of microwave antennas, diffraction plays a major part in preventing the narrow pencil of radiation which is often desired, by generating unwanted side lobes; this is discussed in Section 9-7.1.

spreading in all the direction.

① Ground(surface) wave
② Sky wave propagation - the Ionosphere
③ space wave.

8-2 PROPAGATION OF WAVES

In an earth environment, electromagnetic waves propagate in ways that depend not only on their own properties but also on those of the environment itself; some of this was seen in the preceding section. Since the various methods of propagation depend largely on frequency, the complete electromagnetic spectrum is now shown for reference in Figure 8-11—note that the frequency scale is logarithmic.

Waves travel in straight lines, except where the earth and its atmosphere alter their path. Except in unusual circumstances, frequencies above the HF generally travel in straight lines (except for refraction due to changing atmospheric density, as discussed in the previous section). They propagate by means of so-called space waves. These are sometimes called *tropospheric waves*, since they travel in the troposphere, the portion of the atmosphere closest to the ground. Frequencies below the HF range travel around the curvature of the earth, sometimes right around the globe. The means are probably a combination of diffraction and a type of *waveguide* effect which uses the earth's surface and the lowest *ionized* layer of the atmosphere as the two waveguide walls (see Section 10-1.3). These *ground waves*, or *surface waves* as they are called, are one of the two original means of beyond-the-horizon propagation. All broadcast radio signals received in daytime propagate by means of surface waves.

Waves in the HF range, and sometimes frequencies just above or below it, are reflected by the ionized layers of the atmosphere (to be described) and are called *sky waves*. Such signals are beamed into the sky and come down again after reflection,

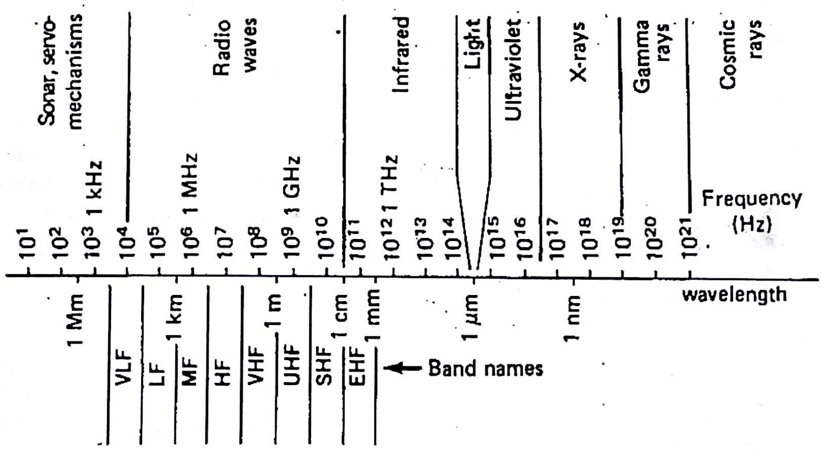

FIGURE 8-11 The electromagnetic spectrum.

returning to earth well beyond the horizon. To reach receivers on the opposite side of the earth, these waves must be reflected by the ground and the ionosphere several times. It should be mentioned that neither surface waves nor sky waves are possible in space or on airless bodies such as the moon.

Two more recently developed means of beyond-the-horizon propagation are tropospheric scatter and stationary satellite communications. Each of these five methods of propagation will now be described in turn.

8-2.1 Ground (Surface) Waves

Ground waves progress along the surface of the earth and, as previously mentioned, must be vertically polarized to prevent short circuiting the electric component. A wave induces currents in the ground over which it passes and thus loses some energy by absorption. This is made up by energy diffracted downward from the upper portions of the wavefront.

There is another way in which the surface wave is attenuated: because of diffraction, the wavefront gradually tilts over, as shown in Figure 8-12. As the wave propagates over the earth, it tilts over more and more, and the increasing tilt causes greater short circuiting of the electric field component of the wave and hence field strength reduction. Eventually, at some distance (in wavelengths) from the antenna, as partly determined by the type of surface over which the ground wave propagates, the wave "lies down and dies." It is important to realize this, since it shows that the maximum range of such a transmitter depends on its frequency as well as its power. Thus, in the VLF band, insufficient range of transmission can be cured by increasing the transmitting power. This remedy will not work near the top of the MF range, since propagation is now definitely limited by tilt.

Field strength at a distance Radiation from an antenna by means of the ground wave gives rise to a field strength at a distance, which may be calculated by use of Maxwell's equations. This field strength, in volts per meter, is given in Equation (8-10), which differs from Equation (8-5) by taking into account the gain of the transmitting antenna.

$$\mathscr{E} = \frac{120\pi h_t I}{\lambda d} \tag{8-10}$$

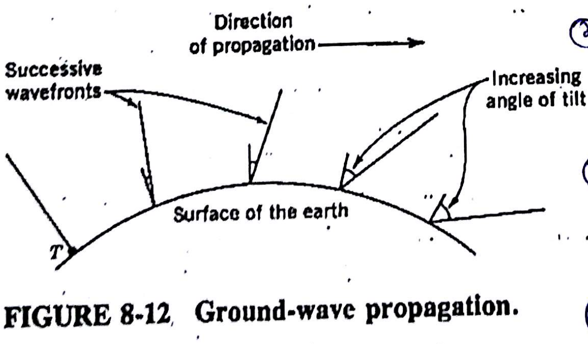

FIGURE 8-12 Ground-wave propagation.

(1) Parallel to Ground.
(2) Induced wave in ground a attenuate E/m wave at a Short range.
(3) used for Low freq (AM wave).
[High freq. attenuate fast]
(4) Few KHZ to few MHZ (5 MHZ).
(5) Large wavelength, (obstruction).
(6) Bend around the corners more efficiently

Antenna size $\frac{\lambda}{4}$

If a receiving antenna is now placed at this point, the signal it will receive will be, in volts,

$$V = \frac{120\pi h_t\, h_r\, I}{\lambda d} \tag{8-11}$$

where 120π = characteristic impedance of free space
h_t = effective height (this is not quite the same as the actual height, for reasons dealt with in Section 9-4) of the transmitting antenna
h_r = effective height of the receiving antenna
I = antenna current
d = distance from the transmitting antenna
λ = wavelength

If the distance between the two antennas is fairly long, the reduction of field strength due to ground and atmospheric absorption reduces the value of the voltage received, making it less than shown by Equation (8-11). Although it is possible to calculate the signal strength reduction which results, altogether too many variables are involved to make this worthwhile. Such variables include the salinity and resistivity of the ground or water over which the wave propagates and the water vapor content of the air. The normal procedure is to estimate signal strength with the aid of the tables and graphs available.

VLF propagation When propagation is over a good conductor like seawater, particularly at frequencies below about 100 kHz, surface absorption is small, and so is attenuation due to the atmosphere. Thus the angle of tilt is the main determining factor in the long-distance propagation of such waves. The degree of tilt depends on the distance from the antenna in wavelengths, and hence the early disappearance of the surface wave in HF propagation. Conversely, because of the large wavelengths of VLF signals, waves in this range are able to travel long distances before disappearing (right around the globe if sufficient power is transmitted).

At distances up to 1000 km, the ground wave is remarkably steady, showing little diurnal, seasonal or annual variation. Farther out, the effects of the *E* layer's contribution to propagation are felt. (See also the next section, bearing in mind that the ground and the bottom of the *E* layer are said to form a waveguide through which VLF waves propagate.) Both short- and long-term signal strength variations take place, the latter including the 11-year solar cycle. The strength of low-frequency signals changes only very gradually, so that rapid fading does not occur. Transmission at these wavelengths proves a very reliable means of communication over long distances.

The most frequent users of long-distance VLF transmissions are ship communications, and time and frequency transmissions. Ships use the frequencies allocated to them, from 10 to 110 kHz, for radio navigation and maritime mobile communications. The time and frequency transmissions operate at frequencies as low as 16 kHz (GBR, Rugby, United Kingdom) and 17.8 kHz (NAA, Cutler, Maine). They provide a worldwide continuous hourly transmission of stable radio frequencies, standard time intervals, time announcements, standard musical pitch, standard audio frequencies and radio propagation notices. Such services are also duplicated at HF, incidentally, by stations such as WWV (Ft. Collins, Colorado) and WWVH (Hawaii) operating at 2.5 MHz and the first five harmonics of 5 MHz.

Since VLF antennas are certain to be inefficient, high powers and the tallest possible masts are used. Thus we find powers in excess of 1 MW transmitted as a rule, rather than an exception. For example, the U.S. Naval Communications Station at North-West Cape (Western Australia) has an antenna farm consisting of 13 very tall masts, the tallest 387 m high; the lowest transmitting frequency is 15 kHz.

8-2.2 Sky-Wave Propagation—The Ionosphere

Even before Sir Edward Appleton's pioneering work in 1925, it had been suspected that ionization of the upper parts of the earth's atmosphere played a part in the propagation of radio waves, particularly at high frequencies. Experimental work by Appleton showed that the atmosphere receives sufficient energy from the sun for its molecules to split into positive and negative ions. They remain thus ionized for long periods of time. He also showed that there were several layers of ionization at differing heights, which (under certain conditions) reflected back to earth the high-frequency waves that would otherwise have escaped into space. The various layers, or strata, of the ionosphere have specific effects on the propagation of radio waves, and must now be studied in detail.

a ~~ter~~ region of the Earth's atmosphere; contains a high concentration of free electrons.

The ionosphere and its effects The ionosphere is the upper portion of the atmosphere, which absorbs large quantities of radiant energy from the sun, becoming heated and ionized. There are variations in the physical properties of the atmosphere, such as temperature, density and composition. Because of this and the different types of radiation received, the ionosphere tends to be stratified, rather than regular, in its distribution. The most important ionizing agents are ultraviolet and α, β, and γ radiation from the sun, as well as cosmic rays and meteors. The overall result, as shown in Figure 8-13, is a range of four main layers, D, E, F_1 and F_2, in ascending order. The last two combine at night to form one single layer.

The D layer is the lowest, existing at an average height of 70 km, with an average thickness of 10 km. The degree of its ionization depends on the altitude of the sun above the horizon, and thus it disappears at night. It is the least important layer from the point of view of HF propagation. It reflects some VLF and LF waves and absorbs MF and HF waves to a certain extent.

The E layer is next in height, existing at about 100 km, with a thickness of perhaps 25 km. Like the D layer, it all but disappears at night; the reason for these disappearances is the recombination of the ions into molecules. This is due to the absence of the sun (at night), when radiation is consequently no longer received. The main effects of the *E layer* are to aid MF surface-wave propagation a little and to reflect some HF waves in daytime.

The E_s layer is a thin layer of very high ionization density, sometimes making an appearance with the *E layer*. It is also called the *sporadic E layer;* when it does occur, it often persists during the night also. On the whole, it does not have an impor-

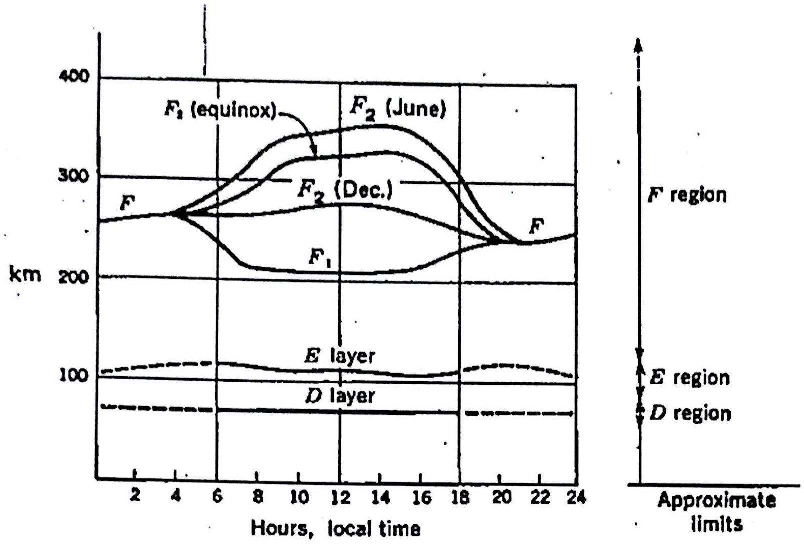

FIGURE 8-13 Ionospheric layers and their regular variations. *(F. R. East, "The Properties of the Ionosphere Which Affect HF Transmission")*

tant part in long-distance propagation, but it sometimes permits unexpectedly good reception. Its causes are not well understood.

The F_1 layer, as shown in Figure 8-13, exists at a height of 180 km in daytime and combines with the F_2 layer at night, its daytime thickness is about 20 km. Although some HF waves are reflected from it, most pass through to be reflected from the F_2 layer. Thus the main effect of the F_1 layer is to provide more absorption for HF waves. Note that the absorption effect of this and any other layer is doubled, because HF waves are absorbed on the way up and also on the way down.

The F_2 layer is by far the most important reflecting medium for high-frequency radio waves. Its approximate thickness can be up to 200 km, and its height ranges from 250 to 400 km in daytime. At night it falls to a height of about 300 km, where it combines with the F_1 layer. Its height and ionization density vary tremendously, as Figure 8-13 shows. They depend on the time of day, the average ambient temperature and the sunspot cycle (see also the following sections dealing with the normal and abnormal ionospheric variations). It is most noticeable that the F layer persists at night, unlike the others. This arises from a combination of reasons; the first is that since this is the topmost layer, it is also the most highly ionized, and hence there is some chance for the ionization to remain at night, to some extent at least. The other main reason is that although ionization density is high in this layer, the *actual air density* is not, and thus most of the molecules in it are ionized. Furthermore, this low actual density gives the molecules a large *mean free path* (the statistical average distance a molecule travels before colliding with another molecule). This low molecular collision rate in turn means that, in this layer, ionization does not disappear as soon as the sun sets. Finally, it must be mentioned that the reasons for better HF reception at night are the combination of the F_1 and F_2 layers into one F layer, and the virtual disappearance of the other two layers, which were causing noticeable absorption during the day.

Reflection mechanism Electromagnetic waves returned to earth by one of the layers of the ionosphere appear to have been reflected. In actual fact the mechanism involved is refraction, and the situation is identical to that described in Figure 8-6. As the ionization density increases for a wave approaching the given layer at an angle, so the refractive index of the layer is reduced. (Alternatively, this may be interpreted as an increase in the conductivity of the layer, and therefore a reduction in its electrical density or dielectric constant.) Hence the incident wave is gradually bent farther and farther away from the normal, as in Figure 8-6.

If the rate of change of refractive index per unit height (measured in wavelengths) is sufficient, the refracted ray will eventually become parallel to the layer. It will then be bent downward, finally emerging from the ionized layer at an angle equal to the angle of incidence. Some absorption has taken place, but the wave has been returned by the ionosphere (well over the horizon if an appropriate angle of incidence was used).

Terms and definitions The terminology that has grown up around the ionosphere and sky-wave propagation includes several names and expressions whose meanings are not obvious. The most important of these terms will now be explained.

The virtual height of an ionospheric layer is best understood with the aid of Figure 8-14. This figure shows that as the wave is refracted, it is bent down gradually rather than sharply. However, below the ionized layer, the incident and refracted rays follow paths that are exactly the same as they would have been if *reflection* had taken place from a surface located at a greater height, called the *virtual height* of this layer. If the virtual height of a layer is known, it is then quite simple to calculate the angle of incidence required for the wave to return to ground at a selected spot.

The critical frequency (f_c) for a given layer is the highest frequency that will be returned down to earth by that layer after having been beamed straight up at it. It is important to realize that there is such a maximum, and it is also necessary to know its value under a given set of conditions, since this value changes with these conditions. It was mentioned earlier that a wave will be bent downward provided that the rate of change of ionization density is sufficient, and that this rate of ionization is measured

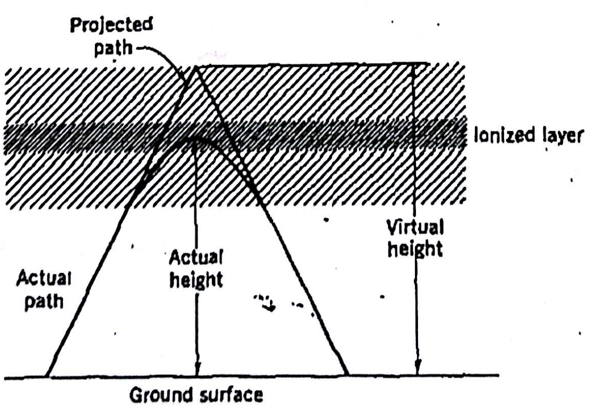

FIGURE 8-14 Actual and virtual heights of an ionized layer.

per unit wavelength. It also follows that the closer to being vertical the incident ray, the more it must be bent to be returned to earth by a layer. The result of these two effects is twofold. First, the higher the frequency, the shorter the wavelength, and the less likely it is that the change in ionization density will be sufficient for refraction. Second, the closer to vertical a given incident ray, the less likely it is to be returned to ground. Either way, this means that a maximum frequency must exist, above which rays go through the ionosphere. When the angle of incidence is normal, the name given to this maximum frequency is *critical frequency;* its value in practice ranges from 5 to 12 MHz for the F_2 layer.

The maximum usable frequency, or *MUF,* is also a limiting frequency, but this time for some specific angle of incidence other than the normal. In fact, if the angle of incidence (between the incident ray and the normal) is θ, it follows that

$$MUF = \frac{\textit{critical frequency}}{\cos\theta}$$

$$= f_c \sec\theta \qquad (8\text{-}12)$$

This is the so-called *secant law,* and it is very useful in making preliminary calculations for a specific MUF. Strictly speaking, it applies only to a flat earth and a flat reflecting layer. However, the angle of incidence is not of prime importance, since it is determined by the distance between the points that are to be joined by a sky-wave link. Thus MUF is defined in terms of two such points, rather than in terms of the angle of incidence at the ionosphere, it is defined at the highest frequency that can be used for sky-wave communication between two given points on earth. It follows that there is a different value of MUF for each pair of points on the globe. Normal values of MUF may range from 8 to 35 MHz, but after unusual solar activity they may rise to as high as 50 MHz. The highest working frequency between a given pair of points is naturally made less than the MUF, but it is not very much less for reasons that will be seen.

The skip distance is the shortest distance from a transmitter, measured along the surface of the earth, at which a sky wave of fixed frequency (more than f_c) will be returned to earth. That there should be a minimum distance may come as a shock. One expects there to be a maximum distance, as limited by the curvature of the earth, but nevertheless a definite minimum also exists for any fixed transmitting frequency. The reason for this becomes apparent if the behavior of a sky wave is considered with the aid of a sketch, such as Figure 8-15.

When the angle of incidence is made quite large, as for ray 1 of Figure 8-15, the sky wave returns to ground at a long distance from the transmitter. As this angle is slowly reduced, naturally the wave returns closer and closer to the transmitter, as shown by rays 2 and 3. If the angle of incidence is now made significantly less than that of ray 3, the ray will be too close to the normal to be returned to earth. It may be bent noticeably, as for ray 4, or only slightly, as for ray 5. In either case the bending will be insufficient to return the wave, unless the frequency being used for communication is less than the critical frequency (which is most unlikely); in that case everything is returned to earth. Finally, if the angle of incidence is only just smaller than that of ray 3, the wave may be returned, but at a distance farther than the return point of ray 3; a ray such as this is ray 6 of Figure 8-15. This upper ray is bent back very gradually,

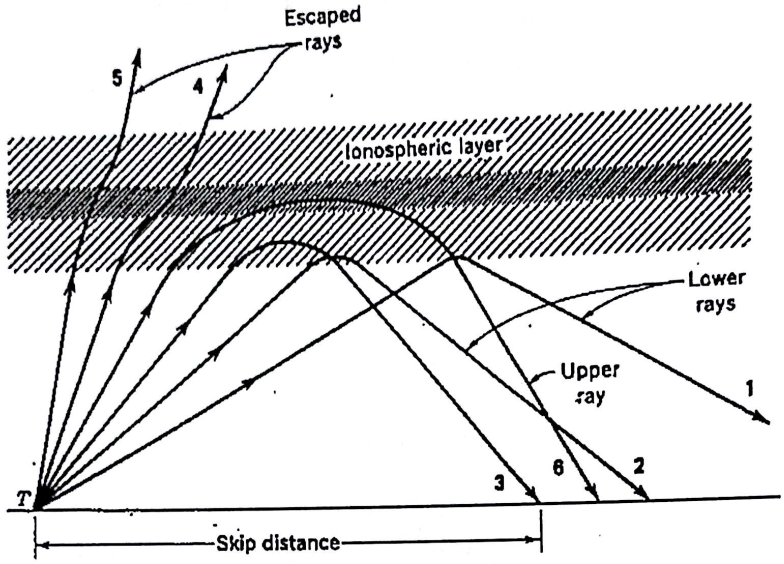

FIGURE 8-15 Effects of ionosphere on rays of varying incidence.

because ion density is changing very slowly at this angle. It thus returns to earth at a considerable distance from the transmitter and is weakened by its passage.

Ray 3 is incident at an angle which results in its being returned as close to the transmitter as a wave of this frequency can be. Accordingly, the distance is the *skip distance*. It thus follows that any higher frequency beamed up at the angle of ray 3 will not be returned to ground. It is seen that the frequency which makes a given distance correspond to the skip distance is the MUF for that pair of points.

At the skip distance, only the normal, or lower, ray can reach the destination, whereas at greater distances the upper ray can be received as well, causing interference. This is a reason why frequencies not much below the MUF are used for transmission. Another reason is the lack of directionality of high-frequency antennas, which is discussed in Section 9-6. If the frequency used is low enough, it is possible to receive lower rays by two different paths after either one or two hops, as shown in Figure 8-16, the result of this is interference once again.

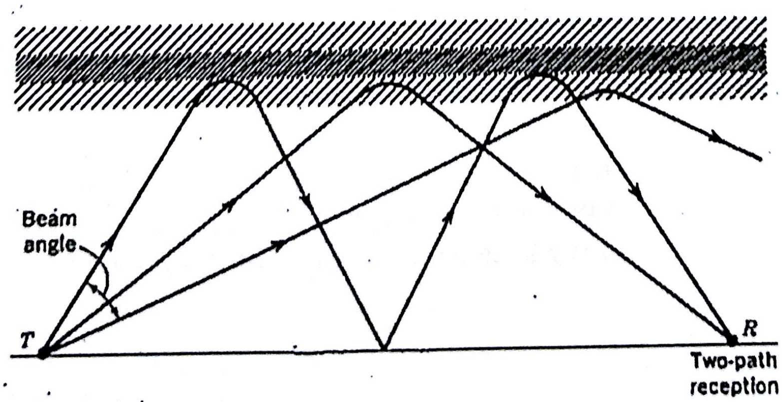

FIGURE 8-16 Multipath sky-wave propagation.

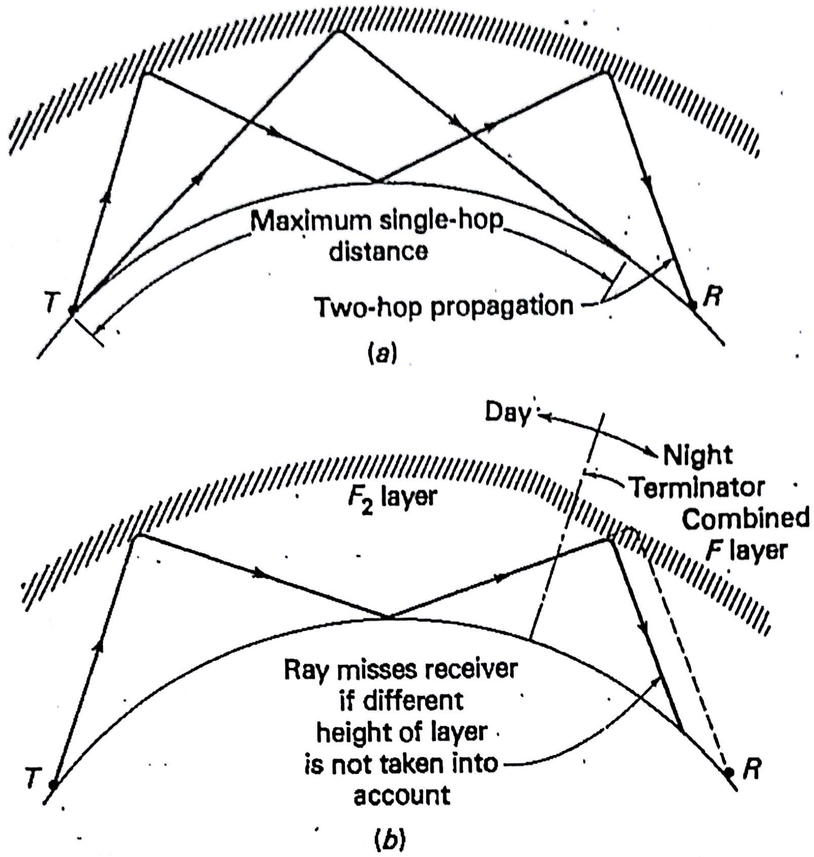

FIGURE 8-17 Long-distance sky-wave transmission paths. *(a)* North-south; *(b)* east-west.

The *transmission path* is limited by the skip distance at one end and the curvature of the earth at the other. The longest single-hop distance is obtained when the ray is transmitted tangentially to the surface of the earth, as shown in Figure 8-17. For the F_2 layer, this corresponds to a maximum practical distance of about 4000 km. Since the semicircumference of the earth is just over 20,000 km, multiple-hop paths are often required, and Figure 8-17 shows such a situation. No unusual problems arise with multihop north-south paths. However, care must be taken when planning long east-west paths to realize that although it is day "here," it is night "there," if "there" happens to be on the other side of the terminator. The result of not taking this into account is shown in Figure 8-17*b*. A path calculated on the basis of a constant height of the F_2 layer will, if it crosses the terminator, undershoot and miss the receiving area as shown—the F layer over the target is lower than the F_2 layer over the transmitter.

Fading is the fluctuation in signal strength at a receiver and may be rapid or slow, general or frequency-selective. In each case it is due to interference between two waves which left the same source but arrived at the destination by different paths. Because the signal received at any instant is the vector sum of all the waves received, alternate cancellation and reinforcement will result if there is a length variation as large as a half-wavelength between any two paths. It follows that such fluctuation is more likely with smaller wavelengths, i.e., at higher frequencies.

Fading can occur because of interference between the lower and the upper rays of a sky wave; between sky waves arriving by a different number of hops or different paths; or even between a ground wave and a sky wave especially at the lower end of the HF band. It may also occur if a single sky wave is being received, because of fluctuations of height or density in the layer reflecting the wave. One of the more successful means of combating fading is to use space or frequency diversity (see Section 6-3.2).

Because fading is frequency-selective, it is quite possible for adjacent portions of a signal to fade independently, although their frequency separation is only a few dozen hertz. This is most likely to occur at the highest frequencies for which sky waves are used. It can play havoc with the reception of AM signals, which are seriously distorted by such frequency-selective fading. On the other hand, SSB signals suffer less from this fading and may remain quite intelligible under these conditions. This is because the relative amplitude of only a portion of the received signal is changing constantly. The effect of fading on radiotelegraphy is to introduce errors, and diversity is used here wherever possible.

Ionospheric variations The ionosphere is highly dependent upon the sun, and hence its conditions vary continuously. There are two kinds of variations. The *normal* ones have already been described as diurnal and seasonal height and thickness changes. *Abnormal* variations are due mainly to the fact that our sun is a *variable star*.

The sun has an 11-year cycle over which its output varies tremendously. Most people are unaware of this, because light variations are slight. However, the solar output of ultraviolet rays, coronae, flares, particle radiation and sunspots may vary as much as fiftyfold over that period. The extent of solar disturbance is measured by a method of sunspot counting developed by Wolf in the eighteenth century. On this basis, a pronounced 11-year (±1 year) cycle emerges, and perhaps also a 90-year supercycle. The highest activities so far recorded were in 1778, 1871 and 1957, which was the highest ever.

The main sun-caused disturbances in the ionosphere are *SIDs* (*sudden ionospheric disturbances,* formerly known as *Dellinger dropouts*) and *ionospheric storms*. Sudden ionospheric disturbances are caused by solar flares, which are gigantic emissions of hydrogen from the sun. Such flares are sudden and unpredictable, but more likely during peak solar activity than when the sun is "quiet." The x-ray radiation accompanying solar flares tremendously increases ionization density, right down to the *D* layer. This layer now absorbs signals that would normally go through it and be reflected from the *F* layer. Consequently, long-distance communications disappear completely, for periods up to 1 hour at a time. Studies with earth-based radioheliographs and from satellites in orbit have provided a large amount of data on solar flares, so that some short-term predictions are becoming possible. Two other points should be noted in connection with SIDs. First, only the sunlit side of the earth is affected, and second, VLF propagation is actually improved.

Ionospheric storms are caused by particle emissions from the sun, generally α and β rays. Since these take about 36 hours to reach the earth, some warning is possible after large sunspots or solar flares are noticed. The ionosphere behaves erratically during a storm, right around the globe this time, but more so in high latitudes because of the earth's magnetic field. Signal strengths drop and fluctuate quite rapidly.

However, using lower frequencies often helps, since the highest ones are the most affected.

Finally, the sporadic E layer previously mentioned is also often included as an abnormal ionospheric disturbance. When present, it has the twin effects of preventing long-distance HF communications and permitting over-the-horizon VHF communications. The actual and virtual heights of this layer appear to be the same. This confirms the belief that the layer is thin and dense, so that actual reflection takes place.

Various national scientific bodies have ionospheric prediction programs which issue propagation notices of great value. Among them are the notices of the Central Radio Propagation Laboratory of the United States and the Ionospheric Prediction Service of the Australian Department of Science and Technology.

8-2.3 Space Waves >40MHZ

Space waves generally behave with merciful simplicity. They travel in (more or less) straight lines! However, since they depend on line-of-sight conditions, space waves are limited in their propagation by the curvature of the earth, except in very unusual circumstances. Thus they propagate very much like electromagnetic waves in free space, as discussed in Section 8-1.1. Such a mode of behavior is forced on them because their wavelengths are too short for reflection from the ionosphere, and because the ground wave disappears very close to the transmitter, owing to tilt.

Radio horizon The radio horizon for space waves is about four-thirds as far as the optical horizon. This beneficial effect is caused by the varying density of the atmosphere, and because of diffraction around the curvature of the earth. The radio horizon of an antenna is given, with good approximation, by the empirical formula

$$d_t = 4\sqrt{h_t} \tag{8-13}$$

where d_t = distance from transmitting antenna, km

h_t = height of transmitting antenna above ground, m

The same formula naturally applies to the receiving antenna. Thus the total distance will be given by addition, as shown in Figure 8-18, and by the empirical formula

$$d = d_t + d_r = 4\sqrt{h_t} + 4\sqrt{h_r} \tag{8-14}$$

A simple calculation shows that for a transmitting antenna height of 225 m above ground level, the radio horizon is 60 km. If the receiving antenna is 16 m above

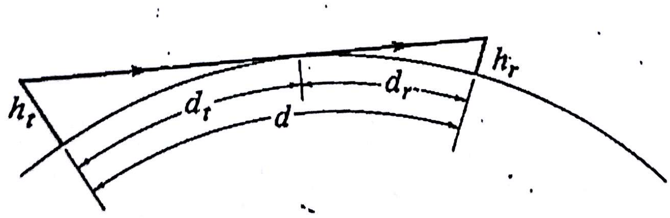

FIGURE 8-18 Radio horizon for space waves.

ground level, the total distance is increased to 76 km. Greater distance between antennas may be obtained by locating them on tops of mountains, but links longer than 100 km are hardly ever used in commercial communications.

General considerations As discussed in detail in Section 8-1.2, any tall or massive objects will obstruct space waves, since they travel close to the ground. Consequently, shadow zones and diffraction will result. This is the reason for the need in some areas for antennas higher than would be indicated by Equation (8-14). On the other hand, some areas receive such signals by reflection—any object large enough to cast a radio shadow will, if it is a good conductor, cause back reflections also. Thus, in areas in front of it a form of interference known as "ghosting" may be observed on the screen of a television receiver. It is caused by the difference in path length (and therefore in phase) between the direct and the reflected rays. This situation is worse near a transmitter than at a distance, because reflected rays are stronger nearby. Finally, particularly severe interference exists at a distance far enough from the transmitter for the direct and the ground-reflected rays to be received simultaneously.

Microwave space-wave propagation All the effects so far described hold true for microwave frequencies, but some are increased, and new ones are added. Atmospheric absorption and the effects of precipitation must be taken into account. So must the fact that at such short wavelengths everything tends to happen very rapidly. Refraction, interference and absorption tend to be accentuated. One new phenomenon which occurs is *superrefraction,* also known as *ducting.*

As previously discussed, air density decreases and refractive index increases with increasing height above ground. The change in refractive index is normally linear and gradual, but under certain atmospheric conditions a layer of warm air may be trapped above cooler air, often over the surface of water. The result is that the refractive index will decrease far more rapidly with height than is usual. This happens near the ground, often within 30 m of it. The rapid reduction in refractive index (and therefore dielectric constant) will do to microwaves what the slower reduction of these quantities, in an ionized layer, does to HF waves; complete bending down takes place, as illustrated in Figure 8-19. Microwaves are thus continuously refracted in the duct and reflected by the ground, so that they are propagated around the curvature of the earth for distances which sometimes exceed 1000 km. The main requirement for the formation of atmospheric ducts is the so-called temperature inversion. This is an increase of air temp with height, instead of the usual decrease in temp of 6.5°C/km in the 'standard atmosphere'. Superrefraction is on the whole more likely in subtropical than in temp. zone (zone).

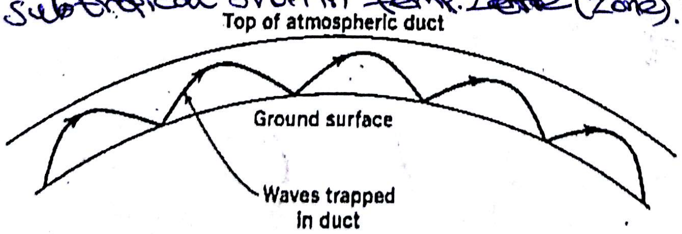

FIGURE 8-19 Superrefraction in atmospheric duct.