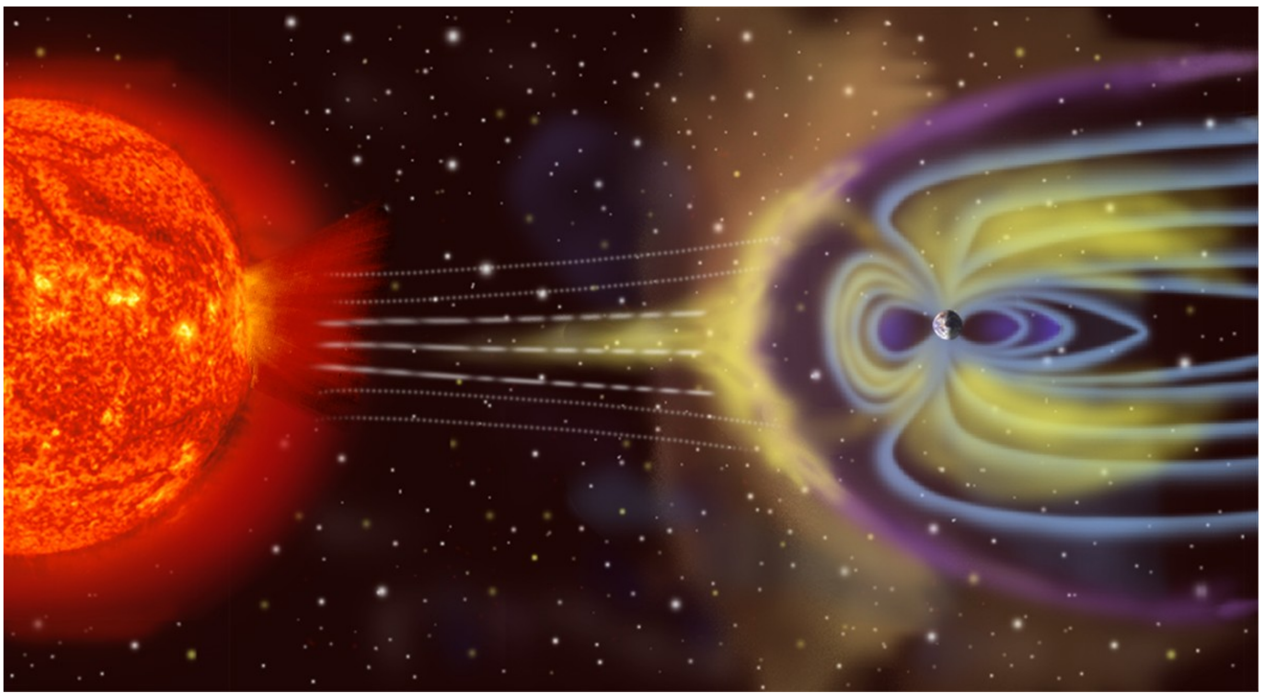
THE MAGNETOSPHERE

C.T. Russell

Institute of Geophysics and Planetary Physics (IGPP) and Department of Earth and Space Science (ESS), University of California, Los Angeles California

In *The Solar Wind and the Earth*, edited by S. -I. Akasofu and Y. Kamide, pp 73-100.



Abstract

The magnetic Field of the earth to a large extent shields it from the continual supersonic outflow of the sun's ionized upper atmosphere. However, some of the mass, momentum and energy of this solar wind gains entry into the magnetosphere, powering current systems, geomagnetic storms and auroral displays. Other planets too have magnetospheres. Some of these magnetospheres are due to intrinsic magnetic fields as in the case of the earth. Others are due solely to the interaction of the planet with the solar wind as in the case of a comet.

Introduction

If there were no solar wind, the physical processes occurring far above the earth in the tenuous upper atmosphere and ionosphere might be well understood and the subject of little interest. However, there is a continual supersonic outflow of ionized gas or plasma from the sun that confines and distorts the terrestrial magnetic field into a cavity with a long tail extending hundreds of earth radii in the antisolar direction. This cavity, called the magnetosphere, is very responsive to changing solar wind conditions. Physical processes occurring therein modulate the energy flow carried by the solar wind to the earth. At times the magnetic cavity acts as a shield deflecting the incident energy; at other times it acts as an accelerator, driven by the solar wind, creating charged particle beams that hit the neutral upper atmosphere, causing it to light up with the brilliant forms of the polar aurora.

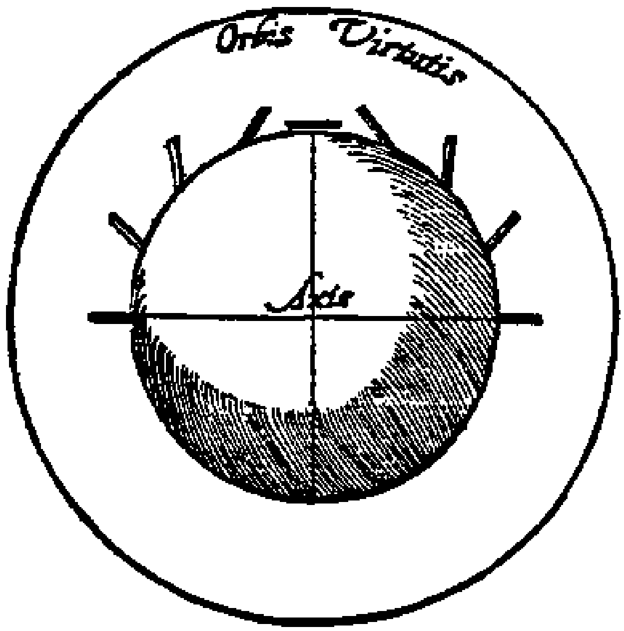


Fig.1

That the earth has a magnetic field has been known for centuries. Gilbert showed as early as 1600, in his treatise *De Magnete*, the predominately dipolar character of the terrestrial magnetic field. As Figure 1 shows, he wisely avoided extrapolating his magnetic field lines into the "Orbis Virtutis," which we now know as the magnetosphere proper, and know to be highly distorted from a dipole at high altitudes. The lower border of the magnetosphere, the ionosphere, was not discovered until 1883 by Balfour Stewart. It was recognized early that the terrestrial magnetic field was not constant. Not only were there secular changes occurring with time scales of centuries, which were presumably due to sources within the earth, but also changes on short time scales of days to hours, which were presumably due to current sources in the upper atmosphere. Although the idea that solar corpuscular radiation, that is, radiation due to ionized solar particles rather than photons, was responsible for this short-term activity in the geomagnetic field and in some way caused the aurora in the high-latitude regions, it was not until 1930 that a satisfactory theory was proposed.

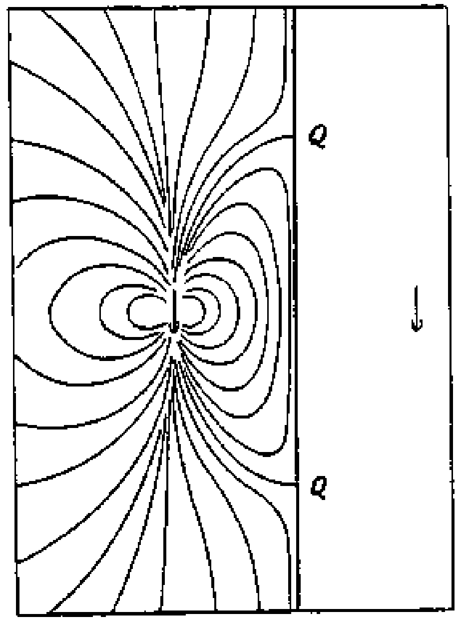


Fig.2

At this time Chapman and Ferraro proposed that a transient outflow of ionized gas from the sun consisting of positive ions and electrons but having no net charge, i.e., a plasma, compressed the terrestrial magnetic field. Figure 2 shows the effect of this plasma, advancing as an infinite planar front against an initially dipolar field. The plasma, being highly conductive, excludes field lines from its interior. This exclusion is mathematically equivalent to the creation of an equal and parallel dipole moment, i.e., the field due to a simple current loop, at a distance behind the front that is equal to the distance the earth's dipole is ahead of the front. If the front of the gas were not planar, as would necessarily occur as the front moved past the earth, the strength of this image dipole and its position would both change. In the case of the real magnetospheric configuration, this image dipole description is only approximate. Nevertheless, image dipole models are still often used today because of their mathematical simplicity. The evolution of this front of ionized gas is envisioned by Chapman and Ferraro, and their idea of leakage of some of the plasma into the magnetosphere is shown in Figure 3. These particles would then become trapped in the earth's dipole field, resulting in a depression of the magnetic field on the surface of the earth. This model explained the major features of the changes of the terrestrial magnetic field known as a geomagnetic storm. Although the major elements of this model are unchanged today, it was not until the advent of spacecraft that the model could be verified or the physical processes that controlled the geomagnetic storm understood.

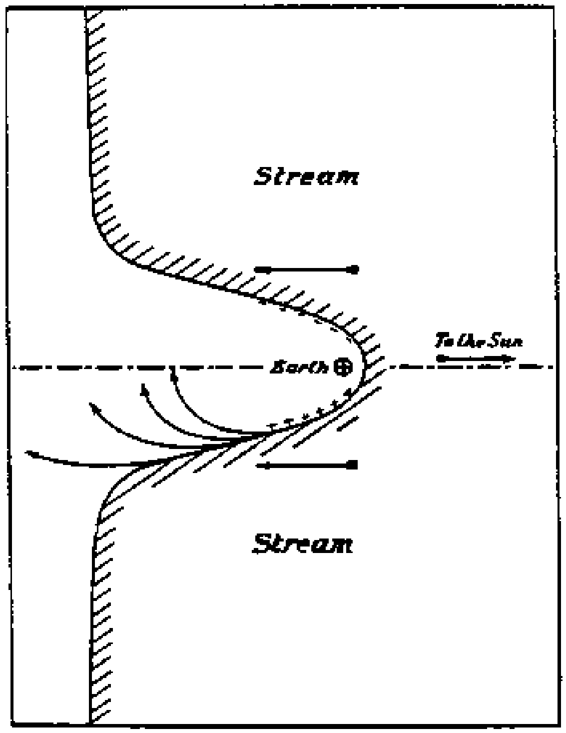


Fig.3

One important change occurred in our understanding of the solar plasma outflow before the space age. This change did not alter our understanding of the geomagnetic storm but was crucial in our understanding of the magnetosphere. In 1951 L. Biermann, as a result of the studies of comet tails, postulated that the solar plasma flowed continually out from the sun, not just in solar bursts. Parker in 1958 showed theoretically how this could occur and coined the name solar wind. Shortly thereafter, spacecraft were flown to altitudes beyond the influence of the earth's magnetic field. The continuous flow of the solar wind was confirmed.

In principle, once we know the properties of the solar wind and the strength of the earth's magnetic dipole moment, we have all the information we need to solve the solar wind-magnetosphere interaction problem. However, the solution of this problem requires an understanding of the physics of plasmas that we did not possess in the late 1950s and early 1960s and to some extent do not even possess today. For example, the size and shape of the magnetopause is dependent not just on the normal stresses exerted by the solar wind dynamic and thermal pressure but also on tangential stress or drag on the cavity. This drag in turn depends on the stability of the boundary between the flowing magnetosheath plasma and the stationary magnetospheric plasma and it depends on the direction of the interplanetary magnetic field. When the external (magnetosheath) and internal (magnetospheric) magnetic fields are parallel, there is minimum drag. When they are antiparallel, there is maximum drag. In short, the system is complex and is controlled by poorly understood processes. Hence the theoretician needs some guidance in his endeavors.

There are three ways in which theoretical efforts can receive this guidance. The first is the more classical approach of observation of the magnetospheric system under a variety of conditions. In fact, the body of this review will be based upon information gathered in this way. Our observations of the magnetosphere are of necessity made on spacecraft. Thus, these observations are single-point measurements in space at any given time. Only under the most ideal circumstances do we have accidental arrangements of satellites providing simultaneous data in key regions of the magnetosphere. One of the goals of the proposed ISTP (International Solar Terrestrial Program) mission is to position satellites in these key areas on a more regular basis.

The second technique is to simulate the solar wind interaction in the laboratory. This technique has the advantage that the whole system can be probed simultaneously. In practice, the simulation is transient but is repeated accurately many times with probes in varying positions so that the global picture is obtained. Of course, this is what we attempt to do with spacecraft observations, but with space data the external conditions cannot be controlled and often are unknown. There is one major drawback to laboratory simulations. One cannot simply make a miniature magnetosphere in the laboratory and expect it to behave like the terrestrial magnetosphere because the various plasma lengths that are important in the solar wind interaction change differently as field strength plasma density, temperature, etc., are altered to fit the solar wind and magnetosphere into a laboratory device. What the laboratory simulationist attempts to do, therefore, is to perform a "limited simulation." In such a simulation, relative scale lengths are maintained but the absolute value of the various "nondimensional parameters" in the problem are not held constant. This limitation has been thought to be serious enough that, perhaps unfairly, laboratory simulations have not played a key role to date in the advance of our understanding of the magnetosphere. Nevertheless, their existence has proven useful and is a reassuring adjunct to the space data. Their utility is demonstrated by Figure 4 which shows a three-dimensional wire model of the magnetospheric magnetic field based on laboratory simulations by Podgorny and co-workers.

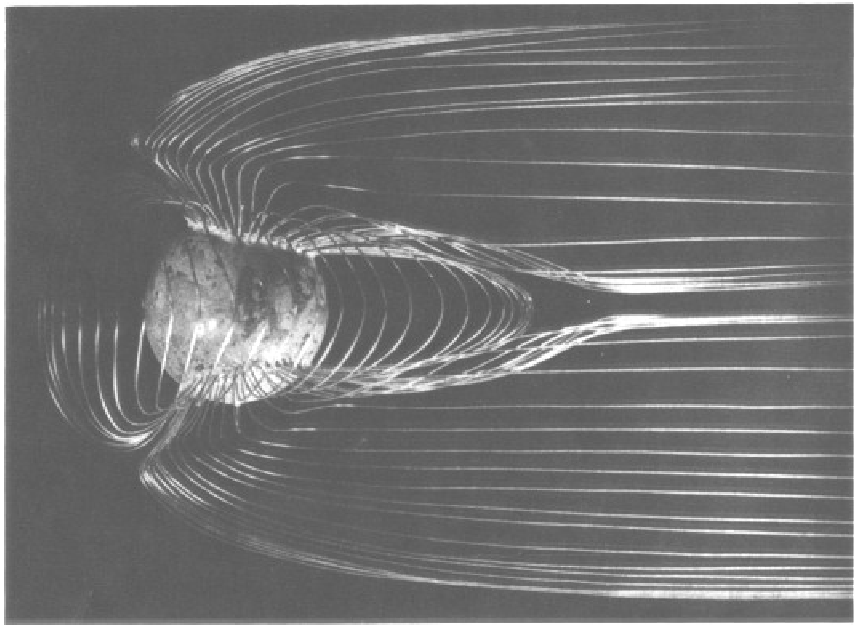


Fig.4

The third technique is computer simulation. In principle, the computer can simulate any plasma condition and any size system. However, in practice, because of the finite size of computer memories and the finite time available on these computers, compromises have to be made. Such compromises include choosing the mass ratio between the electrons and ions to be a number much closer to unity than is in fact true, dividing the system into the minimum number of cells possible to simulate the process of interest, imposing periodic boundary conditions, and restricting the dimensionality of the problem among others. As computer speeds have become greater, sizes larger and costs lower, computer simulations are playing a more important role in magnetospheric physics.

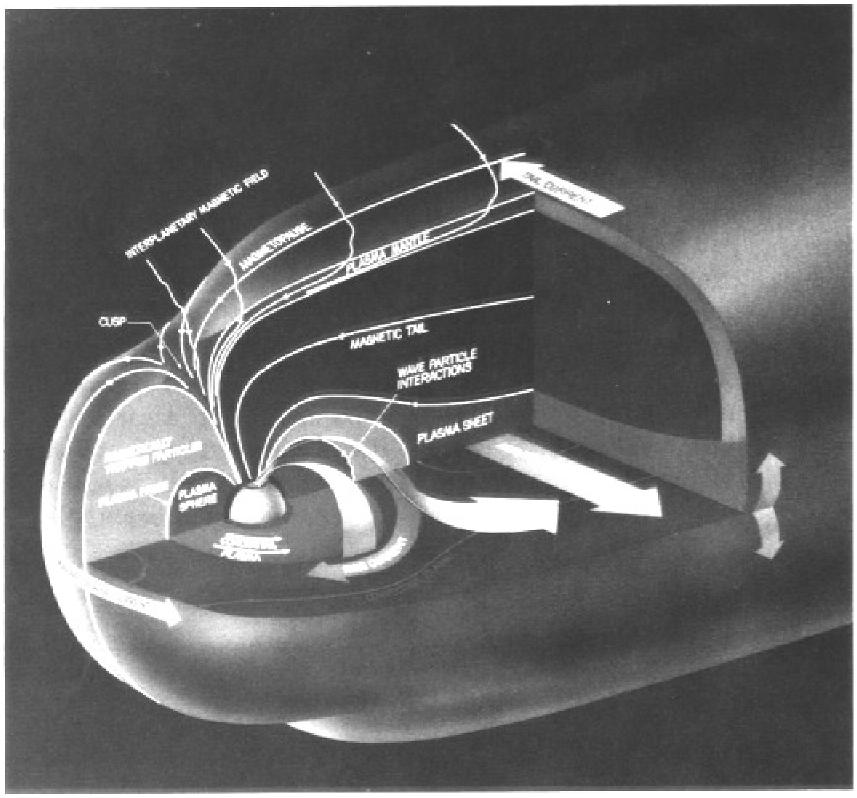


Fig.5

Figure 5 shows a cut-away drawing of the magnetosphere based on knowledge gained from these in situ investigations of spacecraft, the laboratory simulations and the computer simulations. This figure shows the magnetic field, plasma regions and cross-field current systems. Not shown are the currents that flow along magnetic field lines and couple the various regions together.

Bow Shock

The wind of ions and electrons that flows from the sun, the solar wind, travels about 400 km per second, almost 1 1/2 million km per hour. This velocity is faster than a pressure wave can travel upstream toward the sun to deflect the solar wind around the magnetospheric cavity. In order that the solar wind flow be deflected around the magnetosphere, a shock wave forms, much like that in front of a supersonic aircraft. Figure 6 shows the bow shock in front of a noon-midnight meridian cross section of the magnetosphere. Since it is also similar to the bow wave in front of a boat traveling through water, it is called the bow shock wave. This bow shock wave is of great interest because it forms without interparticle collisions in the usual sense. The requisite deflection and heating of the particles that must occur at a shock is accomplished by interactions of the charged particles with magnetic and electric fields which in turn arise because of the motions of these particles. The various processes that occur depend on the conditions in the solar wind. The behavior of the shock is understood under some solar wind conditions, but not all.

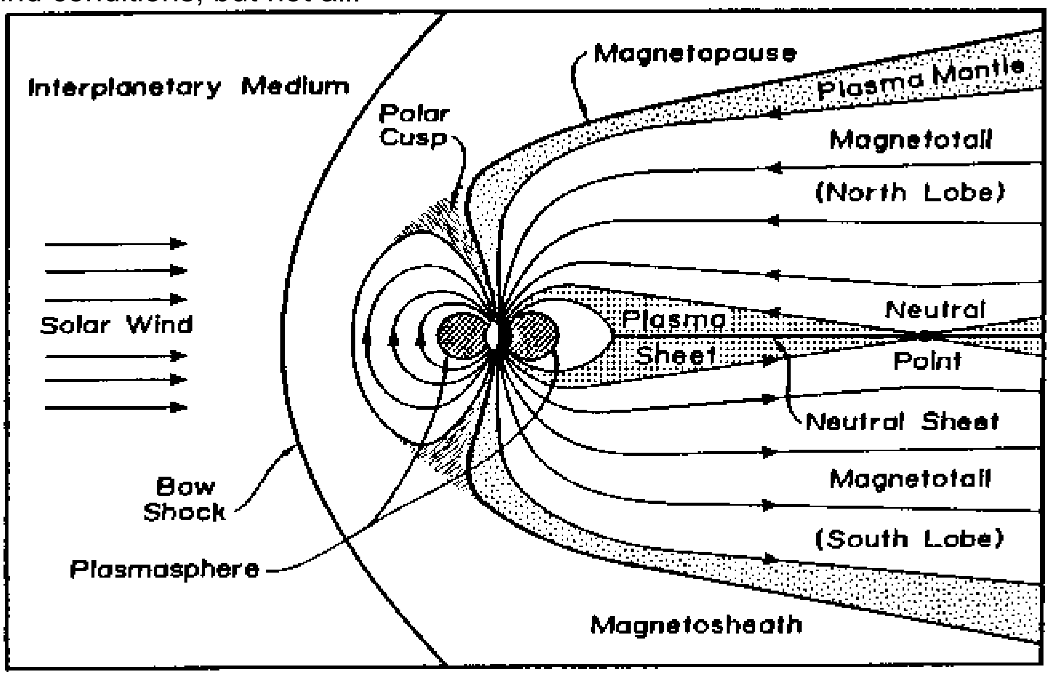


Fig.6

One of the intriguing phenomena of the bow shock is the reflection of a small fraction of the ions back toward the sun. This reflection energizes the particles and can provide much of the required heating at the bow shock. Under certain conditions these reflected ions can move counter to the solar wind flow along the magnetic field. Thus, over some of the region upstream of the bow shock there are beams of ions moving away from the shock. These beams are not in general stable. The beam is disrupted and fluctuations grow larger, feeding on the energy of the beam. These waves, in turn, are swept back to the shock, interact with it and further complicate the structure of the bow shock. The net result of all these processes is a region rich in plasma physical phenomena. The richness of phenomena and the accessibility of this region to investigation, coupled with the importance of these phenomena to both astrophysics and basic plasma physics, as in fusion machines designed to supply future energy needs, keeps it a region of high scientific interest.

The Radiation Belts

The discovery of the earth's radiation belts, also called the Van Allen belts after their discoverer, was at the time surprising; but in retrospect it should not have been a surprise at all. The aurora were thought to be the result of the direct entry of energetic particles into the magnetosphere and the "Stormer" orbits of these energetic particles in a dipole magnetic field were well understood. Particles could gain entry into the high-latitude regions but not the low-latitude regions. Of course, if ionized particles are created in the inaccessible inner regions, it is obvious that they remain trapped there. The ring current associated with geomagnetic storms was also thought to be due to charged particles circulating in the earth's dipole. Perhaps the intensity of the radiation belts and their intimate control (at most energies) by the solar wind was rightly a surprise, but we certainly should have been prepared to observe some trapped particle flux.

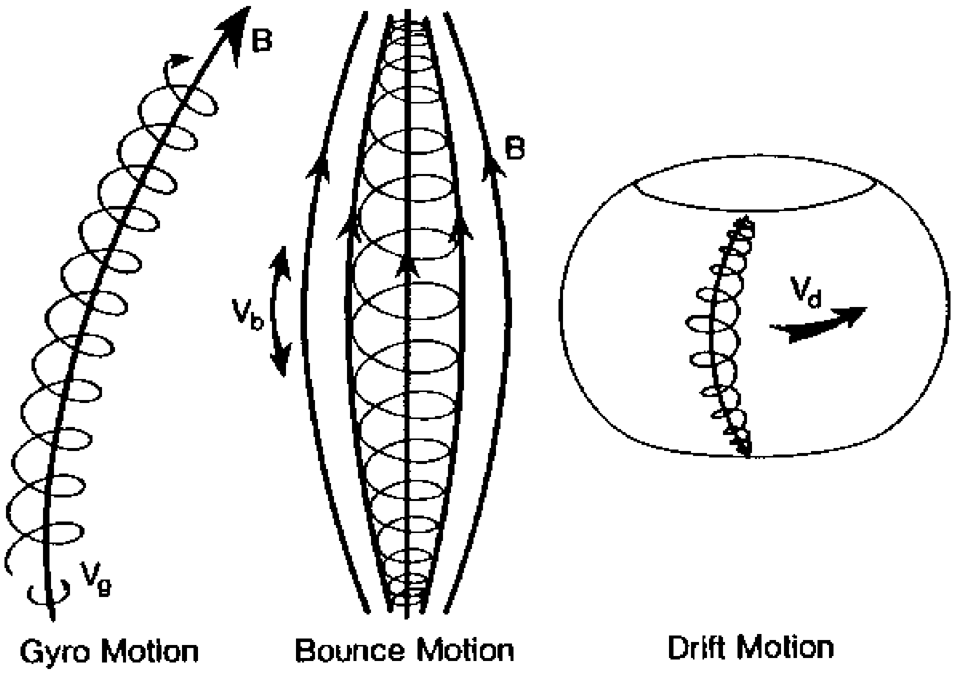


Fig.7

The motion of charged particles in the earth's magnetic field is complicated but is easily visualized by dividing it into three components. As shown in Figure 7, particles gyrate about magnetic field lines with a frequency that is proportional to the field strength. As they do this, they are also bouncing back and forth along the magnetic field while being reflected by the magnetic mirror force associated with the increasing strength of the magnetic field with decreasing altitude. The bounce frequency is determined by how fast the particles are moving along the magnetic field and by the length of the magnetic field line. The distance of a magnetic field line from the center of the earth when it crosses the earth's magnetic equator is called its L-value. Thus, particles at low L-values, on short field lines, bounce more rapidly than particles on long field lines at L-values in the distant magnetosphere. The third component of this motion is the circulation of particles about the dipole axis. While they are gyrating and bouncing, the particles drift. They drift all the way around the magnetic cavity until they are back where they started. This third periodic motion has a frequency determined by the gradients in the magnetic field and curvature of the field lines as well as particle energy. The drift motion usually takes from minutes to hours; the bounce motion seconds to minutes and the gyro motion milliseconds to seconds. Electric fields present in the magnetosphere can significantly distort the motions of particles at the low energies below a few keV. Distortions of the magnetic field are equally important in the distant magnetosphere.

One of the simplest ways of populating the radiation belts and one of the most popular mechanisms after Explorer 1 discovered the belts is the decay of neutrons created by extremely energetic cosmic rays hitting the upper atmosphere. The electrons and protons created in this way might remain trapped for hundreds of years in some regions of the magnetosphere.

If the magnetic or electric field in the magnetosphere is unsteady, the particles will scatter and not retrace exactly the same path as they gyrate about magnetic field lines, bounce along them or drift across them. In particular, if field lines are perturbed at the frequency with which particles drift around the magnetic cavity, they will not return to the same place and will have drifted in or out depending on the relative phase of the perturbation and the particle motion. Even though this process randomly scatters particles in and out, it can lead to a net motion if there are gradients in the particle distribution. If A and B are two adjacent volumes with more particles in A than in B, and if we scatter particles equally in A and B, more particles will move from A to B than from B to A because there are more particles in A to be scattered. So, too, in the magnetosphere particles will be scattered or radially diffuse inward or outward depending on the radial gradients with a rate dependent on the amplitude of oscillations in the magnetic (or electric) field at the drift frequency of the particle. This process is somewhat complicated by the fact that particles gain energy when they diffuse inward and lose energy when they diffuse outward. Thus, keeping proper track of particle energy is necessary in calculating radial gradients in the magnetosphere. Because of the existence of this radial diffusion, particles can get into the magnetosphere from the solar wind. This process appears to be more important at higher energies than at lower energies.

At energies below a few tens of keV, the electric field must be taken into account in describing the motion of charged particles. The fact that this electric field has large temporal variations means that at low energies, especially in the outer magnetosphere, particle trajectories often switch from being closed and totally inside the magnetospheric cavity to being open to direct entry from (or exit to) the solar wind or, equally important, open to direct entry from the magnetotail. Thus, at the lowest energies in the distant magnetosphere the charged particle population is extremely variable.

Finally, we note the existence of the earth's ionosphere which consists of electrons and ions but with energies far below those of the radiation belt particles. If it were possible to supply energy to these particles either by heating or through electric fields along the magnetic field (so-called parallel electric fields), then these cold ionospheric particles could become an important source for the radiation belts. Because it was difficult to imagine how such heating could occur or how parallel electric fields could arise in the collisionless magnetospheric plasma, the ionospheric source was not generally appreciated until quite recently.

Once particles enter the radiation belts, they do not remain trapped there forever. Above we mentioned that particles can diffuse inward toward or outward from the earth, if the electric or magnetic field in the magnetospheric cavity changes with time at the same frequency as that with which the particle drifts around the earth. This resonance is effectively a collision which alters the motion of the particle. Ordinary particle-particle collisions, which in the case of charged particles are called Coulomb collisions, are quite rare in the radiation belts, except at the foot of field lines where particles can hit the atmosphere. One type of particle-particle collision that is thought to be important for protons is charge-exchange in which a fast proton gains an electron from a cold hydrogen atom in the earth's atmosphere. The fast proton becomes a fast neutral particle and escapes from the field line about which it had been gyrating. This does not decrease the number of protons in the radiation belts but does remove energy.

Generally, though, wave-particle resonances provide the collisions for untrapping the radiation belt particles. This is especially true for electrons that cannot charge exchange. Very early in the history of the radiation belts it was noticed that electrons on magnetic field lines that extend beyond 1.5 earth radii decayed from the belts much more quickly than they should have due to particle-particle collisions. It was soon realized that electromagnetic noise generated by lightning (called whistlers) would propagate up through the magnetosphere along field lines and resonate with the gyromotion of these particles causing them to travel more nearly parallel to the magnetic field. Electrons traveling more along the field go further before they are reflected. Hence they are more likely to hit the atmosphere and be lost. Thus, for a while it was thought that the amount of terrestrial lightning controlled the level (or flux) of energetic electrons in the radiation belts. Figure 8 illustrates the phenomenon of lightning-generated whistlers.

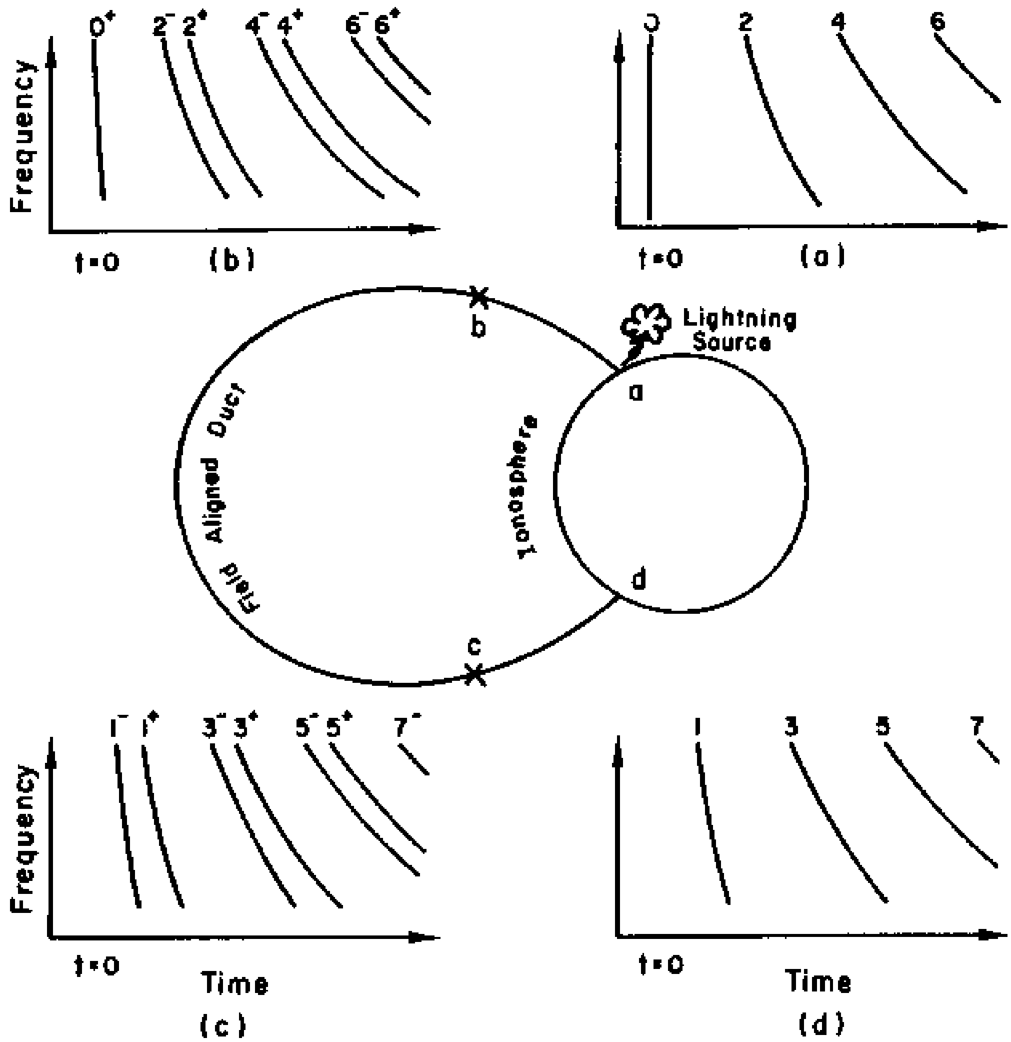


Fig.8

Unfortunately, the frequency spectrum of whistlers was not exactly right for causing the observed electron losses. There had to be another source of collisions. The source turned out to be other electromagnetic waves, not generated in the atmosphere but generated by the particles themselves. These waves had been detected by receivers on the ground for many years. One of the types of emissions was called the dawn chorus, which sounded much like the chirping of birds. Figure 9 and Figure 10 illustrate the frequency-time behavior of some of these signals. This emission and many others were soon detected deep in the magnetosphere when satellites were instrumented to listen for them. We now have even recorded such emissions in the magnetosphere of Jupiter.

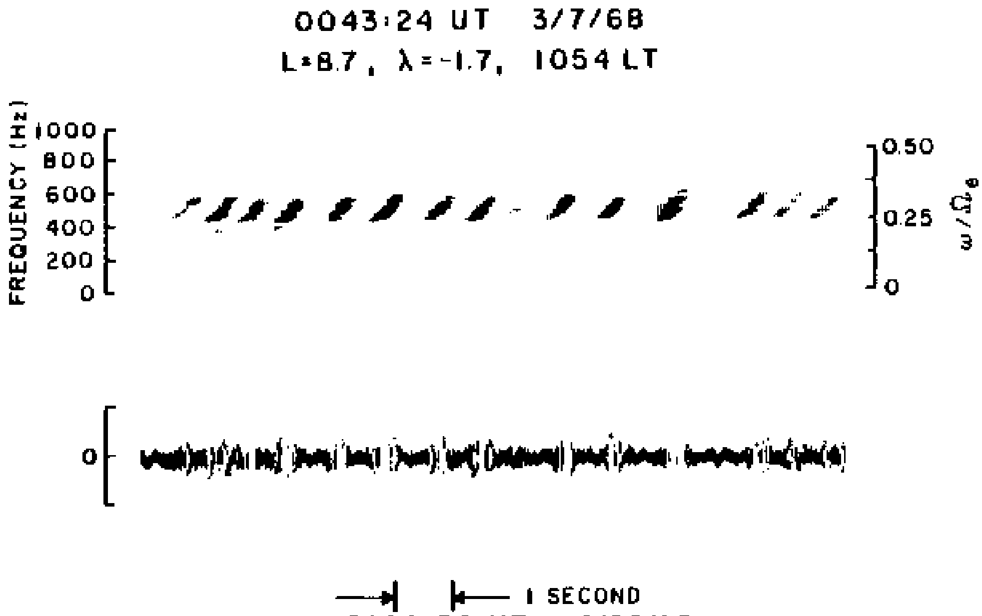


Fig. 9

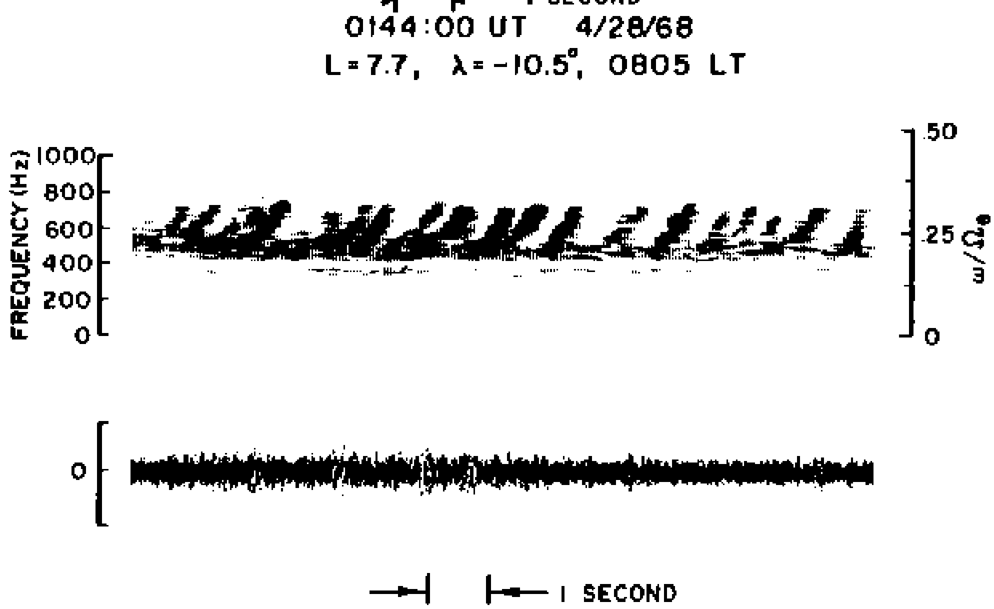


Fig. 10

An examination of the means by which these emissions are generated helps to illustrate how the magnetosphere works to regulate itself. If a righthanded circularly polarized electromagnetic wave propagation along a field line encounters an electron gyrating about the field and moving toward the wave, the electron could experience a field at its period of gyration if the frequency of the wave and the velocity of the electron along the field are just right. Depending on the phase of the interaction, the electron will either become more parallel to the field or more perpendicular. If it becomes more parallel, it loses energy and the wave gains energy; if it becomes more perpendicular, it gains energy and the wave loses energy. Either can happen and does happen just by chance. However, if there are more electrons moving mainly perpendicular to the field than there are moving mainly parallel, then this scattering process will give a net diffusion toward motion along the field. Thus, there will be a net loss of energy by the electrons and a net gain by the wave. Therefore given a specific distribution of electron motions, certain wave types and certain wave frequencies will be amplified. Some of these waves propagate out of the magnetosphere into the atmosphere and are lost, and some reflect back and get amplified again. If the particle fluxes are strong enough, they will amplify the waves enough to overcome the loss due to propagation into the atmosphere. If not, then the waves just decay away and the particles are stably trapped. The flux of charged particles above which this amplification will take place and cause the removal of the particles is called the stable trapping limit and the process is called a wave-particle instability; specifically here, it is called the electron-cyclotron resonance instability, or the whistler loss-cone instability. In general, the magnetosphere attempts to regulate itself by becoming unstable to some process if its parameters exceed some limit. Thus, the magnetosphere attempts to keep itself close to the norm despite quite varied external conditions.

The Plasmasphere

An important source of charged particles for the magnetosphere is the earth's ionosphere, but in general these particles are not nearly so energetic as those particles that come from the solar wind. The ionosphere is created principally by the ultraviolet radiation from the sun being absorbed by the earth's upper atmosphere and thereby being ionized. The temperature of the ionosphere is cool (~2000 K) by magnetospheric standards but hot enough that it could "evaporate" up magnetic field lines and populate the magnetosphere with a moderately dense population of cold electrons and ions, perhaps 100 to 100,000 cm 3 depending on distance. However, this effect is seen only close to the earth.

The lightning-generated whistlers, thought originally to control the fluxes of energetic electrons, provide a useful diagnostic of the density of the magnetospheric plasma because they travel more slowly the denser the plasma becomes. Furthermore, one can use the dispersion of the whistler, that is, its travel time as a function of frequency, to tell where the whistler traveled. Thus, even without spacecraft it was possible to tell that the magnetospheric cavity was filled with cold ionospheric plasma only out to about four earth radii, above that distance the plasma density suddenly dropped.

It was soon postulated that the location of this drop in plasma density or plasmapause separated a region in which the plasma rotated about the earth, the plasmasphere, from a region in which the plasma moved but did not rotate in closed paths around the earth. The reason for this outer region of circulation was the drag on the cavity by the solar wind. If the drag were simple viscosity, then two eddies would be set up at dawn and at dusk. However, if the drag were caused by electrical connection between the solar wind and the magnetosphere, the plasma in this outer region could drift through the magnetopause out into the solar wind and be lost. There was much early opposition to the concept of electrical connection between the solar wind and the magnetosphere. Most preferred to assume the magnetopause was an equipotential surface which flow lines did not cross. However, as we discuss below, the solar wind electric field does couple to the magnetosphere and does control cold particle drift. The electric field that does this is called the convection electric field. The electric field that is associated with the particles which at small radial distances drift around the earth is called the corotation electric field. The plasmapause occurs approximately where the corotation electric field equals the convection electric field. The electrical potential difference applied to the magnetosphere by the solar wind is about 40,000 V.

While it has long been appreciated that the cold plasma in the plasmasphere had its origins in the ionosphere, it has only become recently appreciated that the ionosphere is an important source of the more energetic component of the radiation belts with energies of a few keV. Some process, perhaps electric fields along the magnetic field or intense wave-particle resonances in the ionosphere, is energizing the cold helium and oxygen ions and accelerating them to the equatorial regions. Thus, the chemical composition of much of the magnetosphere is not so much like the composition of the solar wind as originally thought but is more dominated by heavy ions in the lower ionosphere.

Magnetosphere Dynamics

One of the earliest explanations for the source of the aurora was made shortly after the advent of the space era by J. W. Dungey who proposed that the magnetized plasma in the solar wind merged with the magnetized plasma in the terrestrial magnetosphere so that the magnetic field of the earth and the sun became connected, or reconnected, since this process was supposed to start and stop. This reconnection model provided a simple explanation for the convection electric field discussed above, provided a source for the energization of auroral particles, and predicted intimate control of magnetospheric processes by the orientation of the magnetic field in the solar wind, i.e., the interplanetary magnetic field. As shown in Figure 11, this model predicts a neutral point or line on the front of the magnetosphere at which magnetic energy is released to the plasma and one in the magnetotail again at which plasma is energized. While in fact this model was proposed as a steady state model which can be treated simply in terms of electric fields and plasma drifts, it is convenient to think in terms of magnetic field lines moving. Thus, in Dungey's model, field lines in the solar wind of direction opposite to the terrestrial field join with the earth's field when they are convected against it.

This early insight into magnetospheric behavior did not achieve universal acclaim and it is only recently that it has become generally accepted. The model most easily explains the entry and behavior of solar flare particles in the magnetotail and polar caps; the response of geomagnetic activity to the solar wind velocity and the magnitude and direction of the interplanetary magnetic field; the variations in magnetospheric parameters such as the size of the polar cap, position of the polar cusp and field strength in the magnetotail; and the structure of the magnetopause.

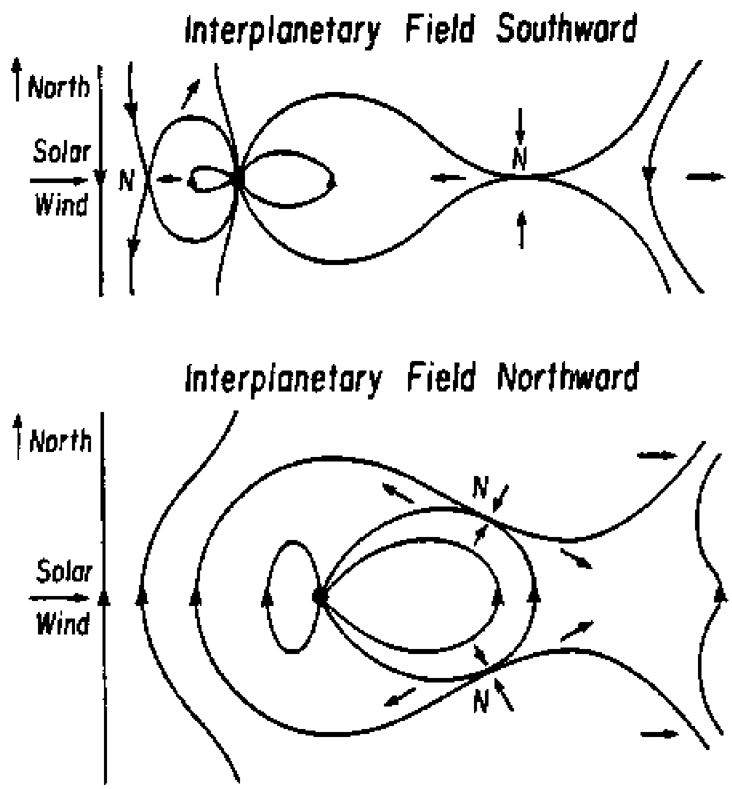


Fig.11

Although the Dungey reconnection model is couched in steady state terms, the real magnetosphere is seldom in a steady state. First, the direction of the interplanetary magnetic field is highly variable and the reconnection rate is extremely sensitive to this direction. When the interplanetary magnetic field is even slightly northward, reconnection seems to cease. Secondly, the magnetosphere has difficulty keeping the rates of all its processes in step. Thus when reconnection is initiated on the dayside, magnetic flux is opened and transported to the tail before more closed flux can convect in from the tail and the magnetopause has to move in toward the earth. The magnetic flux in the tail will then build up until reconnection in the tail starts up to convert the open field lines to closed field lines. If this conversion process is too efficient, more closed flux is supplied than needed by the magnetosphere and the reconnection point or neutral point moves off down the tail. In short, the magnetosphere, because of reconnection, is very dynamic. This cyclic process of flux transfer to the tail and back to the magnetosphere is manifested in the auroral zone as a magnetospheric substorm with brightening and motion of the aurora as well as energization and injection of magnetospheric particles.

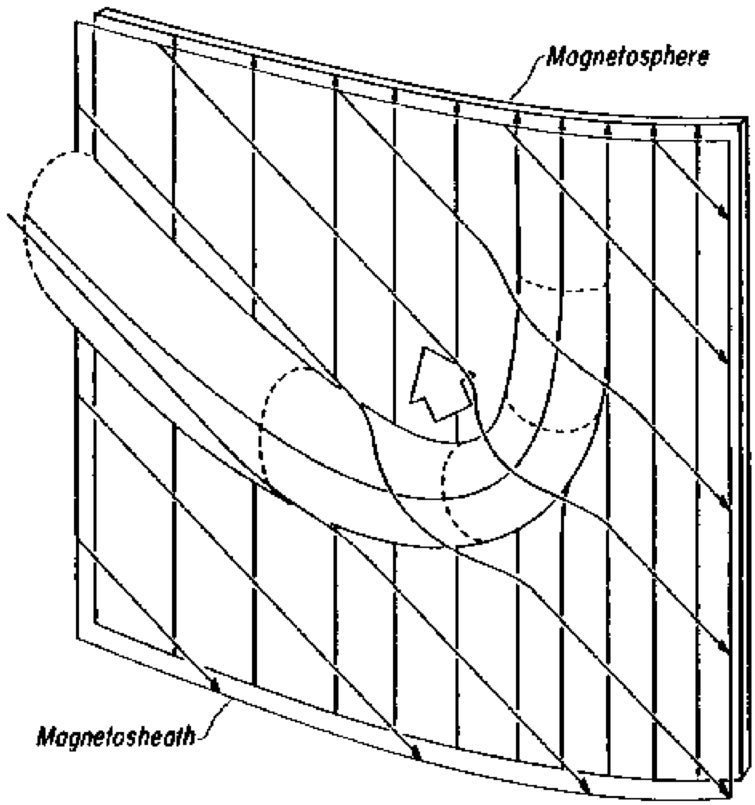


Fig.12

If the magnetosphere is dynamic, the magnetopause is even more so because, not only are the processes operating at the magnetopause sensitive to external conditions, but also the magnetopause moves very rapidly in response to solar wind and magnetospheric pressure changes. These variations plus the extreme thinness of the boundary tax the ability of spacecraft to probe the magnetopause adequately. However, the recent International Sun- Earth Explorer mission (ISEE) has solved the motion problem with the use of two co-orbiting closely spaced satellites and high temporal resolution three- dimensional plasma measurements. For the first time it has been possible to measure the magnetopause boundary's velocity and thickness and resolve the flows associated with the reconnection process. One of the discoveries of the ISEE mission was the Flux Transfer Event whose possible interpretation is shown in Figure 12. Here a tube of magnetic field from the solar wind has become connected to a terrestrial tube of magnetic field lines. The flow of the solar wind past the earth's magnetopause is viscous. The solar wind exerts drag or tangential stress on the magnetopause. The reconnection process on the dayside magnetopause discussed above is one such way; momentum transfer associated with boundary oscillations such as those driven by the Kelvin-Helmholtz ("wind-over-water") instability is another. These stresses must be transmitted to the earth, or at least its ionosphere, otherwise the magnetospheric plasma would soon all be blown away. These stresses are transmitted to the ionosphere via current systems which flow along magnetic field lines and close in the ionosphere across field lines. Electric current across field lines causes a force on the plasma through which it flows, and thus the ionospheric plasma is dragged by magnetospheric stresses against the resistive force of the collisional neutral atmosphere. One of the most evident effects of this drag is a flow of the ionospheric plasma from noon to midnight over the polar cap. We note that stresses exerted by internal magnetospheric processes behave in the same way and have similar effects on the ionospheric plasma. Thus, explosive reconnection in the magnetotail can push the ionosphere plasma toward the dayside at auroral and subauroral latitudes just as the solar wind pulls plasma across the polar cap.

While the magnetic field lines in the magnetosphere are very good conductors for the electric currents flowing parallel to them, they are not perfect conductors. At present, much controversy ensues as to what is the reason for the breakdown in conductivity. Are the plasma waves present due to the various magnetospheric instabilities sufficient to cause the resistivity or do the currents become resistive because the current density exceeds the threshold for wave generation itself? If the currents become unstable, does the potential drop along the field extend for large distances (thousands of kilometers) or is it confined to a narrow layer (termed a double layer) that may be less than a kilometer thick? Despite not knowing the answers to these questions, it is obvious that potential drops do occur along field lines in these currents and that these potential drops have serious consequences for the magnetosphere. In particular, it is believed that auroral arcs are caused by beams of electrons accelerated to a few keV in energy by such potential drops. Potential drops which accelerate electrons down can accelerate ions up the field lines and help populate the radiation belts, at least at lowest energies.

As stated above, the reconnection process is sensitively dependent on the orientation of the magnetic field in the solar wind which in turn is quite variable. This variability leads to varying magnetospheric stresses and varying rates of transfer of magnetic field lines from one part of the system to another. The magnetosphere appears to get out of balance in this process, and reach a metastable state, one from which it is ready to fall. This fall is called a substorm and is probably triggered by the onset of explosive reconnection in the near magnetotail. The substorm is associated with particle energization, additional stresses and accompanying magnetic field-aligned electric currents on the nightside, the brightening and motion of auroral forms, and a host of other phenomena that could form the basis for an entire review article themselves. Suffice it to say that substorms are one of the most fundamental phenomena of the terrestrial magnetosphere and one of the most controversial.

Other Magnetospheres

All the planets that have been explored to date have magnetospheres. In the case of Mercury, Jupiter and Saturn, the magnetosphere is due to the solar wind interacting with a magnetic field arising from sources internal to the planet. In the case of Venus and probably Mars, the magnetic field creating the magnetosphere originates in the solar wind. Very few observations exist for satellite magnetospheres. The moon's surface is magnetized but the overall lunar magnetic moment is insufficient to reflect the solar wind and it hits the surface and is absorbed. Thus, a wake is formed behind the moon which is a solar wind void, but this region has little resemblance to a magnetosphere. Io, the innermost Galilean satellite of Jupiter, may have its own magnetic field but the present evidence is incomplete. Thus we will restrict our discussion to the planets.

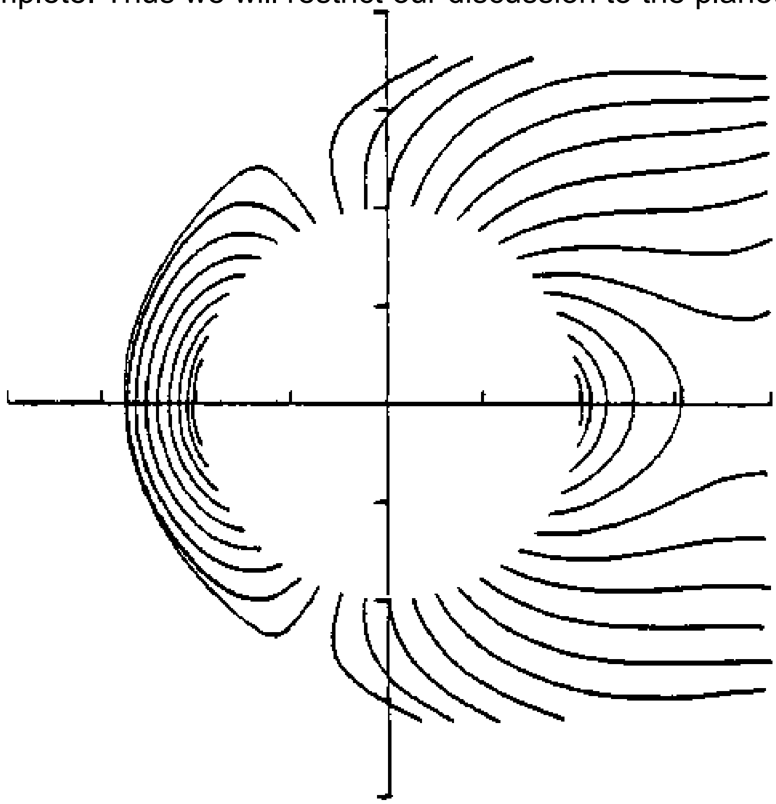


Fig.13

Mercury, which is closest to the sun, has a magnetic moment over 2000 times smaller than that of the earth. The higher density of the solar wind at Mercury together with the smallness of its magnetic moment results in a much smaller magnetosphere. The distance to the subsolar point is one-twentieth of that in the earth's magnetosphere. In fact, even though Mercury is much smaller than the earth it fills much of the magnetosphere, as illustrated in Figure 13. Thus, there does not appear to be an analogue to the earth's radiation belt at Mercury. The fact that Mercury has little or no atmosphere means that it also has little or no ionosphere and the dynamics of the magnetosphere of Mercury should be quite different from the terrestrial magnetosphere. Nevertheless, an apparent substorm was reported during the first encounter of Mariner 10 with Mercury.

Jupiter, in contrast, has a magnetic moment 20,000 times larger than that of the earth. Since it is five times the distance from the sun, the solar wind is 25 times less dense, and the resulting distance to the subsolar point on the Jovian magnetopause is about 100 times larger than the terrestrial distance. If we could see Jupiter's magnetosphere in the night sky, it would appear larger than our moon. Even though the radius of Jupiter is 11 times that of the earth, the planet fills only a very small portion of the magnetosphere in comparison with the earth. The Jovian magnetosphere has two very important differences from the terrestrial magnetosphere. First, Jupiter rotates more rapidly and has a much stronger magnetic field so that the corotation electric field is stronger. Thus, rotational forces are much more important than at earth and the magnetospheric plasma rotates with the planet over most of the magnetosphere. Second, Jupiter has four major moons inside its magnetosphere together with a small moon and a dust ring close to the planet. The moons and dust ring can both absorb charged particles from and release charged particles to the magnetosphere. The major source of particles appears to be volcanic eruptions on the moon Io. The large size of the Jovian magnetosphere allows energetic particles to be accelerated to extremely high energies and can store large fluxes of these particles. Thus, the Jovian magnetosphere presents a radiation hazard to the spacecraft that visit Jupiter.

Saturn's magnetic moment, while much smaller than Jupiter's, is still 600 times larger than that of the earth. The magnetic moment of Saturn is aligned nearly along the rotation axis and thus its magnetosphere does not undergo the daily changes of shape of the terrestrial magnetosphere caused by the alternate pointing of the magnetic dipole axis more toward then more away from the sun. Saturn's magnetosphere also contains moons that are both sources and sinks of particles. The largest satellite Titan has an atmosphere but has not yet been probed closely enough to see if it has volcanoes like Io's. Saturn is famous for its extensive ring system and these rings absorb the radiation belt particles as they diffuse inward toward the planet. Thus, the energetic particle flux does not build up to the levels it does at Jupiter. The observed fluxes are in fact similar to those in the terrestrial magnetosphere.

The magnetosphere of Uranus was the subject of much speculation before the arrival of the Voyager 2 spacecraft in January 1986 because the rotational axis of Uranus is nearly in the ecliptic plane. For many years, as Uranus orbits the sun, the rotation axis points nearly at the sun. It was thought that if Uranus had an intrinsic magnetic field, the magnetic dipole axis might be nearly along this rotational axis. In this case the magnetosphere of Uranus might have its north polar cap facing the sun and its south polar cap directly away from the sun. One-quarter Uranian year later the dipole would be perpendicular to the solar direction and the magnetosphere would resemble that of the earth again. No one anticipated the actual nature of the intrinsic magnetic field of Uranus. The dipole axis was not closely aligned with the rotational axis but at an angle of 60^0 to it. This causes the magnetosphere to undergo large variations in orientation relative to the solar wind flow in the course of the 17-hour Uranian day, providing a complexity to magnetospheric processes that rivals the complexity of the Jovian magnetosphere.

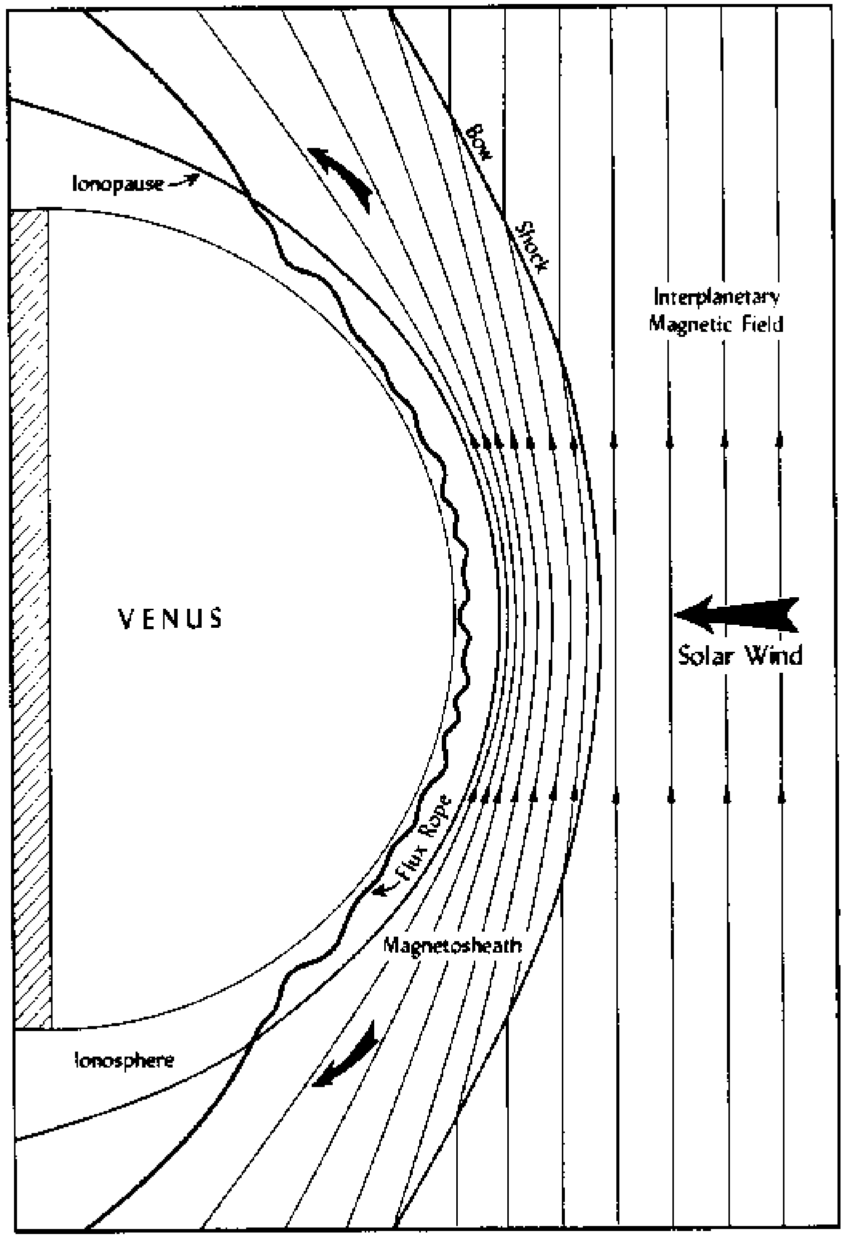


Fig.14

Although Venus is similar in size to the earth and has a spin rate that, while smaller than that of the earth, is not so small that we would expect dynamo action to cease, it does not have a measurable intrinsic magnetic field. Venus does have an atmosphere and an ionosphere. When the solar wind pressure is low, the ionosphere of Venus is sufficient to deflect the solar wind and exclude the magnetic field of the solar wind from the ionosphere. Above the ionosphere the magnetic field of the solar wind is compressed and draped around the planet putting a lid on the ionosphere and forming a long magnetic tail. As shown in Figure 14, the interaction of Venus with the solar wind in many ways resembles that of the earth with the solar wind. However, at Venus the obstacle to the solar wind is approximately the size of the planet rather than a factor of

10 bigger as in the case of the earth. Bundles of magnetic field from the region above the ionosphere appear to form twisted ropes of magnetic field flux. These flux ropes then sink into the ionosphere so that the ionosphere becomes filled with narrow filaments of moderately strong magnetic field separated by magnetic field-free regions or regions of extremely weak magnetic field. When the solar wind pressure becomes strong, the solar wind magnetic field is apparently pushed into the ionosphere and the ionosphere becomes more uniformly and highly magnetized.

Mars has not been probed as thoroughly as Venus. Whether Mars has an intrinsic magnetic field is a subject of intense controversy. It is clear that the largest possible magnetic moment that Mars could possess is much smaller than that predicted for a Martian dynamo. It seems to this author that Mars has no intrinsic magnetic field and that the interaction with the solar wind is similar to that of Venus. However, until we probe the solar wind interaction with Mars at much lower altitudes than we have to date, we will not have an unambiguous answer to this problem.

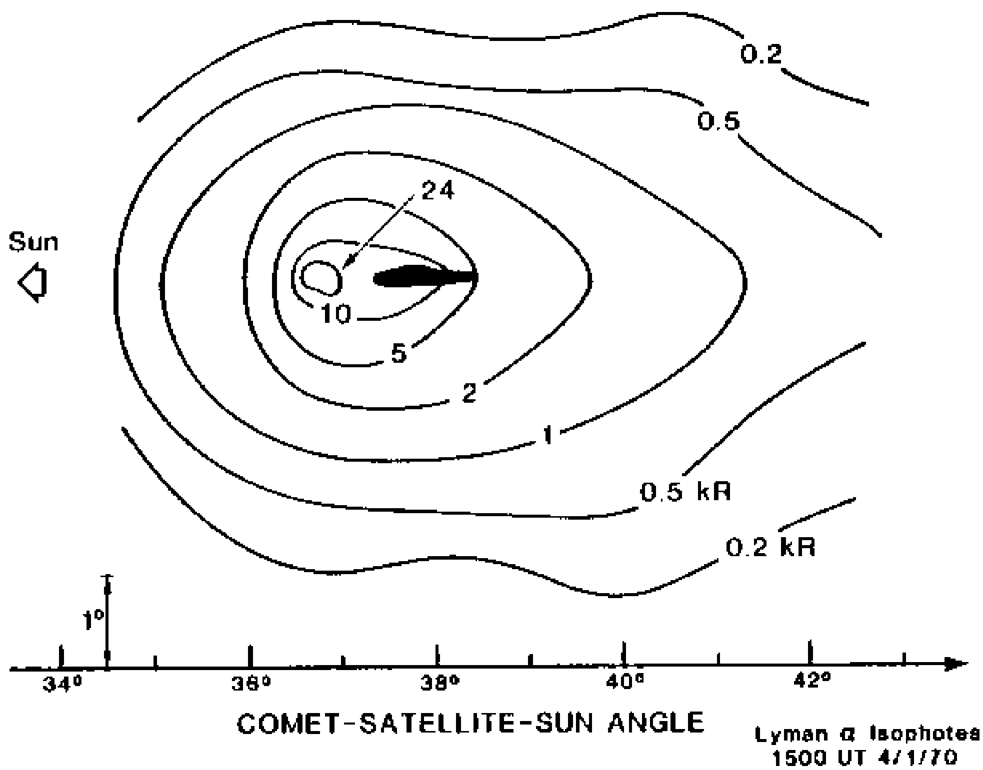


Fig.15

Last, there are comets. For our purposes comets may be thought of as illustrated in Figure 15, which shows contours of the light scattered from the neutral hydrogen cloud surrounding comet Bennett in 1970. Superimposed on these contours is the visible coma and ion tail of the comet which is much smaller.

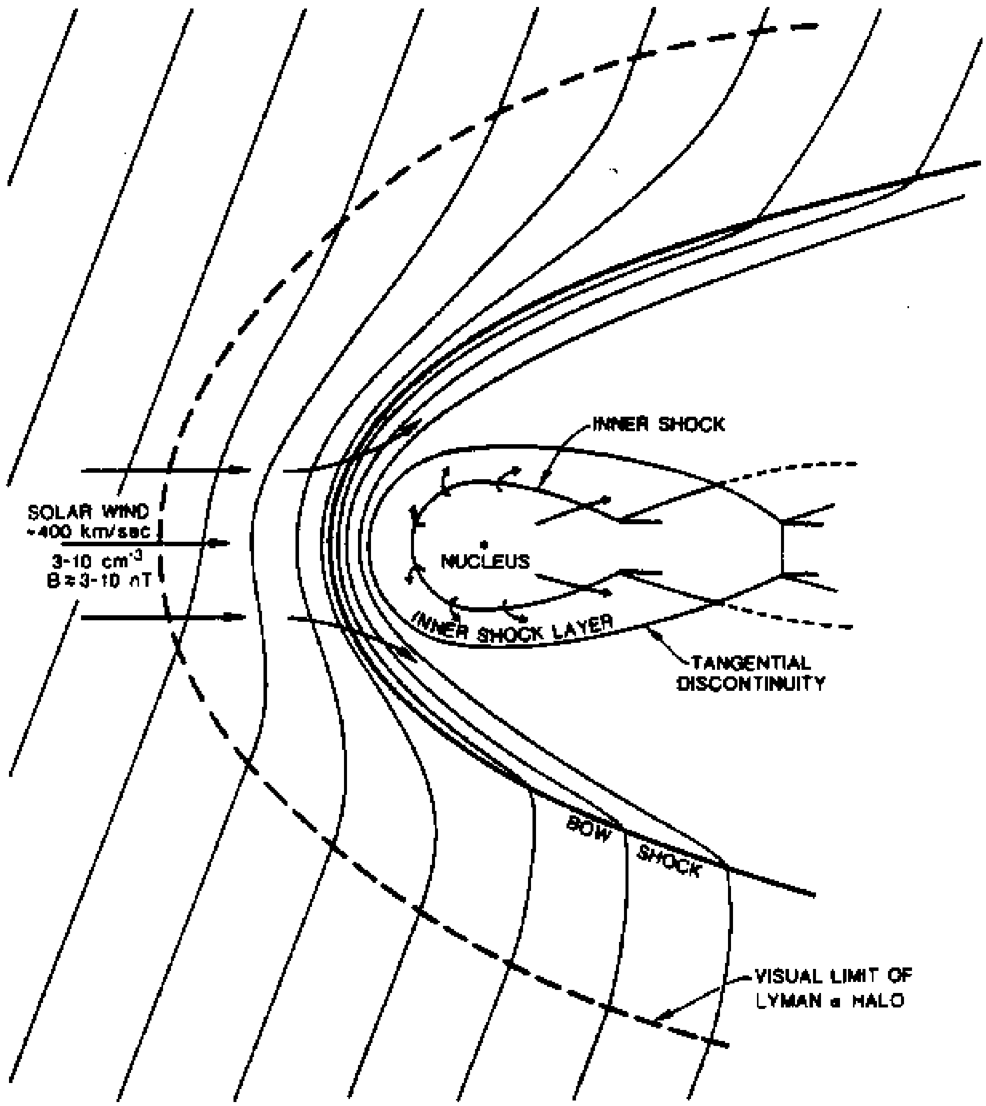


Fig.16

Figure 16 shows a schematic of how the solar wind interacts with such a comet. When the solar wind enters the cloud or corona surrounding the comet, it picks up cometary ions and the solar wind flow becomes heavier, and due to conservation of momentum, it slows down. This slowing down bends the field lines as shown. If the interaction is strong enough, a bow shock will form to deflect the flow around the comet. Venus also adds ions from its planetary atmosphere to the solar wind in much the same way leading to a planetary magnetotail. Studies of the interaction of the solar wind with Venus over the course of a solar cycle suggest that Venus acts most like a comet at solar maximum, and least like a comet at solar minimum.

The Future

While our understanding of the earth's magnetosphere has expanded rapidly over the last two decades, there remains much to be done. The distant magnetotail and the polar cusp have scarcely been explored, and we lack a detailed understanding of these regions. At mid-altitudes over the auroral oval field lines appear to develop electrical potential drops whose source we do not understand. We do not yet agree on all the various mechanisms leading to loss of particles from the radiation belts, particularly over the auroral oval. Some of the answers to these questions may lie in data acquired during the International Magnetospheric Study period from 1977-1979 but not yet analyzed. Some of the questions require new data. The International Solar Terrestrial Program is designed to acquire much of this information. One satellite, Wind, would gather data in the solar wind and monitor the plasma density velocity and temperature and the interplanetary magnetic field strength and direction which controls magnetospheric response. Simultaneously, an equatorial spacecraft would measure the waves and particles at the magnetospheric equator that lead to the auroral precipitation while a Polar spacecraft and a Tail spacecraft (OPEN-J) measure the high-altitude polar cusp and distant magnetotail respectively. The simultaneous probing of each of these regions of the magnetosphere is essential to continued progress in magnetospheric research, for our spacecraft can be in only one location at a time while processes in the magnetosphere affect the entire cavity.

There is also a need to repeat the ISEE-1 and -2 mission but in other regions of space. ISEE-1 and -2 were launched into identical orbits in October 1977 with variable spacing in that orbit. This variable spacing permitted time-of-flight measurements of boundary motions and waves which have proven invaluable in finally determining velocities, thicknesses and current densities of the bow shock and magnetopause and determining wave modes in the solar wind. However, ISEE-1 and -2 probe only the near- equatorial magnetosphere and not the high-latitude regions. Further, there are some boundaries and waves for which orientation cannot be determined from a single spacecraft and for which four spacecraft flying in formation are needed. Such a mission, code-named Cluster, will also be included in the ISTP mission.

Finally, we note that our exploration of planetary magnetospheres is in the state our exploration of the terrestrial magnetosphere was a decade ago. Today in the terrestrial magnetosphere we are mainly attempting to determine how and why processes occur. In planetary magnetospheres we are still discovering what occurs, but our experience with the terrestrial magnetosphere is speeding the understanding of these processes. The exploration of the terrestrial magnetosphere has been a vigorous and exciting program. With the problems left before us and the advent of planetary exploration we look forward to a continuation of this vigor and excitement for many years to come.

FURTHER READING

ACHIEVEMENTS OF THE INTERNATIONAL MAGNETOSPHERIC STUDY (IMS), Edited by B. Battrick and E. Rolfe, European Space Agency, Noordwijk, The Netherlands, 1984.

SOLAR TERRESTRIAL PHYSICS, Edited by R. L. Carovillano and J. M. Forbes, D. Reidel Publishing Co., 1983.

PHYSICS OF THE JOVIAN MAGNETOSPHERE, A. J. Dessler, Cambridge University Press, 1983.

MAGNETIC RECONNECTION IN SPACE AND LABORATORY PLASMAS, Edited by E. W. Hones, Jr., Geophysical Monograph 30, American Geophysical Union, Washington, D.C., 1984.

VENUS, Edited by D.M. Hunten, L. Colin, T.M. Donahue and V.L. Moroz, The University of Arizona Press, 1983.

QUANTITATIVE ASPECTS OF MAGNETOSPHERIC PHYSICS, L.R. Lyons and D.J. Williams D. Reidel Publishing Co., 1984

THE EARTH'S MAGNETIC FIELD: ITS HISTORY, ORIGIN AND PLANETARY PERSPECTIVE, R. T. Merrill and M. W. McElhinny, Academic Press, 1983.

MAGNETOSPHERIC CURRENTS. Edited by T. A. Potemra, Geophysical Monograph 28, American Geophysical Union, Washington, D.C., 1984.

SOLAR MAGNETOHYDRODYNAMICS, E.R. Priest, D. Reidel Publishing Co., 1982.

SOLAR AND PLANETARY MAGNETIC FIELDS, C.T. Russell, In The Solar System: Observations and Interpretations, edited by M. G. Kivelson, Prentiee Hall, 1986.

SPECIAL ISSUE ON INTERNATIONAL MAGNETOSPHERIC STUDY, Edited by D.J. Southwood and C.T. Russell, in Review of Geophysics and Space Physics, Vol. 20, 1982.

COLLISIONLESS SHOCKS IN THE HELIOSPHERE: A TUTORIAL REVIEW, Edited by R. G. Stone and B.T. Tsurutani, Geophysical Monograph 34, American Geophysical Union, Washington, D.C., 1985.

COMETS, Edited by L. L. Wilkening The University of Arizona Press, 1982.