T43-2-1

RADIO PROPAGATION AT FREQUENCIES IN THE STANDARD BROADCAST BAND

by Philip L. Sullivan

There are four basic ways by which a radio wave can propagate, or travel, from transmitter to receiver. These are illustrated in Figure 1.

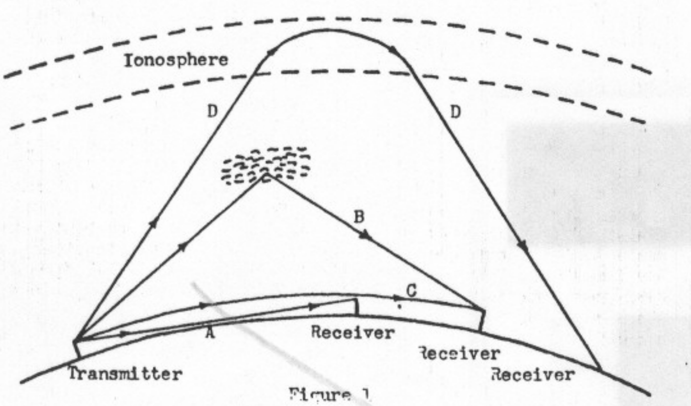

Figure 1

Path A is direct, line-of-sight, propagation called "space wave." Path B involves reflection of the signal from an irregularity in the troposphere; this is called "tropo-scatter." Path C, known as "ground wave" propagation, is a wave guided along and by the earth's surface. Path D involves a signal traveling to and then returning from the ionosphere; this is referred to as "sky wave" propagation. The first two paths are not important in the reception of distant broadcast band stations (the space wave is useless beyond the horizon, and tropo-scatter is important only at high frequencies). Therefore, we will only consider the characteristics of the last two propagation modes.

Ground Wave Propagation

The wave that propagates along the surface of the earth does so as a guided wave, guided in a way analogous to a wave guided through a waveguide or along a transmission line. The waveguide encloses the wave on all sides by good, metallic conductors, whereas in the case under consideration there is a conductor (the earth) on only one side. The earth has a much lower conductivity; nevertheless, many of the same ideas apply.

Two factors contribute to the loss in signal strength during ground wave propagation. The first is the distance factor due to dispersion of the signal as it radiates out from the antenna. This $1/d^2$ factor enters into the signal strength formulas no matter what the means of propagation. The second factor accounts for the amount of signal power lost by absorption by the earth because of the finite and variable conductivity of the earth. At large enough distances, combining both factors, we get the following expression for the received energy of the ground wave:

$$E_{\text{ground wave}} = \frac{K\sigma}{f^2d^2}$$

where K is a constant of proportionality involving transmitter power, antenna gain, etc., f is the frequency in Mhz., σ is the ground conductivity in mA/m units, and d is the distance in miles. This formula is an approximation of more general (and more complicated) formulas, but it is adequate for the values of σ, f and d that normally occur in the study of broadcast band propagation.

From the above equation for received signal strength, it is apparent that better ground wave reception will be achieved over ground with a higher conductivity, or looking at it the other way round, greatest weakening of the ground wave is over terrain having lowest conductivity. Typical values of the ground conductivity in North America range from 1 to 5 mA/m in dry, desert regions and mountainous areas; from 5 to 10 in the Great Lakes region; and from 10 to 20 in the Great Plains and Prairies, with some areas in the Dakotas and just north having values up to 30. For seawater, the ground conductivity is about 5000 units, meaning that ground wave reception over seawater is much better than for a corresponding distance over land. Over fresh water, σ is not as high but is still considerably above that of average soil and results in very good daytime reception over the Great Lakes.

In summary, ground wave propagation is the primary means by which we receive broadcast band stations at distances up to a few hundred (and occationally 1000) miles in the daytime.

Sky Wave Propagation

The important mode of propagation known as "sky wave" is due to the existance of a region in the upper atmosphere called the ionosphere. Although sky wave is the principal (and often only) means of propagation from a distant transmitter, it suffers from many variations and irregularities. To understand the mechanism of sky wave propagation and of the disturbances that affect it, we must first examine the ionosphere.

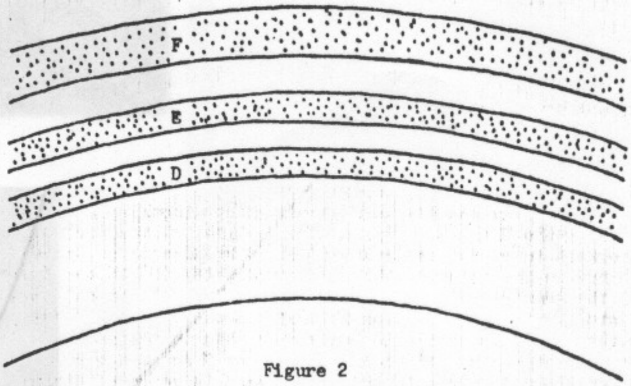

Figure 2

The ionosphere consists of several layers of ionic gases (fig. 2), the ionization being mainly produced by ultraviolet radiation from the sun. The lowest layer, at a height of 30 to 50 miles, is the D layer. Above this is the E layer from 60 to 80 miles. The D and E layers exist only in the daytime (except for a weak residual E layer at night). Above the E layer is a region known as the F layer. At night, this consists of a single layer between 150 and 200 miles up. At broadcast band frequencies, the principal effect of the D and E layers is absorption of the signal. Since these are the lowest layers, they prevent the signal from reaching the higher layers where refraction would otherwise take place. This is why propagation of broadcast band signals is virtually non-existent in the daytime. The layer that is most important for our consideration is the nighttime F layer.

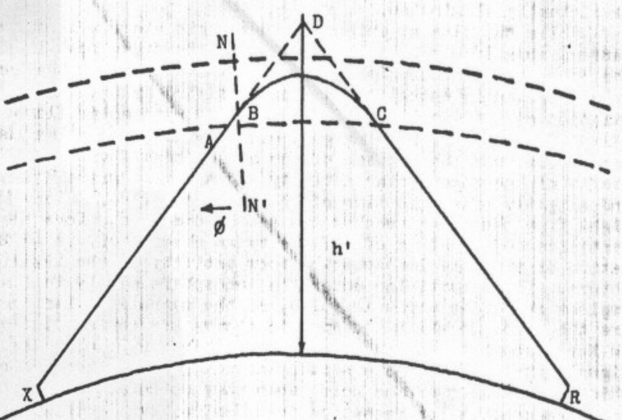

Figure 3

Many discussions of sky wave propagation consider the ionosphere to reflect the radio waves. Strictly speaking, the signals are refracted, not reflected, as shown in figure 3. A signal from the transmitter (X) enters the ionosphere at a point A at an angle Ø from the normal (N-N'). The effect of the ionization is to reduce the index of refraction (n) within the F layer to a value less than that of the atmosphere below it; ($n_F = 1 - 81N/f^2$, where N is the number of electrons per cm.^{2} and f is the frequency in khz., $n_{air} = 1$). Using Snell's law from optics, we find that when a wave enters a region with a lower n it is bent away from the normal. Since n continues to decrease with increasing height the actual path of the signal is the curve XABCR in figure 3. For most purposes, though, the problem can be treated by pretending the signal is sharply reflected from point D. The height, h', of D is called the "virtual height" of the ionosphere. For the nighttime F layer, h' is essentially constant at 190 miles.

Besides refraction, the other way in which the ionized gases of the ionosphere affect a radio signal is by absorption. As previously mentioned, this is the principal effect of the D and especially the E layers (in fact knowledge of when the sun's rays will strike the E layer over a certain location enables one to predict the fade out time of a sky wave signal) but it does occur to some extent in the F layer. Using a simplified model we find that the attenuation per unit length is proportional to $N\varepsilon/(\varepsilon^2 + 4\ {}^2f^2)$ where N is again the electron density and ε is the collision frequency of the ions. Thus a higher frequency will have a lower attenuation. Unfortunately for the broadcast band however, the earth has a magnetic field. Solving the equations of motion of an electron in a magnetic field we find that this motion has a resonant frequency f_o given by $f_o = B_o e/2\pi m$, where B_o is the earth's magnetic field and <u>e</u> and <u>m</u> are the charge and mass of the electron. Evaluating this expression gives f_o, often called the "gyrofrequency," a value between 1400 and 1500 khz., right in the top of the broadcast band. A signal at or near f_o will have a larger portion of its energy absorbed by this resonant system and thus an attenuation higher than would otherwise result occurs in the higher part of the broadcast band. This attenuation is in addition to that normally expected when the magnetic field is not considered.

Variations and irregularities in sky wave propagation are linked to changes in the ionosphere. These can be grouped into four classes: (1) daily, (2) seasonal, (3) those following the sunspot cycle, and (4) irregular.

The daily changes have already been discussed as the appearance of the D and E layers in the daytime. Throughout the year the most important seasonal effect on broadcast propagation is the change in the value of the electron density (N) in the F layer. In the summer, the value of N in the F layer is higher than in the winter - since attenuation is proportional to N (see above), sky waves propagate better in the winter than in the summer. Years of sunspot maxima also result in higher N and hence sky wave propagation is worse in such years than in years of minima. Sunspots come in 11 year cycles, the last maximum being 1968 and the last minimum 1974-75. In addition to deterioration of propagation conditions during the summer and in years of sunspot maxima, the level of the atmospheric noise is generally higher at these times too and this further degrades reception.

Irregular changes in the ionosphere are by their very nature unpredictable. They do, however, have considerable effect on sky wave propagation. At broadcast band frequencies, the most important of these disturbances are the ionospheric storms. Much like ordinary storms at lower altitudes, these storms are periods of turbulence in the ionosphere when the normal, separate, relatively calm layers cease to be distinct and the ionosphere becomes a jumbled mass of ionized gas in rapid and random motion. Such a trubulent system is unable to properly refract a radio wave and instead scatters and absorbs it. Propagation along any path passing through a region of ionospheric storm activity will be virtually eliminated.

While the time at which an ionospheric storm will occur is random, some idea can be obtained as to the times most likely. The cause of the ionopsheric storm is the arrival in the ionosphere of highly energized particles emitted by the sun. The emission of these particles is associated with solar flares - large clouds of hot glowing gas, often millions of miles long - that erupt from the surface of the sun. These flares are the result of increased solar activity and are most prevalent near and slightly after the time of the sunspot maxima. Since the light from the flare reaches the earth some time before the particles do, sighting of a flare by an observatory is an indication of probable ionospheric storm activity in the immediate future. These particles enter the ionosphere mainly in the regions above the magnetic poles, so the magnetic polar regions are the areas in which ionospheric storms are centered. For DXers in North America, an ionospheric disturbance will black out signals arriving from the north, including those from central Asia which must come over the pole (when they do come). This clears many frequencies for reception from Central and South America. The corresponding effect in the southern hemisphere makes it almost impossible to hear South African stations from New Zealand. Besides disrupting normal propagation, these storms also cause the gas in the upper atmosphere to glow; this glow is the "Aurora borealis" (or australis) and for this reason propagation conditions resulting from ionospheric storms are often called "auroral conditions." Once started, such a condition may last 2 or 3 days.

In summary: sky wave propagation involves the return of a signal to the earth from the ionosphere. It is the method by which broadcast band signals travel over large distances (from a few hundred to several thousand miles), but, because of many changes in the ionosphere, the characteristics of this means of propagation are constantly changing. Of great importance is the existence in the daytime of the absorbing D and E layers, allowing sky wave propagation only when the path from transmitter to receiver is mostly or entirely dark. Summers and years of sunspot maxima are the worst times for this type of propagation, whereas winters and sunspot minima are best.

Propagation by Several Modes

So far, we have discussed ground wave and sky wave propagation as if only one of these modes was present at a given time. For daytime reception this is true: ground wave alone is present. At night, just because the sky wave becomes possible and greatly dominates over ground wave, this does not mean that the ground wave disappears or is weaker - it is essentially constant throughout the day and night.

The presence of two possible paths for a signal to travel to a given receiver location might merely result in a greater signal strength received, but this isn't always true. The difference in the lengths of these paths results in a difference in phase between the two arriving signals. The phase of the ground wave is essentially constant so any change in the path differences will come from the sky wave variations. A phase shift of 0°, 360°, 720°, etc. (corresponding to a path length difference of 0, 1 wavelength (λ), 2λ, etc.) results in the addition of the two signals: while a phase shift of 180°, 540°, 900°, etc. (a path length difference of $\frac{1}{2}\lambda$, $1\frac{1}{2}\lambda$, $2\frac{1}{2}\lambda$, etc.) results in the subtraction of the two signals. Thus, a change of $\frac{1}{2}\lambda$ (from 300 to 900 feet for the BCB) in the path length difference can result in signal strength variations equal to twice the strength of the weaker of the two components. This variation in path length is so small that it occurs very readily due to normal fluctuations in the ionosphere, resulting in "fluttering" in the signal from many semi-local stations (out to perhaps 200 miles) at night, since for these stations the sky and ground waves are of comparable strength.

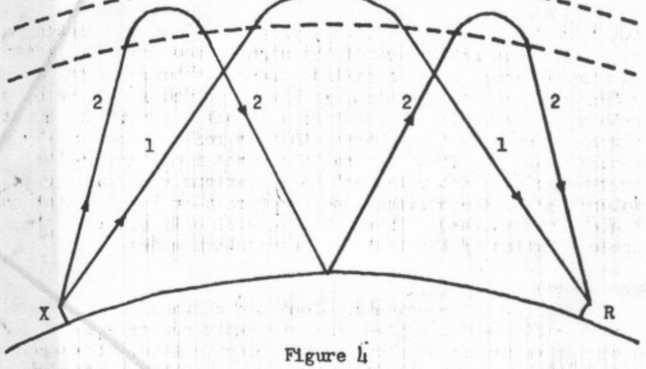

Figure 4

Similar effects can occur at greater distances due to the presence of "multi-hop" sky wave propagation (fig. 4). Path 1 is the normal one-hop sky wave that has already been discussed. Path 2 is an example of two-hop propagation: refraction from the ionosphere, reflection from the ground at point A, and another trip to the ionosphere and back, finally arriving at the receiver. Again the same signal arrives by two different paths, possibly resulting in some cancellation of the signal. While not shown, other possible paths of three or more hops also exist, and scattering of waves as they pass through lower ionospheric layers causes spreading of the idealized paths, further complicating the problem. In general, the signal arriving by the fewest number of hops will be strongest. The largest distance that can be covered by a single hop is about 2000 miles. Beyond this distance, the two-hop signal predominates with the three-hop being the chief component of interference. Beyond 4000 miles the two-hop disappears and the three-hop is dominant, and so on. For multi-hop propagation, the ground conductivity at a point of reflection is important in much the same way that it is for ground wave propagation. For example, TA and TP paths are often good for reception over long distances, but long cross-country reception is more difficult. The presence of several simultaneous modes of propagation and the deviation from ideality of the received signals makes day to day reception on the BCB quite variable and difficult to predict accurately, but hopefully the above discussion will have given you an idea of the general physical phenomena that control BCB propagation.