Global Magnetospheric Plasma Convection

Stefan Eriksson
The cover illustration shows the Weimer model electric potential pattern in the ionosphere for southward interplanetary magnetic field conditions. Two plasma convection cells appear, one in the dawnside and one in the duskside ionosphere, with antisunward plasma drift over the polar cap regions due to the coupling of the geomagnetic field with the interplanetary magnetic field in the solar wind [Weimer, 2001].
To my parents, my brother, and Yurika.
Abstract

This thesis deals with the global aspects of plasma convection in the magnetosphere as measured by the low-altitude polar orbiting Astrid-2 and FAST satellites. The major focus is on the electric field measurements, but they are also complemented by magnetic field, ion and electron particle data, which is fundamental for the understanding of the electrodynamics of the high-latitude auroral ovals and polar cap, which are the regions analysed here. The essential subject of this thesis is the so-called magnetic reconnection process that drives plasma convection in the Earth’s magnetosphere. It is shown that the ionospheric convection, being intimately coupled to the magnetospheric convection, responds in about 15-25 min depending on geomagnetic activity after the arrival of the solar wind at the magnetopause. It also responds on a longer time scale, around 55-75 min, which is interpreted as the unloading of solar wind energy previously stored in the large-scale current system of the magnetotail. These results have been found previously using ionospheric parameters such as the auroral electrojet AL index. What is new is that these same results are reproduced by using a discrete set of cross-polar potential measurements. Using an extensive set of electric and magnetic field data combined with particle precipitation data from the FAST satellite, it is shown that the reconnection process can also be applied to explain features of sunward plasma convection in the polar cap with a likely antiparallel merging site in the lobe magnetopause region. The lobe reconnection is found to depend strongly on IMF $B_y$ and to coexist with dayside subsolar merging. Finally, a comparison is performed between the Weimer electric field model and Astrid-2 electric field data. Empirical electric field models are important in understanding the complete convection pattern at any one time, something, which cannot be provided by measurements from single satellites.

Keywords: Satellite measurements, electric fields, magnetosphere, magnetic reconnection, plasma convection, lobe cell convection, empirical electric field models.
List of Contents

List of Papers .................................................................................................................. i
1. Introduction ................................................................................................................ 1
2. Space Plasma Physics ............................................................................................... 2
3. Sun-Earth Environment ............................................................................................ 9
4. Magnetospheric Response to the Solar Wind ......................................................... 19
5. Lobe Cell Convection ............................................................................................... 23
6. Empirical Electric Field Models .............................................................................. 24
7. Electric Field Measurements .................................................................................... 26
8. Summary of Papers .................................................................................................. 28
9. References ................................................................................................................ 30
Acknowledgements ......................................................................................................... 34
Appendix A: Earth’s Geomagnetic Field .................................................................... 35
Appendix B: Papers 1-4 ............................................................................................... 36
List of Papers

This thesis consists of the following papers authored by Stefan Eriksson, listed in chronological order:


These papers are referred to in the thesis as “Paper 1” – “Paper 4.”
1. Introduction

This study concentrates on the response of the global magnetospheric convection to the solar wind, using electric field data from the FAST and Astrid-2 satellites. The global convection of plasma is an important element for the understanding of how flows in the magnetotail couple to the ionosphere and also for the understanding of the highly dynamic processes occurring during magnetic storms and substorms that bring plasma from the distant tail into near-Earth regions, such as the ring current. It is believed that the process of magnetic reconnection is the main driver for this convection, mainly concentrated to the dayside magnetopause. However, it seems that the same process is occurring in the lobes of the magnetotail. An analysis of the electrodynamic state of the ionosphere is performed during lobe reconnection using data from the FAST satellite. Single satellite measurements of the electric field along their orbits are clearly insufficient for mapping of the complete ionospheric convection at any one time, which is why several empirical models have been developed to bridge these gaps. Astrid-2 data is employed to examine the degree of correspondence between one such model, the Weimer model, and the in situ measured electric field.

The thesis is separated in two introductory sections, where the plasma state is defined and further put into the context of fluid theory (section 2). The second part introduces the Sun and the Earth environments and how these two systems are believed to interact with each other (section 3) using the concepts introduced in section 2. Then follows three sections, which are directly connected with the material of the four papers and appearing in respective order. A brief introduction to electric field measurements is finally given before the papers are summarized.
2. Space Plasma Physics

Space plasma physics is a field of study, which concerns the physics of the near-Earth space plasma environment that is accessible for in situ measurements by space probes. This environment, including the solar wind, the Earth’s magnetosphere, and ionosphere as well as the plasma environment around other solar system bodies, forms a gigantic plasma laboratory in which the plasma density and temperature vary over many orders of magnitude. But what is a plasma?

2.1 Definition of Plasma

Plasma is loosely speaking the name given to the fourth state of matter. As the temperature of a neutral gas increases, collisions between atoms and molecules alike will separate particles into negatively charged electrons and positively charged ions. The neutral gas becomes ionized to some degree. As long as the gas is able to remain in a state consisting of free charges of some density, we refer to it as a plasma. A plasma is then usually defined as a quasi-neutral ionized gas, with approximately equal electron and ion densities, $n_e=n_i$. It also exhibits a collective behavior, such as the ability to carry electric currents. A fundamental property of a plasma is its capability to shield out a locally applied electric potential from the external plasma with the degree of shielding depending on the plasma density and temperature. A positive electric potential applied to a cold plasma with negligible thermal motion would attract negative charges and repel positive charges until the external plasma would see no charge imbalance. In a finite temperature plasma, however, the thermal motion of particles surrounding a positive potential is sufficient for some particles to escape and the shielding becomes imperfect. These particles are located at a distance from the potential where the average electron thermal energy approximately equals the electrostatic potential energy. This characteristic length is called the Debye length,

$$\lambda_D = \sqrt{\frac{\varepsilon_0 k_B T}{n_e e^2}}$$

(1)

where $\varepsilon_0$ is the permittivity of free space, $k_B$ the Boltzmann constant, $e$ is the elementary charge, and $T$ and $n_e$ are the electron temperature and density, respectively. A necessary condition for local concentrations of charges and applied electric potentials to be shielded out is that the physical dimension $L$ of the system satisfies $L>>\lambda_D$, and that there are enough particles present in a volume of radius $\lambda_D$. Finally, in order for the gas to qualify as a plasma, it is required that the number of collisions between the charged and neutral particles are so few that the motion of the charged particles is dominated by electromagnetic forces.

Plasmas are found in most parts of the universe, such as the surrounding of the planets of the solar system and the Sun, and also in interplanetary and interstellar space as a very tenuous plasma. The Earth’s surface and lower atmosphere are but a few places where plasmas do not exist naturally. Understanding the behaviour of space plasmas thus becomes of increasing concern to us as humanity endeavours into space.
2.2 Particle Dynamics - Single Particle vs MHD Description

The charged particles of plasmas are predominantly affected by electric and magnetic forces. A particle of charge \( q \) and velocity \( v \) in an electric field \( \mathbf{E} \) and magnetic field \( \mathbf{B} \) will experience the so-called Lorentz force, \( \mathbf{F} = q\mathbf{E} + q\mathbf{v} \times \mathbf{B} \). The Lorentz force acting on a particle of mass \( m \) leads to the equation of motion,

\[
m \frac{dv}{dt} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) + \mathbf{F}
\]

where \( \mathbf{F} \) symbolizes additional forces such as gravity, which in most situations is negligible.

A particle in a homogeneous and static magnetic field with \( \mathbf{E}=0 \) will move in a circular orbit perpendicular to \( \mathbf{B} \), since the force \( q\mathbf{v} \times \mathbf{B} \) acts perpendicularly to both \( \mathbf{v} \) and \( \mathbf{B} \). The radius of the circular orbit may be calculated as,

\[
r_c = \frac{v_i}{\omega_c} = \frac{mv_i}{qB}
\]

which is referred to as the Larmor- or gyro-radius, where \( \omega_c \) is the particle gyro-frequency. Since any motion along \( \mathbf{B} \) is unaffected by the \( q\mathbf{v} \times \mathbf{B} \) force and remains constant, the complete trajectory of a charged particle in absence of electric fields will take the form of a helix.

A finite electric field \( \mathbf{E} \) parallel to \( \mathbf{B} \) will accelerate particles of opposite charge in opposite directions along \( \mathbf{B} \), until the secondary electric field set up by the charge separation itself cancels \( \mathbf{E} \), assuming there are enough particles present.

An electric field \( \mathbf{E} \) perpendicular to \( \mathbf{B} \) introduces a drift velocity \( \mathbf{u} \) of the gyrating particles as follows. The total particle velocity \( \mathbf{v} \) may be considered a superposition of two velocities, \( \mathbf{v} = \mathbf{v}_c + \mathbf{u} \), where \( \mathbf{v}_c \) represents the circular motion around \( \mathbf{B} \). If we assume that \( \mathbf{u} \) on the average is constant over one gyro-period or that \( du/dt \) is negligible, the equation of motion (Eq.2) becomes,

\[
0 = \mathbf{E} + \mathbf{u} \times \mathbf{B}
\]

Taking the cross-product of Eq.4 with \( \mathbf{B} \) results in an expression for the drift velocity \( \mathbf{u} \) perpendicular to \( \mathbf{B} \) as introduced by \( \mathbf{E} \),

\[
\mathbf{u} = \frac{\mathbf{E} \times \mathbf{B}}{B^2}
\]

The effect of the electric field is to accelerate the charged particles during each half gyro-period, thereby increasing \( \mathbf{v}_c \) and the gyro-radius \( r_c \), only to decelerate the particles during the second half of each gyro-period, thus decreasing \( \mathbf{v}_c \) and \( r_c \). The center of gyration for ions and electrons will move in the same direction with the same drift velocity \( \mathbf{u} \) across the magnetic field, often referred to as the \( \mathbf{E} \times \mathbf{B} \)-drift.

An inhomogeneous magnetic field with a gradient \( \nabla B \) perpendicular to \( \mathbf{B} \) introduces a drift of the gyro-center across the magnetic field as well. Since a charge moving in a magnetic field is equivalent to a current \( I \), it sets up a magnetic moment \( \mathbf{\mu} \) directed antiparallel to \( \mathbf{B} \). The plasma is therefore a diamagnetic medium, which acts to reduce the local magnetic field. The magnitude of \( \mathbf{\mu} \) may be derived as,

\[
\mu = \frac{mv_c^2}{2B}
\]
which is practically a constant as long as the magnetic field or the perpendicular kinetic energy of the charged particle changes slowly over a gyro-period. The diamagnetic moment sets up a force,
\[ F = -\mu \nabla B \]  
and an expression for the drift velocity perpendicular to \( B \) due to it is found by replacing \( E \) by \( F/q \) in Eq. 5. The so-called \( \nabla B \)-drift is then expressed as,
\[ \mathbf{u} = \frac{\mu \mathbf{B} \times \nabla B}{qB^2} \]  
from which an electric current arises, since ions and electrons drift in opposite directions. The two plasma drifts of Eqs. 5 and 8 have been used successfully in describing the global motion of plasma in the Earth’s magnetosphere. The \( \mathbf{E} \times \mathbf{B} \)-drift dominates for low-energy plasmas, while the \( \nabla B \)-drift control the particle motion for higher energies (>10 keV).

Keeping track of separate particle orbits in a plasma is often impossible in the presence of electric and magnetic fields. Additional collective effects, such as currents and particle pressure in the plasma, set up forces that act on spatial scales longer than any individual particle gyro-orbit. The collective behaviour of charged particles is instead analysed using the statistical concept of fluid dynamics. Here, the motion of single particles is replaced by the average motion of many particles in a finite volume element of plasma. Additional quantities such as density, average velocity, temperature, and pressure are introduced as averages taken over the volume element, which are governed by the laws of conservation of mass, momentum, and fluid energy. Unlike the neutral fluid treatment of statistical mechanics, the plasma is a fluid consisting of charged particles and as such depends on the electric and magnetic fields, as we saw from the drift velocities (Eqs. 5 and 8). For example, the motion of a conducting fluid across a magnetic field sets up an electric field that drives currents in the plasma. The current in turn modifies the magnetic field, which affects the original motion. The Maxwell equations describe this coupling between charged particle motions and electromagnetic fields and must therefore be taken into account. The complete set of equations that describe the macroscopic fluid motion of the electrically conducting plasma in presence of electromagnetic fields is referred to as magnetohydrodynamics, or MHD for short. Many large-scale phenomena in space plasmas may be understood in the framework of MHD, such as the solar wind, and the motion of plasmas inside planetary magnetospheres (c.f. section 3.0).

An MHD plasma is assumed to consist of only electrons and ions of one species that forms two sets of fluid equations. These equations may be reduced to the “single-fluid” equations. First, requiring conservation of mass leads to the equation of continuity, 
\[ \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \]  
where \( \rho = n_i m_i + n_e m_e \) and \( \mathbf{u} = (n_i m_i \mathbf{u}_i + n_e m_e \mathbf{u}_e) / \rho \) corresponds to the center of mass velocity of the plasma element. The momentum fluid equation is retrieved by adding the ion and electron equations of motion (Eq.2). Assuming quasi-neutrality and negligible collisions with neutrals we get,
\[ \rho \frac{d\mathbf{u}}{dt} = \mathbf{j} \times \mathbf{B} - \nabla p \]
where $j = n_i q_i u_i + n_e q_e u_e$ is the current density and $\nabla p$ is the pressure force on the plasma element due to thermal motions of the particles.

The Maxwell equations consist of,

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad \text{(Faraday’s law)} \quad (11)$$

$$\nabla \times \mathbf{B} = \mu_0 j + \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} \quad \text{(Ampère’s law)} \quad (12)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (13)$$

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0} \quad \text{(Poisson’s equation)} \quad (14)$$

where $\mu_0$ is the permeability of free space and $c$ is the speed of light in vacuum. The so-called displacement current (the second term on the right-hand side of Ampère’s law) is usually considered negligible for most MHD applications. If charge neutrality holds, then the space charge $\rho_q = n_i q_i + n_e q_e$ does not enter the equations. This implies that $\nabla \cdot j = 0$, which describes a system of closed currents. Two more equations are needed to close the system and these are usually given by Ohm’s law,

$$j = \sigma (\mathbf{E} + \mathbf{u} \times \mathbf{B}) \quad (15)$$

where $\sigma$ is the electrical conductivity, and a thermodynamic equation of state which relates the pressure $p$ with mass density $\rho$.

For a fluid description to be valid, we require that all particles present in a fluid element at time $t$, must also be present at a later time $t+T$. In a regular neutral fluid, such as water or a neutral gas, the particles are attracted to each other by collisions with neighbouring particles. In a fully ionized plasma, particles interact over much longer distances due to the long range Coulomb electric field force, which acts to scatter the particles rather than keeping them fixed to the fluid element. In a plasma, it is the magnetic field that serves as the attracting “glue,” at least in the direction perpendicular to $\mathbf{B}$. This leads to the following restrictions on the plasma for MHD to hold. With $L_\perp$ and $T_\perp$ being the characteristic length and time over which the fields and plasma change, then $L_\perp >> r_c$ and $T_\perp >> T_c$, where $r_c$ and $T_c$ correspond to the ion gyro-radius and period, respectively. These conditions imply that viscous friction and heat-transfer effects perpendicular to $\mathbf{B}$ are small. The temperature must therefore be low enough so that the plasma flow velocity is larger than the thermal velocity. The resistivity due to thermal motions is then so low that currents may flow unimpeded. In the limit of no resistivity, or infinite electrical conductivity, the fluid motion is defined by Eq. 4. This limit is sometimes called ideal MHD and is reasonably good when describing the solar wind and large-scale structures in the magnetosphere (c.f. section 3.0). On the other hand, it is not applicable to rapid time variation phenomena or small-scale structures such as in the aurora (c.f. section 3.3).

### 2.3 Magnetic Reconnection

A concept, which is useful within the framework of MHD, is the electrodynamic state where the magnetic field may be considered as tied to the plasma flow. This limit is found for example by studying the so-called magnetic induction, or dynamo, equation.
This equation is derived by taking the curl of Ohm’s law (Eq.15) and substituting the electric field and the current by using Faraday’s law and Ampère’s law, respectively:

\[ \frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}) + \eta \nabla^2 \mathbf{B} \quad \text{(dynamo equation)} \quad (16) \]

where \( \eta = 1/\mu_0 \sigma \) is the magnetic diffusivity. The so-called magnetic Reynold’s number is defined as the ratio of the first to the second term of the right-hand side of Eq. 16, with \( L \) as the characteristic length over which the fields change,

\[ R_m = \mu_0 \sigma u L \quad (17) \]

In the limit of infinite conductivity, the magnetic diffusivity is negligible and the second term of Eq. 16 vanishes, which implies that the total magnetic flux \( \Phi = \int \mathbf{B} \cdot d\mathbf{S} \) crossing a surface \( \mathbf{S} \) bounded by a closed curve at some initial time, will remain constant in time as the plasma fluid moves through the system and the location of \( \mathbf{S} \) and/or its shape change (Alfvén and Fälthammar, 1963). All plasma elements initially connected by a magnetic flux tube of cross section \( \mathbf{S} \) then remain linked by the same flux tube in time as the plasma drifts through space and we can think of the magnetic flux tube as moving with the plasma.

A direct consequence of the frozen-in magnetic field concept is that boundaries form between plasma regimes of different properties, since the plasma can mix easily along \( \mathbf{B} \), but not perpendicularly. The magnetic field is then tangential on either side of the boundary, but will in general be of different magnitude and direction with a current sheet at the boundary according to Ampère’s law. If the plasma flow on either side is negligible, the dynamo equation relaxes into a pure diffusion equation. Assume that \( \mathbf{B} \) is oppositely directed on either side of an infinite plane, but of same magnitude and locally parallel to the plane. The magnetic flux will then diffuse toward the current sheet where it annihilates and magnetic field energy is converted into heat. In time, the gradient in \( \mathbf{B} \) decreases, which results in a reduced diffusion rate as the local plasma pressure builds up. The process is therefore self-limiting in nature and stops when the local plasma pressure balances the magnetic field pressure. Now, assume the presence of an electric field \( \mathbf{E} = -\mathbf{u} \times \mathbf{B} \) well away from the current sheet that creates a plasma inflow \( \mathbf{u} \), which brings magnetic flux toward the boundary. A steady-state between magnetic field annihilation and plasma inflow may be set up and the current sheet thickness \( d = 2L \) derived as \( L = 1/\mu \mu_0 \sigma \). The magnetic Reynold’s number \( R_m \) in the current sheet then equals unity, which means that the MHD frozen-in flux condition breaks down. However, mass conservation requires that there must be an outflow as long as there is a plasma inflow and a second dimension gets introduced that sets a limit to the extension of the diffusion region. This simplified two-dimensional picture results in the so-called x-line reconnection configuration of Sweet (1958) and Parker (1957), which is illustrated in Figure 1.
Figure 1 Sweet-Parker magnetic reconnection configuration. $B_i$ and $u_i$ are the magnetic field and plasma flow velocity convecting toward the diffusion region of dimensions $L$ and $l$, while $B_o$ and $u_o$ are the magnetic field and plasma velocity leaving the merging region [Kivelson and Russell, 1995].

The magnetic field approaches zero only along a single line, called the neutral line, and not in an entire plane. The plasma outflow is driven by the same external electric field that drives the inflow and magnetic flux is carried away from the diffusion region rather than being annihilated. Two elements of plasma initially on the same magnetic flux tube flowing toward the diffusion region may be located on two different flux tubes at a later time flowing away from the diffusion region! We can imagine a process by which the initial magnetic field is “cut” and later “reconnected.” The boundary is no longer closed, since a normal component of $B$ is produced and plasma is suddenly allowed to mix along $B$, exchanging mass, momentum and energy between the two regions. Several models of this so-called reconnection process have evolved since the 1950’s x-line model of Sweet and Parker. Petschek (1964) suggested that most of the plasma does not need to flow through the diffusion region itself to be accelerated and introduced shock planes in the external convection region, connected to the diffusion region where MHD breaks down. A shock is a discontinuity where the plasma properties such as $B$, $u$, and mass density change abruptly with plasma compression and dissipation taking place at the shock as long as the plasma flow across the shock is finite. The plasma is then thought to be accelerated via the $j \times B$ force due to the current sheet that develops in the shock plane with the maximum outflow speed from the interaction region being limited to the Alfvén velocity,

$$v_A = \frac{B}{\sqrt{\mu_0 \rho}}$$

which is the speed by which information about the diffusion region is transferred to the external plasma. Reconnection may then be considered a global process, since also the external system must be considered in deriving the position of these shock planes. A shortcoming of the reconnection process is the assumption of steady-state. Moreover, the present models describing the process cannot exactly address when or where reconnection will occur. Today, three factors are believed to influence the probability of reconnection onset and the rate by which reconnection takes place. The relative orientation of the interacting magnetic fields or the magnetic shear is one factor, but whether or not the fields need to be completely antiparallel is still an open question. The flow speed is another factor. The lower the flow velocity, the longer the time of interaction. The last factor that needs to be considered is the plasma $\beta$, which
is defined as the ratio of plasma to magnetic field pressures. A lower $\beta$ is considered favorable for reconnection onset.
3. Sun-Earth Environment

Today we know that the Sun is more than a source of light and heat for the planets of the solar system. It is also emitting a continuous stream of electrons and light ions such as hydrogen and helium from the solar corona, due to the pressure difference between the Sun and interstellar space. This charged particle stream forms a tenuous and hot plasma called the solar wind, which is reasonably well described as an ideal MHD fluid. The Earth is to a large extent protected against this plasma flow by its magnetic field, which is generated by the thermal convection of its liquid interior. The geomagnetic field deflects most of the solar wind around the Earth, much like a rock in a stream, and creates a cavity in the solar wind that we call the magnetosphere (see Figure 2).

![Figure 2](image)

**Figure 2** The Earth’s magnetosphere and large-scale current systems [Kivelson and Russell, 1995].

However, the solar wind experiences an abrupt transition that slows it down well before reaching the outermost magnetospheric boundary, called the magnetopause. The transition is a so-called collisionless shock termed the bow shock, which is set up as a result of the magnetospheric obstacle in the high-speed solar wind flow. The region between the magnetopause and the bow shock is referred to as the magnetosheath with plasma flow speeds lower than both the sonic speed and the Alfvén speed (Eq. 18). The position of the bow shock upstream of the magnetopause is generally located beyond 10 Earth radii ($R_E \sim 6300$ km), but changes with changing solar wind parameters such as the flow speed. The average properties of the solar wind at the Earth are shown in Table 1 below.
Table 1 Observed average properties of the solar wind plasma near the Earth’s orbit [Kivelson and Russell, 1995]

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proton density</td>
<td>$6.6 \text{ cm}^{-3}$</td>
</tr>
<tr>
<td>Electron density</td>
<td>$7.1 \text{ cm}^{-3}$</td>
</tr>
<tr>
<td>He$^{++}$ density</td>
<td>$0.25 \text{ cm}^{-3}$</td>
</tr>
<tr>
<td>Flow speed</td>
<td>$450 \text{ km/s}$</td>
</tr>
<tr>
<td>Proton temperature</td>
<td>$1.2 \times 10^5 \text{ K}$</td>
</tr>
<tr>
<td>Electron temperature</td>
<td>$1.2 \times 10^5 \text{ K}$</td>
</tr>
<tr>
<td>Magnetic field</td>
<td>$7 \text{ nT}$</td>
</tr>
</tbody>
</table>

3.1 The Solar Wind and the Interplanetary Magnetic Field

Embedded in the solar wind is a weak magnetic field, which normally is nearly aligned with the ecliptic plane (the plane of the Earth’s orbit around the Sun), and directed at approximately $45^\circ$ to the Sun-Earth line. The origin of this so-called interplanetary magnetic field (IMF) is the solar corona where the electrical conductivity is so high that the solar magnetic field satisfies the frozen-in magnetic field condition. As a plasma element leaves the corona it drags the magnetic field radially outwards and forms a magnetic flux tube. By the time the plasma element reaches the Earth, the direction of the equatorial magnetic field is close to $45^\circ$ from the radially directed Sun-Earth line, due to the solar rotation that moves the footpoint of the magnetic field on the Sun in longitude. The resulting equatorial IMF configuration is that of an Archimedean spiral, which is usually referred to as the garden-hose spiral. A solar dipole magnetic field with “unipolar” regions above some latitude creates an interplanetary current sheet due to the opposite directions of the unipolar fields originating in the two solar hemispheres above and below the equatorial plane. If the solar dipole is tilted with respect to the equatorial pole, then the current sheet will also be tilted. This is the explanation for the observation of patterns of identical IMF polarity at the Earth with toward and away sectors, that repeats every 27 days, which is the solar rotation period observed at Earth. The component of IMF perpendicular to the ecliptic plane is thought to originate from local phenomena taking place in transit from the Sun, such as interactions between high-speed solar wind flows overtaking regions of lower speed. The IMF may in general then be written as, $\mathbf{B}=(B_x,B_y,B_z)$, where $B_x$ and $B_y$ are in the ecliptic plane and $B_z$ is perpendicular to it. $\mathbf{x}$ points toward the Sun, $\mathbf{z}$ is perpendicular to the ecliptic plane, and $\mathbf{y} = \mathbf{z} \times \mathbf{x}$. This is the so-called geocentric solar ecliptic (GSE) coordinate system. At times of high solar activity, the Sun sometimes emits large quantities of plasma in so-called coronal mass ejections that form fast flowing magnetic clouds in interplanetary space. Large $B_z$ components of the IMF are usually encountered within such high density plasma clouds.

3.2 The Magnetosphere

The geomagnetic field is approximately described as a dipole field within about $6 \ R_E$ with its axis tilted almost $11^\circ$ relative to the Earth’s rotation axis. The rotation axis is in turn inclined by $23.5^\circ$ to the ecliptic pole. In studies of the interaction between the IMF and the magnetosphere, it therefore makes more sense to use the geocentric solar magnetospheric (GSM) coordinate system, where $\mathbf{z}$ is antiparallel to the Earth’s dipole axis. As the solar wind flows around the Earth, it compresses the dayside
magnetopause that separates the solar wind plasma from the magnetosphere and stretches the nightside geomagnetic field into the so-called magnetotail, similar in shape to that of cometary tails. The shape and position of the entire magnetopause results from a pressure balance between the solar wind and the magnetosphere. The average position of the dayside magnetopause is located where the solar wind kinetic momentum flux balances the geomagnetic field pressure. For a typical solar wind momentum flux, the mean dayside magnetopause is found at about $10 \, R_E$ (Fairfield, 1971).

The magnetopause is generally, but not always (c.f. section 3.2.2), considered a closed boundary with no magnetic field component normal to the surface. Thus, applying Ampère’s law across the boundary results in an eastward dayside magnetopause current and a dawnward magnetopause tail current (see Figure 2). In satisfying the MHD momentum equation, this current produces the $\mathbf{j} \times \mathbf{B}$ force that deflects the magnetosheath plasma.

The magnetotail is separated in two regions by a finite boundary in its mid-plane where the so-called cross-tail or neutral sheet current flows (see Figure 2). The two parts are called the northern and southern tail lobes with the geomagnetic field directed toward and away from the Earth, respectively. The lobe magnetic fields originate in the regions centred around the magnetic poles of the northern and southern hemispheres, which are referred to as the polar caps. The cross-tail current is embedded in the hot plasma region of the plasma sheet and current continuity closes the tail magnetopause currents with the cross-tail current. Moreover, the tail lobes contain a plasma of lower density than in the magnetosheath (see Table 2).

<table>
<thead>
<tr>
<th>Density ($\text{cm}^{-3}$)</th>
<th>Magnetosheath</th>
<th>Plasma mantle</th>
<th>Tail lobe</th>
<th>BPS</th>
<th>CPS</th>
<th>Neutral sheet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ion temperature ($K$)</td>
<td>$10^6$</td>
<td>$10^6$</td>
<td>$10^6$</td>
<td>$10^7$</td>
<td>$5 \times 10^7$</td>
<td>$5 \times 10^7$</td>
</tr>
<tr>
<td>Electron temperature ($K$)</td>
<td>$5 \times 10^5$</td>
<td>$5 \times 10^5$</td>
<td>$5 \times 10^5$</td>
<td>$5 \times 10^5$</td>
<td>$10^7$</td>
<td>$10^7$</td>
</tr>
<tr>
<td>Magnetic field ($nT$)</td>
<td>5</td>
<td>25</td>
<td>25</td>
<td>20</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Debye length ($m$)</td>
<td>18</td>
<td>40</td>
<td>400</td>
<td>400</td>
<td>280</td>
<td>200</td>
</tr>
<tr>
<td>Alfvén speed ($\text{km/s}$)</td>
<td>49</td>
<td>550</td>
<td>5500</td>
<td>1400</td>
<td>300</td>
<td>44</td>
</tr>
<tr>
<td>Convection speed ($\text{km/s}$)</td>
<td>200</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 2 Average plasma properties for the different regions of the magnetosphere, where the BPS and CPS are abbreviations for the plasma sheet boundary layer and the central plasma sheet, respectively [Olsson, 1997].

By magnetic flux conservation between the polar cap and tail lobe in one hemisphere and assuming that no flux crosses either the tail magnetopause or the neutral sheet, we get a tail magnetopause radius of about 20-30 $R_E$, depending on the polar cap size and
tail magnetic field strength. The hot plasma of the plasma sheet (see Table 2), with density between that of the magnetosheath and the tail lobes, is found on the closed field lines of the stretched out dipole field that reaches down to the auroral ionosphere (c.f. section 3.3). The ion population of the plasma sheet is a mixture of solar wind and ionospheric particles, evidence that the magnetosphere is not entirely closed.

### 3.2.1 Magnetospheric Boundary Layers

In situ observations in space tell us that the different regions of the magnetosphere are not generally separated by infinitesimal discontinous boundaries, but rather by boundary layers of finite size in which the field and plasma from the two regions mix. The plasma sheet and tail lobe interface is separated by the plasma sheet boundary layer where counter-streaming ions are frequently observed. Spacecraft observations of the magnetopause have also detected three other types of boundary layers (see Figure 3).

![Figure 3](image)

*Figure 3* The closed dayside low-latitude boundary layer (LLBL), the cusp and the mantle layers are shown to the left in a noon-midnight meridian plane of the magnetosphere. The arrows mark the antisunward convecting IMF connected to the geomagnetic field. The duskside flank LLBL is shown to the right [Kivelson and Russell, 1995].

These are called the low-latitude boundary layer (LLBL), the entry layer (or the cusp), and the plasma mantle. The LLBL is generally found earthward of the dayside magnetopause and along the near-Earth magnetopause flanks. It contains a mix of magnetosheath and magnetospheric plasma and is most probably found on closed field lines (e.g. Mitchell et al., 1987). The entry layer, also known as the cusp, is the region around the magnetic null of a closed model magnetosphere (see Figure 4a) with plasma characteristic of the magnetosheath but with lower flow speeds. The plasma mantle is found tailward of the cusp with tailward plasma flow. Flow speed, density, and temperature decrease with distance from the magnetopause into the tail lobes. The magnetic field of the cusp, plasma mantle, and tail lobes are most probably connected to the magnetosheath magnetic field.

### 3.2.2 Closed vs Open Model Magnetosphere

The IMF has a strong influence on the Earth’s geomagnetic field. Two simple cases of this effect may be studied by ignoring the plasmas of the solar wind and the Earth and simply superposing the dipole field and the IMF. Assuming first an IMF strictly directed in the positive z-direction, or a northward IMF, results in the so-called *closed* model magnetosphere illustrated in Figure 4a.
Figure 4 Superposition of a geomagnetic dipole field with the IMF giving (a) a closed model magnetosphere for northward IMF and (b) an open model magnetosphere for southward IMF. N1 and N2 refer to the two neutral points (Figure 4a) or the two neutral lines (Figure 4b) forming. The magnetic field is labelled “1” for the IMF, “2” for the closed geomagnetic field, and “3” for the magnetic field lines that have one foot connected to the IMF and the other with the geomagnetic field [Parks, 1991].

The field is dipole-like within a limiting radius $r^*$, with both ends of the field attached to the Earth, and IMF-like at large radial distances, where both ends of the field are attached to the Sun. Two neutral points, N1 and N2, appear. In assuming an IMF strictly in the negative $z$-direction, a southward IMF, we get the so-called open model magnetosphere. This is shown in Figure 4b. Instead of two neutral points located above each magnetic pole, we now get two neutral lines in the magnetic equatorial plane, N1 and N2, directed in the $y$-direction, just as in the x-line configuration of reconnection (c.f. section 2.3). The magnetic field line labelled 1 has both its ends in the Sun and the one labelled 2 has both ends attached to the Earth, as in the previous case. However, there is now a third type labelled 3 having one end attached to the Sun and the other to the Earth. Magnetic field lines of this type are referred to as being open.

3.3 The Ionosphere

The ionosphere is a transition region formed between the neutral atmosphere and the fully ionized magnetosphere and acts both as a source and a sink of plasma for the magnetosphere. It reaches from about 60 km altitude and gradually merges into the lower magnetosphere. The outer boundary of the plasmasphere (c.f. section 3.4), the plasmapause, defines its upper boundary in the equatorial plane, though most phenomena occur below ~1000 km, where the plasma density maximizes. The major source of ionization is the solar ultraviolet radiation, but at high latitudes precipitation along magnetic field lines of energetic particles of magnetospheric origin can contribute significantly to the ionization. The plasma densities are thus generally lower during nighttime than during daytime, whereas magnetospheric particle precipitation is most intense in an oval shaped belt around both magnetic poles, which is connected to the plasma sheet and the plasma sheet boundary layer. This belt is called the auroral oval, where precipitation causes significant ionization that result in a higher conductivity than that found in the polar cap region. The characteristic light emission of the aurora is produced when neutral atoms or molecules are excited by collisions with predominantly precipitating electrons. On returning to their ground state, photons of characteristic energies are emitted.
The ionosphere in the approximate altitude range from the so-called E-region peak at about 120 km and below is dominated by a collisional plasma where the ion-neutral collision frequency \( \nu_{\text{i-ne}} \) is larger than the ion gyro-frequency \( \omega_{gi} \) and the opposite being true for electrons with \( \nu_{\text{e-ne}} \ll \omega_{ge} \). The electrons then move in the local \( \mathbf{E} \times \mathbf{B} \) direction with ions moving along \( \mathbf{E} \) if the collision frequency is large enough or at a small angle between \( \mathbf{E} \) and \( \mathbf{E} \times \mathbf{B} \). Electrons thus support a so-called Hall current in the antiparallel \( \mathbf{E} \times \mathbf{B} \) direction, while the ions carry a perpendicular Pedersen current along \( \mathbf{E} \). The Hall current, caused by the electron drift in the \( \mathbf{E} \times \mathbf{B} \) direction and flowing within the layer of enhanced auroral conductivity is called the auroral electrojet.

### 3.4 Solar Wind-Magnetosphere-Ionosphere Coupling

The Earth’s magnetosphere is most probably never entirely closed. Evidence of its open nature is for example the observation of so-called polar rain in the polar cap. Polar rain is a highly field-aligned electron population of the solar wind and is observed in the polar cap, which is directly connected to the Sun (Fennell et al., 1975; Fairfield and Scudder, 1985).

The high-latitude electrodynamics of the Earth is studied by analysis of its two separate regions of interaction, the solar wind-magnetosphere and the magnetosphere-ionosphere interface, respectively. Below about 200 km altitude, the ionospheric electric and magnetic fields are related by Ohm’s law, \( \mathbf{j} = \sigma (\mathbf{E} + \mathbf{u}_n \times \mathbf{B}) \), where \( \mathbf{u}_n \) is the neutral wind velocity due to the Earth’s rotation and \( \sigma \) is the ionospheric conductivity tensor. Plasmas in the ionosphere typically have different conductivities along the magnetic field and the two perpendicular directions due to the magnetic field. In other words it is highly anisotropic. The magnetospheric plasma above about 2000 km is generally collisionless and the current \( \mathbf{j} \) and electromagnetic fields are related by Eq. 4 and the momentum equation (Eq. 10). Gravity is assumed to have a negligible effect on the motion. Taking the cross product of the momentum equation with \( \mathbf{B} \) results in an expression for the perpendicular current given as,

\[
\mathbf{j}_\perp = \frac{\rho \mathbf{B} \times \frac{d\mathbf{u}}{dt} + \mathbf{B} \times \nabla p}{\mathbf{B}^2}
\]

The intermediate 200-2000 km region is a transition region, but the frozen-in condition usually holds there for the plasma motion. The coupling geometry of the high-latitude ionosphere is shown in Figure 5, where the geomagnetic field lines are connected with the magnetosheath and solar wind for a southward IMF. The region of direct connection defines the open polar cap.
The electric field in the solar wind plasma is given by $\mathbf{E}_{sw} = -\mathbf{u}_{sw} \times \mathbf{B}_{sw}$ in the Sun-fixed inertial frame of reference and directed from dawn to dusk for a southward IMF. $\mathbf{E}_{sw}$ maps down along the open equipotential field lines to form the ionospheric $\mathbf{E}_i$, if we assume that no electric field exists along the magnetic field. Magnetic flux conservation in a converging magnetic field results in an ionospheric electric field $\mathbf{E}_i > \mathbf{E}_{sw}$ that drives antisunward ionospheric plasma across the polar cap from noon to midnight. If the neutral wind is negligible, then $\mathbf{j}_i = \sigma \cdot \mathbf{E}_i$ and $\mathbf{j}_i \cdot \mathbf{E}_i > 0$, since the Pedersen current $\mathbf{j}_i$ is parallel to $\mathbf{E}_i$. The ionosphere thus corresponds to an electric load in an electric circuit, the source of which is the solar wind. Assuming that the solar wind slows down when in contact with the polar cap, $\frac{du}{dt} < 0$, and that $\nabla p$ is negligible, results in a solar wind current $\mathbf{j}_{sw}$ which is antiparallel with $\mathbf{E}_{sw}$ (see Eq. 19). The solar wind acts as an MHD generator with field-aligned currents caused by $\nabla \cdot \mathbf{j}_{sw} = 0$ feeding energy into the ionospheric load. An expression of the field-aligned current in terms of the ionospheric properties can be derived from the divergence of $\mathbf{j}$ as,

$$j_{fac} = \nabla \cdot (\Sigma_{\perp} \cdot \mathbf{E}_i)$$

(20)

where $\Sigma_{\perp}$ is the height-integrated perpendicular conductivity tensor. Expanding this relation shows that the field-aligned current in the ionosphere is related to spatial variations in both $\mathbf{E}_i$ and $\Sigma_{\perp}$.

Dungey (1961) was the first to suggest the open model magnetosphere (c.f. sections 2.3 and 3.2.2). The closed magnetopause is opened up by the existence of an x-line, which is supposed to form in a limited region on the dayside magnetopause where the frozen-in flux condition breaks down (see Figure 4b). If the merging of the IMF and the geomagnetic field continues indefinitely, then the entire magnetosphere would eventually be connected to the IMF. A second neutral line is required to form in the neutral sheet of the magnetotail where an open field line from each hemisphere reconnects and provide a return of flux with the resulting earthward convecting plasma of the plasma sheet. The convecting plasma sets up a dawn-dusk electric field inside the closed magnetosphere that corresponds to the applied potential difference across the polar cap. The $\nabla B$ force deflects the sunward convecting plasma around the Earth as it reaches the inner magnetosphere (see Eqs. 7 and 8). Ions move westward and
electrons eastward, creating the partial-ring current that closes via the field-aligned currents referred to as the Region 2 current system. The ring current that forms earthward of the partial-ring current circulates around the Earth and produces a symmetric magnetic perturbation on the ground (c.f. Appendix A). The higher-latitude field-aligned currents that feed the ionospheric load from the solar wind-magnetosphere interaction region are called the Region 1 currents (see Figure 6).

![Figure 6](image)

**Figure 6** Solar wind-magnetosphere-ionosphere coupling geometry, illustrating the sunward return plasma flow within the magnetospheric cavity and the Region 1 and Region 2 field-aligned current systems [Kelley, 1989].

The mapping of the magnetospheric electric fields from the “open” and “closed” regions to the ionosphere results in the simplified two-cell ionospheric convection pattern (see Figure 7) with antisunward $\mathbf{E} \times \mathbf{B}$ flow in the polar cap and a sunward return flow at lower latitudes.

![Figure 7](image)

**Figure 7** Idealized two-cell plasma convection pattern mapped from the magnetosphere to the ionospheric magnetic local time and corrected geomagnetic latitude coordinate system [Kelley, 1989].

Iijima and Potemra (1976) found that the Region 1 and 2 currents on the average form a typical statistical pattern in the auroral oval region (see Figure 8). The Region 1 currents are typically found near the convection reversal region that also mark the poleward boundary of the auroral oval, where the Pedersen currents connect the Region 1 and Region 2 field-aligned currents. The electrojet is directed along the
ionospheric equipotentials of the two-cell convection pattern, with a westward electrojet on the dawnside and an eastward electrojet on the duskside. The geomagnetic perturbation observed on the ground beneath the electrojet is given by the “right-hand rule.” The instantaneous global perturbation due to the two-cell electrojet pattern is named DP-2 for “disturbance polar of the second type.”

![Figure 8 Statistical field-aligned current pattern](image)

Magnetic reconnection, though not fully understood in three dimensions, is considered the dominant driving mechanism for the global magnetospheric convection. However, when studying the potential drop across the polar cap and the IMF, a smaller potential of about 20 kV or so still exists even after several hours of northward IMF (Wygant et al., 1983). Axford and Hines (1961) suggested a different solar wind-magnetosphere interaction process that results in a similar plasma convection. This so-called “viscous interaction” assumes that solar wind momentum is transferred across a closed magnetopause with no normal component of $B$. The exact nature of this interaction is poorly understood, since the magnetosheath and/or a narrow region inside the magnetopause would need to support collisions perpendicular to $B$, but the Kelvin-Helmholz instability set up between sheared flows has been suggested as an explanation (see e.g. Miura, 1984). The “viscous interaction” process is sometimes invoked in understanding observations of the LLBL (e.g. Mitchell et al., 1987; Siscoe et al., 1991) with plasma sometimes flowing antisunward in a narrow region just inside the magnetopause and plasma flowing sunward radially inwards with a stagnation region resulting inbetween. The convection electric field is not the only large-scale electric field within the magnetosphere. The Earth’s rotation causes a motion of the atmosphere and the ionospheric plasma, which is tied to it by charged-neutral particle collisions. The motion of the ionospheric plasma sets up an electric field $\mathbf{E} = -\mathbf{u} \times \mathbf{B}$ directed radially earthward in the equatorial plane, where $\mathbf{u} = \mathbf{\omega} \times \mathbf{r}$ is the corotation velocity and $\mathbf{\omega}$ is the corotation angular velocity. The corotation field decreases with radial distance, since $B \propto 1/r^3$, with the convection field dominating at distances beyond say 4 $R_E$ while the corotation field dominates within 3-4 $R_E$, a region called the plasmasphere (see Figure 2). The ring current just beyond the cool and dense plasmasphere causes a weak polarizing electric field due to the charge separation of westward $\nabla B$-drifting
ions and eastward drifting electrons. This field shields the convection electric field from the lower latitudes of the plasmasphere.
4. Magnetospheric Response to the Solar Wind

So far we have assumed that the solar wind-magnetosphere coupling is in steady state. However, since the solar wind plasma and its magnetic field generally change in time, so does the magnetospheric plasma convection and the large-scale current systems, which are set up by the interaction. A stronger coupling is quantified, e.g., by an increased solar wind voltage applied across the magnetopause or by stronger magnetospheric currents. Whenever the currents change, they will affect the geomagnetic field according to Ampère’s law and we say that the geomagnetic activity increases.

4.1 Solar-Terrestrial Dynamics

The geomagnetic activity is quantified by various indices based on ground measurements of the geomagnetic field, which is normally separated into the three components H, D, and Z (c.f. Appendix A). The indices indicate the deviation from a quiet level, which often must be defined as an average of some of the lowest days of activity during a certain period of time or as the mean over a month, since the activity never ceases entirely. Some of the most common indices used in studies of the solar wind-magnetosphere interaction are the magnetic storm Dst index, the westward electrojet AL index, and the total auroral electrojet AE index (c.f. Appendix A). The Dst index is due mainly to the negative horizontal disturbance field from the longitudinally symmetric westward ring current. The ring current is sometimes greatly enhanced by the injection of charged particles during prolonged periods of southward IMF. This process is called a magnetic storm. The injection is believed by some researchers to be caused by fluctuations of enhanced sunward convection in the tail (e.g. Kamide et al. 1997; Chen et al., 1994). There is also a systematic longitudinally asymmetric horizontal disturbance centred on the duskside, which is believed (Crooker and Siscoe, 1981) to be caused mainly by the system of high-latitude field-aligned currents. This disturbance is indexed by the so-called ASY-H (Iyemori and Rao, 1996), or ASYM index (Clauer et al., 1983). Another highly dynamic process that most often occur when the IMF turns southward is the sequence of events leading to the auroral substorm that taps energy from the magnetotail into an increased westward electrojet centred around local midnight with resulting increased auroral displays. The energy is guided by field-aligned currents into the ionosphere through the diversion of the near-Earth cross-tail current referred to as the substorm current wedge (Clauer and McPherron, 1974). The enhanced westward electrojet corresponds to the ionospheric DP-1 current system that stands for “disturbance polar of the first type.” The three phases that constitute a substorm are the growth, expansion and recovery phase (c.f. sections 4.2 and 4.3). From all of the above, it can be seen that there is a common solar wind variable that all geomagnetic activity depends on, directly or indirectly, and that is the Bz component of the IMF.

4.2 Magnetospheric Plasma Circulation

There exist two processes for the explanation of the solar wind induced magnetospheric plasma convection. These are the magnetic reconnection and the viscous interaction processes (c.f. section 3.4). Antisunward flow on closed field lines have been observed in the LLBL and explained in terms of some viscous interaction process (e.g. Mitchell et al., 1987; Traver et al., 1991).
Concentrating now on the open model magnetosphere illustrated in Figure 4b. A change in IMF direction $\mathbf{B}_{sw}$ or a change in $\mathbf{u}_{sw}$ should be affecting the rate of dayside merging at N1 (c.f. section 3.2.2). Assume that the rate increases at N1. This results in new open flux on the dayside and a larger fraction of $\mathbf{E}_{sw}$ will map along open flux tubes to the ionosphere where it eventually leads to an increased potential across the polar cap and an increased antisunward flow toward midnight. A lower reconnection rate at N2 than at N1 results in increased magnetic flux inflow from the dayside and thus increased magnetotail currents, that corresponds to a loading of magnetic energy into the tail. It also leads to a slower sunward return flow. The opposite is generally true if the reconnection rate at N2 is higher than at N1 with higher sunward return flow than antisunward polar cap flow. This is the general picture illustrated in Figure 9 adapted from Lockwood et al. (1990). The total ionospheric convection electric field thus consists of a superposition of convection electric fields produced at two sources. One is related to the convection flow driven by dayside reconnection at N1 and the other by the release of magnetotail energy at the nightside reconnection rate (N2). The total potential across the polar cap is according to this picture a superposition of electric fields from these two sources, where the contribution from the dayside source depends on the mapping of the solar wind electric field. These two systems are equivalent to the ionospheric DP-2 and DP-1 Hall current systems. The two initial phases of the substorm may be explained using this simple model. The growth phase would correspond to N1>N2 and an expanding polar cap, while the expansion phase corresponds to N1<N2 and a contracting polar cap.

![Figure 9](image)

**Figure 9** Dayside driven reconnection plasma flows (thin arrows) with reconnection rate at N1 being larger than the nightside reconnection rate at N2 (left), resulting in an expanding polar cap (thick arrows). The opposite situation (right) with nightside reconnection rate at N2 larger than the dayside rate, the polar cap contracts [Lockwood et al., 1990].

### 4.3 Direct vs Loading/Unloading Response

How fast does the ionosphere respond to a southward turning of the $B_z$; instantaneously or by a characteristic delay? It is already assumed that the nightside DP-1 convection/current system of substorms is delayed during a growth phase when magnetic energy must be deposited in the tail and reach a certain threshold before onset may be triggered, with the ensuing energy deposit into the auroral midnight ionosphere. The trigger mechanism may be external (e.g. northward turning of IMF), some kind of internal instability, or a combination of both (Rostoker et al., 1987). This is the so-called loading-unloading model of the magnetospheric response to the solar wind with characteristic response times around 60 min or more. This implies
that the DP-1 system lags the DP-2 system. Then, does the DP-2 system respond immediately to a southward turning of $B_z$ at the magnetopause? Observations and theoretical analysis (see Papers 1 and 2 and references therein) of the dayside global magnetosphere-ionosphere coupling (c.f. section 3.4) suggest that it does not and that there is a delay of 10-20 min, which is believed to partly originate from the self-inductance $L$ of the field-aligned currents (c.f. section 3.4). Sanchez et al. (1991) used an electric circuit analogue of the current loop, where the applied solar wind voltage drop $\Phi_{sw}$ may be described in terms of the ionospheric potential drop $\Phi_i$ as,

$$\Phi_{sw} = \Phi_i + L dI/dt = (1 + \Sigma L d/dt) \Phi_i$$

where the effective ionospheric conductance $\Sigma = 2\Sigma_p$. The time constant was found to be approximately 20 min. This time response corresponds to the large-scale ionospheric cross-polar potential drop and is usually referred to as the directly driven model of magnetospheric response with energy being directly deposited into the ionosphere. Rostoker et al. (1987) give a thorough description of both the directly driven and the loading-unloading models.

### 4.4 Predicting an Average Magnetospheric Response

The magnetospheric response to the solar wind is usually studied using a so-called linear prediction technique that treats the solar wind-magnetosphere system as a black box under the assumption that a linear relationship exists between a solar wind input parameter $I(t)$ at time $t$ and a magnetospheric output parameter $O(t')$ at time $t' = t + \Delta t$, with $\Delta t$ as a general time delay. The transfer function $H(t)$ of the black box is determined from a time series analysis of $I(t)$ and $O(t)$ and the Fourier transform of $H(t)$ results in the impulse response of the system that can be used to predict the future output of the system. Some commonly used solar wind input parameters are the IMF $B_s$, the IMF electric field $vB_s$, and the solar wind energy input $\varepsilon$-parameter (Akasofu, 1980). Here, $B_s = |B_z|$ when $B_z < 0$ and zero otherwise, and $\varepsilon = vB^2 \sin^4(\theta/2)\mu_0 / \mu_0$, where $\theta$ is the IMF clock angle and $l^2_0$ is an effective cross section of the solar wind-magnetosphere coupling region. The magnetospheric output parameter is most often a geomagnetic activity index, such as the AL, AE, or Dst, since these are measured continuously and can be retrieved easily as time series. Continuous time series coverages of both input and output parameters have so far been necessary for the successful application of the linear prediction filter technique as described above. Bargatze et al. (1985) used the input $vB_s$ parameter and the output AL parameter to analyse the magnetospheric response for different levels of average geomagnetic activity and found two major time delays at 15-20 min and around 60 min. These were interpreted as the directly driven and loading-unloading responses of the DP-2 and DP-1 convection, respectively (see Figure 10).
The ionospheric convection as quantified by the cross-polar potential drop is not available as time series, only as discrete measurements, and another approach becomes necessary. Its response to the solar wind is analysed here for the first time using a novel statistical technique (Paper 1, Paper 2) by studying the time lag series $\Delta t$ of the correlation coefficient between a set of discrete potential measurements and a discrete input solar wind parameter. The individual response for an ensemble of substorms to the solar wind $vB_s$ was analysed by a corresponding technique and described by Blanchard and McPherron (1995) using the AL index. A striking resemblance is observed when comparing the two response modes of the $vB_s$ to AL time series with the two separate response modes of the discrete $\varepsilon$ to $\Phi_{pc}$ and $\varepsilon$ to ASY-H time series (Blanchard and McPherron, 1995; Paper 1).
5. Lobe Cell Convection

The concept of tail lobe reconnection was first suggested by Russell (1972) as merging of a purely northward IMF with the southward directed magnetic field of the open tail lobe just poleward of the cusp (see Figure 11 from Crooker and Rich, 1993).

![Figure 11](image)

**Figure 11** Lobe cell generation by merging of open geomagnetic field lines with the IMF. The shaded area marks the region of draped lobe field lines over the dayside [Crooker and Rich, 1993].

Crooker (1979) later proposed that such a merging site generates plasma circulation in one tail lobe with sunward convection in the polar cap ionosphere and that the sense of circulation depends on IMF $B_y$, clockwise on the duskside and anticlockwise on the dawnside. Lobe cell convection appears to occur also for southward $B_z$ (Burch et al., 1985; Coley et al., 1987; Gosling et al., 1986, 1991). Lu et al. (1994) observed significant sunward flows near the polar cap boundary for $B_z < 0$ as long as $|B_y| > |B_z|$. This implies a merging site on the tail lobe flanks away from the noon-midnight meridian, where the geomagnetic field has a dominant $B_y$ component and where it is antiparallel to the IMF as is also reported in Paper 3. It seems that a condition for sunward ionospheric convection to result from tail lobe merging is that the local Alfvén speed is greater than the magnetosheath antisunward flow speed (e.g. Gosling et al., 1996). Paper 3 reports the existence of a $B_y$-dependent lobe convection cell within the dayside merging cell based on a clear and extensive data set not previously reported.
6. Empirical Electric Field Models

The instantaneous global convection electric field cannot be measured. There are obviously not enough satellites present in situ and remote ground measurements by radars cannot measure the entire polar cap region in a continuous mode. If we want to estimate the complete spatial distribution of the high-latitude ionospheric potential at any one time, we must use an empirical model of the electric field that combines theory with past snap-shot observations. Several such models exist, most of which show how the IMF and solar wind control ionospheric potentials or how the level of geomagnetic activity responds to the solar wind.

![Figure 12](image.png)

The two-cell convection pattern given by the empirical Weimer [2001] electric potential model for the solar wind parameters shown. The model is based on the electric potentials from 2645 polar cap passes of the Dynamics Explorer 2 satellite and compared with 780 Astrid-2 orbits [Paper 4].

The empirical Heppner and Maynard (1987) model is based on OGO 6 electric field measurements in the dawn-dusk plane and extended by DE 2 measurements. The main objective is to represent observations of typical convection electric fields during IMF B_z<0 by a minimum number of ionospheric convection patterns ordered by IMF B_y. The assimilative mapping of ionospheric electrodynamics (AMIE) model (Richmond et al., 1998 and references therein) derives snap-shot maps of quantities such as conductances, electric fields, and currents from high-latitude observations. Another snap-shot technique make use of global UV-imager data from polar orbiting space craft combined with in situ measurements of electric field, currents, and particles to derive global snap-shots of electric fields, currents, conductivities, and Joule dissipation (Blomberg and Marklund, 1988; Marklund et al., 1988, 1991). The IZMIRAN electrodynamic linear regression model (IZMEM) provides global patterns of ionospheric electric fields and geomagnetic perturbations to name but a few quantities (Papitashvili et al., 1994). Statistical maps of ionospheric potentials were also derived using a technique of least error fit of spherical harmonic coefficients with data from multiple DE 2 satellite passes (Weimer, 1996). This model calculates the potential as a function of IMF B_y and B_z, solar wind velocity and density, as well as the Earth’s dipole tilt angle. All these models lack details associated with the
substorm DP-1 system. The Weimer 1996 model has been improved by adding terms for the solar wind electric field and the solar wind dynamic pressure and with the optional substorm AL parameter (Weimer, 2001). An example of the ionospheric potential pattern for the Weimer 2001 model is shown in Figure 12 and on the front page. The solar wind input parameters are shown explicitly.
7. Electric Field Measurements

Electric fields in the ionosphere are measured in situ with satellites or remotely from the ground using radar measurements of the ionospheric plasma flow. There are two techniques for measuring the in situ electric field. Either directly, using the so-called double probe technique, or indirectly by measuring the particle $\mathbf{E} \times \mathbf{B}$ drift motion across the magnetic field. The double probe technique provides an electric field by measuring the potential difference between two probes in the plasma and dividing by their separation distance. The particle instrument emits an electron beam into the plasma that after one gyration around the magnetic field is displaced a short distance by the $\mathbf{E} \times \mathbf{B}$ drift. The electric field is found indirectly knowing this distance. A thorough discussion of the relative advantages is given by Lindqvist (1997). The electric field data used in this thesis come from the FAST and the Astrid-2 satellites, both of which apply the double probe technique. The probes must extend a distance much longer ($\approx 10\lambda_D$) than the local Debye length (Eq. 1) from the central body in order to avoid the disturbances introduced by currents drawn to it. Other effects that affect the final wire boom lengths include so-called wake effects due to the main satellite body. The Debye length at Astrid-2 and FAST altitudes of about 1000 km is estimated to $\lambda_D \approx 1$ cm. The probe to probe wire boom lengths for Astrid-2 and FAST are 7 m and 56 m, respectively, which means that the condition of having longer booms than the local Debye lengths is satisfied. The probes are extended and kept at a fixed distance from the central platform by the centrifugal force induced by the satellite rotation. FAST measures two components of the electric field $\mathbf{E}'$ in the spin plane of the moving satellite frame $K'$. The spin axis is directed approximately perpendicular to the satellite velocity $\mathbf{v}$ and the model magnetic field $\mathbf{B}$. Astrid-2, however, is a Sun-pointing satellite due to the distribution of its solar arrays in the spin plane. Neither satellite measures all three components of $\mathbf{E}'$, only the spin plane electric fields, and an assumption is necessary to handle the problem of the missing axial component. Two options are commonly used, either assuming equipotential field lines $\mathbf{E} \cdot \mathbf{B} = 0$, or setting the axial component $E_{ax} = 0$. The latter is applied here, since the former assumption in the limit of negligible axial magnetic field gives an infinitely large $E_{ax}$. This assumption is acceptable when the spin plane of the satellite is close to the expected direction of the convection electric field. This is especially true for Astrid-2. The axial field of FAST, however, is basically perpendicular to the satellite velocity and will contribute negligibly in finding the along-track electric potential,

$$\Phi = -\int \mathbf{E} \cdot \mathbf{v} dt$$

(22)

with steady state assumed and $\mathbf{E} = - \nabla \Phi$.

The in situ electric field $\mathbf{E}' = \mathbf{E} + \mathbf{v} \times \mathbf{B}$ is measured in a reference frame moving with the satellite velocity $\mathbf{v}$ and therefore includes the induced electric field $\mathbf{v} \times \mathbf{B}$ due to the satellite motion. The Sun-fixed electric field $\mathbf{E}$ is found by subtraction of $\mathbf{v} \times \mathbf{B}$ from the measured field $\mathbf{E}'$. This electric field is the sum of the corotation electric field and the convection electric field. When studying the convection electric field alone, we must subtract the corotation electric field and thus end up in an Earth-fixed frame. The cross-polar potential is now obtained by integrating the convection electric field along the satellite orbit (Eq. 22), where we assume that the convection field is totally shielded out by the plasmasphere in the equatorial plane. This implies that the potential $\Phi$ ideally should return to zero at low latitudes and this provides the
necessary boundary condition when integrating the convection electric field. The potential drop applied to the magnetosphere will, however, be underestimated by a single satellite moving across the polar regions since the probability of passing through each large-scale potential extremum in a single pass is rather low.
8. Summary of Papers

Paper 1: The cross-polar potential drop and its correlation to the solar wind.

This paper analyses the response of ionospheric convection using a discrete set of cross-polar potential drop from the FAST satellite to the solar wind input parameters IMF $B_z$, the solar wind reconnection electric field $E_r$, and the empirical solar wind energy input $\varepsilon$-parameter, where $E_r$ incorporates the empirical dependence of $vB_z$ on clock-angle $\theta$. The state of the solar wind is measured by the Wind spacecraft. A resulting response of the calculated correlation coefficients at 15 min time lag is interpreted as the DP-2 directly driven response, while the two responses at 55 min and 105 min time lags are explained in terms of the DP-1 loading-unloading response. The convection response, in particular for the $\varepsilon$-parameter, is further compared with the response of the asymmetric disturbance field using the ASY-H index. A striking resemblance is observed with the two response modes reported by Blanchard and McPherron (1995), which were interpreted as the directly driven DP-2 and the unloading DP-1 responses.

Paper 2: Magnetospheric response to the solar wind as indicated by the cross-polar potential drop and the low-latitude asymmetric disturbance field.

This paper also presents an analysis of the ionospheric response of convection and low-latitude asymmetric disturbance field to the solar wind electric field. Here, we use Astrid-2 electric field data for the derivation of the cross-polar potential drop. The optimized ionospheric responses are found to have a single response at 25 min for the potential, but three multiple responses at 35, 65, and 80 min for ASY-H, respectively. This is clearly different from the results of Paper 1, with three responses for the potential at 15, 55, and 105 min, but only one broad response for ASY-H between 60 and 75 min. We suggest that these differences may be explained by the higher geomagnetic activity for the Astrid-2 data set as compared with the FAST data. A similar tendency appeared in the Bargatze et al. (1985) study, although they used the input $vB_z$ parameter and the AL index as the linear prediction output parameter.

Paper 3: Lobe cell convection and field-aligned currents poleward of the Region 1 current system.

Paper 3 examines the simultaneous presence of dayside merging, lobe reconnection and viscous type of cells and the collocation of four field-aligned currents with the convection reversal boundaries for this ionospheric convection pattern. The electrodynamics is investigated by using electric field, perturbation magnetic field, and particle precipitation data from the FAST satellite in the dawn-to-dusk meridian plane. It is clear from this study that the IMF is predominantly directed in the $y_{GSM}$ direction and the sign of $B_y$ determines the location of the lobe cell, with $B_y>0$ ($B_y<0$) producing a duskside (dawnside) lobe cell within the dayside merging cell. The presence of low-latitude boundary layer and/or boundary plasma sheet particles precipitating on newly opened field lines convecting in the sunward directed part of the lobe cell implies a $B_y$-dependent northward twist of the plasma sheet as was suggested by Siscoe and Sanchez (1987).
Paper 4: Comparing the Weimer 2000 electric field model with Astrid-2 observations.

The final paper compares the electric fields observed by Astrid-2 with those derived from the electric potential model of Weimer (2001). The average latitudinal offsets between the convection reversals of the model and measured two-cell convection patterns are less than 2° corrected geomagnetic latitude, although some cases showed as much as 9° offsets. The solar wind density, electric field or the IMF do not seem to influence these offsets. The model is found to work better for southward IMF conditions than for northward IMF, when using the correlation coefficient as the quality measure of the agreement. However, the model does seem to give quite reasonable results for a few northward IMF cases. The measured mean magnitude of the electric field above 55° corrected geomagnetic latitude is in general 25% larger than that predicted by the model, which is most likely due to the underestimated electric fields near the convection reversal boundaries, since the model is based on fits of spherical harmonics of the order four or lower. From studying clear southward and northward IMF turnings during the polar cap passes, it seems that the global convection pattern is adjusted faster for southward IMF than for northward IMF events. Caution is necessary when timing the IMF to any specific satellite pass.
9. References


Acknowledgements

I would like to thank Prof. Göran Marklund, Dr. Charles W. Carlson, and the space physics groups of both the Alfvén Laboratory in Stockholm and the Space Sciences Laboratory at University of California in Berkeley for providing me with the opportunity to perform this coordinated study of data from the Astrid-2 and FAST satellites. Thank you also Tomas for your encouragement and support here and there along the way. This has been a learning experience indeed.
Appendix A: Earth’s Geomagnetic Field

The (H,D,Z) components of the geomagnetic field B are measured relative to the geographic (X,Y,Z) system, where X and Y point to geographic north and east, respectively. The projection of B on the horizontal plane gives the H-component. It is positive in the direction of the magnetic south pole. D is the declination angle that H makes with X.

The longitudinally symmetric westward ring current produces a southward magnetic field perturbation that suppresses the northward directed dipole geomagnetic field. The longitudinal average depression DB of the ring current, including the dusk-centred partial ring current, relative to the dipole field are used to derive the geomagnetic activity index Dst. A stronger ring current results in a more negative value for Dst, which is observed during so-called magnetic storms, that occur during prolonged intervals of southward IMF. The approximate strengths of the auroral electrojets are indexed by the AL, AU, and AE indices. They are found from measurements of the H-component from a worldwide chain of auroral zone stations, where AU is defined as the maximum positive disturbance at any one time (i.e. the upper envelope of all H-traces) and AL as the maximum negative disturbance, or the lower envelope. The dawnside westward electrojet should ideally be reflected in the AL index while the duskside eastward electrojet by the AU index if all stations are located in the auroral oval. This is not the case however, since the auroral oval generally moves equatorward for southward IMF and poleward for northward IMF and caution is necessary in using these indices. The AE index is defined as AE=AU-AL.