

Geomagnetic Responses to the Solar Wind and to Solar Activity

LEIF SVALGAARD
Stanford University

This paper presents a unified overview of our present knowledge of the geomagnetic response to the dynamic solar wind. Physical understanding rather than observational details is emphasized. Following some historical notes, the formation of the magnetosphere and the magnetospheric tail is discussed. The importance of electric fields is stressed and the magnetospheric convection of plasma and magnetic field lines under the influence of large-scale magnetospheric electric fields is outlined. Ionospheric electric fields and currents are intimately related to electric fields and currents in the magnetosphere and the strong coupling between the two regions is discussed. The energy input of the solar wind to the magnetosphere and upper atmosphere is discussed in terms of the reconnection model where interplanetary magnetic field lines merge or connect with the terrestrial field on the sunward side of the magnetosphere. The merged field lines are then stretched behind Earth to form the magnetotail so that kinetic energy from the solar wind is converted into magnetic energy in the field lines in the tail. Localized collapses of the crosstail current, which is driven by the large-scale dawn/dusk electric field in the magnetosphere, divert part of this current along geomagnetic field lines to the ionosphere, causing substorms with auroral activity and magnetic disturbances. The collapses also inject plasma into the radiation belts and build up a ring current. Frequent collapses in rapid succession constitute the geomagnetic storm. The merging model emphasizes the importance of the interplanetary magnetic field and especially the north-south component because the merging efficiency is strongly dependent on the amount of southward flux. The solar sector structure with its organized magnetic field and embedded high-speed plasma streams is identified as the source of the recurrent geomagnetic disturbances while flare-associated interplanetary shock waves are the source of most violent and sporadic geomagnetic storms. An appendix contains numerical estimates of some relevant physical quantities related to intensities of fields and currents in the magnetosphere and the ionosphere.

HISTORICAL NOTES

In 1843, Swabe discovered the 11-yr sunspot cycle from 17 yr of regular observations of the Sun commencing in 1826. Following this, in 1852 Sabine announced his discovery of a strong positive correlation between the number of sunspots and the disturbance variation of the declination of the geomagnetic field measured in Toronto, Canada, during the years 1841 to 1848, not covering even one full sunspot cycle. It was concluded on this limited statistical evidence that the geomagnetic environment was strongly influenced by solar activity. Over a century of subsequent mon-

itoring of solar and geomagnetic activity have confirmed these early conclusions, although the first indication of an explicit event on the Sun with direct terrestrial response was observed as early as 1859 by the renowned solar astronomer Carrington. While observing a large spot group on the Sun, he saw an intense outburst of white light from the sunspot group. The event lasted only a few minutes, but at the same time all three components of Earth's magnetic field recorded at Kew Magnetic Observatory became abruptly disturbed, followed about 18 hr later by a great geomagnetic storm that surpassed in intensity and duration all previous observations. For several

days auroral displays of almost unprecedented magnificence were observed and telegraph communication was widely interrupted because of currents induced in the wires.

While Carrington cautiously proposed a connection between this solar and the terrestrial events, it was difficult for the scientific world to accept any such idea. In 1905, Maunder drew attention to the 27-day recurrence pattern of the magnetic activity and Chree removed every doubt about the existence and significance of this 27-day period. Because the synodic rotation period of the Sun is also near 27 days, the 27-day recurrence period was additional evidence that its ultimate cause is resident in the Sun. Chree and Stagg noted in 1927 that

The exhibition of a 27-day interval in groups of days of all types, from the most highly disturbed to the quietest, seems to imply that there is no exceptional phenomenon on highly disturbed days, but merely increase in the activity of some agent always more or less active. If magnetic disturbance is due to radiation from the Sun, then . . . the radiation must always be going on.

Chapman and Ferraro in a series of papers in the 1930's examined theoretically the effect of a plasma stream emanating intermittently from the Sun and impinging on Earth to interact with Earth's magnetic field and causing geomagnetic storms. Their basic ideas were largely correct except that, as pointed out by Chree and Stagg and later by Bartels, the geomagnetic field is always somewhat disturbed, indicating a continuous rather than intermittent mode of interaction. Activity never ceases completely and auroras can always be seen somewhere. The realization and general acceptance that the Sun continuously emits a tenuous, magnetized plasma which at all times interacts with Earth and its magnetic field has come slow and had to await direct in situ probing by spacecraft in 1962. From studies of movements and directions of comet tails, Bierman in 1951 proposed that the Sun emits "corpuscular radiation" in essentially all directions at essentially all times, and Parker in 1958 proposed a hydrodynamic model of the solar corona from which the material flowed out as a natural consequence of the million degree temperature of the corona. Parker named this phenomenon the "solar wind," by which name it has been known ever since. But

final acceptance of the existence of an essentially continuous solar wind came first after measurements made on board the Venus probe Mariner 2 in 1962. The principal features of the solar wind as reported by Neugebauer and Snyder were:

- (1) A detectable solar wind was present at all times.
- (2) The average solar wind speed was 500 km.
- (3) The speed varied between 300 and 860 km and was correlated with geomagnetic activity.
- (4) The average proton density was 5/per cm³.
- (5) Several streams of high-speed plasma were found to reoccur at 27-day intervals.
- (6) The plasma was found to possess a weak magnetic field.

The discovery of the magnetized solar wind and the concept of a continuous interaction of the wind with the terrestrial magnetic field are the basis for our understanding of the geomagnetic response to solar activities.

THE MAGNETOSPHERE

In the presence of a weak interplanetary magnetic field, the solar wind plasma behaves as a supersonic continuum fluid over scale lengths that are large compared with the proton gyroradius (typically 100 km for solar wind plasma near Earth). Earth's magnetic field thus presents an obstacle to the solar wind flow. To a first approximation the solar wind flow around this obstacle can be treated fluid dynamically. The magnetic pressure in the dipolar geomagnetic field falls off as $(r^{-3})^2 = r^{-6}$ and eventually becomes comparable with the directed gas pressure p of the solar wind. Close to the geomagnetic field, there is a region where the magnetic pressure $B^2/2\mu_0$ (where B denotes the magnetic flux density and μ_0 is the permeability of free space) is much larger than p , but in the free solar wind p is much larger than the magnetic pressure of the weak interplanetary field. The boundary between these two regions is called the magnetopause and the region inside the magnetopause that confines the geomagnetic field is called the magnetosphere.

Because the magnetic pressure of the geomagnetic field varies rapidly with distance, the magnetopause can be adequately represented by a tangential discontinuity in which there is no solar wind plasma on the magnetosphere side of the

magnetopause and no magnetic field on the solar side. In this approximation, the gas pressure p in the solar wind must balance the magnetic pressure $B^2/2\mu_0$ just inside the magnetopause, and solar wind particles are specularly reflected from the magnetopause. From these assumptions the shape and size of the magnetopause can be computed using an iterative method to solve what is essentially a free-boundary problem: both the boundary and the conditions that determine it are to be found.

A standing shock front or bow wave would be expected at some distance upstream in the solar wind. This is because the geomagnetic field is an obstacle in a supersonic (more precisely, super-Alfvénic) flow. A transition to subsonic flow is necessary for the solar wind to flow smoothly around Earth as required by the zero flow velocity normal to the magnetopause. A supersonic solar wind cannot receive knowledge of the obstacle ahead so the wind must undergo an upstream shock transition to subsonic flow. The position and shape of this bow shock can be calculated using conventional equations of fluid dynamics for a solid obstacle of the same shape as the magnetopause.

The region between the shock and the magnetopause is called the magnetosheath and contains shocked solar wind plasma with increased density and temperature and also somewhat disturbed interplanetary magnetic field. Given the interplanetary field, the average configuration of the magnetic field in the magnetosheath can finally be computed assuming that field lines move with the streaming plasma and taking the boundary condition that the field normal to the magnetopause b_n vanishes. For an interplanetary field directed along a 45° spiral angle, the calculated geometry and extent of the magnetosphere and magnetosheath regions on the day side of Earth is shown in figure 1. Several comparisons of theory and measurements made in space have confirmed the adequacy of the continuum fluid model for predicting even quantitatively the location and shape of both the magnetopause and the bow shock wave and for explaining the observed properties of the flow of the solar wind plasma in the magnetosheath. In fact, the agreement between theory and observation is surprisingly good, considering both the

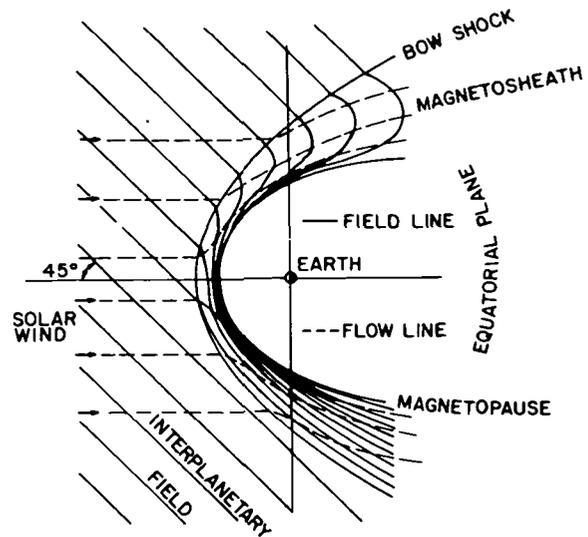


FIGURE 1.—Flow lines of the solar wind around the geomagnetic field confined within the magnetosphere. Interplanetary magnetic field lines corresponding to a spiral of 45° are draped around the magnetopause. The geomagnetic dipole is assumed to be perpendicular to the plane of the figure and to the solar wind flow.

gross simplifications that are necessary to make the problem tractable and the lack of a rigorous justification for applying fluid concepts to a collisionless, weakly magnetized plasma.

The treatment of the solar wind as a cold plasma flow leads to the formation of a magnetosphere that is open in the antisolar direction with its flanks stretching asymptotically to the solar wind flow direction. At great distances from Earth, the dynamic flow pressure on the magnetopause tends to zero together with the magnetic field inside the magnetosphere. In the more realistic case, where the solar wind pressure includes both the directed dynamic pressure of the flow and the more nearly isotropic thermal pressure due to nonzero plasma temperature, the magnetosphere will be closed in the antisolar direction at some distance from Earth. In this case the magnetosphere is expected to extend in the solar wind flow direction (corrected for the small aberration resulting from the orbital movement of Earth around the Sun) to three or four times the standoff distance on the sunward side of Earth. This extension, the magnetospheric tail, has also

been observed to exist by in situ spacecraft measurements.

The observed properties of the tail are, however, not understood in terms of the fluid dynamic approach, which was so successful in describing the sunward regions of the magnetosphere. Figure 2 summarizes the observational results. Field lines in the tail beyond about 10 Earth radii are roughly parallel to the Sun/Earth line. The tail itself approximates a long cylinder. In the northern half of the cylinder the field lines are directed toward the Sun, and in the southern half their direction is away from the Sun. The length of the tail and its eventual termination is not well known but is at least several hundred Earth radii, and is therefore very much larger than predicted. It is important to note that the tail field lines all come from fairly small regions around the magnetic poles inside the classical auroral zones. High fluxes of kiloelectron volt plasma are observed in the so-called plasma sheet separating the oppositely directed fields in the tail lobes. The thickness of this plasma sheet varies greatly with geomagnetic activity but is typically 5 Earth radii, and the sheet extends most of the way down the tail. The plasma sheet surrounds a region of very weak fields, the neutral sheet, where the tail field reverses. To maintain the tail configuration of oppositely directed field lines, a current must flow in the neutral sheet across the tail. Figure 3(a) shows a north-south cut through the magnetotail. Figure 3(b) shows a schematic cross section of the tail. The field directions above and below the neutral sheet require a tail current flowing in the sheet from dawn to dusk.

That the tail is much longer than predicted by the continuum fluid model is obviously the result of forces (external or internal) exerted on the magnetic field to stretch out the field lines. We do not know precisely what these forces are. The pressure of the quiet solar wind is about an order of magnitude larger than the tension in the tail, so it is natural to assume that interactions between the solar wind and the magnetosphere at the magnetopause provide the necessary *tangential stresses* to pull out the tail in the antisolar direction.

Turbulence in the solar wind could produce such interactions because it ripples the magnetopause with a phase velocity exceeding the Alfvén

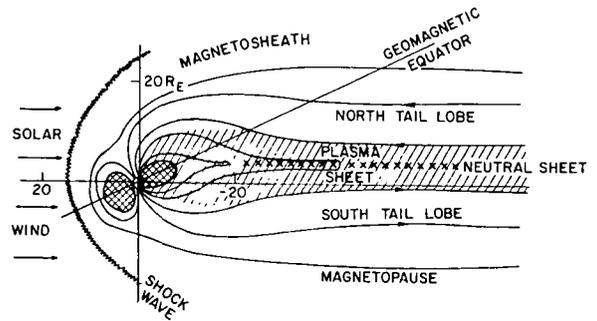


FIGURE 2.—Observed properties of the magnetotail. The distant tail is approximately aligned with the solar wind flow direction independent of the inclination of the geomagnetic equator to the ecliptic plane. Field lines in the northern tail lobe are directed toward the Sun, and field lines in the southern tail lobe are directed away from the Sun. The plasma sheet separates the two tail lobes and the field reversal takes place in the neutral sheet, which then contains a very weak net northward magnetic field. The inner part of the magnetosphere (crosshatched) contains plasma of mainly terrestrial origin. This plasmasphere corotates with Earth, while the rest of the magnetosphere stays roughly fixed in relation to the Sun/Earth line.

speed, thereby generating waves that propagate into the magnetosphere. Another possibility is that the magnetopause is not a perfect separation of interplanetary and geomagnetic field lines. If field lines cross the magnetopause, then the solar wind “may blow away the magnetic lines of force like smoke from a chimney.” However, we can in this case not relate the magnetopause to a boundary separating different field lines because these cross the magnetopause. Moreover, solar wind plasma may penetrate the boundary and equalize the concentration on both sides of the boundary. In the case of an isotropic velocity distribution of the solar wind particles, the plasma concentration along magnetic field lines would be constant and there would be no near-stationary magnetopause. But since the directed energy for solar wind particles greatly exceeds their thermal energy, we have a very highly anisotropic velocity distribution and the majority of the particles will be reflected back by a region of increasing magnetic field. This region where the magnetic field intensity increases rapidly could then be considered to be the magnetopause. Energetic particles from solar flares penetrate easily into the magnetosphere because of the much higher degree of

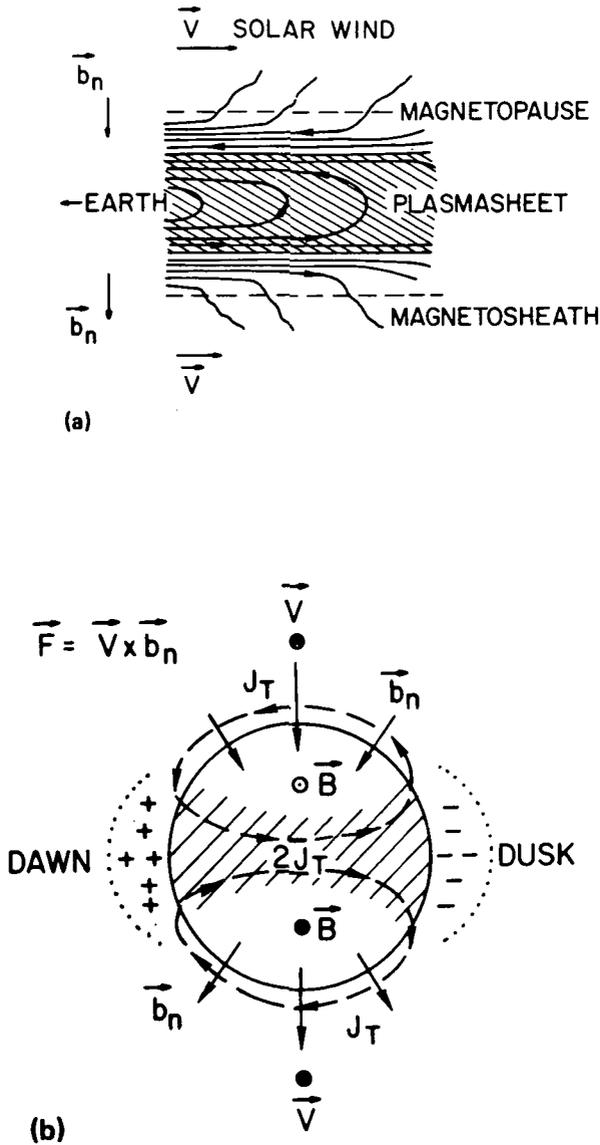


FIGURE 3.—Cross sections of the magnetotail. (a) North-south cut through the magnetotail. Field lines in the central plasma sheet connect with field lines from the other tail lobe. Field lines outside the plasma sheet connect to the interplanetary magnetic field thus providing a field component b_n normal to the magnetopause. (b) Cross section of the magnetotail as viewed from Earth. The plasma sheet is indicated by shading in the middle of the tail. The electromotive force, $V \times b_n$, of the magnetospheric dynamo drives a current J_T around each tail lobe and accumulates positive space charge on the dawn side magnetosphere and negative space charge on the dusk side. The electric field resulting from the charge separation is discharged through the cross tail current $2J_T$ keeping the two lobes apart.

isotropy of these particles which do not recognize a magnetopause. In a sense the magnetopause could be considered "magnetoporous" to magnetic field lines and isotropic particles.

ELECTRIC FIELDS AND CONVECTION

A plasma always sets itself in motion such as to oppose any external electric field in order that there be no electric field in the rest frame of the plasma. Switching on an electric field causes the particles to drift so that they do not see any electric field. One might say that collisionless plasmas abhor electric fields, so that

$$\mathbf{E} + \mathbf{v} \times \mathbf{B} = 0 \tag{1}$$

or, alternatively,

$$\mathbf{v} = \mathbf{E} \times \frac{\mathbf{B}}{B^2} \tag{2}$$

where \mathbf{E} is electric field strength, \mathbf{B} is magnetic flux density, and \mathbf{v} is the resulting plasma drift velocity. Similarly, magnetic field lines in a highly conducting plasma move with the plasma because the electromotive force around any closed loop must vanish and, hence, the flux through the loop cannot change. We can therefore, to a good approximation, consider field lines to be a permanent part of the ionospheric and magnetospheric plasma and also to the conducting interior of

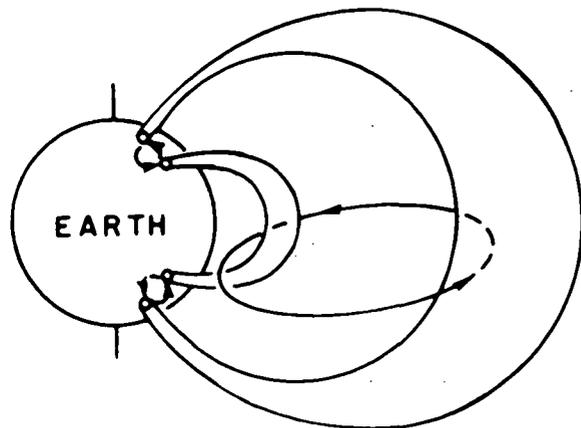


FIGURE 4.—Interchange of tubes of magnetic field lines. The inner tube can be stretched to go into the position of the outer tube, but the outer tube shortens upon moving to the position of the inner tube. In the absence of dissipative forces, no work is done by interchanging flux tubes.

Earth. But this is not true in the neutral atmosphere, and, as a result, two magnetic tubes of force may be interchanged as shown in figure 4. The inner flux tube must be stretched to go into the position of the outer tube, which requires work, but the outer tube shortens upon moving to the position of the inner tube and gives up just as much energy as the other consumes. So there is no tendency for the tubes to interchange or to resist interchange. Moving the flux tubes amounts to interchanging the plasma in the tubes.

Field lines passing through the ionosphere are embedded in a plasma that is highly conducting, and a potential difference between any two points in the ionosphere must exist everywhere along the two field lines containing these points. This is because the field lines are approximately equipotential because of the plasma lying along any of them, and therefore a potential difference between two points in the ionosphere must be maintained all along the magnetic field lines. This means that there is an electric field between these two field lines, and the plasma tied to the field lines must then drift with a velocity

$$\mathbf{v} = \mathbf{E} \times \frac{\mathbf{B}}{B^2}$$

in order that there be no electric field in the rest frame of the plasma. This drift is called *convection of permanent field lines* in the presence of an electric field and has proven to be of fundamental importance in the dynamics of the magnetosphere.

Within the *E* region (90- to 150-km altitude) of the ionosphere, electrons drift freely, but the motion of ions is strongly impeded by collisions with neutral particles because the relations between the collision frequency ν and the gyrofrequency ω are such that $\nu_{\text{electron}} < \omega_{\text{electron}}$ and ω_{ion} . Therefore the ions move essentially with the neutral gas except for a small drift parallel to the electric field in the sense of a direct (Pedersen) current that discharges this field. The electrons still satisfy equation (1) and can be considered still frozen to the field lines. The drift of the electrons results in a Hall current that flows perpendicular to the electric and the magnetic fields. Throughout the *E* region the Hall conductivity is much larger than the Pedersen conductivity, so

that in this region the major ionospheric currents can be considered as being Hall currents to a fair approximation. This is important because it enables us to infer the approximate direction and (with an estimate of the conductivity) the magnitude of electric fields in the ionosphere, and because magnetic lines of force are almost equipotentials, also roughly to determine the distribution of electric potential in the magnetosphere.

Although the Pedersen current is not important in producing magnetic variations, it is significant in that it is dissipative. The energy dissipation, which can be considered to be the result of friction between the charged and the neutral constituents of the atmosphere, is so effective that electric fields in the magnetosphere that are not maintained by some driving mechanisms are discharged in a few seconds. Constantly maintained convective motions in the magnetosphere are, therefore, normally accompanied by a substantial amount of ionospheric heating.

If interplanetary and geomagnetic field lines are connected across the magnetopause, there will be a component \mathbf{b}_n of the magnetic field normal to the magnetopause as shown in figure 3(a). The electromotive force, $\mathbf{F} = \mathbf{V} \times \mathbf{b}_n$, where \mathbf{V} is the solar wind velocity, caused by the solar wind flow along the magnetopause, drives electric currents of intensity J_T as indicated in figure 3(b). The current builds up a positive space charge on the dawn side of the magnetopause and a negative space charge on the dusk side and completes its circuit by the current across the tail in the neutral sheet where the magnetic field is very weak. In a sense we can regard the magnetosphere as a very large lossy capacitor that acts as a load for the solar wind electric generator. The dawn and dusk sides are the two capacitor plates, and the magnetosphere, particularly the plasma sheet, is the dielectric between them. Geomagnetic and auroral activity constitute loss mechanisms, or resistive elements, or maybe at times short circuits.

The existence of this large-scale magnetospheric electric field directed from dawn to dusk has been verified by a variety of techniques including satellite, rocket, and balloon observations. This magnetospheric electric field has been found to be a permanent feature of the magnetosphere and it is now generally accepted that it plays a central role

in magnetospheric processes. The separated charges causing this electric field are located in a thin layer immediately adjacent to the magnetotail surface. A boundary layer of plasma less dense than the magnetosheath plasma and flowing antisunward at less than magnetosheath flow speed has been observed by satellites; it exists at all times on both the morning and evening sides and probably extends completely around the surface of the tail. Plasma from this boundary layer drifts into the tail, thereby maintaining the plasma sheet. Once these particles are on tail field lines in the plasma sheet they feel the influence of the magnetospheric electric field and drift toward Earth as the result of the net northward magnetic field across the plasma sheet and the dawn/dusk electric field. This drift under the influence of the electric field accelerates the plasma particles adiabatically because of the increasing magnetic field as the plasma comes closer to Earth. If the energy gain is large enough the plasma may penetrate deep into the ionosphere before mirroring back and may be precipitated due to Coulomb scattering, collisions, and wave-particle interaction.

These considerations can be summarized by

noting that plasma flows down the tail near the tail surface and back again toward Earth in the plasma sheet within the tail. This large-scale circulation of the plasma is commonly referred to as the deep magnetospheric convection and is expressed in terms of convection of permanent magnetic field lines. Figure 5 shows a schematic of these convective motions of the magnetic field lines and associated particles in the equatorial plane of Earth. This convective circulation is often described in rather loose terms by saying that magnetospheric field lines are carried by the solar wind from the day side, over the polar caps, and into the night side magnetosphere, wherefrom they return to the day side having their foot-points flowing through the subpolar or auroral zone ionosphere.

Because of viscosity, the neutral atmosphere largely rotates with Earth. In the lower ionosphere the neutral atmosphere interacts with the ions by collisions to set the ionosphere in corotational motion. In the frame of reference of the rotating Earth, the ionospheric plasma at subauroral zone latitudes is not appreciably affected by the deep magnetospheric convection and is approximately

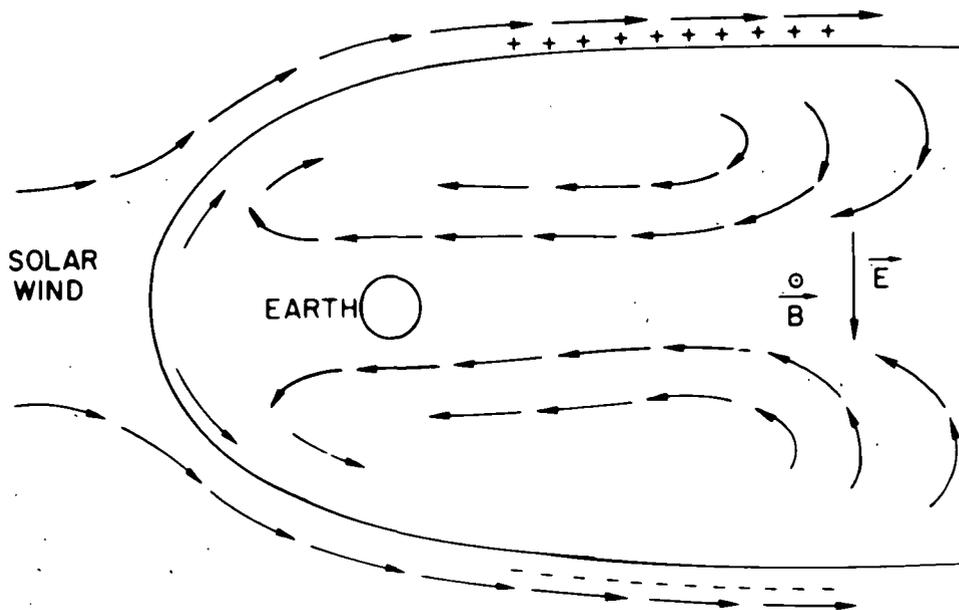


FIGURE 5.—Large-scale magnetospheric circulation of plasma and permanent field lines in the equatorial plane. Solar wind plasma flows down the tail near the magnetopause and toward Earth in the plasma sheet within the tail.

at rest so the electric field is zero. The electric field in a nonrotating frame of reference then becomes

$$\mathbf{E}_c = -\mathbf{v}_c \times \mathbf{B}$$

where \mathbf{v}_c is the corotation velocity and \mathbf{B} is the magnetic field of Earth. For a dipolar \mathbf{B} , the magnitude of the ionospheric corotational electric field is

$$E_c = 0.014 \cos \theta (1 + 3 \sin^2 \theta)^{1/2} \quad \text{V m}^{-1}$$

In the approximation that the magnetic field lines are equipotentials, the ionospheric corotational electric field persists along field lines into the magnetosphere causing the inner magnetosphere to corotate with Earth. This inner part of the magnetosphere contains cold (~ 1 -eV) plasma that has evaporated from the day side ionosphere onto the corotating magnetic field lines.

Even if Earth's rotation and the solar wind were turned off, the upper atmosphere would move because of thermal and tidal effects from the Sun and the Moon. The motions couple to the ionospheric plasma through collisions to set it in motion, and the resulting currents partially polarize the ionosphere to create an electric field. The precise effect of this field depends on the large-scale upper atmospheric wind system, which is poorly known; but in any case, the electric field at a given location has a 24-hr variation because of the diurnal solar heating and ionization of the upper atmosphere. The existence of these ionospheric dynamo currents was suggested by Balfour Stewart in 1882 to account for the observed small (0.1 percent) diurnal variations of the geomagnetic field, the so-called Sq variations. Direct low-latitude magnetic and electric field measurements by rocket and radar techniques have proved the existence of the Sq currents, explaining the first geomagnetic variations to be physically understood.

The relative importance of the ionospheric electric fields produced by rotation of Earth, by tidal motions of the upper atmosphere, and by interaction of the magnetosphere with the solar wind is illustrated in figure 6. At latitudes below 45° , the dynamo and magnetospheric electric field strength are much less than the corotation field strength so that the plasmasphere clearly rotates

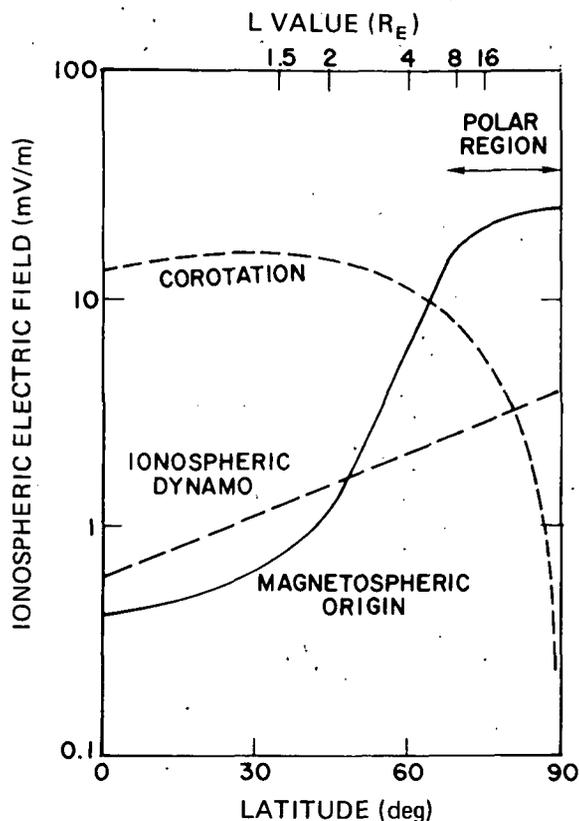


FIGURE 6.—Survey of the relative importance of ionospheric electric fields of different origins as a function of latitude: At low latitudes the corotation and ionospheric dynamo electric fields dominate, while electric fields of magnetospheric origin are most important in the polar regions.

with the Earth. At high latitudes the ionospheric electric field is dominated by magnetospheric processes that cause the plasma to flow in the antisolar direction in the polar cap and toward the Sun at somewhat lower latitudes.

The high latitude electric field has recently been directly observed by low altitude spacecraft and also from active experiments injecting barium vapor into the *F* layers of ionosphere where it is ionized by sunlight; the electric field can then be inferred from the $\mathbf{E} \times \mathbf{B}$ drift of the sunlit barium cloud.

Figure 7(a) shows the electric field observed on a polar pass of the OGO 6 satellite after subtraction of the $\mathbf{V} \times \mathbf{B}$ fields from both the motion of the satellite and the rotation of Earth. The field seems to be quite uniform across the polar cap directed toward the evening side. Field reversals

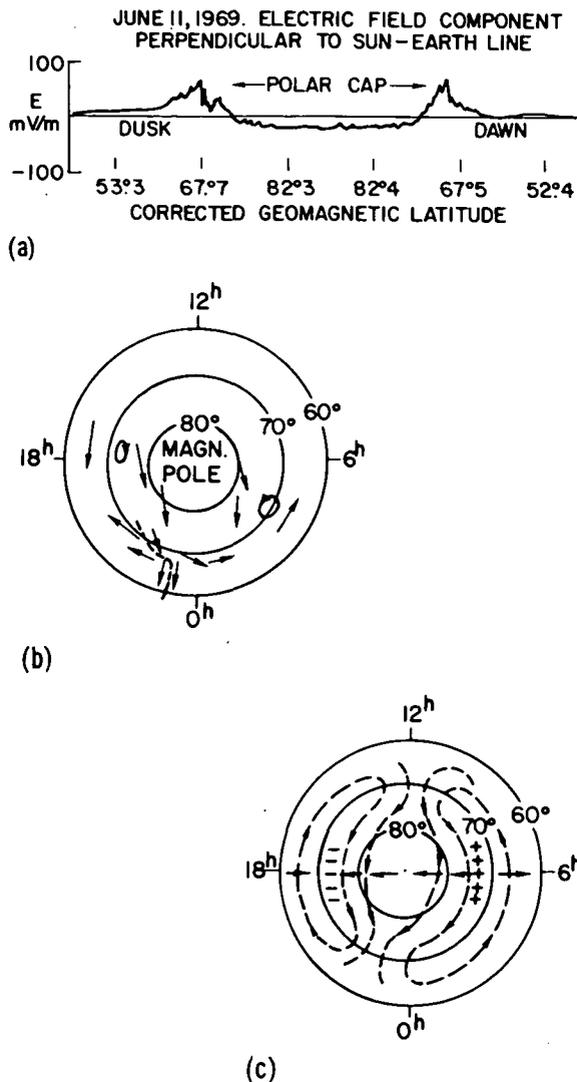


FIGURE 7.—Electric fields and convection. (a) Ionospheric electric field dusk-dawn components measured by OGO 6 satellite passing over the north polar region. A rather uniform electric field is found in the polar cap with reversals near the auroral zones. (b) Typical drifts of Ba⁺ clouds in the F₂ layer in a coordinate system of corrected geomagnetic latitude and local magnetic time. The direction to the Sun is from the magnetic pole to the tick mark labeled 12^h. (c) Summary of electric fields and convection pattern in the polar regions. The direction of the electric field (a) is shown as a series of arrows along the 6^h to 18^h meridian. Regions of positive space charge (source) and negative space charge (sink) are shown at the electric field reversals. Hall currents circulating around these regions are indicated by dashed curves. The geomagnetic field is nearly vertical over the polar regions, directed downward over the northern pole.

are seen at the polar cap boundary. Figure 7(b) shows typical drifts of Ba⁺ clouds released in the F₂ layer plotted in a coordinate system of corrected (taking into account the nondipolar parts of the field) geomagnetic latitude and local magnetic time. The Ba⁺ ions drift antisunward over the polar cap and toward the Sun at lower latitudes in accordance with the expected convection pattern. A schematic summary of the high latitude electric fields and the associated convection is given in figure 7(c).

The convection pattern can be described as consisting of two vortices, one in the morning and one in the evening. Because usually it is the electrons and not the atmospheric ions that participate in the convection in the lower ionosphere, the result is a Hall current in the E region flowing in the opposite direction to the convection flow. Because the electric field is strongest at auroral latitudes surrounding the polar cap (see fig. 7(a)) and because the ionospheric conductivity is highest there, the Hall currents can become quite concentrated and intense at latitudes around and just below 70° and are referred to as the auroral electrojets. Figure 8(a) shows a schematic of the two-celled current system with the electrojets indicated by heavy arrows, while figure 8(b) is an example of current vectors as inferred from magnetometers on the ground. Such configurations would be expected if the convection is in balance, that is, when the return flow in the auroral zone equals the antisunward flow over the polar cap.

FIELD LINE MERGING

There is an increasing understanding that most geomagnetic and related activity results from non-balance of the convection rates on time scales less than typical reaction times of various parts of the coupled magnetosphere-ionosphere system. Understanding the processes that govern the convection rates in different regions within the magnetosphere is therefore extremely important but is largely lacking or at best phenomenological and qualitative in nature. The necessary tangential stresses on the magnetopause to stretch the field lines back into the tail could be provided or at least aided by connecting interplanetary magnetic field lines to geomagnetic field lines. This connection or merging of field lines could take place at

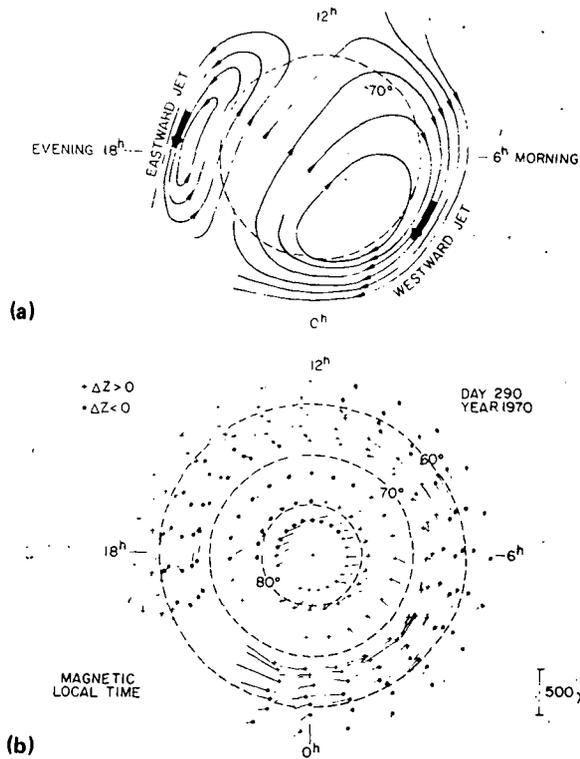


FIGURE 8.—Currents. (a) Schematic overhead equivalent currents flowing in the polar ionosphere. Equivalent currents are not necessarily real currents but simply model currents at constant altitude that could produce the observed magnetic variations on the ground. The current system is plotted as a function of corrected geomagnetic latitude and local magnetic time and is constructed assuming that the current pattern is fixed in space and time with Earth rotating below it. (b) Observed current vectors at a chain of ten polar region magnetic observatories. For a given hourly interval the average directions of the equivalent currents are plotted as lines originating in the observing stations having a length proportional to the observed magnetic perturbation. By plotting these current vectors for successive hourly intervals we can construct the total equivalent current system. The data were chosen for a day where geomagnetic activity was moderately high and nearly constant throughout the day, to minimize temporal variations of the current strength. The sign of perturbations of the vertical component Z of the geomagnetic field is given at each point as a plus for positive and a dot for negative disturbances. Construction of equivalent current systems is a commonly used tool in geomagnetic physics. Interpretation of the current systems is often difficult and the distinction between equivalent and real currents is not always emphasized. Other examples of equivalent ionospheric currents are shown in figure 22.

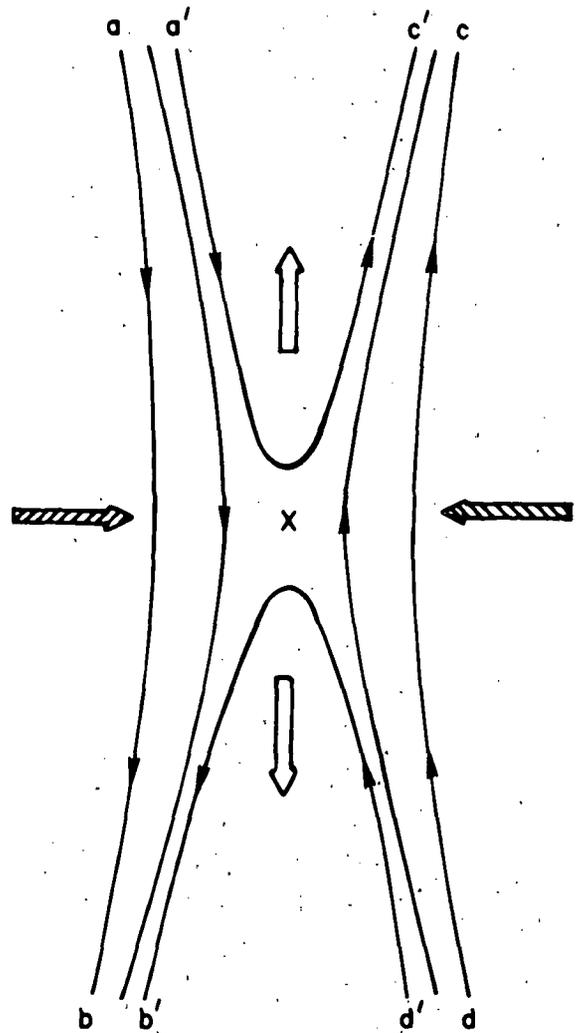


FIGURE 9.—Reconnection of oppositely directed magnetic field lines embedded in a plasma. If the plasma is compressed (shaded arrows) field lines merge at the X-type neutral point and plasma flows away (open arrows) from the reconnection region carrying the connected field lines. Field lines ab and cd eventually assume the new configuration $a'c'$ and $b'd'$.

an X-type magnetic neutral point. As plasmas with oppositely directed magnetic fields are pressed together as illustrated in figure 9, pairs of magnetic field lines such as ab and cd , identified via the plasma frozen to them, flow toward a point where the magnetic field vanishes in an electric discharge. At that point the field lines merge to form a new pair of lines $a'c'$ and $b'd'$. The plasma is squeezed out and accelerated away from the neutral point, aided by the tendency of the new

field lines to reach a lower energy state by shortening themselves. Exactly how the merging takes place is poorly understood, but the process can be made to work in laboratory plasmas. As the plasma on the newly merged field lines flows away from the neutral points more field lines can be merged, and so on.

If the interplanetary magnetic field has a southward component, the geometry at the subsolar point of the day side magnetopause is that of an X-type neutral point as indicated in figure 10(a). The interplanetary field lines and the geomagnetic field lines merge at *A*, and the magnetosheath plasma flow carries the field lines in the antisolar direction. The numbers 1 to 7 in figure 10(a) indicate successive positions of an interplanetary field as it connects to the geomagnetic field. Even if the field lines are not strictly antiparallel, merging can still occur but with lower efficiency, so field lines connected across the magnetopause can be a permanent feature not exclusively dependent on the presence of a southward field. Merging of field lines has the effect that we must distinguish three classes of magnetic field lines near Earth: interplanetary field lines, such as *AA'* in figure 10(b), which are unlinked with the geomagnetic field lines; open field lines, such as *BB'*, which

link the two fields; and closed terrestrial field lines, such as *C* and *D*, which are not linked to the interplanetary magnetic field. The use of the descriptive terms open and closed geomagnetic field lines refers in an incorrect but obvious manner to an important topological property of the field line. On open field lines, solar wind particles and electric fields have direct access to Earth, and ionospheric plasma can directly escape into interplanetary space. It is much more difficult for particles to diffuse across field lines onto closed field lines, and once they are there, the particles are trapped and cannot easily be removed. This trapping region on closed field lines is indicated by crosshatching on figure 2 and coincides roughly with the outer part of the plasmasphere.

When interplanetary field lines have just merged on the day side with the previously outermost closed terrestrial field lines, magnetosheath plasma suddenly gets access to these field lines and can penetrate to low altitudes into the ionosphere before mirroring back. Some of the plasma precipitates and causes a subvisual band of 6300-Å emission. Satellite observations both at low altitude and also out in the magnetosphere show the existence of large fluxes of magnetosheath plasma on geomagnetic field lines near the

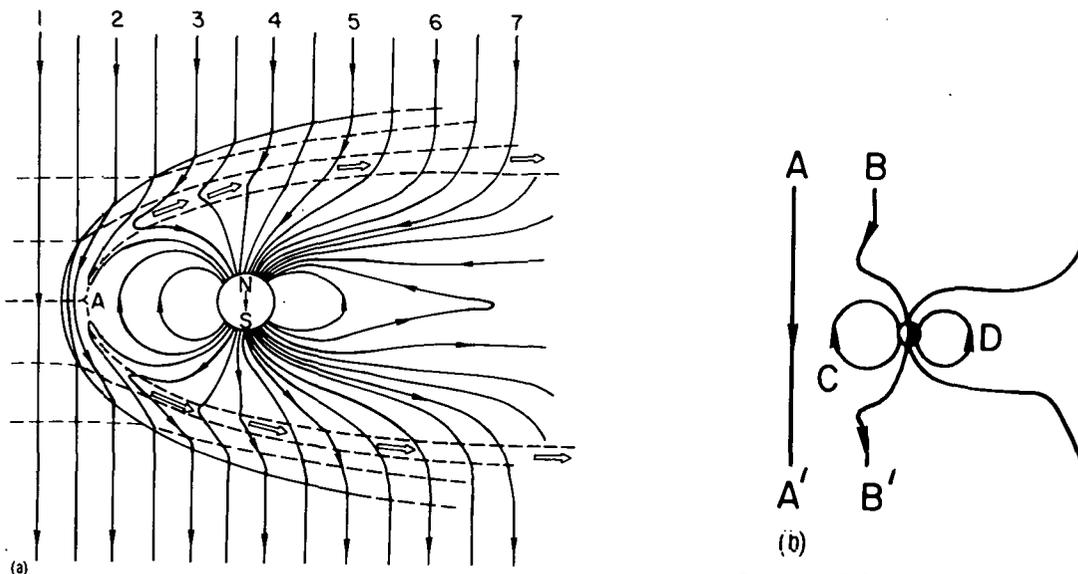


FIGURE 10.—Magnetic field lines. (a) Successive stages (1 to 7) in the linkage of a southward-directed interplanetary magnetic field line with the terrestrial field as the linked lines are carried past Earth by the magnetosheath flow (open arrows). (b) Classes of magnetic field lines with different terrestrial relationships: *AA'* is an unlinked interplanetary field line; *BB'* is an open terrestrial field line connected to the interplanetary field; *C* and *D* are closed terrestrial field lines not linked to any external field.

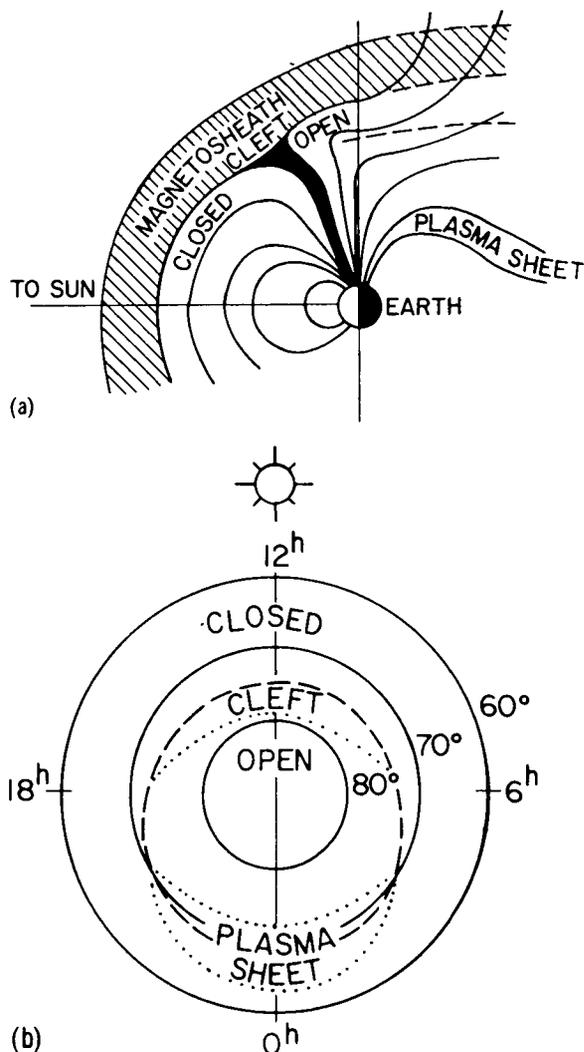


FIGURE 11.—Magnetospheric cleft. (a) The position of the magnetospheric cleft in a north-south section of the magnetosphere. Various magnetospheric regions are indicated. The cleft is shown as the heavy black funnel-shaped region at the boundary between open and closed day side field lines. (b) The boundary on the ground (in corrected geomagnetic latitude and local magnetic time coordinates) between the regions of the closed and open field lines is indicated by the dashed oval-shaped curve, which is closer to the pole on the day side than on the night side. The plasma sheet maps down to the night side oval tapering out as we approach the day side.

day side boundary between open and closed field lines. The region containing this plasma is called the magnetospheric cleft or the polar cusp and is shown in Figure 11(a) as a funnel-shaped connection between the magnetosheath and Earth. As

indicated on Figure 11(b) the cleft has a large longitudinal extent adjacent to most of the day-side polar cap boundary. The field lines extending into the plasma sheet are in a similar manner located near the night side polar cap boundary. The observed properties of the plasma in the magnetospheric cleft strongly support the idea that terrestrial field lines there do connect to the solar wind magnetic field. The location of the cleft has also been found to depend on the strength of the north-south component B_z of the interplanetary magnetic field. A strong southward B_z persisting for some time causes an equatorward movement of the cleft as if more terrestrial field lines have been "peeled" off and transported into the tail. This *erosion* of the geomagnetic field on the day side is closely related to B_z : particle observations of position of the cleft show that a persistent 6 γ southward B_z for 45 min is enough to move the cleft 5° equatorward. The amount of magnetic flux added to the tail during that interval can then be estimated to be about 10 percent of the total southward flux impinging on the magnetosphere.

We have discussed how the merging of the geomagnetic field lines with southward-directed interplanetary field lines provides a normal component of the magnetic field across the magnetosphere and therefore a potential difference across the magnetotail. The currents around the tail then tend to accumulate positive space charges along the dawn side of the magnetopause and negative space charges along the dusk side (fig. 3(b)). The resulting electric field drives an electric current from dawn to dusk in the "neutral sheet" and is also responsible for the downtail convection of the newly merged magnetic tubes of force containing magnetosheath plasma. When these field tubes reach the distant tail and meet the corresponding ones from the opposite hemisphere, reconnection is again likely to take place because two plasmas with oppositely directed fields are being pressed together. After the reconnection in the tail, the field tubes are convected back toward Earth because of the northward component across the neutral sheet. During this convective motion, the field lines resume a more dipolar configuration, as they approach Earth, and the kinetic energy of the plasma increases because of increasing magnetic field and progressive shortening of

the field lines. Magnetic energy stored in the stretched-out field in the tail is then converted into kinetic energy of the charged particles. Electrons precipitated into the atmosphere where the field lines from the plasma sheet and the cleft reach Earth cause auroral displays along an oval-shaped belt, the auroral oval, around the magnetic pole. Figure 12 (a) shows a noon-midnight cross section of the magnetosphere indicating the relationship between the auroral oval and the cleft, the plasma sheet, and the outer boundary of the trapping region. The auroral oval is a permanent feature even during extremely quiet conditions. As geomagnetic activity increases, the oval expands away from the pole as seen in figure 12(b). In view of the merging model we would explain this by saying that when more field lines are piled up in the tail and the polar cap therefore is large corresponding to an expanded oval, then the magnetosphere contains more energy and any release of that might result in enhanced geomagnetic disturbance. As we shall see, activity in itself tends to expand the oval further.

SUBSTORMS

At times the flux transport to and back from the tail can take place smoothly and balanced. Fluctuations in B_z are then just manifested as fluctuations in the convection and in particular in the ionospheric electric currents and their magnetic effects. An example of such correlated fluctuations is shown in figure 13(a). There seems to be about 30 min delay in the ionospheric response, which is reasonable for such a large circuit as the magnetosphere. At other times, the response to enhanced tail flux as the result of a steady southward B_z is much more dramatic. Intense magnetic and auroral activity may develop. Figure 13(b) shows a sudden southward turning of the interplanetary field followed by the magnetic signature of enhanced convection. The auroral electrojets were intensified for some time after the southward turning, and just before 7^h UT, magnetograms from auroral zone stations (fig. 14) near local midnight showed a rapid decrease of the horizontal component: a magnetic substorm is now progressing. At the same time a quiet auroral arc along the midnight portion of the auroral oval suddenly brightened and started to

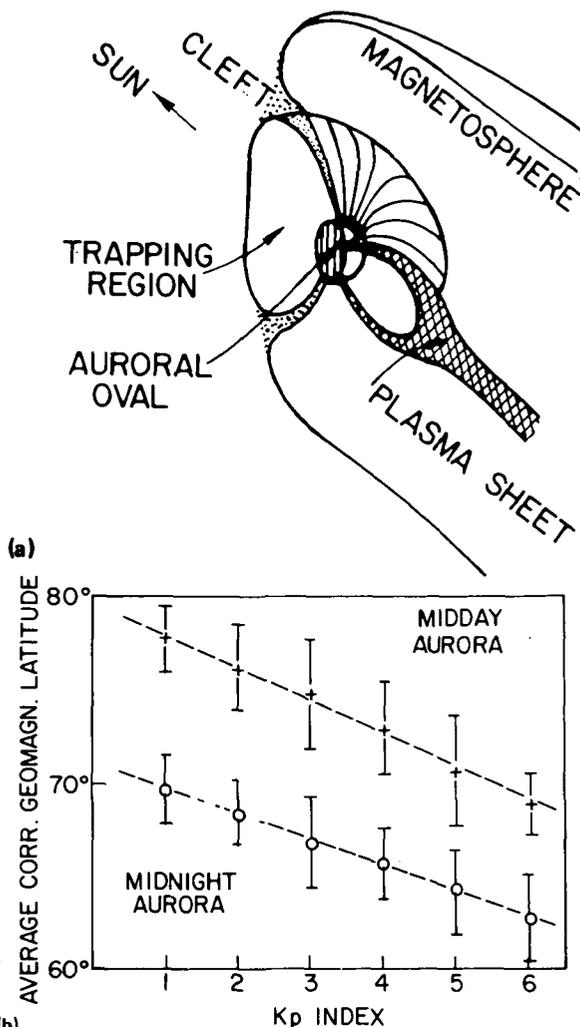
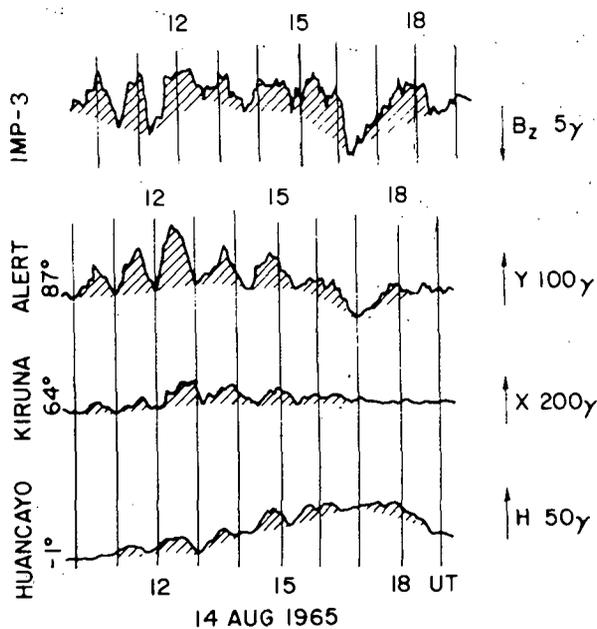


FIGURE 12.—Auroras. (a) Noon-midnight cross section of the magnetosphere showing the auroral oval as the region where the cleft and the plasma sheet intersect the ionosphere. (b) Average corrected geomagnetic latitude of auroras in the midday and midnight parts of the auroral oval as function of geomagnetic activity as given by the K_p index. Both parts of the oval move toward lower latitude as the activity increases.

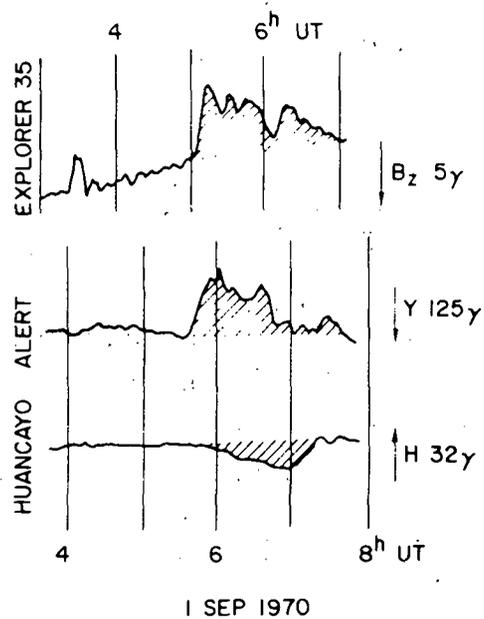
move rapidly poleward while new bright auroral forms were forming behind it. This is the onset of an auroral substorm. We may understand the phenomenon by considering the effect of an increased dawn-dusk electric field due to the increased magnetic flux in the tail. The earthward convection of the plasma in the plasma sheet increases, thereby removing plasma from the sheet in an earthward motion. This progressive thinning of the plasma sheet, together with the added mag-



(a)

FIGURE 13. (a) Coherent fluctuations in the north-south component of the interplanetary magnetic field (as viewed from IMP 3) and in the horizontal component of the geomagnetic field at Alert near the pole (87° corrected geomagnetic latitude), at Kiruna in the auroral zone (64°) and at Huancayo near the equator (-1°). The fluctuations on the ground seem to be delayed approximately 45 min. (This day (August 14, 1965) is also shown in the bottom panel of figure

netic pressure in the tail, increases the reconnection rate drastically with resulting increased plasma flow both toward Earth and also toward the distant tail away from the reconnection point. The process may be described as a local collapse or disruption of the magnetotail current because there is no plasma to carry it. The magnetic configuration in the near-Earth tail changes suddenly to a more dipolar configuration from a stretched "taillike" state. The plasma moving rapidly toward Earth is partly injected into the trapping region and partly spirals down along fieldlines into the auroral oval ionosphere where precipitating electrons cause brilliant, rapidly moving auroras. Thus, the disrupted magnetotail current establishes a new circuit from the dawn side tail to the dawn side auroral oval along the geomagnetic field lines, flows then in the ionosphere to the dusk side oval and finally up to the dusk side magnetotail



(b)

21(b), where fluctuations in the east-west component of the interplanetary magnetic field correlate with fluctuations in the vertical component of the geomagnetic field at Thule (86° after a delay of approximately 30 min.) (b) Response of the geomagnetic field at Alert and Huancayo to a sudden southward turning of the interplanetary field. The responses have the opposite sign of the responses shown in (a) because of the different time of day (about 9^h).

as shown in figure 15. An intense westward current develops in the midnight auroral ionosphere, and the ionization of the ionosphere is greatly enhanced by precipitating plasma particles.

In lower latitudes the magnetic effect of the currents along the field lines is seen as magnetic bays on the magnetograms. Birkeland suggested in 1913 that an intense westward ionospheric current connected via field-aligned currents to a current circuit located at great distance beyond Earth could explain the magnetic variations associated with substorms or "elementary disturbances" as he called them. Recent rocket and satellite observations do indicate that the concept of field-aligned electric currents is fundamental in understanding magnetic substorms: disruptions of the magnetotail divert part of the magnetotail current down through the ionosphere and temporarily relax the load on the magnetosphere con-

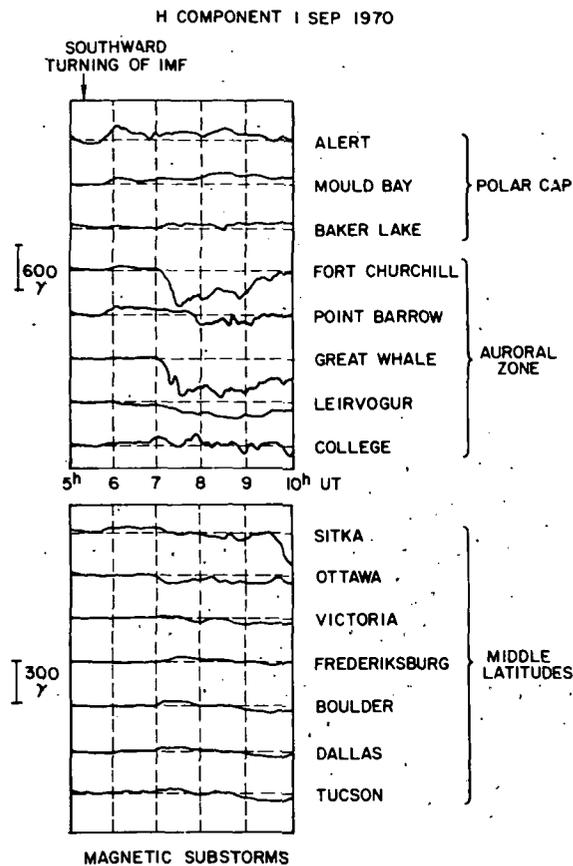


FIGURE 14.—Horizontal component magnetograms from several observatories for the interval following the southward turning of the interplanetary magnetic field shown in figure 13(b). In the polar cap the horizontal component in the direction of the corrected geomagnetic pole is increased after the event. This is indicative of an enhancement of the cross-polar-cap convection. In the midnight sector of the auroral oval (Fort Churchill and Great Whale stations), a magnetic substorm becomes evident at about 7^h UT. At middle and lower latitudes a positive perturbation at the same time is seen at, for example, Boulder and Tucson. The complex variations can be explained as the effects of the (real) current system shown in figure 15. The uniform midlatitude positive perturbation is an indication of eastward current flow at large distances. A disruption (disappearance) of a part of the (westward) magnetotail current is equivalent to temporarily superposing such an eastward current.

verting magnetic energy in the tail to heating and ionization of the upper atmosphere. Often the tail collapse progresses in a stepwise fashion as if several localized disruptions take place successively; the whole process can exhibit extraordinary complexity and diversity with a series of rapidly

moving and very bright looplike auroral displays. The rapid earthward movement of the plasma leads to jetlike injection of hot plasma into the trapping region. This injection may be described as a convection under the influence of an intense induction electric field corresponding to the rapid changes in magnetic configuration when the near-Earth tail field becomes more dipolar.

Once injected, the particles will drift around Earth because of gradient and curvature of the magnetic field. The drift direction depends on the charge of the particles, and electrons tend to move toward the morning side, while protons are drifting toward the evening side as sketched in figure 16(a). The drifting particles constitute a net westward ring current. The magnetic field produced by this current is opposite to the dipole field (see fig. 16(b)) and is observed as a decrease of the horizontal component H at the ground in low and middle latitudes. Furthermore, a strong ring current deforms the magnetospheric field in the trapping region and therefore changes the structure of the inner magnetosphere. In particular, it shrinks the inner radius of the trapping region and shifts the auroral oval toward the equator. The injected particles are rapidly lost again to the atmosphere, partly due to various instabilities as they interact with the plasmasphere. To build up a strong ring current, a number of successive injections is required or, stated differently, a number of substorms must occur in rapid succession.

GEOMAGNETIC STORMS

Identification of the basic magnetospheric processes driven by the continuous and continuously changing solar wind has been the clue to our understanding of the magnetospheric response to the more violent manifestations of solar activity: solar storms. A solar storm starts with a solar flare in magnetically complex active region. Intense X-ray, UV, radio, $H\alpha$, and in rare cases even white light emissions mark the beginning of the storm. The solar atmosphere over the active region is violently disturbed; shock waves are generated and travel through the solar wind plasma, and part of the solar atmosphere is ejected into interplanetary space at high speed. When the shock front reaches Earth, the geomagnetic field

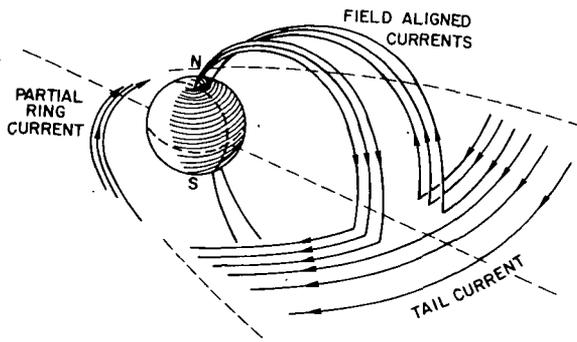
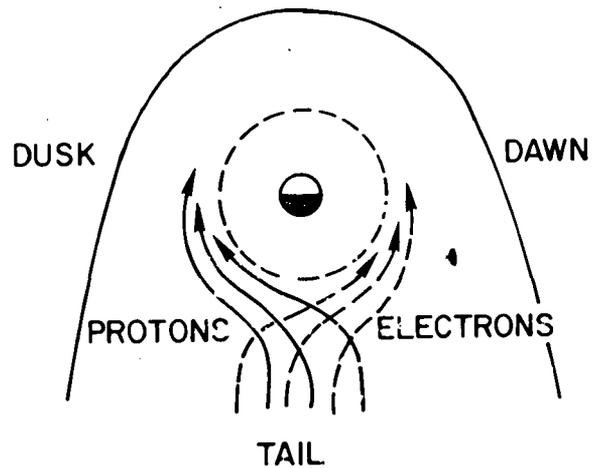
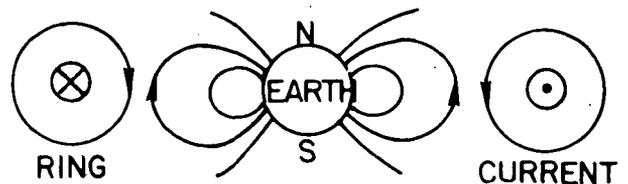


FIGURE 15.—Currents within the magnetosphere during a magnetospheric substorm. The magnetotail current is disrupted and the magnetospheric currents establish a new circuit down the field lines to the ionosphere and back again to the tail. The intensity of the ring current becomes enhanced. There are some indications that currents also flow along field lines from the ring current to the ionosphere (this circuit is not shown in the figure).

is suddenly exposed to a shocked solar wind with increased speed, density, temperature, and magnetic field, resulting in a sudden compression of the magnetosphere. Thus the magnetic field intensity inside the magnetosphere increases suddenly. Ground magnetograms show this sudden storm commencement almost simultaneously over the globe. Figure 17(a) shows the effect of the passing of an interplanetary shock wave where the solar wind pressure increased by a factor of 8 and stayed high for many hours after the shock. The horizontal component at Honolulu increased suddenly by 30γ , maintaining the increase during the initial phase of the storm for about 9 hr. When the shock-driving plasma reached the magnetosphere and the turbulent interplanetary field had developed a strong southward component, the energy input to the compressed magnetosphere increased rapidly by enhanced merging of field lines on the front side. A number of substorms followed in rapid succession, each of them increasing the strength of the ring current, causing the main phase decrease of the field. When the solar wind returns to its quiet state and most of the magnetic energy stored in the magnetotail has been released by the intense substorm activity, the storm enters its recovery phase with the field slowly returning to its normal value. This is because the ring current particles injected into the



(a)



(b)

FIGURE 16. (a) Injection of plasma from the tail into the trapping region. The protons tend to drift westward, while the electrons tend to move eastward. The net result is a westward ring current as shown in panel. (b) The ring current and its magnetic effect, which is opposite the dipole near Earth.

trapping region and compressing the plasmasphere are steadily being lost and the inner magnetosphere is returning to its quiet state as shown in figure 17(b).

Geomagnetic storms show a considerable variety. Some storms have no clear indication of the sudden onset and no initial compression of the magnetosphere but the main phase progresses essentially in the same way as for storms with a sudden storm commencement and a well-developed initial phase. This may be related to the diversity of interplanetary shocks. At times there is no great change in the solar wind pressure across the shock but instead the magnetic field parameters change drastically, or in other cases a rarefaction region follows the shock with resulting expansion of the magnetosphere instead of the usual compression. The geometry of the shock front in connection with the position on the Sun of the solar storm seems to determine the overall structure of

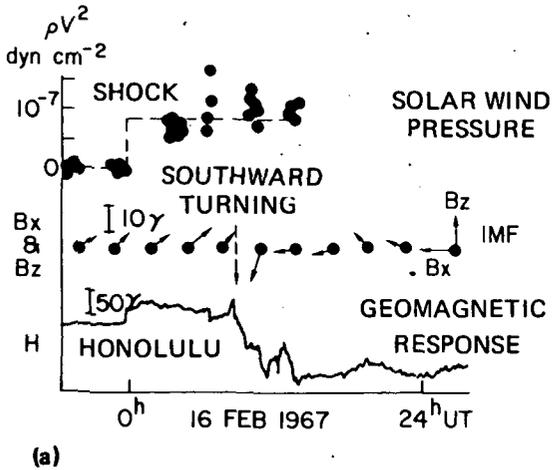
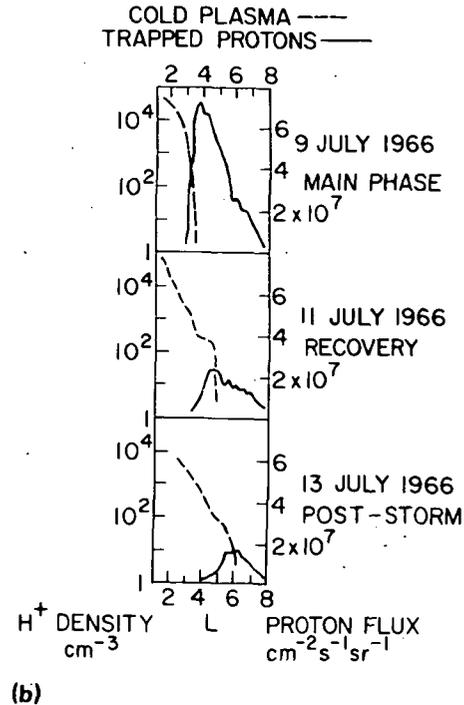


FIGURE 17.—Results of a geomagnetic storm. (a) A geomagnetic storm on February 16, 1967, following an interplanetary shock. The solar wind pressure increased eightfold, compressing the geomagnetic field. The interplanetary magnetic field in the north-south plane is shown in the center panel. After a southward turning of the field the main phase decrease in the horizontal component H at Honolulu is observed. (b) Changes in the size of the plasmasphere (dashed line) and the flux of protons (solid line) in the trapping region during a geomagnetic storm. The H^+ density in the plasmasphere decreases abruptly at a geocentric distance of 3 Earth radii during the main phase, while significant density is found out to more than 5 Earth radii in the poststorm phase. The “ L ”

the magnetospheric storm. Solar storms in the eastern part of the solar disk produce geomagnetic storms with a sudden commencement but not with a large main phase. Western storms cause in general very complicated magnetic storms sometimes with multiple onsets, while storms near the central meridian usually cause typical geomagnetic storms with a well-defined sudden commencement, initial compression phase, and a large main phase decrease. Figures 18 and 19 show further examples of geomagnetic storms. In figure 18 horizontal component magnetograms from low-latitude and auroral zone stations are superposed separately to bring out the difference in the storm morphology in the two regions. The impulsive occurrence of substorms in high latitudes is clearly evident, while sudden storm commencement, a main phase, and the recovery phase can be discerned in the low-latitude records. The figure also illustrates the definition of the D_{st} magnetic



parameter on the abscissa is characterizing the field lines on which the plasma is trapped. For $L = 3$ the field line crosses the geomagnetic equatorial plane at a geocentric distance of 3 Earth radii. High fluxes of trapped protons are found at $L = 4$ during the main phase; later the fluxes are much smaller and have moved out to $L = 6$.

index as the average difference between the actual field and its quiet undisturbed level for the low-latitude stations. The AE index is defined as the field difference between the upper and lower envelopes of the superposed high-latitude records. The variation of these two indices during September 1957 is shown in figure 19. The variability of the low-latitude storm signature D_{st} and the impulsive nature of the high-latitude substorm index AE is evident.

The plasma driving the interplanetary shock is highly turbulent and so, in particular, the north-south component of the interplanetary magnetic field, B_z , is quite irregular both spatially and temporally and may develop quite large southward values. Thus, during the passage of the turbulent plasma, many substorms are expected to occur, especially when the magnetosphere is compressed and the tail field therefore is increased. In the quiet solar wind, the interplanetary magnetic field

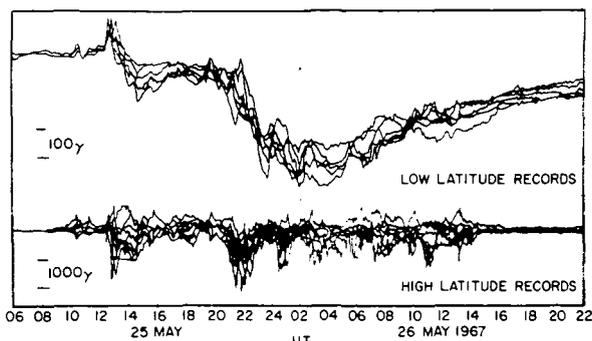


FIGURE 18.—Horizontal component magnetograms for a magnetic storm on May 25 and 26, 1967. The traces are superposed for a number of low-latitude stations and for a number of auroral zone stations separately. The quiet level before the storm has been used as a common zero level. The difference between the actual field intensity and the zero level for the low-latitude stations defines the equatorial ring current index D_{st} . The difference (in gammas) between the upper and lower envelopes of the superposed high-latitude record defines the auroral electrojet index AE.

vector is mainly in the solar equatorial plane and the average B_z is usually small. It is important, however, to note that the dipole axis generally is not perpendicular to the solar equatorial plane but is inclined to it at an angle, which has both diurnal and semiannual variations. Even if the interplanetary field had a constant B_z perpendicular to the solar equatorial plane, there would still be a varying component that was antiparallel to the geomagnetic dipole so that diurnal and semiannual modulations of the field line merging efficiently would be expected. On the other hand, the radially outflowing solar wind forming the magnetosphere aligned with the Sun-Earth line would tend to diminish these modulations. It is at present not clear what the relative importance of all these effects is, but semiannual and diurnal modulation of geomagnetic activity are, in fact, observed.

SECTOR STRUCTURE EFFECTS

While it has long been clear that large geomagnetic storms are closely related to solar storms in conspicuous active regions on the Sun, the solar source of the lesser geomagnetic disturbances is not easily distinguished. The pronounced 27-day recurrence tendency of moderate geomagnetic activity strongly suggests some semipersistent solar regions or features responsible for the activity.

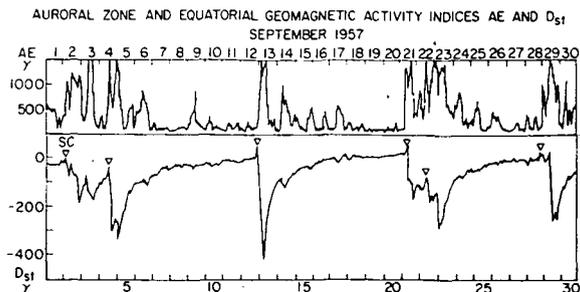


FIGURE 19.—Variations of the AE and D_{st} index during the very disturbed month of September 1957. Sudden storm commencements (SC) are marked by open triangles.

The magnetic field structure in the solar wind also shows marked 27-day recurrence, in some cases for several years. The interplanetary magnetic field tends to be directed predominantly toward or away from the Sun along the basic spiral configuration for intervals of several days at a time. The tendency for these intervals of organized polarity to recur with a period near 27 days has led to the concept of a long-lived interplanetary magnetic sector structure that rotates with the Sun. Regions with opposite polarity are separated by quite narrow sector boundaries that may sweep by Earth in a few minutes. The sector structure implies that the solar wind within each magnetic sector emanated from a coronal region of similarity organized magnetic polarity. Often the solar wind parameters have an organized structure within each sector. The flow speed and the magnetic field strength tend to be low near the sector boundary, rising to a maximum 1 or 2 days after the boundary, and then declining toward the end of the sector. If the sector is very broad, that is, lasting for, say, 14 days, this organized structure may be found twice within the sector, suggesting a time scale of about a week for the basic structure, corresponding to 90° of solar longitude. Near a sector boundary, where the field changes direction, we may expect it to be somewhat disturbed and turbulent, thereby increasing the probability of substorm occurrence or at least of readjustments of the state of the magnetosphere. The increased solar wind speed and the enhanced magnetic field following the sector boundary increase the energy input to the magnetosphere, hence we would expect geomagnetic activity to be organized in a similar manner within a sector. Figure

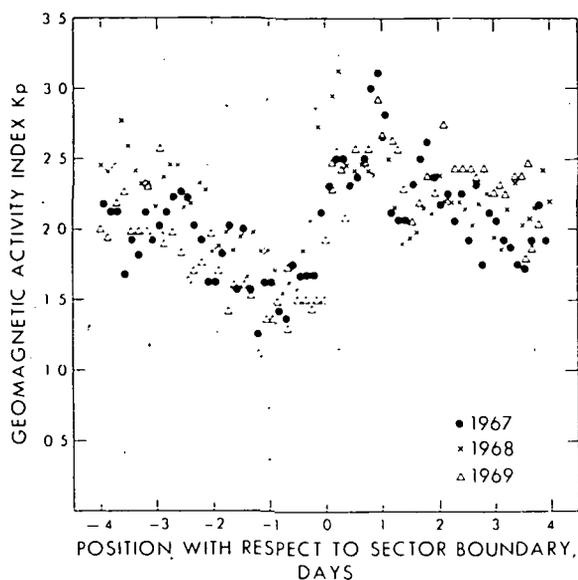


FIGURE 20.—Average response of the geomagnetic activity index K_p to passage of an interplanetary sector boundary. The response is shown separately for three different years as the response averaged for all sector boundaries occurring in each year.

20 shows that this is indeed the case. The geomagnetic field is usually most quiet just before the boundary and increases to a maximum approximately 1 day after the boundary. We therefore identify the source of the long-lived 27-day recurrent geomagnetic activity with the magnetic sector structure and ultimately with the corresponding large-scale organization of the magnetic fields on the Sun.

The direct responsiveness of the magnetosphere to the ever-changing interplanetary magnetic field environment is maybe best illustrated by the recently discovered effect of the east-west or azimuthal component B_y of the interplanetary field on the geomagnetic field at very high latitudes in the heart of the polar caps. The effect is most easily seen in the vertical component Z very near to the magnetic poles. Figure 21(a) shows the average variation during the day of Z at Vostok in the southern and Resolute Bay in the northern polar cap, in both cases about 600 km from the corrected geomagnetic pole. The hourly means of Z are divided into three classes depending on the average value of B_y during the hour. If the east-west component B_y is small, there is very little variation of Z because the two stations are

near the center of the electrojet system, but for nonzero B_y significant perturbations of the vertical component are observed at both stations. The perturbations are of opposite sign when B_y changes sign and are observed in the opposite part of the day in opposite hemispheres. Because positive B_y is associated with sectors with magnetic polarity away from the Sun and negative B_y is associated with toward polarity and because the vertical component is positive when directed toward Earth, we can summarize the effect by noting that central polar cap Z perturbations are predominantly directed away from Earth during sectors with polarity away from the Sun, and toward Earth during sectors with magnetic polarity directed toward the Sun. From Figure 21(b) it may be seen that this remarkable correlation is not only seen in a statistical sense for long-period variation but also extends to individual fluctuations as short as 30 min or less during the interval 10^h to 22^h UT.

A note about coordinate systems: The x -axis points toward the Sun. In magnetospheric coordinates the xz plane contains the geomagnetic dipole. In ecliptic coordinates the xy plane contains the ecliptic. The third axis completes the normal right-handed orthogonal system. When discussing the interaction with the magnetosphere, the interplanetary magnetic field is normally expressed in magnetospheric coordinates. For our purpose the distinction is not important.

There seems to be a delay of about 20 min before the response of the polar cap field. The figure clearly demonstrates that the sector structure may exhibit a high degree of variance and that the polar cap Z component responds to variations of the sector structure on a time scale of a few tens of minutes.

Further analysis of this response has shown that at a somewhat larger distance from the magnetic poles the horizontal components begin to respond to variations of B_y . The effects can be described as the magnetic effects of an ionospheric current flowing around the magnetic pole at a corrected geomagnetic latitude of 80° to 82°, as indicated on figure 22. The sense of the current is clockwise for negative B_y and counterclockwise for positive B_y . Passage of a sector boundary thus causes an abrupt reversal of the current.

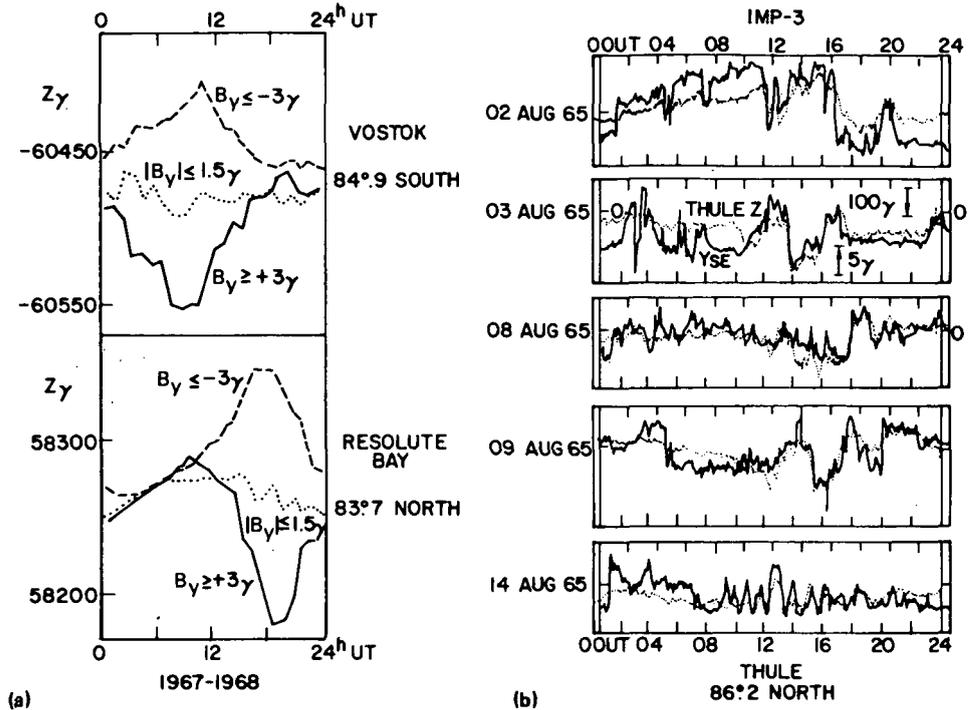


FIGURE 21. (a) Diurnal variation of the vertical component Z at Vostok and Resolute Bay during 1967 and 1968. All hours where the hourly average of the interplanetary east-west component (solar magnetospheric coordinates) B_y was less than -3γ were averaged for each UT hourly interval to yield the dashed curves. When B_y is greater than $+3\gamma$, the solid curves result, while the dotted curves were computed for times where B_y was near zero ($|B_y| \leq 1.5\gamma$). (b) Corresponding fluctuations of the Z component at Thule (dotted trace plotted positive downwards) and the east-west component (solar ecliptic coordinates) Y_{SE} of the interplanetary magnetic field (solid trace). The fluctuations are well correlated in the interval 10^h to 24^h UT with the fluctuations on the ground delayed about 25 min.

The physical reason for the existence of this polar cap current is presumably some modification of the convection pattern caused by the azimuthal component of the interplanetary field, but no clear picture of the precise nature of the effect and of its mechanism has emerged yet. One thing is, however, clear, namely that the magnetosphere is directly affected by the interplanetary field; the existence of this response is also a good indication that geomagnetic and interplanetary field lines are connected.

CONCLUDING REMARKS

A tremendous advance in our understanding of the properties of the solar wind and its interaction with the terrestrial environment has been achieved

in recent years through intensive observational and theoretical programs. Enough observational evidence has been in hand to guide the theory along realistic paths, and enough theory has been developed to interpret data that are characteristically incomplete in coverage. The explorative phase of magnetospheric research is coming to an end, and the basic magnetospheric processes are identified. The basic structure of the magnetosphere—the bow shock, the magnetosheath, the magnetopause, and the magnetotail—has been unveiled. The importance of the continuous interaction between the solar wind and the magnetosphere is realigned and the concept of the magnetospheric substorm constitutes a basic framework for our understanding of the major disturbances within the magnetosphere.

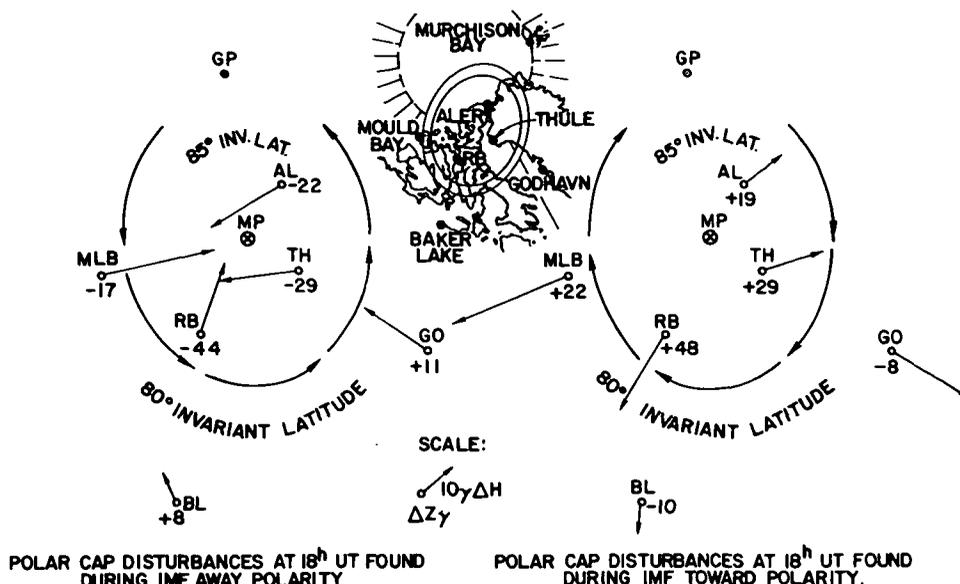


FIGURE 22.—Typical polar cap magnetic disturbances observed for the two opposite polarities of the east-west component B_y of the interplanetary magnetic field. Two synoptic maps are shown with disturbance vectors corresponding to positive B_y (normally within the “away” sector) at the left and to negative B_y (“toward” sector) at the right. The vectors showing the horizontal perturbations are drawn from the positions of each six northern polar cap stations. An insert shows the geographical locations of these stations. Signed numbers next to the station circles denote the Z perturbations. The positions of the geographical pole (GP) and of the corrected magnetic pole (MP) are indicated. Parts of equivalent currents that could produce the magnetic variations are sketched. The perturbations (and the current) reverse when B_y reverses sign.

The interplanetary magnetic field—although having an energy density two orders of magnitude less than the solar wind plasma—is essential in controlling the solar wind interaction with Earth. It gives the collisionless plasma fluid properties over scale lengths comparable to (or less than) the size of our planet. The interplanetary field connects with the geomagnetic field to provide efficient solar wind/magnetosphere coupling to drive the magnetospheric dynamo. Solar wind kinetic energy is then converted into magnetic energy stored in the magnetotail. Instabilities in the system release part of the stored energy and convert it into kinetic energy of magnetospheric plasma particles. The upper atmosphere acts as a sink for this kinetic energy as it is converted into radiation and heating.

ACKNOWLEDGMENTS

I thank John M. Wilcox for critical reading of the manuscript. This work was supported in part

by the Office of Naval Research under Contract N00014-67-A-0112-0068, by the National Aeronautics and Space Administration under Grant NGR 05-020-559, and by the Atmospheric Sciences Section of the National Science Foundation under Grant GA-31138.

APPENDIX—ESTIMATES OF SOME RELEVANT PHYSICAL QUANTITIES FOR THE SOLAR WIND INTERACTION WITH THE GEOMAGNETIC FIELD

The electromotive force, $\epsilon \sim \mathbf{w} \times \mathbf{b}_n$, supplied by the solar wind to the magnetospheric dynamo is of the order

$$\epsilon = w b_n$$

where w is the solar wind speed. The normal component b_n of the magnetic field connecting the magnetospheric tail and the interplanetary field can be estimated by assuming that the magnetic flux M_p from the polar cap is connected to the

interplanetary field along the surface A_T of the tail. With a polar cap radius r_P and a polar cap field B_P , we get

$$M_P = \pi r_P^2 B_P$$

Taking the length of the tail as S_T , we have

$$A_T = \pi R_T S_T$$

where R_T is the radius of the tail. Hence,

$$\begin{aligned} b_n &= \frac{M_P}{A_T} \\ &= \frac{r_P^2 B_P}{R_T S_T} \end{aligned}$$

With $r_P = 15^\circ = 1.7 \times 10^6$ m, $B_P = 55\,000\gamma = 0.055 \times 10^{-4}$ Wb/m², $R_T = 20R_E = 1.3 \times 10^8$ m, and $S_T = 500R_E = 3.2 \times 10^9$ m, we get $b_n = 3.7 \times 10^{-10}$ Wb/m² = 0.37γ . One Earth radius is $R_E = 6.38 \times 10^6$ m. Taking the solar wind speed as $w = 420$ km/s = 4.2×10^5 m/s, we find

$$\varepsilon = 1.6 \times 10^{-4} \quad \text{V/m}$$

The total potential difference across the tail then becomes

$$\begin{aligned} \Phi &= \varepsilon \pi R_T \\ &= 6.4 \times 10^4 \quad \text{V} \\ &= 64 \text{ kV} \end{aligned}$$

and the electric field in the polar cap is

$$\begin{aligned} E_i &= \frac{\Phi}{2r_P} \\ &= 20 \times 10^{-3} \quad \text{V/m} \\ &= 20 \quad \text{mV/m} \end{aligned}$$

We can also write

$$\begin{aligned} \Phi &= w b_n \pi R_T \\ &= \frac{w M_P \pi R_T}{A_T} \\ &= \frac{w M_P}{S_T} \end{aligned}$$

The field strength in the near Earth tail (before too much flux has leaked out) can be estimated to be

$$\begin{aligned} B_T &= \frac{M_P}{\frac{1}{2}\pi R_T^2} \\ &= 2B_P \frac{r_P^2}{R_T^2} \\ &= 19 \times 10^{-9} \quad \text{Wb/m}^2 \\ &= 19 \quad \gamma \end{aligned}$$

The typical quiet time convection velocity over the polar cap can be obtained from

$$v_c = \mathbf{E} \times \frac{\mathbf{B}}{B^2}$$

as

$$v_c = \frac{E_i}{B_P} = 360 \quad \text{m/s}$$

The time to convection the footpoints of the tail field lines across the polar cap is now

$$\begin{aligned} t_c &= \frac{2r_P}{v_c} \\ &= 9250 \text{ s} \\ &= 2.6 \text{ hr} \end{aligned}$$

In that time the interplanetary end of the field line moves wt_c which then is also an estimate of the length of the tail:

$$\begin{aligned} S_T &= wt_c \\ &= \frac{w^2 r_P B_P}{E_i} \\ &= 3.8 \times 10^9 \quad \text{m} \\ &= 600R_E \end{aligned}$$

For a line current (auroral electrojet) at height h over the ground to give a magnetic substorm effect of $B_A = 1000\gamma = 10^6$ Wb/m², the current strength must be of the order

$$i_A = \frac{2\pi h B_A}{\mu_0}$$

Taking $h = 110$ km = 1.1×10^5 m, we get $i_A = 550\,000$ A. If n_T is the current density of the tail current estimated by treating each half of the tail as a solenoid: $n_T = B_T/\mu_0$, we find that the extent of the tail current disruption is of the order of

$$\begin{aligned} k_d &= \frac{i_A}{n_T} \\ &= 3.7 \times 10^7 \quad \text{m} \sim 6R_E \end{aligned}$$

Assuming that the energy in this part of the tail was stored as magnetic energy, we get

$$\begin{aligned} U_d &= \frac{B_T^2}{2\mu_0} (\text{volume}) \\ &= \frac{B_T^2}{2\mu_0} \frac{\pi R_T^2}{2} k_d \\ &= \frac{B_T \pi R_T^2 i_A}{4} \end{aligned}$$

But we have also $U_d = \frac{1}{2} Li^2$, so the inductance

of the circuit becomes

$$L = \mu_0 \frac{B_T}{B_A} \frac{R_T}{4h} = 890 \text{ H}$$

The resistance R in the circuit is essentially that of the ionosphere: $R = \Phi/i_A = 0.12 \Omega$, so the time constant of the circuit can be estimated as

$$\begin{aligned} t &= \frac{L}{R} \\ &= 7.4 \times 10^3 \text{ s} \\ &\approx 2 \text{ hr} \end{aligned}$$

This shows us that the magnetotail certainly contains enough energy to drive a substorm which lasts, say, 1 hr. The energy dissipated in the ionosphere alone by the substorm current is of the order

$$\begin{aligned} P &= i_A \Phi \\ &= 3.5 \times 10^{10} \text{ W} \end{aligned}$$

Taking into account also the current in the southern hemisphere, we get a total rate at which work is being done of the order of 10^{11} W. If the substorm lasts for 1 hr, the total amount of energy dissipated in the currents is then about 3×10^{14} J. The additional energy deposited in the auroral substorm by the precipitating electrons can be estimated from the auroral luminescence and is about 2×10^{14} J. Therefore the total substorm energy dissipation amounts to 5×10^{14} J corresponding to an earthquake of magnitude 6.7 on the Richter scale.

We can estimate the total magnetotail current J_T by setting the average magnetic field in the tail to $B_T/2$. We do this because the field decreases down the tail as more and more field lines are connected to the solar wind and leak out of the tail (see fig. 3(b)). Hence the average current density:

$$\begin{aligned} n_T &= \frac{1}{2} n_T \\ &= \frac{B_T}{2\mu_0}, \end{aligned}$$

so that

$$\begin{aligned} J_T &= J_{\text{northern}} + J_{\text{southern}} \\ &= 2S_T n_T \\ &= \frac{S_T B_T}{\mu_0} \\ &= 5 \times 10^7 \text{ A.} \end{aligned}$$

The total amount of energy drawn from the solar

wind by the current J_T over a potential difference Φ is then

$$\begin{aligned} P_s &= J_T \Phi \\ &= 3 \times 10^{12} \text{ W} \end{aligned}$$

The energy deposited in a substorm corresponds to about 2 min of solar wind input. We see that substorms are not major collapses of the magnetosphere but rather have the character of minor internal adjustments to changing external conditions.

The kinetic energy of the solar wind falling on the magnetosphere is essentially

$$K = \pi R_T^2 w (\frac{1}{2}) n m_p w^2$$

where $m_p = 1.67 \times 10^{-27}$ kg is proton mass and $n = 5$ protons/cm³ = 5×10^6 m⁻³ is the number density. We find $K = 1.6 \times 10^{13}$ W, which is 5 times the energy in the magnetotail. From energy considerations, the solar wind thus seems capable of driving the magnetospheric dynamo and maintaining the magnetotail.

BIBLIOGRAPHY

- Magnetospheric research has been a very rapidly developing field during the last decade or more. Some useful standard works on the subject are listed below in decreasing order of obsolescence.
- Chapman, S., 1964, *Solar Plasma, Geomagnetism and Aurora*, Gordon & Breach Sci. Pub.
- Matsushita, S., and W. H. Campbell, eds., 1967, *Physics of Geomagnetic Phenomena*, vol. II, Academic Press, Inc.
- Carovillano, R. L., et al., eds., 1968, *Physics of the Magnetosphere*, D. Reidel, Dordrecht, Holland.
- Williams, D. J., and G. D. Mead, eds., 1969, "Magnetospheric Physics," *Rev. Geophys.*, 7, p. 1.
- McCormac, B. M., ed., 1970, *Particles and Fields in the Magnetosphere*, D. Reidel, Dordrecht, Holland.
- McIntosh, P. S., and M. Dryer, ed., 1972, "Solar Activity Observations and Predictions," *Prog. Astronaut. Aeronaut.*, 30.
- Dyer, E. R., et al., eds., 1972, *Solar-Terrestrial Physics/1970*, D. Reidel, Dordrecht, Holland.
- Poeverlein, H., 1972, "The Earth's Magnetosphere," *Handbuch der Physik*, vol. XLIX/4, ed., S. Flügge, Springer Verlag, Berlin-Heidelberg-New York.
- Akasofu, S. I., and S. Chapman, 1972, *Solar-Terrestrial Physics*, Oxford University Press, Inc.
- Dyer, E. R., ed., 1972, *Critical Problems of Magnetospheric Physics*, National Academy of Sciences, Washington, D.C.

DISCUSSION

SCHMERLING: I am having some difficulty bridging the sharp discontinuity between one speaker and another, and I wonder if somebody can help me by providing a 1 AU matching transform. In particular, what bothers me is that in one view—and that is primarily the view of the Sun in the interplanetary medium—what is important is the field structure in the ecliptic plane, and what appears to be important for triggering some of the terrestrial events is whether the field, as it arrives at Earth out of the ecliptic plane, is north-south or south-north. More specifically, I can look at that picture you have drawn on the board and imagine that with precisely the same kind of ecliptic plane projection I can have north-south or south-north fields, depending simply on whether some of the structure is a little bit above or a little bit below the ecliptic plane.

SVALGAARD: Part of the answer is that the important thing is the fluctuation of the field. A field line is not really like a straight line; it is wiggling all around. And so, as seen from Earth, that field line is carried past us, and it appears as a wiggly line that changes direction—it runs east, it runs west, it runs north, and it runs south. And when it “decides” to go southward, the energy input to the magnetosphere, because of the connection of the field lines across the magnetopause, goes up, and if it is fluctuating enough, then it goes southward a lot and you have a lot of input to the magnetotail.

MANKA: It seems to me that you discussed a lot of mechanisms that might provide energy input, ultimately, into the atmosphere. If the solar wind flow velocity is related to position in the sector structure, there is a direct plasma energy input and then you have a magnetic connection.

It seems to me that, in a sense, basically you are dealing with whether it is a plasma energy input, ultimately, or a field input. You also have the interplanetary electric field that will be related to the magnetic field strength and the flow velocity. When it gets to Earth, if the interplanetary electric field creates polarization and crosstail field, then you could convert that field energy

into a plasma input via currents down the field lines, or we may have the magnetospheric electric field itself mapping down the magnetic field lines and then driving currents in the atmosphere.

So it seems to me that a possible approach might be to try to track through the sequences and see whether it is the field or the plasma which is, in a sense, the cause, and which is the effect. Do you have any feel for this? Which of these processes might dominate? Which one might be a key one in relationship to the magnetic sector structure?

SVALGAARD: That is a difficult question to answer straightaway, but I think that (to be very brief) the kinetic energy of the solar wind plasma is, via this reconnection, stored up as magnetic energy in the tail, and then instabilities in the tail sooner or later release that energy, and so we have a conversion of plasma kinetic energy into magnetic energy, and then later from that magnetic energy again into plasma energy. It is that latter plasma that has the effect on Earth. There is very little solar plasma that comes directly from the solar wind and goes directly down, down to the ground.

So one could say that the solar wind acts from the Sun on the sunward side of Earth, but then it is the tail that really gives the action on the night side, and I think the crucial thing here is to note that the energy is stored up in stretched magnetic field lines of the tail, and that stretching out is presumably done by the magnetic field of the solar wind.

MARKSON: There have been studies that indicate that, on one hand, the Moon's position may have something to do with weather, and also that the Moon's position may have something to do with geophysical parameters, such as Stolov's studies relative to the position from the ecliptic. I wonder if you could comment on how important this might be and how it might happen.

SVALGAARD: The Moon passes through the tail, and, therefore, might upset the balance in the tail. However, the tail is extremely large and the Moon is very small, and I think the consensus right now is that the Moon has very little, if any, effect at all. Maybe in another 55 yr or so there will be a conference on lunar influences on the weather!