A BRIEF HISTORY OF IONOSPHERIC STUDIES

Robert R. Brown, NM7M

Editorial Note: The following is certainly the most scholarly propagation article yet published by *Proceedings.* This is not surprising because Bob Brown, NM7M, is also Dr. Robert R. Brown, Emeritus Professor of Physics, University of California, Berkeley. Bob spent a career as high latitude geophysicist focusing on the structure and behavior of the polar regions of the ionosphere when under bombardment by auroral electrons and solar protons. Since his retirement, Bob has become a widely acclaimed columnist in several magazines where he shares the knowledge of ionospheric researchers with radio enthusiasts throughout North America. The following article outlines our understanding of the ionosphere as it developed over the last century. It also describes our current view of this highly complex medium and describes the directions of current research. The article is written in terms familiar to every professional working in the field today. However, unlike most scholarly papers, Bob has carefully defined each professional term the first time that it is used. This article deserves careful study by every radio enthusiast interested in long distance radio. It will provide all but the most casual reader with a much clearer picture of the ionosphere and how it behaves.

INTRODUCTION

At this late date in history, almost 100 years after the first observations of radio propagation, we can look back to see where we've been and look ahead to where the interesting developments will be taking place. At present, we're in the "plasma and fields era", the ionosphere now recognized as being but one part of a coupled system, thermosphere-ionosphere-magnetosphere, which finds its origin at the sun. Thus, we're learning about the top-side of the ionosphere and its interaction with regions inward and outward. But our interest in communications is really tied to the first 50 years of radio, the exploration of the ionosphere below the peak of the F-region and what might be termed the "photo-chemical era".

In that earlier era, it was shown that the ionosphere consists of free electrons, positive ions, atoms and molecules, resulting from photo-ionization and photo-dissociation by ultra-violet (UV) radiation from the sun. There was one complication, however; the ionosphere is immersed in the earth's magnetic field which constrains electrons and positive ions to gyrate around the local field lines. As a result, the local properties of the ionosphere are no longer identical in all directions and radio propagation varies with direction relative to the geomagnetic field, particularly at lower frequencies.

Theoretical discussions showed that inertia and magnetic forces on electrons affect their motions when exposed to radio waves and that results in propagation which is frequency-dependent. In particular, the way in which a radio wave propagates through the ionosphere depends on how its own frequency (F) compares with several other frequencies determined by the local properties of the medium itself.

The most important frequency for propagation purposes is the local plasma frequency (fn) which depends on the number density of electrons at a given height. The

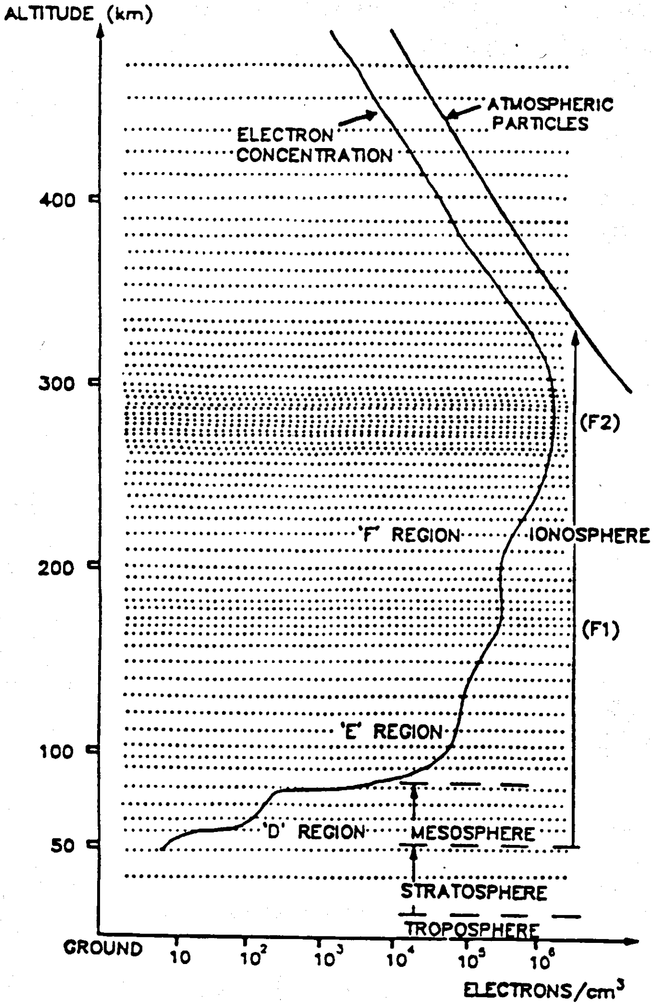

Figure 1 - Daytime variation of the electron density with altitude. (From McNamara, 1994)

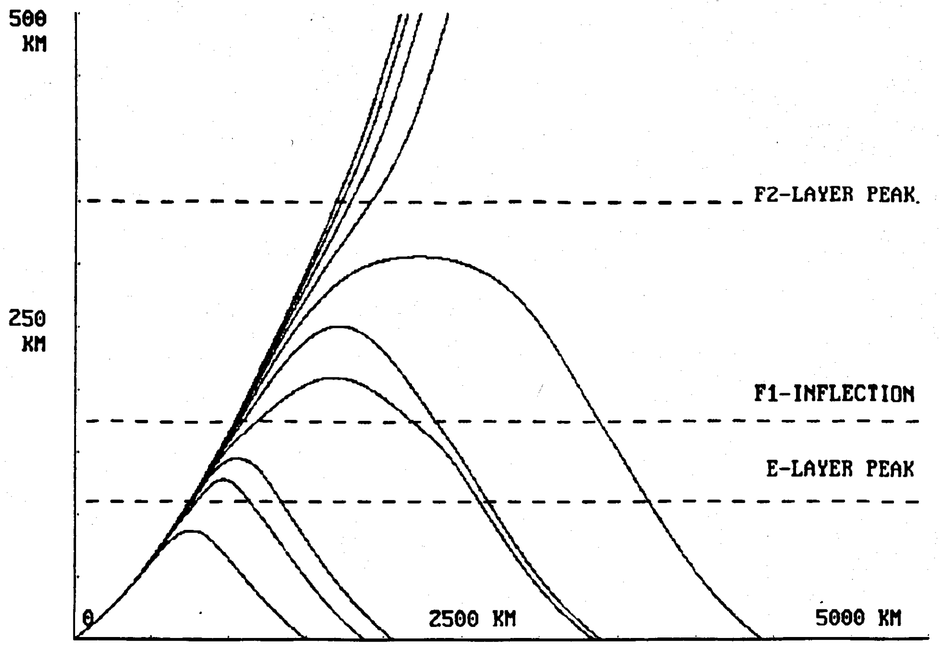

Figure 2 - Refraction and penetration of the ionosphere by waves from 3 to 30 MHz in 3 MHz steps.

collision frequency (fc), giving the rate of collisions between electrons and nearby constituents, is another one. Finally, there's the electron gyro-frequency (fH) which gives the rate at which electrons spiral around the local field lines. To put the various frequencies in perspective, the plasma frequency fn ranges from 3 to 15 MHz over the earth, the electron collision frequency fc ranges from 1 to 10 MHz with decreasing altitude and finally the electron gyro-frequency fH averages about 1 MHz at ionospheric heights.

As to the significance of those frequencies, signal absorption depends on the electron collision frequency in the lower ionosphere, the plasma frequency determines the frequency limit for successful propagation on a path while a comparison of the electron gyro-frequency and the signal frequency determines the extent to which the ionosphere is anisotropic, waves propagating differently according to direction relative to the magnetic field and polarization. At the upper end of the HF spectrum (3-30 MHz), the plasma frequency is the most important of all three while the collision and the gyro-frequency dominate at the lower end of the range. So much for the theoretical formalities.

On the experimental side, the historical record shows that our knowledge and understanding in the photo-chemical era proceeded outward in direction and upward in frequency. Thus, we came to know the first hundred kilometers above the earth, now termed the D-and E-regions of the ionosphere, are largely controlled by the degree and duration of solar illumination. In the course of time, the frequencies in common useage rose into the HF range, 3 to 30 MHz, and our radio experience touched even higher and different regions, say the two parts of the F-region. With that, it became apparent that the sun played an even larger role, affecting growth and decay of ionization in the F-region in different ways than at lower altitudes and also by its slowly changing activity, as judged by the rise and fall of sunspot counts in its cyclical behaviour.

Beyond that, more dynamic influences of solar activity became evident in the first 50 years, sudden ionospheric disturbances (SID) on the sunlit side of the earth affecting the D- and E-regions with solar flares, D-region black-outs of propagation across the polar caps following larger and more energetic solar flares and even global disruptions of F-region propagation with geomagnetic storms. While a good fraction of those sporadic disturbances occurred close to or within the peak activity in the solar cycles, they were not limited just to those times, thus adding an element of unpredictability to HF radio propagation that was not there at lower frequencies. So much for the early history.

HF PROPAGATION

While the formation of the ionosphere is of interest to those working in physics and chemistry, others who follow communications are more interested in when propagation takes place and how it would be disturbed by external influences. Of course, the "when and how" of propagation follows from electromagnetic theory, wave refraction by an

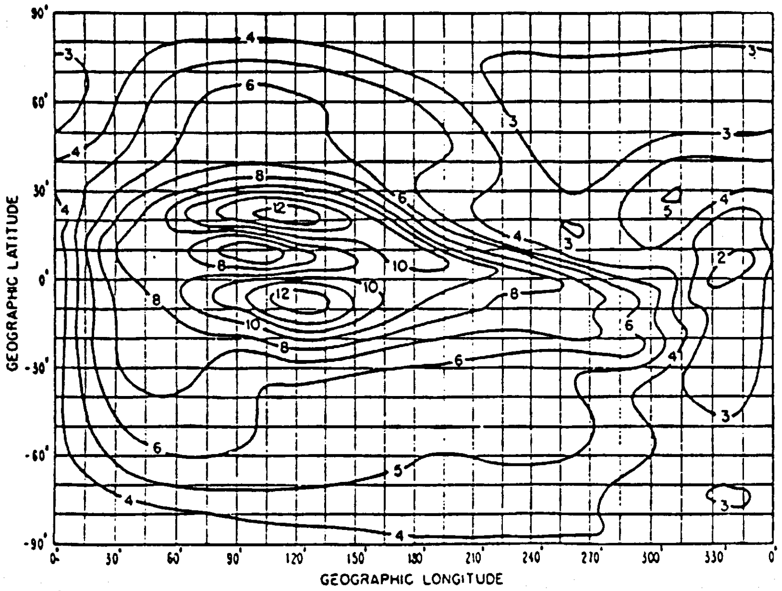

Figure 3 - Global foF2 map for March 1976 at 0600 UTC (From Davies, 1990)(The sun is located at 90° longitude and 0° latitude)

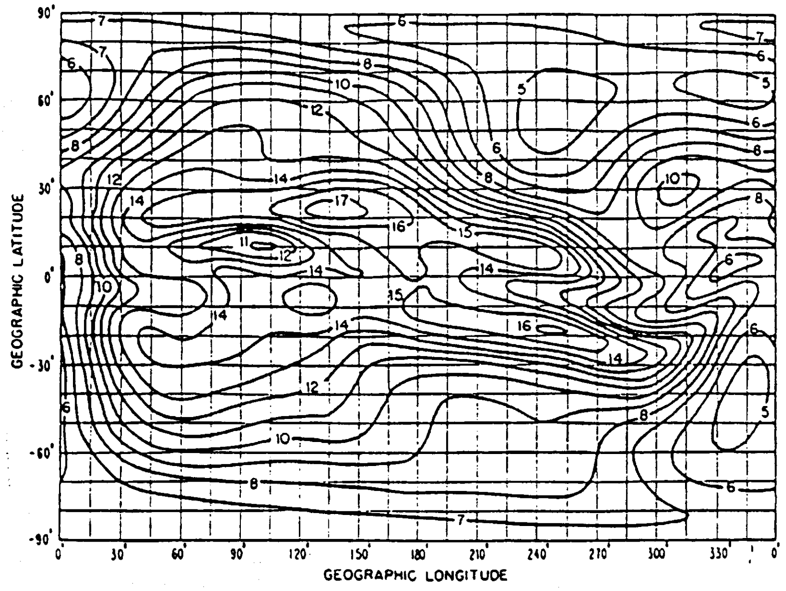

Figure 4 - Global foF2 map for March 1979 at 0600 UTC (From Davies, 1990) (The sun is located at 90° longitude and 0° latitude)

ionized medium, as well as the global distribution of ionization at a given time. An excellent discussion of those ideas can be found in the new book, Radio Amateurs Guide to the Ionosphere, by Dr. Leo McNamara (Krieger Publishing Co., 1994).

A complete treatment of oblique ionospheric propagation would require not only a full mapping of the limiting or critical frequency for reflection at vertical incidence (foF2) as a function of latitude, longitude but also the date, time of day and sunspot number. Beyond that, each foF2 map would require details of the vertical distribution of ionization throughout the various regions. By way of illustration, a representative electron density profile is shown in Figure 1.

Without being too specific, it is safe to say that for waves of a given radiation angle, the distances covered on propagation paths are greater the higher the frequency until the waves penetrate the layer and escape. This is illustrated in Figure 2 where ray-tracing has been used for (ordinary) waves launched at a 5 degree radiation angle, the HF frequency range (3-30 MHz) being covered in steps of 3 MHz for a simple F-layer with foF2=6 MHz. Here, it is seen that the lower frequencies do not penetrate the ionosphere as much as the higher ones, particularly near the critical frequency for penetration.

In the late '30s, those ideas lead to the concept of a maximum useable frequency (MUF) for a path. Further, since the height and critical frequency of the ionosphere depend on the time of day and the level of solar activity, with the aid of foF2 maps of the globe for different times and levels of solar activity, it should be possible to predict radio propagation off in the future, for the same months, at similar times of day and in other solar cycles. Two such maps are shown in Figures 3 and 4, the first one showing the state of the ionosphere at solar minimum in 1976 and the second one at the following solar maximum in 1979.

Those two maps were representative for conditions at the spring equinox, a time when the earth is illuminated symmetrically with respect to the equator and the sunlit region extends from 0 to 180 East geographic longitude. The fact that significant ionization extends beyond the sunlit region is due to the slow rate of electron loss at F-region altitudes, on the order of 300 km. In addition, the lack of symmetry of the map with respect to the geographical equator is due to the geomagnetic control of electrons in the ionosphere. Thus, a better or more symmetrical organization of foF2 maps would be found using geomagnetic rather that geographic latitudes. However, for practical communications purposes, the foF2 maps are given in more convenient geographical coordinates.

Before WW-II, vertical soundings of the ionosphere were taken at only a small number of stations. But with the start of hostilities in WW-II, such observations were expanded greatly so as to provide reliable radio communications on a world-wide scale. By the 1950's, maps such as shown in Figures 3 and 4 were readily available. But beyond those steps, the "reach" of propagation predictions was extended by Newbern Smith in the USA and independently by K. W. Tremellen in the U.K. Their approach was to consider paths longer than one hop and apply the critical frequency idea at each end of a path. Thus, a tentative MUF was worked out for the first hop on a path and then again for the last hop.

Experience seemed to show that a path would fail most often at one end or the other so paths with two or more hops were considered to be open for oblique propagation at a frequency below the lower of the MUFs for hops at the two ends. That scheme is called the "control-point method" and continues in use today although if one were pressed for a more reliable estimate of an over-all MUF, MUFs of any intermediate hops would have to be considered as well and the MUF for the entire path taken as the lowest one for all the hops.

MODERN DEVELOPMENTS

The global maps for foF2 are but one example of the wealth of information that had been collected and organized as a result of the expansion of ionospheric sounding after WW-II. With that database available, the American Central Radio Propagation Laboratory published ionospheric maps on a regular basis and made them available to parties interested in making their own predictions of HF radio propagation. In a historical sense, those maps represent the culminating documentation of the photo-chemical era of ionospheric study.

But there were disruptions of that well-ordered scheme, intrusions which are now recognized as relevant to the new era where plasmas and fields dominate the discussion of the ionosphere. One example, found J. H. Dellinger in 1937, was that some solar flares cause radio black-outs from the effects of bursts of solar X-rays penetrating to the D-region of the sunlit hemisphere. Other disturbances studied in that era had similar origins: the bombardment of the ionosphere by electrons from aurora or protons from some energetic solar flares. Finally, scientists in the late 1930s began to note that there were cases of ionospheric disturbances associated with magnetic storms.

In a sense, the understanding of the origin of those effects on radio propagation awaited the developments in theory and experiment which came well after WW-II. In the photo-chemical era, the cases of particle bombardment at auroral and polar latitudes were discussed in terms of the particles being guided in their motions by the geomagnetic field, then viewed as a symmetrical dipole. In essence, the particles were agents, guided by magnetic field, which "mapped" downward other disturbances in regions at much greater heights. While those events were of interest in their own right, the study of their distant origins and mapping along field lines ushered in a new era, now dealing with the coupling between the plasma in the ionosphere and the distant magnetosphere, and even the interplanetary magnetic field.

PLASMAS AND FIELDS

The figure of speech, "action-at-a-distance", may not seem appropriate to characterize the origin of the basic ionosphere, solar photons liberating ionospheric electrons after going across the open region between the sun and the earth, but there is some truth to it. But the coupling between the ionosphere and the magnetosphere is different, involving the continuous interaction between plasma (protons and electrons) and fields, both electric and magnetic, across that same vast expanse. And that is the case even though it begins back at the sun where field strengths are large and ionization complete and ends in the ionosphere where the earth's magnetic field is weak and the ionization of its atmosphere only slight.

The physics of the new plasma and fields ionospheric era is purely classical, the combination of Maxwell's Electromagnetism and Newton's Mechanics, and is termed "magneto-hydrodynamics" in some circumstances and plasma physics in others. While ionospheric effects appear simple when mapped down to low altitudes, only involving small numbers of particles or ions in the ionosphere, the problems are complex in formulation and geometry. To see the latter, just look at Figure 5 which shows the earth and its magnetosphere, as understood at the present time.

This figure gives a snapshot of the field line topology, the earth's magnetic field compressed by the solar wind on the dayside and drawn out on the nightside. Historically, the geomagnetic field had been thought to be like that of a dipole. But with observations of the solar wind by spacecraft, the dipole model gave way to one where the geomagnetic field is confined to a cavity carved out of the stream of interplanetary plasma going by, something like the wake behind an object fixed in a stream.

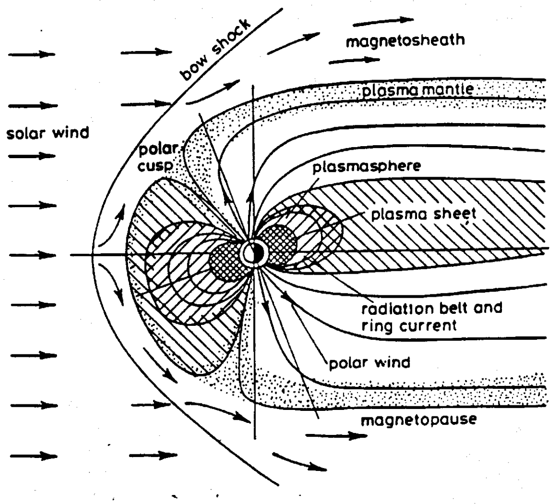

Figure 5 - The earth's magnetosphere (From Davies 1990)

The first observations which suggested this new configuration of the earth's magnetic field came early in the Space Age, a Pioneer spacecraft going out from earth carrying nothing more sophisticated than a rotating search coil. The magnetic field in close to the earth in the sunward direction was orderly, decreasing as expected with a dipole model. However, out around 8-10 earth radii distance from the geocenter, the field started to fluctuate in magnitude and direction, suggesting a disordered magnetic regime. After that, the field became more orderly again, when the spacecraft passed into the region of the interplanetary field.

What is now termed the magnetotail has something of a similar history but involved spacecraft in highly elliptical orbits, crossing the region behind the earth time after time. From those two sets of satellite observations, the present model finally emerged: a turbulent region behind the bow shock formed by the solar wind and an extensive magneto-tail behind the earth. Without getting into processes in the magneto-tail, at least for the moment, Figure 5 shows the geomagnetic cavity carved out of the solar wind has two field-line configurations. The first is a low-latitude toroidal arrangement of dipole-like field lines and the second a high-latitude arrangement where field-lines rise from the polar regions and extend into the geomagnetic tail on the nightside of the earth. These two field-line geometries have the effect of dividing the terrestrial ionosphere into three regions, one mapped by low-latitude field lines, the other mapped by field lines and plasma processes in the magnetotail and a transition region between the two regimes on the front of the magnetosphere, denoted as the polar cusp, where solar plasma may enter directly into the ionosphere.

The early development of ionospheric physics took place when the dipole model was still considered the correct description of the geomagnetic field. While not appreciated at the time, two of the major ionospheric disturbances known then find their understanding in the present field model. Thus, polar black-outs were known before WW-II but it was not until the International Geophysical Year (IGY) of 1957 that the geographical extent and character of those black-outs was established.

The principal instrument responsible for untangling the black-out problem was the riometer (Relative Ionospheric Opacity METER). It is a radio receiver which monitors galactic radio noise around 30 MHz. Thus, its record shows the signal strength of the radio galaxy as it passes overhead. With any radiation bombardment of the atmosphere that penetrates to the D- or E-regions, increased ionospheric absorption occurs and the galactic radio noise reaching ground level is reduced.

Those polar black-outs, now known as polar cap absorption (PCA) events, follow some solar flares and result from energetic protons accelerated at the flare site reaching the vicinity of the earth. The earth's magnetic field acts as a spectrometer, allowing only protons of a certain minimum momentum (or energy) to reach a point on the earth and deflecting all others away, back into space. The cut-off energy as a function of geomagnetic latitude could be calculated with the dipole model and suggested ionospheric absorption would be the greatest at the magnetic poles where the protons would simply slide down the vertical field lines.

The riometer network over the northern polar cap during the IGY showed otherwise, the ionospheric absorption showing little, if any, latitude variation during PCA events. After much intellectual anguish, it became apparent that the dipole model had to be abandoned. But what to take its place? It took some time to find the present model.

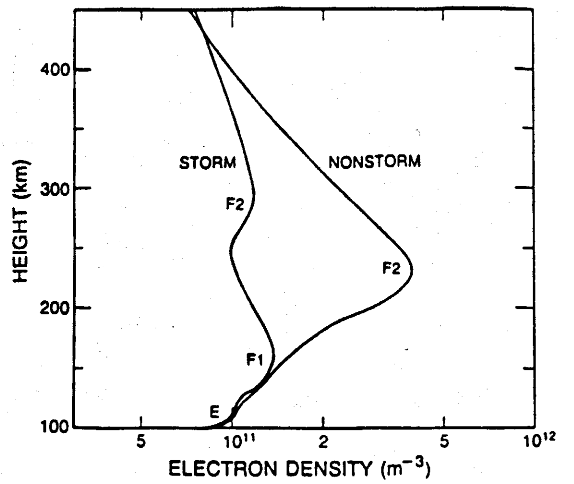

Figure 6 - Electron density profiles for quiet and storm conditions (From Davies, 1990)

The auroral displays in the optical part of the spectrum and auroral ionospheric absorption (AA events) in the radio spectrum resulting from such activity also had some clues within them. In particular, with the equatorward expansion of auroral displays at the onset of magnetic storms, it was found that any polar cap absorption in effect at the time expanded in area also. In terms of the field topology shown in Figure 5, those effects would be understood if more non-dipole field lines were in the geographic polar caps and auroral activity located on field-lines just below the transition between dipole and geotail field-lines.

The physics of the magnetosphere is difficult and often defies one's intuition. Thus, the explanations of polar cap expansion and auroral activity involve the concept of "field-line merging". For polar cap expansion, the physics of the problem indicates that solar wind plasma coming up to the earth carries magnetic field-lines with it and on encountering the front of the magnetosphere, the interplanetary field lines may actually "merge" with terrestrial field lines if the interplanetary field has a significant southward component. On that basis, dipole-like field lines would "merge" with inteplanetary field lines and then be carried back into the magneto-tail as the solar wind continues to advance, thus enlarging the polar cap.

Other merging or "reconnection" processes are invoked back in the magnetotail and result in some field lines moving in toward the front of the magnetosphere. Those same processes are considered to energize low-energy electrons in the magnetotail, sending them earthward and giving rise to auroral displays. Of course, all this takes place in three dimensions, not just the two dimensions in Figure 5, and the result is a convective motion of field lines and associated plasma, from both the solar wind and the magnetotail directions.

There are other disturbances in the F-region during geomagnetic storms and in contrast with PCA and AA events, they reach into regions of lower latitudes as well as a wide range of longitudes, lowering the critical frequency foF2 there. While the effects of storms on foF2 are complicated, varying with local time, season and latitude, their study was carried out using data organized according to geomagnetic latitude of the observing site and the time after the sudden commencements of magnetic storms.

Those studies showed that major magnetic storms could reduce foF2 at sub-auroral latitudes by as much as 30% and reach 5-10% reductions at mid-latitudes in the day after the onset of a storm. In a qualitative sense, the latitude variation of foF2 reduction can be understood in terms of processes being initiated on field lines as solar plasma compresses the magnetosphere, ionospheric effects being greatest on higher latitude field lines which map closer to the front of the magnetosphere. And by the same token, the fact that a wide range of longitudes are affected means that convective processes are involved as well.

An excellent example of the storm depression of foF2 is shown in Figure 6, an electron density profile over St. John, Newfoundland in April 1965 from the records of ground-based and satellite-borne ionosondes. The loss of ionization at the F2-peak and the presence of a lower F1-peak in ionization resulted in a reduction in critical frequency and a shortening of hop lengths across Newfoundland during the disturbance.

Of the three forms of ionospheric disturbance, solar proton or auroral electron bombardment and ionospheric storms from magnetic disturbances, the latter ties together theories and understandings developed during the earlier photo-chemical era and with the more modern ones developed during the current era involving "plasmas and fields". Thus, the present view is that the undisturbed ionosphere is really confined within the low-latitude, dipole-like field line regime seen in Figure 5. If viewed from above one of the poles, the ionospheric plasma contained in the toroid would extend out to about 6 earth-radii, corresponding to a geomagnetic latitude limit of 65 degrees.

The ionosphere in that region is created close to the earth by solar UV but photo-chemical processes determine the ionic composition and distribution in altitude. At low altitudes, the atmosphere is well mixed and the positive ions are mainly those of nitrogen and oxygen molecules. Around E-region altitudes, the photo-dissociation of oxygen becomes important and the principal ion in the F-region is that of atomic oxygen. But at those altitudes, collisions are infrequent and the gases are no longer well mixed; instead, they are sorted out by mass in the gravitation field, the lighter ones extending further out from the earth. That affects the ionic composition as the lightest atmospheric constituent, atomic hydrogen, undergoes charge exchange with atomic oxygen ions, thus creating the high ionosphere populated with electrons and protons.

The electrons and protons are constrained to gyrate around the magnetic field lines, sometimes likened to loops or rings held on a metal wire. The ionized gas or plasma will flow upward along the magnetic lines of force until the plasma gas pressure is equalized along the entire line of force, extending from the northern to the southern hemisphere. The ionosphere as we know it is located at the feet of the magnetic field lines while the region containing plasma at greater altitudes is called the plasmasphere. The outer limit of the plasmasphere is called the plasmapause and plasma densities fall by one or two orders of magnitude in the 2-4 earth-radii between the plasmapasuse and the magnetopause, the limit of ordered geomagnetic field lines.

So much for the description of the undisturbed ionosphere. Now, when it comes to disturbances, in this model they are to be understood to result from field line merging advancing inward to around the plasmapause, even further inward if the pressure of the solar wind is great enough. With merging, the electrons and protons are no longer held on field lines which are closed or connected with the conjugate hemisphere. Instead, the particles will diffuse outward and be lost for ionospheric processes as they gyrate on field lines which are now "open", going back toward the magnetotail and then connected to the interplanetary field.

Ionospheric disturbances during magnetic storms involve decreases in MUFs on paths and are always greatest at the higher latitudes. In that regard, the penetration of field line merging provides a natural explanation of the events, ionospheric electrons on the feet of field lines essentially released to the interplanetary medium instead of being held on closed field lines. And when the pressure from the solar wind subsides, the original dipole-like field lines will then become connected again and the plasmasphere will be refilled by photo-chemical processes at ionospheric heights. Of course, refilling takes place as the earth rotates and at a rate which depends on solar activity, taking days to bring the ionosphere back to its pre-storm condition.

CURRENT ISSUES

While a smooth, slowly-changing ionosphere describes the basic understanding of propagation during the photo-chemical era, studies in that era also dealt with disturbed regions which tended to be large and with long time-scales. Such large events will continue to be of interest on into the future. Even in the photo-chemical era there were indications of roles played by smaller regions and on shorter time-scales. This interest was generated by radio scientists noticing flutter effects in propagation at auroral and equatorial latitudes.

At auroral latitudes, the suggestion was obvious, indeed "eye-catching", with auroral displays showing a wide range of activity through their motions and intensity variations. With the knowledge that energetic electrons were coming down field lines, ionizing atoms and molecules as well as exciting visible emissions, it took no stretch of anyone's imagination to conclude that irregular patches of ionization might exist with those same spatial and temporal scales.

The rapid auroral flutter of voice signals on paths going through auroral displays confirmed the matter. Those events are part of what is termed an "auroral sub-storm" and have their origin in field-line merging many earth radii back in the magnetotail, in the direction away from the sun.

Particle precipitation, like that at auroral latitudes, is not found at equatorial latitudes yet radio signals traversing those regions show "flutter", i.e., deep and rapid interference fading, particularly in the evening hours and at the equinoxes. In that regard, there are two important equatorial effects: an equatorial anomaly of the ionosphere, with higher than normal heights and critical frequencies, and an equatorial electrojet of current at E-region altitudes.

The equatorial electrojet, recognized and understood before WW-II, is the more important of the two effects as it is closely related to the source of the second. Thus, there are strong ionospheric currents driven by an east-west electric field or "dynamo" due to the interaction of tidal winds and charged particles, created by photo-ionization, with the local magnetic field. At equatorial latitudes the geomagnetic field has a small dip-angle with respect to the horizontal direction; in addition, the geomagnetic field is in the north-south direction. As a result, ionospheric electrical conductivity in the east-west direction reaches large values just north and south of the geomagnetic equator and strong currents of electrons and positive ions are driven by the dynamo.

The equatorial electrojet may show spatial irregularities as a result of a "two-stream instability", an effect in plasma physics due to the large relative velocity of the electron and ion streams. Those irregularities may give rise to a type of sporadic E-layer but the dynamo electric field driving the current system also results in an evening uplift of the F-region as ionization encounters the north-south magnetic field. This is the "fountain effect", producing a higher than normal F-region at low latitudes.

The flutter fading on equatorial paths points to the presence of irregularities at F-region altitudes and those irregularities, like the ones at auroral latitudes, tend to be magnetic-field aligned, now in the horizontal direction. The presence of spatial variations in the electron density means that scattering of radiation plays a larger role in propagation while that of wave refraction is diminished as a result of less "smoothness" in the medium. In those circumstances, waves may be scattered instead of refracted downward, carrying signals farther forward for longer periods of time and making critical frequencies (and MUFs) seem higher for trans-equatorial paths. Of course, that result is not exactly a new idea as the "operational MUF" for high-power (50 kw) transmitters is higher, due to effects from ionospheric scatter, than the classical MUF derived from refraction.

Higher power, shorter wavelengths and ionospheric scattering are the directions ionospheric research has taken recently, the idea being to penetrate the F-region peak and explore top-side ionospheric structures directly from earth using the back-scatter processes. In the 1960's, satellite-borne ionosondes, such as the Alouette series from Canada, were used to explore the top-side ionosphere but only with convential sweep-frequency techniques because of power limitations. That study proved particularly valuable in obtaining the complete electron density distribution of the ionosphere shown in Figure 1. As might be expected, however, it did not reveal any gross structure of a photo-chemical origin such as found below the F2-layer.

Alouette did show top-side irregularities, ionograms with so-called "Spread-F" characteristics; but satellites, in their ever-changing motions, broaden the scope of some ionospheric studies and frustrate others for lack of continuous coverage. At present, the approach has shifted to using high-power radars at fixed locations (e.g., the American installation at Arecibo, P.R. at 30 degrees magnetic latitude and the EISCAT installations in northern Scandinavia) to explore the structure and motions of ionospheric plasma at and above the F-layer peak.

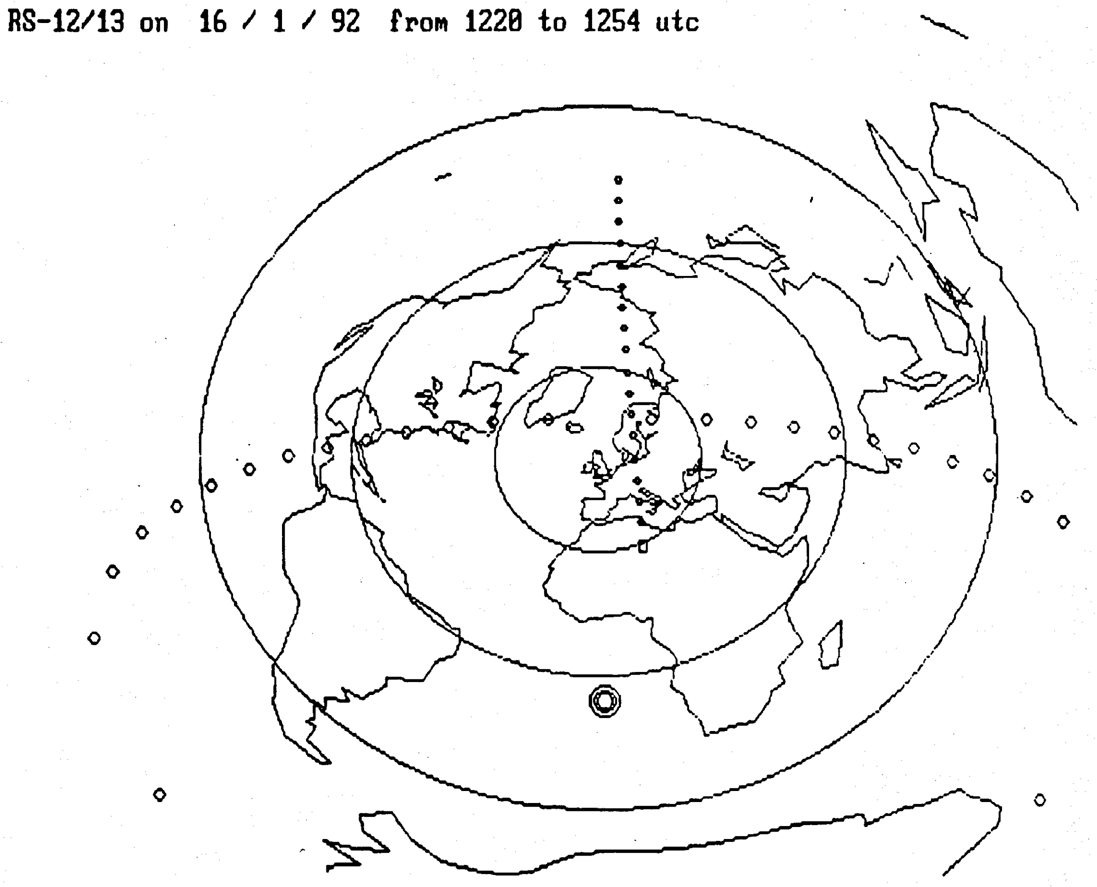

Figure 7 - RS-12 satellite track for January 1992 (From Branegan and Brown, 1993)

One problem of considerable interest is the mid-latitude trough of the F-region, a region of low electron density lying between 50 and 70 degrees magnetic latitude which occurs between dusk and dawn. It was discovered in the recordings of the Alouette satellite and shows a steep gradient in electron density at the F-layer peak. Like the lower boundary of the auroral zone, it is quite variable in location during active periods and the ionospheric tilts from its steep gradients in electron density may produce off-great-circle propagation, a matter of continuing interest and speculation.

From the experimental standpoint, the spatial and temporal resolutions of ground-based sounding programs have been insufficient to show the trough among the features of the quiet ionosphere but its presence does not stand in doubt. At present, theoretical discussion of the mid-latitude trough deals with the convection of ionospheric plasma across the polar cap, driven by processes in more distant regions of the magnetosphere, and possible changes in electron-ion recombination rates which would lower the electron density.

Other ionospheric processes at high-latitudes involve plasma convection from the dayside, showing correlations with the N-S direction of the interplanetary magnetic field (IMF) at the front of the magnetosphere. Thus, the polar F-region has been found to be either characterized by sunward-aligned auroral arcs during quiet conditions when the IMF has a northward component or large patches of enhanced plasma density in anti-sunward motion during active conditions when the field points more southward. Of particular interest is the fact that electron density enhancements up to a factor of ten above quiet conditions have been noted, bringing F-region critical frequencies in the polar cap up from a low level around 2-4 MHz to occasional levels as high at 9-11 MHz.

Those observations were obtained by looking vertically upward with ionosondes in the two polar caps and using IMF data from various satellite observations. Interestingly enough, enhanced patches of ionization have a propagation aspect as well, enabling 29.4 MHz beacon signals from the Russian satellite RS-12 at 1,000 km altitude to be propagated across darkened polar caps for great distances.

The first indication of such unusual propagation came from here in the Northwest when the 29 MHz beacon signals of RS-12 were heard coming over a dark, winter polar cap from just north of Mongolia. This went on over a two-week period, some nine events noted around local noon. A quick check of IONCAP predictions showed the path should not be open at those times, even though the 29 MHz signals came down from 1,000 km altitude and were reflected off the ground some 3,000 km ahead of the satellite. In short, the best model of the polar ionosphere was being violated and a different explanation required for the observations.

An independent verification of those observations was obtained from Scotland where the station logs of GM4IHJ showed no less than twelve cases of similar, distant signals from the RS-12 beacon coming across the polar cap during the same two-week period. And the same local noon feature was observed in Scotland.

It was suggested that the signals represented propagation via drifting patches of F-region ionization and corroboration might be obtained by contacting the US research group conducting ionosonde studies near the geomagnetic north pole. An inquiry was made as to whether their observations showed enhanced patches of ionization during the period in question and the reply confirmed the presence of enhanced F-region ionization as well as foF2 values favorable to oblique propagation by means of them.

An excellent example of these events is shown in Figure 7. On that occasion, January 16,1992, the RS-12 pass was an ascending one and signals picked up in Scotland at 1220 UTC as it came across the southern horizon. The satellite passed over Scandinavia and the northern horizon around 1236 UTC but was last heard at 1252 when the sub-satellite point was just south of the Kamchatka Peninsula (40 N, 170 E). After reflection off the earth, the great-circle distance for that RF path amounted to about 6,000 km.

While those observations covered only a short period, further observations of RS-12 orbits by GM4IHJ in Scotland as well as by L. K. Andersen, a Danish SWL, have extended the coverage from November '92 to mid-March '93 and increased the number of days showing anomalous propagatgion to 38 and with 18 days showing more than one such event.

CONCLUSION

The basic terrestrial ionosphere results from photo-ionization and photo-dissociation of the atmosphere by the solar spectrum. It is made more complicated by the presence of the geomagnetic field and shows temporal structure primarily due to particles, plasma and fields of solar origin reaching the outer boundary of the magnetosphere. While the photo-chemistry of the lower ionosphere and wave propagation at medium and high frequencies are still of interest for communication purposes, in recent years there has been a major shift of priorities and resources in the research community with the result that interest in ionospheric processes now centers at high latitudes and altitudes around the F-region peak.

How those studies are conducted depends on which hemisphere is involved. Thus, in the southern hemisphere, where the population density is low and inhabited land masses more removed from the polar regions, the current emphasis is on expanding coverage by establishing a number of automated observatories at remote sites in Antarctica. Of course, due to power limitations, observations at remote sites are restricted to passive recordings, e.g., overhead ionospheric absorption, magnetic variations and pulsations, etc.

The northern hemisphere, where the population is greater and located closer to the polar regions, offers better logistics so more active programs are pursued and the future will see a continuation of ionosonde observations in the polar cap as well as more radar probing of the topside of the ionosphere. Those programs, however, look upward and not off to the horizon, as with propagation studies. But the prospects for the latter are excellent, particularly in connection with studying the presence of ionization patches in the polar F-region, as found recently from monitoring the RS-12 satellite.

For the useful life of the RS-10 and RS-12 satellites, 29 MHz beacons will be available for studying propagation. Suitable observing sites are another matter as the most favorable locations would be at northern geographic latitudes of 50 degrees or higher. From there, it is possible to look into the polar caps and receive beacon signals returned by ionospheric refraction from ionized patches, even after a ground reflection deeper in the polar cap.

Future possibilities for monitoring those HF satellites are quite promising as there are many sites with permanent population along the coast of northern Russia, Alaska and the Canadian Arctic. With a growing awareness of the interest in the problem, additional satellite recordings of signals from polar patches by amateur operators and SWLs would supplement the vertical ionospheric observations at fixed sites in the Arctic.

Another possibility would be the monitoring of ground-based beacon transmitters. Already, there are a number of 28 MHz beacons operating but they are limited to northern Scandinavia and the U.K. If a number of additional beacons could be established around the Arctic coast, it would be possible to monitor the beacons continuously on trans-polar paths instead of just a limited number of satellite passes each day. That arrangement would provide not only continuous coverage but also show the growth and decay of polar F-region patches by observing signal propagation on individual paths as well as those which intersect in the polar cap.

All of the above amount to an agenda for the next solar cycle, starting in 1996. An “Arctic Initiative” by amateurs and SWLs is an exciting idea to consider. Perhaps it is something that could be nourished and developed even as we go toward solar minimum. Certainly, it would be something at the leading edge of studies in propagation.

BIBLIOGRAPHY

Branegan, J. and R. R. Brown, A New HF Propagation Mode via Electron Density Structures in the Polar F-region, *Proceedings of the AMSAT-UK Colloquium,* Surrey, England, 1993

Buchau, J. and B. Reinisch, *Electron Density Structures in the Polar F-region,* Adv. Space Res. Vol. 11, pp 29-37, 1991

Davies, K., *Ionospheric Radio Propagation,* U.S. Department of Commerce, National Bureau of Standards, Monograph 80, Washington, D.C., 1965

Davies, K., *Ionospheric Radio,* Peter Peregrinus Ltd., London, 1990

Goodman, J. M., *HF Communications Science and Technology,* Van Nostrand Reinhold, New York, 1992

McNamara, L. F., *Radio Amateurs Guide to the Ionosphere,* Krieger Publishing Co., Melbourne, FL, 1994

Whitten, R. C. and I. G. Poppoff, *Fundamentals of Aeronomy*, John Wiley and Sons, New York, 1971