

Utah State University

DigitalCommons@USU

Faculty Honor Lectures

Lectures

5-1-1986

Solar- Terrestrial Physics: A Space Age Birth

R. W. Schunk

Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/honor_lectures

 Part of the [Physics Commons](#)

Recommended Citation

Schunk, R. W., "Solar- Terrestrial Physics: A Space Age Birth" (1986). *Faculty Honor Lectures*. Paper 57.
https://digitalcommons.usu.edu/honor_lectures/57

This Presentation is brought to you for free and open access by the Lectures at DigitalCommons@USU. It has been accepted for inclusion in Faculty Honor Lectures by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.



9.2

Solar-Terrestrial Physics: A Space Age Birth

by
R. W. Schunk



**73rd Faculty Honor Lecture
May 1, 1986
Utah State University
Logan, Utah**

9.2
no. 73

SEVENTY-THIRD HONOR LECTURE DELIVERED AT THE UNIVERSITY

A basic objective of the Faculty Association of Utah State University is, in the words of its constitution,

To encourage intellectual growth and development of its members by sponsoring and arranging for the publication of two Annual Faculty Honor Lectures in (a) the biological and physical sciences, including engineering, called the Annual Faculty Honor Lecture in the Natural Sciences; and (b) the humanities and social sciences, including education, family life, and business administration, called the Annual Faculty Honor Lecture in the Humanities.

The administration of the University is sympathetic with these aims and shares, through the Scholarly Publications Committee, the costs of publishing and distributing these lectures.

Lecturers are chosen by a standing committee of the Faculty Association. According to the Faculty Constitution:

in choosing the lecturers, the committee shall take into consideration the achievements of faculty members in all the various areas of learning represented by the teaching and research of the Institution. Among the factors to be considered shall be outstanding achievement in one or more of the following: (1) creative activity in the field of the proposed lecture; (2) publication of research through recognized channels in the field of the proposed lecture.

R. W. Schunk was selected by the committee to deliver the Annual Faculty Honor Lecture in the Natural Sciences. On behalf of the members of the Association, we are happy to present Professor Schunk's paper.

Committee on Faculty Honor Lecture

Solar-Terrestrial Physics: A Space Age Birth

**by
R. W. Schunk**

73rd Faculty Honor Lecture

May 1, 1986

Faculty Association

and

Utah State University Press

Utah State University

Logan, Utah

Solar-Terrestrial Physics: A Space Age Birth

by
R. W. Schunk*

What is Solar-Terrestrial Physics?

Solar-Terrestrial Physics, in its broadest sense, is concerned with the transport of energy, particles, and fields from the sun to the earth and their consequent effect on the terrestrial environment. Most of the solar energy eventually deposited in our atmosphere, at a rate of approximately a trillion megawatts, arrives in the form of visible light. The study of how this energy affects our environment falls within the purview of *meteorology*, a discipline that has experienced an independent development and that has sufficiently different problems from solar-terrestrial physics that it can be regarded as a separate but neighboring discipline. In contrast, solar-terrestrial physics is concerned with the higher-energy radiations (ultraviolet, x-ray, and gamma-ray) that carry a relatively small amount of power (approximately a million megawatts), but nevertheless have significant and highly variable effects on the terrestrial environment.

In its current evolutionary state, solar-terrestrial physics is not *directly* concerned with the deep interior of the sun, where the energy is generated, nor is it directly concerned with the transport of this energy to the sun's surface. Similarly, at the other end of the sun-earth system, solar-terrestrial physics is not *directly* concerned with the deep interior of the earth, where the electric currents associated with the fluid core generate the earth's strong magnetic field. Very recently, solar-terrestrial physics has embraced paleomagnetism because this subject can shed light on the historic variability of the geomagnetic field, and perhaps thereby provide clues to the past state of the solar-terrestrial system. Also, solar-terrestrial physicists are beginning to ask questions about the effect that solar activity has on weather, and if indeed some connection can be clearly established, the distinction between solar-terrestrial physics and meteorology may become nebulous.

Solar-terrestrial physics, which is now a thriving, well-supported, and widely-known discipline, traces its birth to the late 1950s when rocket and satellite technology ushered in a new age for mankind, the *Space Age*. In this lecture, I will attempt to trace the evolution of the field from its ancestry to the present time, describe what it encompasses, note some major new findings, allude to

*Professor, Department of Physics; Director, Center for Atmospheric and Space Sciences

some societal benefits that are associated with research in this field, and indicate what its future direction appears to be.

Historical Background

Although the birth of solar-terrestrial physics as a discipline can be traced to the beginning of the Space Age, some of the subdisciplines and phenomena have a long and colorful history. Foremost among these are the aurora borealis, sunspots, and the ionosphere. After a brief description of their ancestry, I will describe the influence that the Space Age had on the birth of solar-terrestrial physics and also the impact of the International Geophysical Year, 1957-58.

Aurora Borealis

The aurora borealis, or northern lights, is a natural phenomenon of striking beauty evident in the polar skies, and perhaps is the best known of all solar-terrestrial phenomena. The visual displays of colored light appear in the form of arcs, bands, patches, blankets, and rays, and often the features move rapidly across the night sky. It has been suggested that the earliest records of the aurora can be traced to Stone Age man [Eather, 1980]. References to the aurora appear in the Old Testament, in writings of Greek philosophers including Aristotle's *Meteorologica*, and possibly in ancient Chinese works before 2000 B.C. In most of these early writings, the auroral displays were interpreted to be manifestations of God.

A serious scientific study of auroras began at about 1500 A.D. However, many of the theories put forth by noted scientists were completely wrong. Edmund Halley, who correctly predicted the reappearance of what is now known as Halley's comet, first suggested that the aurora was "watery vapors, which are rarefied and sublimed by subterraneous fire, [and] might carry along with them sulphureous vapors sufficient to produce this luminous appearance in the atmosphere." In 1746, the famous Swiss mathematician Leonhard Euler suggested that "the aurora was particles from the earth's own atmosphere driven beyond its limits by the impulse of the sun's light and ascending to a height of several thousand miles. Near the poles, these particles would not be dispersed by the earth's rotation" [Euler, 1746]. Our own Benjamin Franklin, who was a very respected scientist in his time, thought that the aurora was related to atmospheric circulation patterns [Franklin, 1779]. Basically, Franklin argued that the atmosphere in the polar regions must be heavier and lower than in the equatorial region because of the smaller centrifugal force, and therefore, the vacuum-atmosphere interface must be lower in the polar regions. He then further argued,

may not then the great quantity of electricity, brought into the polar regions by clouds, which are condensed there, and fall in snow, which electricity would enter the earth, but cannot penetrate the ice; may it not, I say, break through the low atmosphere, and run along the vacuum over the air and towards the equator; diverging as the degrees of longitude enlarge; strongly visible where densest, and becoming less visible as it more diverges; till it finds a passage to the earth in more temperate climates, or is mingled with their upper air.

Franklin claimed that such an effect would “give all the appearances of an Aurora Borealis.”

Numerous other theories of the aurora have been proposed over the last 150 years, including reflected sunlight from ice particles, reflected sunlight from clouds, sulphurous vapors, combustion of inflammable air, luminous magnetic particles, meteoric dust ignited by friction with the atmosphere, cosmic dust, currents generated by compressed cosmic ether, thunderstorms, electric discharges between the earth’s magnetic poles, electric discharges between fine ice needles, etc. A comprehensive and fascinating account of the aurora in science, history, and the arts has been written recently by Eather [1980], and additional theories are given there.

Although early auroral theories did not fare very well, observations made during the latter half of the 1700s and throughout the 1800s elucidated many important auroral characteristics. In 1790, the English scientist Cavendish used triangulation and estimated the height of auroras as between 52 to 71 miles [Cavendish, 1790]. In 1852, the relationship among geomagnetic disturbances, auroral displays, and sunspots was clearly established; the frequency and amplitude of these features varied with the same 11-year periodicity [Sabine, 1852; Wolf, 1852]. In 1860, Elias Loomis drew the first diagram of the region where auroras are most frequently observed and noted that the narrow ring is not centered on the geographic pole, but that its oval form resembles lines of equal magnetic dip, thereby establishing the relationship between the aurora and the geomagnetic field. In 1867, the Swedish physicist Angström, who pioneered spectral studies, made the first crude measurements of the auroral spectrum [Angström, 1869]. However, a significant breakthrough in understanding auroral behavior was not achieved until the closing decades of the nineteenth century, when *cathode rays* were discovered and subsequently identified as *electrons* by the British physicist J. J. Thomson. With this knowledge in hand, the Norwegian physicist Kristian Birkeland proposed that the aurora was caused by a beam of electrons emitted by the sun; those electrons reaching the earth would be affected by the earth’s magnetic field and guided to the high-latitude regions to create the aurora. Thus began modern auroral physics.

ANNUAL MEAN SUNSPOT NUMBERS FROM 1610 TO THE PRESENT

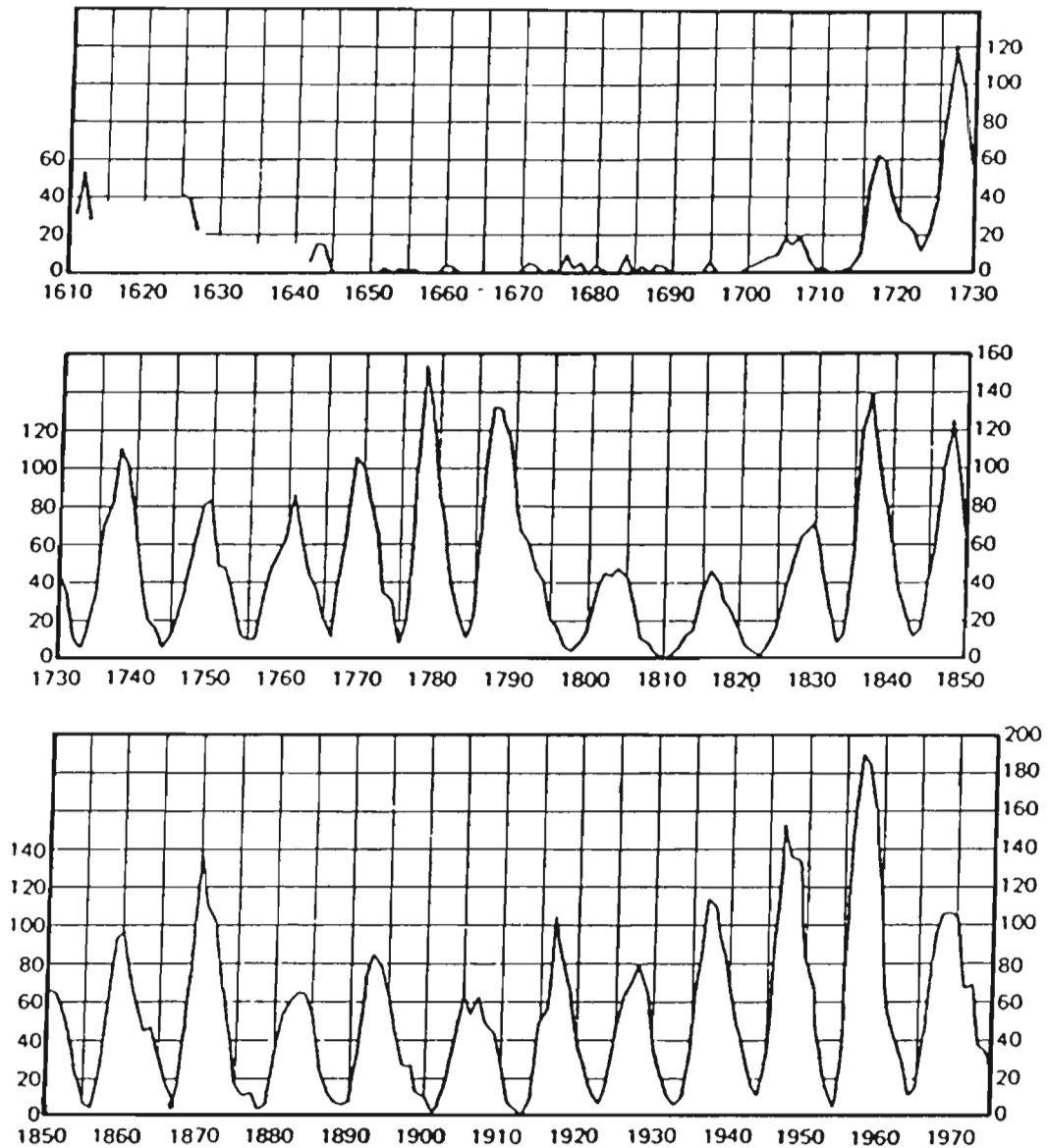


Figure 1. Annual Mean Sunspot Numbers from 1610 to 1976. The numbers before about 1650 are not reliable. From Waldmeier [1961] and Eddy [1976].

Sunspots

Until the discovery of sunspots by Galileo in 1610, the sun was generally thought to be a quiet, featureless object. Galileo not only discovered the dark spots, but he also noted their westward movement, which was the first indication that the sun rotates. In subsequent observations it was quickly established that the number of sunspots varies with time. It was not until more than two centuries later, however, that an amateur astronomer in Germany, Heinrich Schwabe, noted an apparent 10-year periodicity in his 17 years of sunspot observations [Schwabe, 1843]. Shortly after Schwabe's discovery, professional astronomers

set out to determine whether or not the cycle was real. The leader of the worldwide effort was Rudolf Wolf of the Zürich observatory. Wolf conducted an extensive search of past data on sunspots and was able to clearly establish that the number of sunspots varied with an 11-year cycle that had been present since at least 1700 [Wolf, 1856]. In 1890, Maunder [1890] called attention to the 70-year period from 1645 to 1715, when almost no sunspots were observed. Because of this period, which is now known as the Maunder Minimum, it is not clear whether the sunspot cycle is a universal feature of the sun or just a modern phenomenon.

Since the number of visible sunspots varies significantly from day to day because of the solar rotation and because of the growth and decay of both individual spots and groups of spots, annual averages are usually taken. Figure 1 shows the annual mean sunspot numbers from 1610 to 1976. Clearly evident in the figure are the Maunder Minimum Period prior to 1700 and the 11-year cycle since this time.

Ionosphere

The ionosphere is that part of the earth's upper atmosphere where the concentration of free electrons and ions is sufficient to affect the propagation of radio waves. It begins at an altitude of about 70 kilometers and extends to roughly 3000 kilometers, with the peak electron concentration occurring at approximately 300 kilometers. The first suggestion of the existence of what is now called the ionosphere can be traced to Stewart [1882], who was interested in geomagnetism. Prior to this time, it had been well established that there was a direct correlation between the solar cycle and magnetic disturbances on the earth. To account for this strong correlation, Stewart speculated that electrical currents must flow in the earth's upper atmosphere, and that the sun's action is responsible for turning air into a conducting medium. He also concluded that the conductivity of the upper atmosphere is higher at sunspot maximum than at sunspot minimum. This view, however, was not widely accepted and strong counterarguments were put forth in 1892 by Lord Kelvin.

The existence of the ionosphere was clearly established in 1901 when G. Marconi successfully transmitted radio signals across the Atlantic. This experiment indicated that radio waves were deflected around the earth's surface to a much greater extent than could be attributed to diffraction. The following year, A. E. Kennelly and O. Heaviside suggested that free electrical charges in the upper atmosphere could reflect radio waves [Ratcliffe, 1967]. In 1903, J. E. Taylor suggested that solar ultraviolet radiation was the source of the electrical charges, which implied solar control of radio propagation. The first rough measurements of the height of the reflecting layer were made by Lee de Forest and L. F. Fuller at the Federal Telegraph Company in San Francisco from 1912

to 1914. The reflecting layer's height was deduced using a transmitter-receiver spacing of approximately 500 kilometers, which was determined by the circuits of the Federal Telegraph Company [Villard, 1976]. However, the de Forest-Fuller results were not well known, and generally accepted measurements of the height of the reflecting layer were made in 1924 by Breit and Tuve [1925] and by Appleton and Barnett [1925]. The Breit-Tuve experiments involved a "pulse sounding" technique, which is still in wide use today, while Appleton and Barnett used "frequency change" experiments, which demonstrated the existence of downcoming waves by an interference technique. These successful experiments led to a considerable amount of theoretical work, and in 1926 the name "ionosphere" was proposed by R. A. Watson-Watt in a letter to the United Kingdom Radio Research Board [Watson-Watt, 1929].

Birth of Solar-Terrestrial Physics

The rocket technology available at the end of World War II was used by scientists to study the upper atmosphere and ionosphere, and it also paved the way for space exploration via satellites. The great potential of this new technology, coupled with a major advance in ground-based instrumentation, led scientists to realize that a dramatic increase in our knowledge of the terrestrial environment was possible. To take advantage of these new capabilities, the International Geophysical Year (IGY) was organized.

The idea of international cooperation to study the geophysical environment was not new, and in fact the IGY was the third in a series of such endeavors. The International Polar Year of 1882-83 was the first major event in which many nations took part in a multidisciplinary study of the geophysics of the polar regions, with emphasis on the Arctic. During this year, teams of scientists set up polar meteorological, magnetic, and auroral stations and operated them for a full year, obtaining a wealth of geophysical data. Fifty years later, in 1932-33, radio techniques had revolutionized concepts of the upper atmosphere, and this led to the Second International Polar Year. The program for the Second Polar Year was carried out, but unfortunately the outbreak of World War II in 1939 suspended normal international scientific relations and much of the data was not analyzed. At the end of the war, a commission was established to oversee the analysis of the data and a target date for completion was set at December 1950. With this target date rapidly approaching and with the availability of the new rocket technology, scientists in early 1950 began to plan for a Third International Polar Year, but the scope of the program was broadened and it was later renamed the International Geophysical Year 1957-58 [Van Allen, 1984; Nicolet, 1984]. This new cooperative effort was to begin twenty-five years after the Second Polar Year, and the timing was to coincide with the next maximum of the solar cycle.

As a part of the IGY, scientists proposed to launch artificial satellites, and in 1954 the National Science Foundation indicated it would support the launch of a small satellite during 1957-58. In July 1955, the White House announced that the United States would launch “small unmanned earth-circling satellites as part of the U.S. participation in IGY” [Emme, 1961]. The Soviet Union also embraced the idea, and two days later it announced it would launch satellites during the IGY [Thomas, 1961]. However, this announcement was largely overlooked by the American press and public, possibly because it was believed that the Russians didn’t have the required technology.

In the United States, two teams competed for the privilege to launch the “first” artificial satellite. The army team, headed by von Braun, proposed to use an adapted version of the army’s Jupiter rocket to launch the satellite and indicated it could be ready as early as 1956. The navy team proposed to use a nonmilitary rocket, the *Vanguard*, for the satellite launch. The latter proposal seemed more appropriate for an international program and the navy was selected. At any rate, the U.S. military and the American public were rudely awakened when the Soviet Union launched *Sputnik I* in October 1957. This shock, coupled with the subsequent launch of *Sputnik II* in November 1957, convinced the White House to allow the army to pursue a satellite program in parallel with the navy’s *Vanguard* program. Nevertheless, the navy was still given the first try, and only after the navy’s December 1957 launch attempt failed, in a fireball on the launch pad, did the army get the go-ahead. The following month, on January 21, 1958, the “first American” satellite, *Explorer I*, was placed into orbit by von Braun’s army team.

Explorer I carried a small Geiger counter supplied by James Van Allen of the University of Iowa. The instrument was supposed to record the presence of cosmic rays, which are very fast particles from deep space, but surprisingly the instrument showed no response when the satellite was at high altitudes. There seemed to be no logical explanation, but a second instrument flown two months later confirmed the result. Carl MacIlwain, a graduate student working with Van Allen, solved the problem. He suggested that the satellite encountered a region of very intense energetic particles, which saturated the Geiger tube and caused the counting circuits to read zero. Thus, the Van Allen Radiation Belts had been discovered [Van Allen, 1959].

The large international cooperative efforts, the vast amount of geophysical data collected, and the launch of artificial satellites during the IGY led to the birth of solar-terrestrial physics. The subsequent major infusion of money into this area by several countries led to a rapid advance in our knowledge of the earth’s environment. In less than thirty years, the field has already entered its third phase. The first phase of *exploration* was a time when every measurement yielded new and exciting results. The second phase of *understanding* was a time

when detailed measurements were available and theoretical models were able to reproduce the observed features. The current phase of *perturbations* is a time when active experiments are performed in the various regions of the solar-terrestrial system so that the response of the system can be studied.

Solar-Terrestrial Physics as a Discipline

At the present time, the solar-terrestrial system is composed of the following five major regions:

1. Sun
2. Interplanetary Medium
3. Magnetosphere
4. Ionosphere
5. Upper Atmosphere

During the last thirty years, not only has our knowledge of each region increased markedly, but more importantly, we have learned that the different regions are strongly coupled and a complete understanding of the behavior of an individual region cannot be achieved without studying the system as a whole. In the discussion that follows, I will show some of the exciting discoveries that have been made in solar-terrestrial physics.

The Sun

The sun is a star of average mass and luminosity whose remarkably steady output of radiation over several billion years has allowed life to develop on earth. The solar energy is generated from nuclear reactions in a very hot core, which is about 16 million degrees (see Figure 2). This energy is first transmitted through the radiative zone and then the convective zone (outer 200,000 kilometers; solar radius is 700,000 kilometers). The 27-day solar rotation combined with convection produce a dynamo action where intense electric currents and magnetic fields are generated; these fields and currents vary with a 22-year cycle. The magnetic field that is generated is far from being uniform. Stormy, active magnetic fields lead to the formation of *sunspots*, which last from several hours to several months. The stormy magnetic fields choke the flow of energy from below, and consequently, sunspots are cooler than the surrounding area, which accounts for their dark appearance since cooler regions emit less electromagnetic radiation. Sometimes there are powerful explosions in the atmosphere above sunspots, which are called *solar flares*. These bright flashes of light last only a few minutes

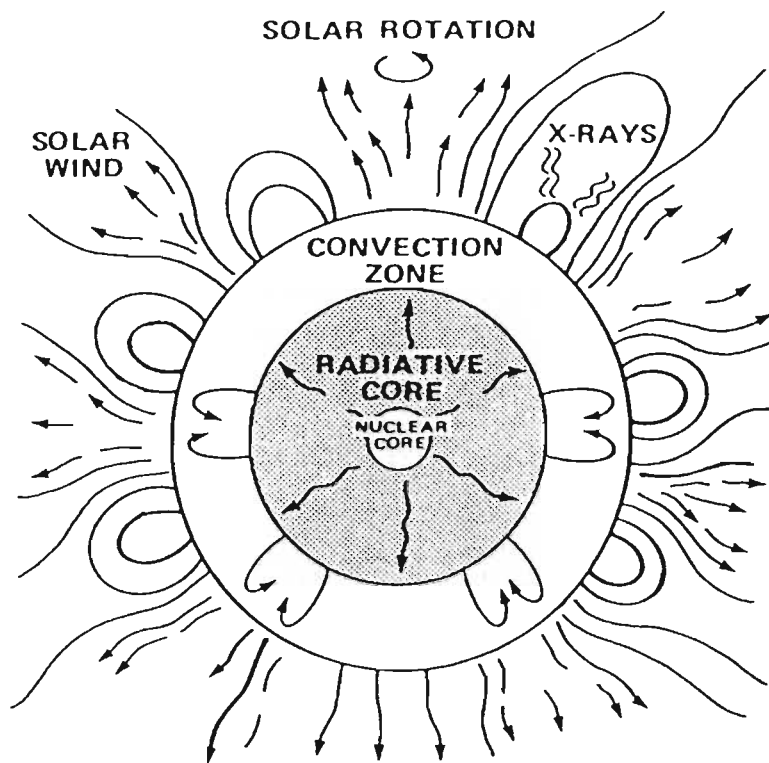


Figure 2. Schematic diagram of the sun's interior region showing the nuclear core, radiative core, and convection zone.

to a few hours, but the explosions send bursts of energetic particles into space that engulf the earth and produce terrestrial magnetic storms, intense auroral displays, and communication disruptions. Another kind of solar explosion is a *prominence*. The prominence extends far into the sun's upper atmosphere and follows the loop of a closed magnetic field line, with the ends of the loop rooted in sunspots. The strong, curved magnetic field traps hot plasma, and because of the intense heating, thermal conduction fronts can race through the loops, raising the temperature to 20 to 30 million degrees.

The sun's atmosphere is composed of three major regions. The *photosphere* is a relatively thin, cool (about 6000 degrees) layer from which the visible radiation is emitted. In the intermediate layer, the *chromosphere*, the temperature is somewhat hotter (about 10,000 degrees), while the upper atmospheric region, the *corona*, is so hot (about 1 million degrees) that it expands continuously into space. This continuous outflow of hot plasma from the corona is called the *solar wind*.

Our knowledge of both the corona and the solar wind increased significantly in the 1970s because of NASA's Skylab mission. On this mission, eight telescopes and cameras were directed at the sun for several months, and the corona was found to be highly structured, as shown schematically in Figure 3. Hot plasma

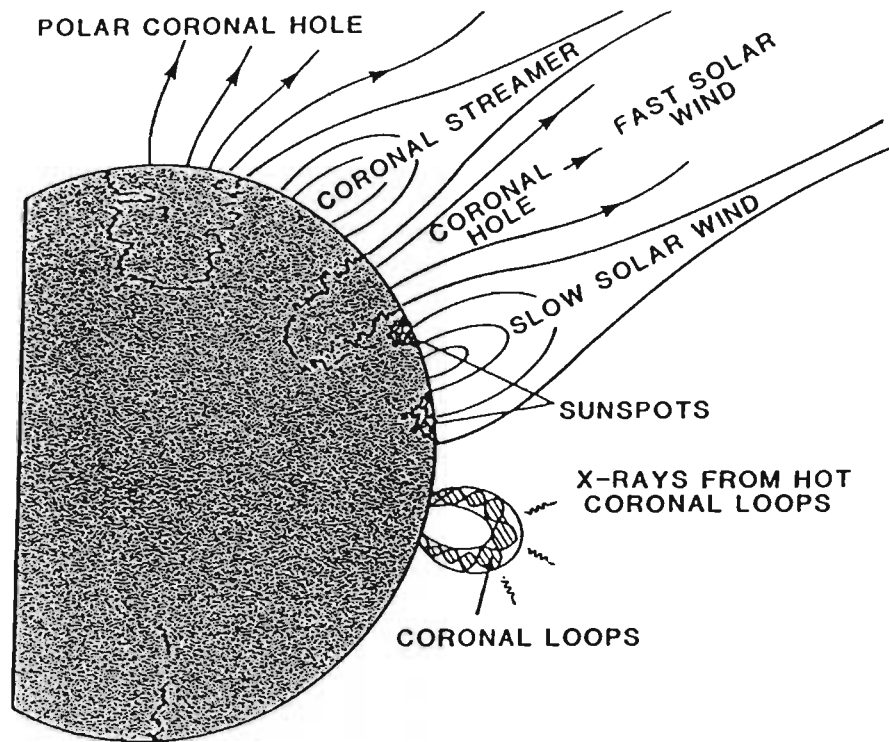


Figure 3. Schematic diagram of some of the recently discovered features in the corona.

can be trapped on strong magnetic field loops, and a very intense x-ray emission is associated with these *coronal loops*. Depending on the strength of the magnetic field, some hot plasma can slowly escape from these loops, forming *coronal streamers* that extend into space. These streamers are the source of the slow component of the solar wind. However, at other places in the corona, the sun's magnetic field does not loop, but extends in the radial direction. In these regions, the hot plasma can easily escape from the corona, which leads to the high-speed component of the solar wind. As a result of this rapid escape, the plasma densities and associated electromagnetic radiation are low, and consequently, these regions have been named *coronal holes*. Typically, coronal holes are transient features which vary from day to day, but during sunspot minimum conditions, extensive coronal holes can exist at the sun's polar regions.

Interplanetary Medium

Prior to the 1950s it was generally believed that interplanetary space was a vacuum, except for the occasional bursts of energetic particles associated with solar flares. However, it is now known that the solar wind acts as a continuous source of plasma for this region. The solar wind flow starts in the lower corona

and the velocity steadily increases as the plasma moves away from the sun, becoming supersonic at a distance of a few solar radii from the sun. At about the same distance, the rarefied solar wind plasma becomes collision-free and electric currents can flow with very little resistance. As a result, the solar magnetic field, which resembles a dipole close to the sun, becomes “frozen” into the solar wind and is carried with it into interplanetary space. As the magnetic field is drawn outward by the solar wind, the sun’s slow rotation acts to bend the field lines into spirals that extend deep into space (see Figure 4). In three dimensions, the spirals can be described by *Alfvén’s ballerina skirt* model [Alfvén, 1981]. The *skirt* represents a sheet of electrical current, and the magnetic field lines on opposite sides of the skirt have opposite polarity. The polarity of the whole system reverses at the beginning of each new 11-year cycle. Thus, our pre-1950 view of interplanetary space being essentially a vacuum has been drastically revised.

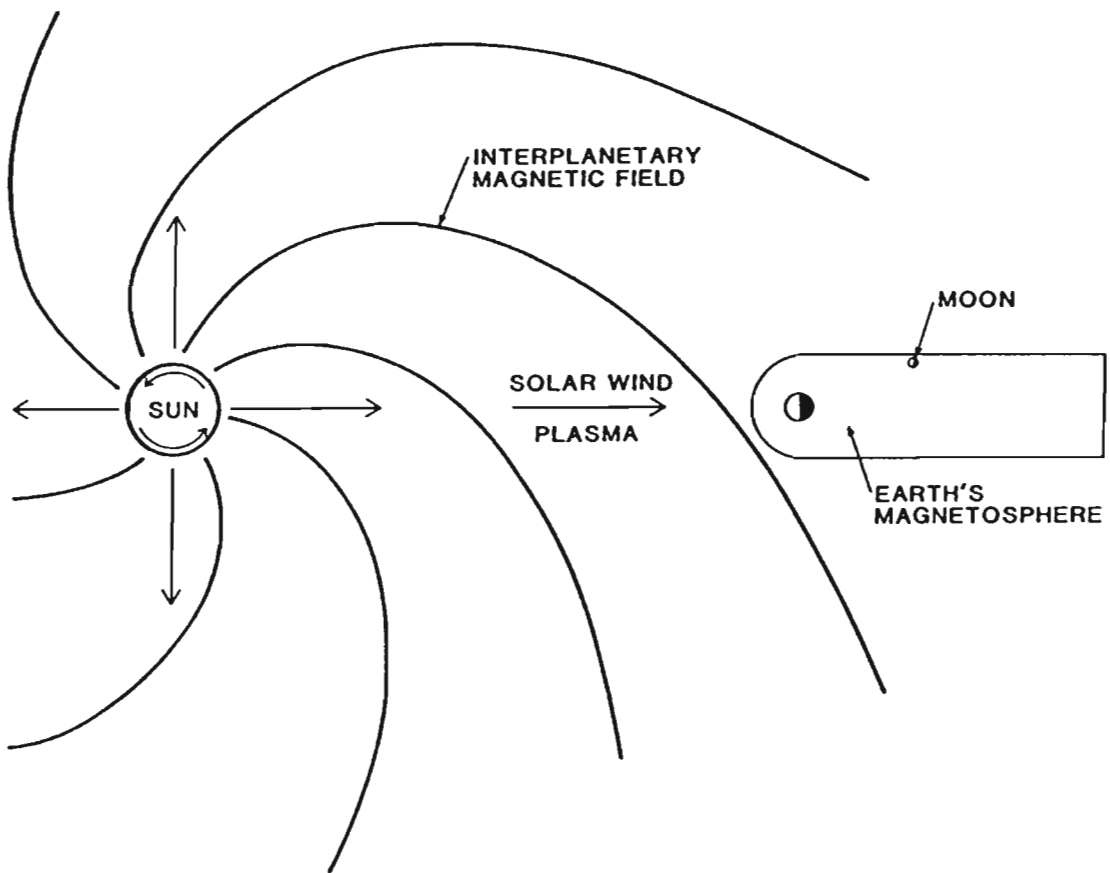


Figure 4. Schematic diagram of the sun-earth system in the plane containing the orbits of the planets.

The solar wind plasma generally takes 2 to 3 days to travel from the sun to the earth and is highly variable on a day-by-day basis. Near the earth, the flow speed ranges from 200 to 900 kilometers per second and the plasma density varies from 1 to 80 particles per cubic centimeter [Brandt, 1970]. As the plasma moves away from the sun, it expands and cools, with the electron temperature decreasing from about one million degrees in the corona to about 100,000 degrees near the earth. The magnetic field strength also decreases with distance from the sun, and near the earth it is of the order of 1 to 10 gamma.

The Magnetosphere

The magnetosphere is the region of near-earth space in which the geomagnetic field plays a dominant role in controlling the physical processes that take place and in which the ionized gas dominates the neutral atmosphere. Our ideas about this vast region have changed markedly since the early 1950s. Prior to that time, it was well known that the earth possessed an intrinsic magnetic field and that at the surface of the earth it had a simple dipole configuration. However, it was generally believed that the dipole form extended deep into space and that except for the ionosphere very little plasma existed in the near-earth environment. After nearly thirty years of satellite measurements, we now know that the earth's magnetic field presents an obstacle to the magnetized solar wind and that the solar wind is deflected around the earth, leaving a magnetic cavity that is shaped like a comet head and tail (see Figure 5). The "head" occurs on the

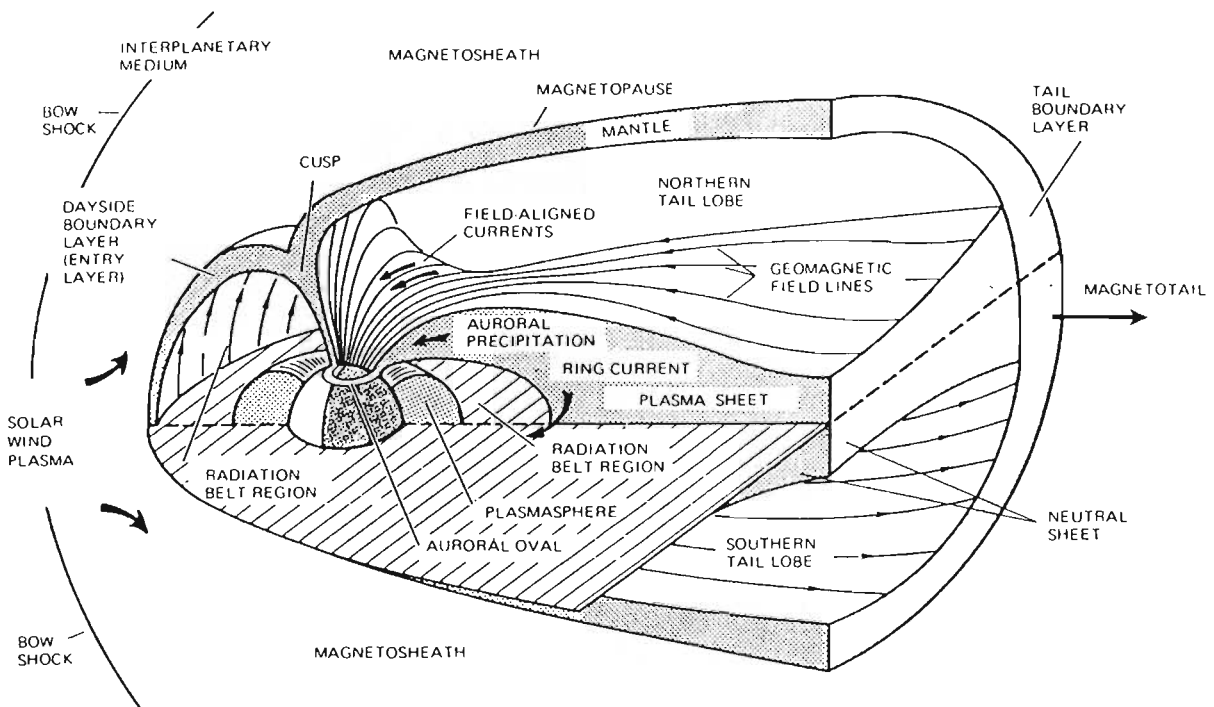


Figure 5. A schematic of the three-dimensional magnetosphere showing the various regions and boundaries. (Courtesy of J. R. Roederer, Geophysical Institute, University of Alaska).

sunward side of the earth where the pressure from the solar wind acts to compress the geomagnetic field, while the solar wind flow past the earth acts to produce an elongated “tail” on the side away from the sun that extends well past the orbit of the moon. This vast region is populated by charged particles of both solar and terrestrial origins, with energies ranging from a few electron volts to hundreds of millions of electron volts.

As the solar wind impinges on the magnetospheric boundary, the solar magnetic field embedded in the flow becomes connected to the geomagnetic field at high latitudes. The subsequent solar wind flow across these “open” field lines constitutes a gigantic magnetohydrodynamic (MHD) dynamo that can generate a voltage drop across the magnetosphere of up to 100,000 volts, electric currents of 10^7 amps, and 10^{12} watts of power. The idea that such a dynamo should exist was first suggested in 1950 by Hannes Alfvén, who also suggested that a discharge could occur through the upper atmosphere and cause the aurora [Alfvén, 1950]. In 1970, Hannes Alfvén was awarded the Nobel Prize in Physics for his work in magnetohydrodynamics and auroral physics, thus becoming the first scientist to receive a Nobel Prize in the solar-terrestrial physics area.

The solar wind is responsible for the formation of well-defined regions and boundaries in the magnetosphere (Figure 5). As the “supersonic” solar wind approaches the earth, the earth’s magnetic field acts as an obstacle to the flow and a free-standing shock wave, called a *bow shock*, is formed [Vasyliunas, 1983]. The shock location is determined by a balance between the solar wind dynamic pressure and the magnetic pressure of the compressed geomagnetic field. As the solar wind passes through the bow shock, it is decelerated, heated, and deflected around the earth in a region called the *magnetosheath*. The bow shock is unique in that it is a “collisionless” shock; the deceleration and heating are caused by particle “collisions” with oscillating electric fields, in contrast to shocks around supersonic aircraft in which particle-particle collisions provide the dissipation mechanism.

The boundary layer that separates the “shocked” solar wind plasma in the magnetosheath from that in the confined geomagnetic field region is called the *magnetopause*. Although the bulk of the solar wind is deflected around the earth, some of it crosses the magnetopause and enters the magnetosphere. Direct entry of solar wind plasma occurs on the dayside in the cusp/mantle region. The solar wind particles entering the cusp can travel along geomagnetic field lines and deposit their energy in the upper atmosphere. Solar wind particles also get into the tail of the magnetosphere by some, as yet, unknown mechanism. These particles populate a region known as the *plasma sheet*. They are not trapped, but have direct access to the earth’s upper atmosphere on the nightside along certain magnetic field lines. By a mechanism not fully understood, the plasma sheet particles are accelerated, stream along the magnetic field lines toward the earth, collide with the upper atmosphere, and produce the auroral displays. This occurs

in both hemispheres in a region called the *auroral oval*. Solar wind particles also work their way closer to the earth and become trapped on closed geomagnetic field lines. As these “high-energy” particles spiral along the magnetic field toward the earth, they encounter an increasing magnetic field strength, are reflected, and then bounce back and forth between the northern and southern hemispheres. These are the Van Allen Radiation Belt particles.

Plasma produced in the earth’s upper atmosphere by photoionization also populates this general region. The plasma is produced in the ionosphere at altitudes between 200 to 1000 kilometers, and then slowly drifts upward to high altitudes along geomagnetic field lines. The region occupied by this relatively cool (1000 degrees) high-density plasma is called the *plasmasphere*. In three dimensions, this region is shaped like a torus [Schunk and Nagy, 1980].

The size and shape of the magnetosphere are determined by the ram pressure of the solar wind, and as this ram pressure varies from day to day, the entire magnetosphere quivers like a mass of jelly. Also, when the magnetosphere becomes overloaded with energy from the solar wind, a magnetic substorm can be triggered. Typically, most of the magnetic field lines in the tail of the magnetosphere have open ends. During magnetic substorms, however, it appears that the plasma sheet is pinched so thin that some of these open field lines reconnect and form closed loops. During reconnection, magnetic energy is converted into particle energy, and the energized particles stream toward the earth, collide with the upper atmosphere, and produce bright auroras. On the average, magnetic substorms occur 4 times a day.

The Ionosphere

The ionosphere is the ionized portion of the upper atmosphere. Its behavior is primarily driven by the diurnal cycle of solar electromagnetic radiation, especially in the ultraviolet and x-ray bands of the spectrum, and at high latitudes by the dissipation of energy from the magnetosphere [Schunk, 1983]. At middle and low latitudes, the principal source of ionization is photoionization of the neutral constituents in the upper atmosphere. The ions produced can then undergo chemical reactions with the neutrals, recombine with electrons, or diffuse along magnetic field lines to either higher or lower altitudes. They can also be forced up or down magnetic field lines by large-scale horizontal neutral winds, which tend to blow in great-circle paths away from the heated atmosphere near noon to the coldest part on the nightside. Representative ion density profiles are shown in Figure 6 for the daytime, mid-latitude ionosphere. In the *D*- and *E*-regions, the heavy molecular ions dominate and their densities are determined by a balance between local production and loss processes. In the *F*-region and topside ionosphere, the lighter atomic ions dominate and diffusion along magnetic field lines

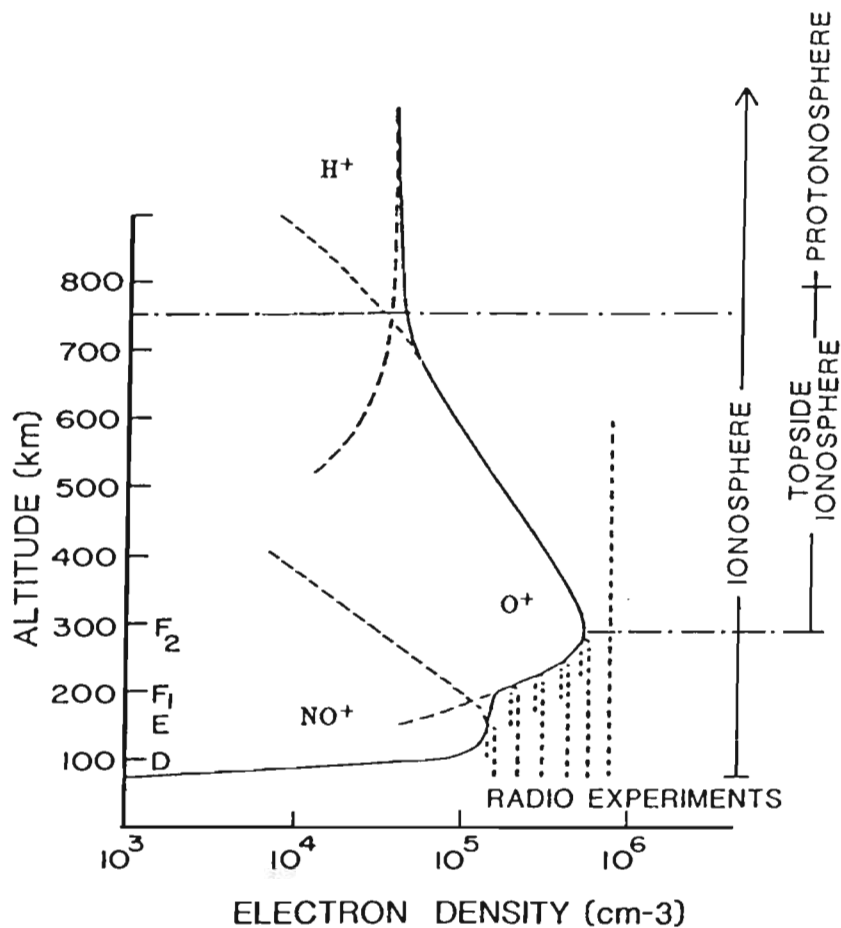


Figure 6. Representative ion density profiles versus altitude for the daytime, mid-latitude ionosphere. [From Schunk, 1983].

is extremely important. The maximum concentration of ionization occurs in the *F*-region, where the peak density is of the order of 10^4 to 10^6 ions per cubic centimeter.

At high latitudes, the ionospheric behavior is more complicated because of magnetospheric effects. Particle precipitation in the auroral ovals acts as a source of both ionization and heat for the ionosphere. Also, the dynamo electric field created by the solar wind-magnetosphere interaction is impressed on the **high-latitude ionosphere**, causing it to undergo a continual horizontal motion **with** speeds as high as 2 kilometers per second. Because of these processes, the ionosphere can be highly structured, and both large-scale patches of ionization and small-scale blobs can easily form. In addition, deep ionization troughs and holes can form, depending on the temporal history of the ionosphere and on the seasonal and solar cycle conditions.

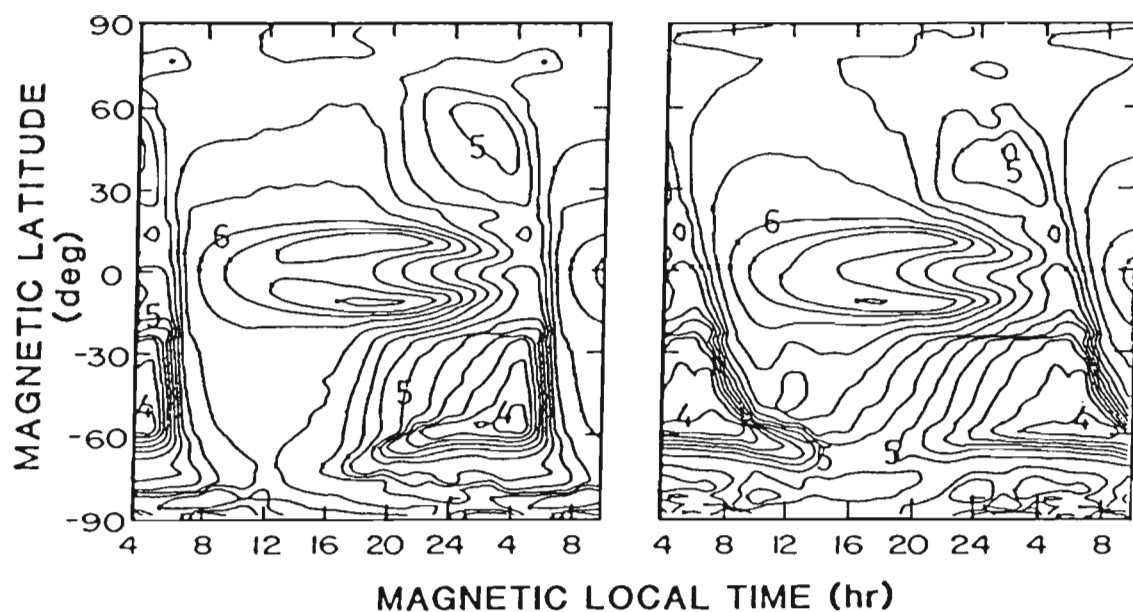


Figure 7. Contours of the peak electron density in the F -region as a function of magnetic latitude and magnetic local time at 2 universal times. The contours are labeled in \log_{10} of the peak density and the contour interval is 0.2. [From Sojka and Schunk, 1985].

At a given instant of time, the “peak” electron density in the F -region can vary by a factor of 1000 over the globe. This is shown in Figure 7, where contours of the peak electron density are plotted versus latitude and longitude at 2 universal times for solar maximum, for June solstice, and for quiet geomagnetic activity conditions [Sojka and Schunk, 1985]. These “snapshots” of the ionosphere indicate that the densities in the summer (northern) hemisphere are generally higher and more uniform than those in the winter (southern) hemisphere. Note that in the winter hemisphere a deep-ionization trough forms at mid-latitudes and that it nearly circles the globe at 1500 universal time. Also note the presence of ionization peaks on both sides of the magnetic equator. These peaks, which are a result of dynamo electric fields generated by atmospheric motion, form at approximately local noon and last until approximately four in the morning.

The Upper Atmosphere

The atmosphere is mixed and the composition of the major constituents is essentially constant up to about 90 kilometers. Above this altitude, diffusion processes are sufficiently strong for a gravitational separation to occur. In addition, photo-dissociation of the N_2 and O_2 molecules is important, and consequently, copious amounts of O and N atoms are produced. The net effect of these processes is shown in Figure 8, where profiles of the neutral densities are shown as a function of altitude for daytime conditions. The dominance of the

molecular constituents at low altitudes and the atomic constituents at high altitudes is clearly evident. Although the neutral densities decrease exponentially with altitude, the neutrals still outnumber the charged particles by about 1000 to 1 at *F*-region altitudes (300 kilometers).

The dynamics of the upper atmosphere during quiet times are primarily controlled by solar heating on the dayside and by Coriolis and ionospheric drag forces. The atmospheric wind tends to blow horizontally from the subsolar heated region around the earth to the coldest region on the nightside. As the wind develops, Coriolis forces act to deflect the flow and the ionosphere provides resistance to the flow. Typically, the wind speeds in the upper atmosphere range from 100 to 300 meters per second for quiet conditions. For active magnetic conditions, on the other hand, magnetospheric processes can significantly affect the global wind pattern. At high latitudes, the ionospheric motion driven by magnetospheric electric fields can provide both a momentum source and a heat source for the upper atmosphere. The momentum source can lead to horizontal wind speeds that are almost supersonic (800 meters per second), and the heat source can produce an upwelling of sufficient strength to destroy the diffusive separation shown in Figure 8. Such a major change in the structure and circulation of the upper atmosphere, in turn, will significantly affect the ionosphere so that important feedback mechanisms exist.

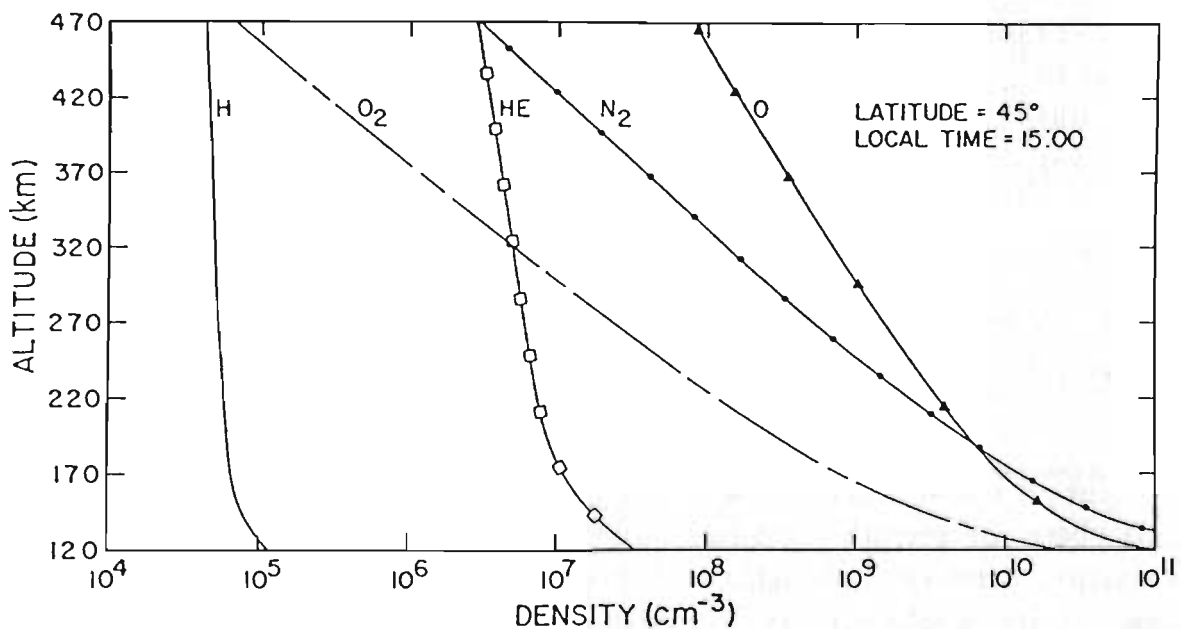


Figure 8. Representative neutral gas density profiles versus altitude for the daytime upper atmosphere at a mid-latitude location. [From Hedin et al., 1977].

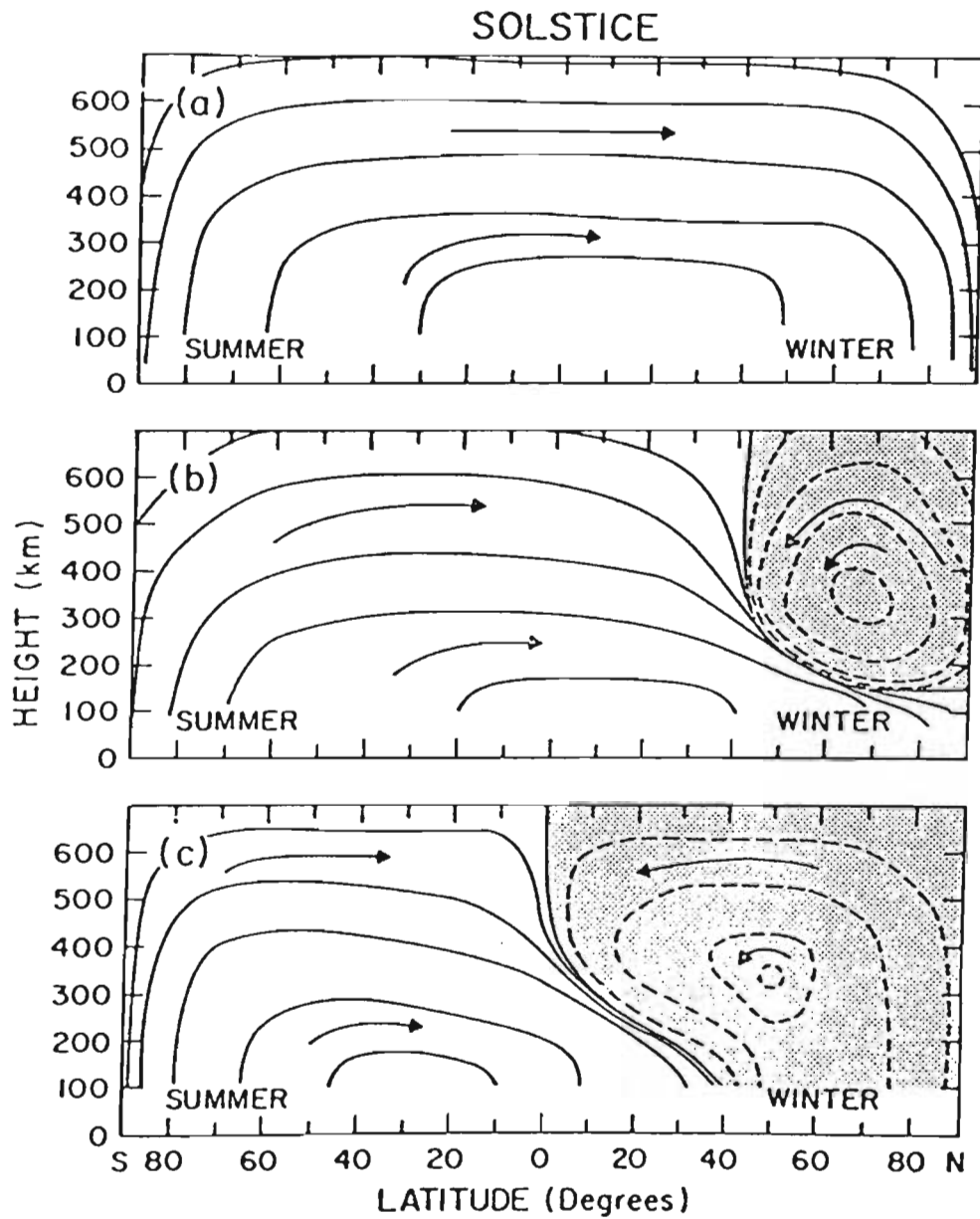


Figure 9. The global circulation of the earth's upper atmosphere for solstice conditions during solar cycle maximum. The terms "quiet," "average," and "storm" refer to the level of geomagnetic activity. [From Roble, 1983].

Figure 9 shows a schematic of the global circulation of the earth's upper atmosphere for solstice conditions during solar cycle maximum. The global circulation is shown as a function of latitude for "quiet," "average," and "storm" levels of geomagnetic activity. For quiet conditions, the upper atmosphere rises in the warm summer hemisphere, flows horizontally across the equator, and then descends in the cool winter hemisphere. For these quiet conditions, the general circulation is primarily controlled by solar heating. However, for higher levels of geomagnetic activity, the energy deposited in the upper atmosphere via magneto-

spheric electric fields and particles is sufficient to disrupt the solar-driven circulation pattern, and for solstice conditions a reverse circulation cell can develop in the winter hemisphere. For storm conditions, this reverse cell can affect the global circulation pattern all the way down to the equator at high altitudes.

Future Direction of Solar-Terrestrial Physics

During the past two decades, nearly every element of the solar-terrestrial system has been sampled with a sufficient frequency so that the local properties of all of the regimes are fairly well characterized. In the future, the direction of the research will be aimed at obtaining a quantitative appreciation of all the significant couplings, time delays, and feedback mechanisms that exist within the system. To accomplish this goal, the solar-terrestrial physics community has convinced NASA, the European Space Agency, and the Institute of Space and Astronautical Science in Japan to undertake a very ambitious program to study the solar-terrestrial system as a whole. The *International Solar-Terrestrial Physics Program* [ISTP] is planned for 1990 and now includes up to six new spacecraft missions, with one or more spacecraft provided by each of the agencies in the international consortium. The overall goal is to place the spacecraft in different regions of the solar-terrestrial system so that the flow of mass, momentum, and energy can be traced from the sun, through interplanetary space, through the magnetosphere, and down into the earth's ionosphere and upper atmosphere. In addition, many ground-based facilities, airborne observatories, balloon and rocket campaigns, and other spaceflight missions will provide data for worldwide collaborative studies. Also, a dedicated data networking system will be established to make data available to the worldwide scientific community within only weeks.

In support of this major experimental initiative, theorists around the world are currently developing time-dependent, three-dimensional models of the different regions of the solar-terrestrial system so that current understanding can be tested against the wealth of data returned by the ISTP program. Initially, the different regions will be studied separately, but as more advanced supercomputers become available in the 1990s, the regional models will be coupled and the system can then be studied as a unit. Ultimately, the theorists would like to obtain a predictive capability so that the day-to-day changes in the solar-terrestrial environment can be forecasted. Currently, however, our ability to forecast "space weather" is very limited, and we are probably where the meteorologists were one hundred years ago with regard to their ability to predict the weather.

In the future, scientists will also be using the solar-terrestrial system as a gigantic laboratory where experiments can be performed without the interference of walls. Much of this kind of research has already begun, and because of

the initial successes it will no doubt continue for at least another decade. Electron and ion beams will be fired into space to create artificial auroras and to study some fundamental plasma instability processes. Also, both neutral and ionized clouds will be released in space so that the effects of man-made perturbations on the natural environment can be studied. Finally, electromagnetic waves will be injected into the solar-terrestrial system at various places in an effort to "trigger" natural auroras and in an effort to understand basic disturbance phenomena.

Societal Impacts

By far, most scientists working on solar-terrestrial physics problems are motivated by the quest for knowledge about the near-earth space environment. However, the research in this field has some extremely important practical aspects. In the coming years, the use of space will increase because of the increasing demand for a wide range of scientific, commercial, and military spacecraft. Numerous communications and surveillance satellites already circle the globe, and within the next decade NASA plans to launch a permanent space station. These space vehicles must function continuously in the near-earth environment, subject to the dynamic variations of the sun, the magnetosphere, the ionosphere, and the upper atmosphere. Such variations may have a significant impact not only on the lifetime of the spacecraft, but also on the performance of the electronic parts and the power systems.

Perhaps the most dramatic example of the effect of changing solar-terrestrial conditions on the lifetime of a spacecraft is the recent demise of Skylab. In 1974, Skylab was placed in what was thought to be a safe parking-orbit and the astronauts returned to earth. The plan was to revisit the spacecraft on an early space shuttle flight and then push it to a higher orbit for safekeeping until it could be refurbished. Unfortunately, the space shuttle was delayed and the coming months saw a rapid increase in solar activity. With the increase in solar activity came a denser atmosphere at Skylab's altitude, which enhanced the drag on Skylab and caused the orbit to decay much faster than anticipated. The news media covered the event with extreme interest because parts of the spacecraft might survive reentry and hit populated regions. With a better knowledge of solar-terrestrial physics, we might have been able to predict the time and place where Skylab would plunge to the earth.

Spacecraft charging is another aspect of the effect of the solar-terrestrial environment on our presence in space. Typically, a spacecraft will charge up to a potential such that the net current to the spacecraft is zero. Since electrons are more mobile than the heavier ions, they tend to collect on the spacecraft first, giving rise to a negative spacecraft potential. The potential rises until the ions are

attracted to the spacecraft at the same rate that the electrons collect on it. Basically, the maximum potential is determined by the electron energy and the ion abundance. In the ionosphere, where most manned flights occur, the electron temperature is low and the ambient plasma densities are sufficiently high so that ion attraction to the spacecraft is easy, and consequently, spacecraft charging is generally not a problem. However, most communications and surveillance satellites are placed at geosynchronous altitudes (22,000 miles), where they can encounter very energetic electrons and where ion collection is difficult owing to low-plasma densities. As a result, these satellites can charge to very high potentials, and potential differences can be set up between different parts of the spacecraft, which in turn can lead to spurious pulses that can interfere with the spacecraft electronics and communications, and can even result in the damage of components by arcing or sparking. The possible consequences of this were dramatically demonstrated several years ago when one of our surveillance satellites suddenly went dead. At first, it was thought that the Russians had developed, and successfully tested, a very advanced antisatellite weapon. Since the security of the United States is heavily dependent upon surveillance systems and since the United States had no comparable antisatellite weapon, the loss of this satellite caused great initial concern. However, after the problem was studied by a group of experts, it was concluded that the satellite fell victim to spacecraft charging.

In addition to atmospheric drag and spacecraft charging, the near-earth space environment affects our activities in other ways. For example, communication interruptions frequently occur during magnetic storms. Also, in the early 1970s, when there was a gasoline shortage, NASA proposed to place large solar energy collectors in space with the hope that the collected energy could be beamed to earth via microwaves. However, one of the problems with this concept was that a significant fraction of the energy in the beam would be absorbed in the lower ionosphere. Another problem was that many shuttle flights would be needed to build the solar collectors and the resulting outgassing from these flights would cause a significant pollution problem. Although this energy source is no longer under consideration, the same concerns about pollution have been expressed with regard to the construction of NASA's planned space station. Although many additional examples can be cited, it is clear that a study of solar-terrestrial phenomena is important not only for satisfying our intellectual curiosity, but for understanding our effect on the environment and the limitations imposed on our activities by the environment.

REFERENCES

- Alfvén, H., *Cosmological Electrodynamics*, Oxford University Press, 1950.
- Alfvén, H., *Cosmic Plasma*, D. Reidel Publishing Company, 1981.
- Angström, A. J., Spectrum des Nordlichts, *Ann. Phys.*, 137, 161, 1869.
- Appleton, E. V. and Barnett, M. A. F., Local reflection of wireless waves from the upper atmosphere, *Nature*, 115, 333-334, 1925.
- Brandt, J. C., *Introduction to the Solar Wind*, W. H. Freeman and Company, 1970.
- Breit, G. and Tuve, M. A., A radio method of estimating the height of the conducting layer, *Nature*, 116, 357, 1925.
- Cavendish, H., On the height of the luminous arch which was seen on February 23, 1784, *Phil. Trans. Roy. Soc.*, 80, 101, 1790.
- Eather, R. H., *Majestic Lights*, American Geophysical Union, 1980.
- Eddy, J. A., The Maunder minimum, *Science*, 192, 1189, 1976.
- Emme, E. M., *Aeronautics and Astronautics, 1915-1960*, U.S. Government Printing Office, 1961.
- Euler, L., Recherches physiques sur la cause des queues des comètes de la lumière boreale et de la lumière zodiacale, *Hist. Acad. Roy. Sci. Belles Lett. Berlin*, 2, 117, 1746.
- Franklin, B., *Political, Miscellaneous and Philosophical Pieces*, ed. by Vaughan, 504, Johnson, 1779.
- Hedin, A. E., et al., A global thermospheric model based on mass spectrometer and incoherent scatter data, MSIS, *J. Geophys. Res.*, 82, 2139, 1977.
- Maunder, E. W., The prolonged sunspot minimum, 1645-1715, *MNRAS*, 50, 251, 1890.
- Nicolet, M., Historical aspects of the IGY, in *History of Geophysics*, ed. by C. S. Gillmor, American Geophysical Union, 44, 1984.
- Ratcliffe, J. A., *Proc. Inst. Elec. Eng. (London)*, 114, 1-8, 1967.
- Roble, R. G., Dynamics of the earth's thermosphere, *Rev. Geophys. Space Phys.*, 21, 217, 1983.
- Sabine, E., *Phil. Trans. Roy. Soc.* 142, 103, 1852.
- Schunk, R. W. and A. F. Nagy, Ionospheres of the terrestrial planets, *Rev. Geophys. Space Phys.*, 18, 813-852, 1980.
- Schunk, R. W., The terrestrial ionosphere, in *Solar-Terrestrial Physics*, ed. by R. L. Carovillano and J. M. Forbes, 609-676, D. Reidel Publishing Co., 1983.
- Schwabe, S. H., *Astron. Nachr.*, 20, 495, 1843.

- Sojka, J. J. and R. W. Schunk, A theoretical study of the global *F* region for June solstice, solar maximum, and low magnetic activity, *J. Geophys. Res.*, 90, 5285-5298, 1985.
- Stewart, B., Aurora borealis, in *Encyclopaedia Britannica*, 9th Ed., 36, 1882.
- Thomas, S., *Men in Space*, Vol. 2, 218, Chilton Press, 1961.
- Van Allen, J. A., Radiation belts around the earth, *Sci. Amer.*, 200, 39, 1959.
- Van Allen, J. A., Genesis of the International Geophysical Year, in *History of Geophysics*, ed. by C. S. Gillmor, American Geophysical Union, 49, 1984.
- Vasyliunas, V. M., Large-scale morphology of the magnetosphere, in *Solar-Terrestrial Physics*, ed. by R. L. Carovillano and J. M. Forbes, 243-254, D. Reidel Publishing Co., 1983.
- Villard, O. G., The ionospheric sounder and its place in the history of radio science, *Radio Sci.*, 11, 847-860, 1976.
- Waldmeier, M., *The Sunspot-Activity in the Years 1610-1960*, Schulthess and Co., 1961.
- Watson-Watt, R. A., *Q. J. Roy. Meteorol. Soc.*, 55, 273-301, 1929.
- Wolf, R., *C. R. H. Acad. Sci.*, 35, 364, 1852.
- Wolf, R., Astronomische mittheilungen, *Astron. Mitt Zürich*, No. 1, 8, 1856.

