar ionosphere. This is also known as polar cap absorption (PCA). Attenuation over skywave circuits in excess of 100 dB has been measured. Ionospheric absorption during a PCA event usually occurs in the altitude range of 45 to 80 km (includes the D-layer). The absorption decreases with increasing frequency and increases with increasing propagation path length through the absorbing region and may last for several days, especially when the polar cap is in daylight. SPEs are almost always preceded by a major solar flare and occur most frequently at solar sunspot maximum. Fortunately, SPEs are rare events, which are not typically encountered at solar minimum and are observed approximately once a month at solar maximum.

The dominant ionization irregularities in the polar winter ionosphere are polar arcs and patches. Since the absence of solar radiation (polar winter) results in a weak ambient ionosphere, the appearance of the arcs and patches is much more noticeable than during the rest of the year when the polar area is sunlit. Arcs and patches are areas of enhanced ionization density which generally move quite rapidly in the area. Arcs are narrow, elongated irregularities of enhanced ionization density (up to 1,000 km in length by about 100 km in width) which drift from the dawn to dusk meridians at speeds between 100-250 km/s, and seem to occur most frequently in "quiet" magnetic conditions during solar maximum. Arcs, though, are not considered significant contributors to the polar ionosphere during solar minimum. Patches are large, on the order of 1000 km in diameter, slower (about 1000 m/s), and are believed to originate in a sunlit portion of the subauroral ionosphere. HF propagation via these irregularities generally shows significant delay and Doppler spread values. During polar summer, these and other ionospheric irregularities are usually masked by the dense ionospheric ionization produced by the continuous exposure to solar radiation.

## A1-8. Nuclear effects on HF radio

HF radio using ionospheric propagation is more susceptible to the disrupting effects of nuclear explosions in the atmosphere than is any other radio propagation mechanism in any other frequency band. This is due primarily to catastrophic changes in the structure of the ionosphere. Such changes occur minutes after the explosion and last for from several minutes to several days. When explosions occur on the day side of the Earth, continuing ionization by the sun can restore ionospheric layers in as little as ten minutes. When the explosions occur on the night side, this major ionization source is blocked by the Earth and the effects last until daylight. Decaying effects of the explosion, though, may reappear each night for several nights. For nuclear explosions below about 100km in altitude, the predominant effect on radio waves is signal absorption. See Figure 3-5. Above about 100 km, the major effects are refraction and dispersion.

The primary effects on skywaves from the lower level bursts are: (1) greatly increased attenuation through absorption in the D-layer caused primarily by the X-rays and the free electrons from the burst; (2) loss of F-layer propagation paths, due to depletion of electrons in the layer; and (3) anomalous propagation modes caused by irregular ionization enhancements. Reference 4, Section 4, provides a more detailed and illustrated description of the effects of nuclear bursts on HF radio propagation.

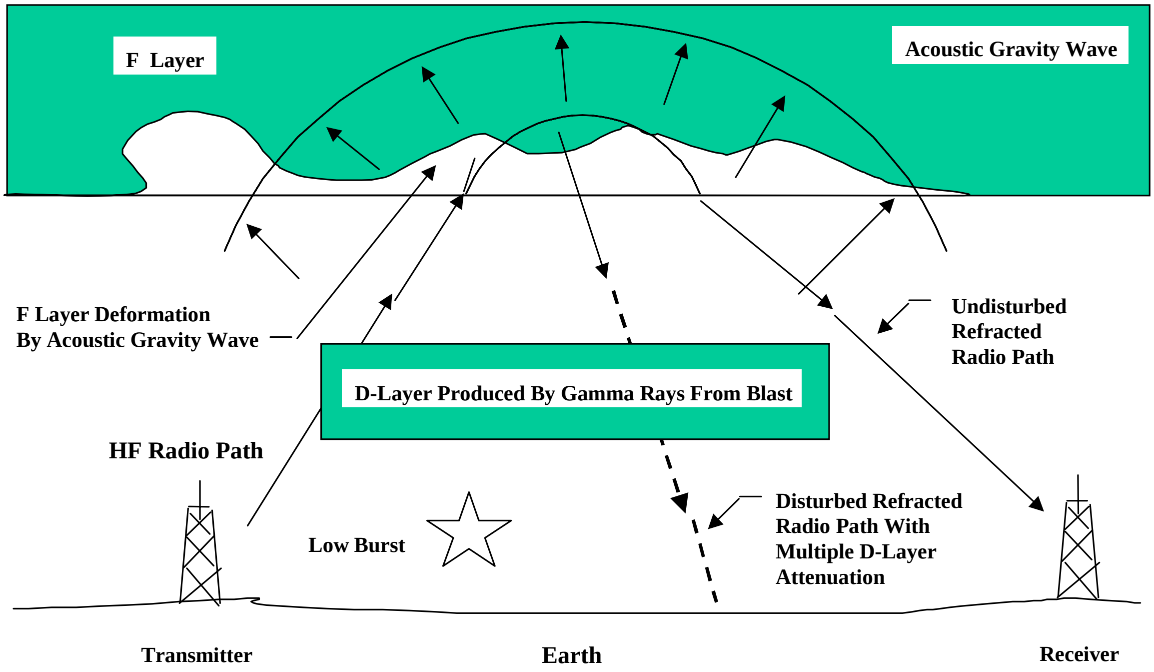


FIGURE A1-5
**Nuclear burst effects**

## A1-9. Multipath propagation

Most HF antennas radiate over a broad vertical angle. As a result, the radiated energy reaches one or more ionospheric layers at different angles of incidence, thus causing the refracted waves to follow different paths. Some paths may involve reflection from the ground and a subsequent return to the ionosphere one or more times before reaching the destination. Other paths may involve refraction from other than the design or preferred layer. The received signal will probably be made up of a number of components arriving via several routes, including one or more skywave paths and a ground-wave path. This is multipath. The arrival times of each of these components differ from each other primarily because of differences in path length.

The multihop paths, being physically longer, substantially delay the waves relative to a single-hop or direct path. This causes the received signal to be dispersed over time. The range, or spread, of multipath propagation times is a function of frequency, path length, geographical location, local time, season and sunspot activity. This range of time differences is called multipath spread. This is particularly a problem for digital transmissions at high data rates, since the multipath spread can cause intersymbol interference. Intersymbol interference is distortion of the received signal by the overlapping of individual pulses to the degree that the receiver cannot reliably distinguish between individual signal elements.

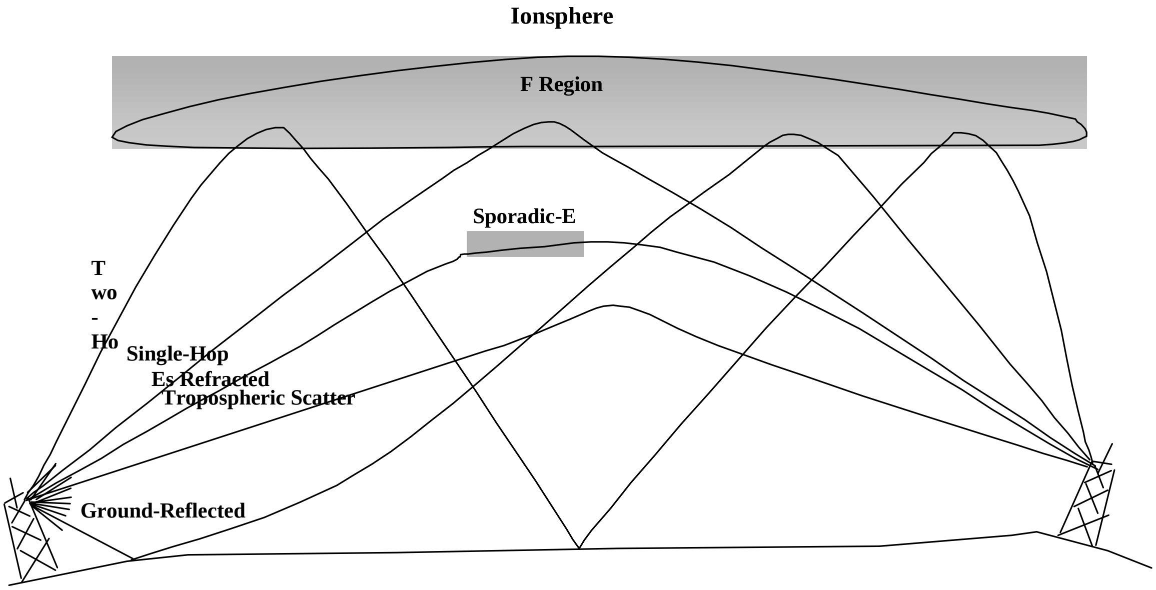


FIGURE A1-5
**Multipath propagation**

Selecting a frequency as close as possible to the MUF is probably the best way to minimize the effects of multipath spread. As the operating frequency is decreased from the MUF, a frequency is reached at which the multipath spread is a maximum. It is noted that for a 2500 km path, the maximum time dispersion has been shown to be about 3 ms; for 1000 km it increases to 5 ms; and for 200 km, it is about 8 ms. The time dispersion can be much greater when conditions are such that the ionosphere contains many irregularities, like those resulting in spread-F.

The route the radio signal travels must also be considered in optimizing the communications links. Each propagation path or mode has its own characteristic group delay. Multipath spread is basically the difference in group delays between the different modes. The received signal strengths of the different modes will normally show considerable differences, thus allowing the receiver-demodulator chain to be able to reject some of the weaker, distortion-causing signals.

## A1-10 Conclusion

In this annex, we have addressed the concepts associated with understanding the propagation media for HF radio communication. We have introduced the various types of radio-wave propagation (*i.e.*, skywave, reflected wave, direct wave, surface wave, ground wave, *etc*.). Since this annex dealt specifically with HF radio-wave propagation we have shown which of these is more meaningful for HF, how these waves propagate, how they may be refracted by the ionosphere, and what times of the day or night propagation by these waves will be best. This annex has introduced other HF propagation factors such as the effects on HF communication by noise (atmospheric and man-made), the concepts of fading, and why HF is considered limited bandwidth. The important concepts of maximum usable frequency (MUF) and lowest useable frequency (LUF) are introduced. The annex has also introduced other ionospheric propagation disturbances such as solar flares and storms. The annex shows how HF communication is effected at difference parts of the earth such as equatorial, mid-latitude, auroral and polar regions. The annex concluded with a introduction into how HF communication might react in a nuclear atmosphere as well as introduced the concepts of multipath propagation. The next annex, Annex 2, of this handbook will address prediction techniques used to estimate communication conditions along these unpredictable skywave signal paths. Annex 2 also addresses the need for predictions of HF communication performance and addresses the relationships, between short-term and long-term prediction.