Earth’s magnetosphere and ionosphere

J. De Keyser, Space Physics Division, Belgian Institute for Space Aeronomy, Brussels
S. Stankov, T. Verhulst, Geophysical Department, Royal Meteorological Institute, Brussels

**ABSTRACT**

The Earth’s magnetic field creates a cavity in interplanetary space, called the magnetosphere. Physical processes in this region of space determine how mass and energy from the solar wind reach the ionosphere, the partially ionized upper atmosphere. Magnetosphere and ionosphere are strongly coupled. Together, they modulate the impacts of solar activity on man and technology. This paper presents a brief overview of the magnetosphere-ionosphere system under quiet conditions, followed by a summary of the most important dynamic effects during disturbed conditions.

**SAMENVATTING**

Het aardmagnetisch veld creëert een holte in de interplanetaire ruimte, de magnetosfeer genaamd. Processen in dit gebied in de ruimte bepalen hoe massa en energie vanuit de zonnewind terecht kunnen komen in de ionosfeer, de deels geioniseerde hoge atmosfeer. De magnetosfeer en de ionosfeer zijn sterk gekoppeld aan elkaar. Samen moduleren ze de effecten van de zonneactiviteit op mens en technologie. Dit artikel beschrijft eerst de magnetosfeer en de ionosfeer tijdens kalme periodes, gevolgd door een bondig overzicht van de belangrijkste dynamische verschijnselen gedurende gegevens magnetische storingen.

**RÉSUMÉ**


**Introduction**

The Sun fills interplanetary space with a continuous magnetized flow of charged particles, known as the solar wind. Earth’s internally generated magnetic field, essentially a magnetic dipole tilted about 10° relative to the rotation axis, carves out a magnetic cavity in interplanetary space, the magnetosphere. The magnetosphere reaches down to the upper atmosphere, the ionized upper part of which is the ionosphere. As the solar wind is far from static, the magnetosphere’s size and outer shape changes continuously. At the same time, the Sun’s electromagnetic radiation output (in particular ultraviolet and X-rays) affects the degree of ionization of the Earth’s upper atmosphere. Therefore, both the Sun’s corpuscular and electromagnetic radiation control the state of the closely coupled magnetosphere – ionosphere system. We will first introduce the basic elements of this system under average conditions (non-disturbed or quiet-time conditions). After that, we present a brief review of the various types of time-dependent perturbations (disturbed or active conditions) that modify the behaviour of the system and that make up much of the phenomena associated with “space weather”.

**Non-disturbed conditions**

**The magnetosphere**

At the dawn of the Space Age, in 1961, the US Explorer 12 spacecraft revealed an abrupt magnetic field change as it left the geomagnetic field and entered the solar wind region [Cahill and Amazeen, 1963], leaving the magnetosphere for the first time.

It is important to realize that the magnetized solar wind (Janssens and Vanlommel, this issue) and the gas in the upper atmosphere and in the magnetosphere are almost completely ionized. Since charged particles are much more mobile along magnetic field lines than across, both media do not easily mix. The magnetosphere therefore constitutes an obstacle to the solar wind flow. Moreover, this flow is supersonic (and even super-Alfvénic) and thus resembles the flow around a supersonic jet. The different magnetospheric regions are shown in Fig. 1. The solar wind transitions from supersonic to subsonic at the bow shock, where it starts to feel the presence of the obstacle and its flow is diverted around the magnetosphere, forming the magnetosheath. The interface between the interplanetary magnetic field (IMF) and plasma of solar origin on one hand, and the geomagnetic field lines and plasma of ionospheric origin on the other hand, is the magnetopause [see, e.g., De Keyser et al., 2005]. The magnetopause carries a current associated with the magnetic field change across the interface. The solar wind pressure (mostly the ram pressure) compresses the magnetosphere at the day side to a subsolar magnetopause standoff distance of 10 ?? (Earth radii), while it extends into a long magnetotail on the night side over more than 200 ?? . This tail consists of two nearly empty lobes, the northern one with a magnetic field pointing towards Earth, the southern one with a field...
away from Earth. Both are separated by a current sheet embedded in a denser plasma sheet.

There are two sources for the plasma inside the magnetosphere. First, the magnetopause is not completely impermeable. Various processes do allow solar wind plasma to get across, forming a boundary layer just inward of the magnetopause, which is called the low-latitude boundary layer on the dayside and the mantle at high latitudes and on the nightside. The field lines in this boundary layer extend down into the cusps, which play an important role in plasma entry. The second source of plasma is the ionosphere, the ionized outer atmosphere. Plasma (mostly H\textsuperscript{+}, He\textsuperscript{+}, and O\textsuperscript{+}) can move up out of the ionosphere along magnetic field lines. At higher latitudes, the outflow is called the polar wind. At lower latitudes, where the dipolar field lines are closed, the cold outflowing plasma forms a reservoir, the plasmasphere [Darrouzet et al., 2009].

Magnetospheric plasmas are very tenuous. One therefore has to use spacecraft to sample the plasma and the fields in situ. The time variability of the magnetosphere, however, renders the interpretation of such measurements quite difficult. Indeed, if a spacecraft detects variations in its environment it flies through, one never knows for sure whether these are due to spatial structure or time variations. Present-day magnetospheric missions therefore are based on often multi-spacecraft constellations (e.g. Cluster, Themis, MMS) that allow to disambiguate between these different interpretations.

The ionosphere

The upper atmosphere of the Earth is partially ionized. The neutral part is the thermosphere, the ionized part is the ionosphere (Fig. 2). The degree of ionization generally increases with altitude, but the atmosphere also becomes less dense. Above 600 km, the medium is essentially collisionless.

There are several sources of ionization. First, there is photo-ionization due to solar ultraviolet and X-rays. The vertical stratification of the Earth’s atmosphere leads to the existence of layers with enhanced plasma densities at specific altitudes. The ionospheric D-layer, at 70 – 80 km altitude, is mainly due to Lyman-·ionizing the nitric oxygen. The E-layer, around 100 – 110 km, is due to X-ray and far ultraviolet ionization of molecular oxygen. The F-layer, above about 180 km altitude, corresponds to the densest plasma, with extreme solar ultraviolet ionizing atomic oxygen. While such ionization is produced during the day, recombination processes lead to a reduction of the electron density. Recombination of the ions or de-excitation of atoms and molecules leads to the emission of airglow. At lower altitudes, plasma produced during day time moves up and fills the magnetic flux tubes that make up the plasmasphere, only to descend again during the night time and maintaining a non-negligible F-layer.

The dynamical nature of the ionosphere requires high-resolution multi-instrument permanent monitoring. Fig. 3 shows the diurnal ionospheric variations above Belgium via a real-time ionospheric monitoring and imaging system, LIEDR [Stankov et al., 2003; 2011]. The first three days, 14–16 March 2013, exhibit ionospheric
quiet-time behaviour [Davies, 1990; Hargreaves, 1992; Schunk and Nagy, 2000]. The top panel shows the plasma frequency (a measure of electron density) as a function of height. During the day the electron density in the F-layer (above 180 km) increases and peaks at noon when the Sun is highest above the horizon, while it decreases at night. At the same time, the height of the density maximum goes up at night; this height is also shown as the green line in the bottom panel. The D-layer, with a far lower density than the F-layer, can also be seen in the top panel slightly above 100 km; it only exists while the Sun is up, and disappears very quickly at sunset. The D-layer, which is lower and even thinner, is not visible here. Further developments of the currently operated ground-based monitoring system will utilize satellite-borne observations, providing more detail on the processes occurring in the topside ionosphere [Verhulst and Stankov, 2014].

Due to the changes in solar illumination there are also seasonal variations. The state of the ionosphere depends on geographic position, first and foremost because geographic latitude determines the solar flux, but in addition the geomagnetic latitude plays a role since it directs the flow of the ionospheric particles as part of the coupling to the plasmasphere and magnetosphere. Because the magnetic axis of the Earth is not aligned with the rotational axis (see Verhulst et al., this issue), the ionosphere varies with both magnetic and geographic latitude and longitude.

A second source of ionization is the bombardment of the atmosphere with particles. Such bombardment occurs in the cusps, for instance, where solar wind particles may reach down into the ionosphere. Above the polar cap, there is the polar rain: fast solar wind electrons that follow the half-open magnetic field lines that are connected to the solar wind and thread the lobes. At the same time the polar wind removes ions and electrons from the ionosphere. In the high latitude ionosphere the accelerated particles associated with polar lights may produce a lot of additional ionization, but the associated heating may also promote ionospheric outflows. Finally, solar energetic particles produced intermittently as a consequence of solar activity and the permanent cosmic ray background bombard the atmosphere; because of their high energies, these particles contribute in particular to ionization at lower altitudes.

A third source of ionization is the high-speed entry of meteoroids into the atmosphere. This source changes during the day and with the time of the year.

The charged particle content of the ionosphere controls its electric conductivity. As such, it plays a modulating role in the magnetosphere – ionosphere coupling. But it also has myriads of other consequences. The presence of charged ions may affect the chemistry in the upper atmosphere. The ionization also has a major impact on radio wave propagation. In fact, in 1901 Marconi established the first trans-Atlantic radio link between Cornwall and Newfoundland; because of the Earth’s curvature, the radio waves involved in Marconi’s experiments could not have travelled directly across the Atlantic but must have been reflected from the conducting ionosphere. The ionospheric electron content slightly affects the propagation speed of radio signals leading, for example, to a small phase shift in global navigation system signals and thus accounts for part of the positioning error. Actually, this effect is exploited to determine the ionosphere’s total electron content. Radio wave sounding with ionosondes is still one of the preferred ways to determine the altitude profile of the plasma content in the ionosphere.

The high mobility of charge carriers along the magnetic field lines ensures an electromagnetic coupling between the magnetosphere and the ionosphere. Any electric potential difference in the magnetosphere has an imprint in the ionosphere and, conversely, structures in the ionosphere affect the magnetosphere. Any electric potential difference in the magnetosphere has an imprint in the ionosphere and, conversely, structures in the ionosphere affect the magnetosphere.

One example of this coupling is the ionospheric convection above the polar caps. This is related to the substorm cycle, which will be discussed in more detail in section 3.1. Most of the time, the solar wind magnetic field connects somehow across the magnetopause to the ionosphere. Thus, field lines rooted in the ionosphere are
Tempêtes solaires et leur impact sur les réseaux électriques

Fig. 4: Electric equipotential lines (which correspond to plasma convection patterns) above the northern hemisphere for southward interplanetary magnetic field (IMF), derived from Cluster EDI measurements over 2001-2006 [Haaland et al., 2007]. The figure shows the typical two-cell convection pattern, with day to night convection at the center and return flows at lower latitudes. A very similar situation arises for other IMF orientations, except for northward IMF for which there is a four-cell pattern.

Fig. 5: The substorm cycle. See the text for an explanation of the process. Adapted from [Russell et al., 2008]

dragged tailward since their open end is carried along with the solar wind flow. As a consequence, the footpoints of such field lines move from the day to the night side along the noon-midnight meridian. Figure 4 shows the electric isopotential lines above the northern polar cap as observed by the Cluster EDI instruments [Haaland et al., 2007]. The figure shows the predominant two-cell convection pattern (only for rather strictly northward IMF, there is a four-cell pattern). The equipotential lines correspond to the plasma streamlines, with day to night convection at the noon-midnight meridian until the flow is diverted sideways and returns at lower latitudes. Due to ion-neutral collisions the plasma convection pattern affects the neutral flow in the thermosphere as well.

A second example is evident in the magnetosphere-ionosphere electric circuit. Electric potential differences in the magnetosphere (for instance, those caused by magnetospheric flows in the geomagnetic field) can be short-circuited through the ionosphere. As soon as the magnetospheric driver disappears, the conductivity of the ionosphere determines the time scale over which these potentials disappear, as in an RC circuit. Such phenomena all require field-aligned currents (FACs), also known as Birkeland currents, while the current circuit must close through horizontal currents in the ionosphere. These currents are very much related to the auroral oval, a ring-like region around both magnetic poles where polar lights or aurora are common. At quiet times, the horizontal ionospheric closure currents constitute the so-called auroral convection electrojets, two weak and steady currents in the D and E regions that flow along the auroral oval towards midnight. During perturbed times, so-called substorms, the cross-tail current is diverted through up- and down-going FACs and a local horizontal closure in the ionosphere. The horizontal current is the auroral substorm electrojet, concentrated near midnight and accompanied by intense aurora. The substorm electrojet is a strong and temporary westward current that grows within minutes and spreads in local time and in latitude as substorm activity increases. There exists also an equatorial electrojet, an eastward current near the magnetic equator due to the electric field associated with the diurnal longitudinal change of the subsolar ionization maximum.

Magnetospheric and ionospheric disturbances

The substorm cycle

The magnetosphere appears to be a system that operates close to its limit of stability. As a consequence, it exhibits a constant quasi-periodic dynamical phenomenon, the substorm cycle [Dungey, 1961] as illustrated in Figure 5. Let us focus on the situation where the interplanetary magnetic field (IMF) points southward (the type of interaction actually occurs for most IMF orientations; it is somewhat different for purely northward field). Then reconnection takes place near the subsolar magnetopause. Although scientists do not yet agree on the details of this process, it can be summarized as follows: If two regions with oppositely directed magnetic field are in contact, the field lines will reconnect so that the system can evolve to a state with lower energy.
Reconnection between the IMF and the geomagnetic field at the dayside magnetopause creates half-open field lines that are dragged tailward at one end with the solar wind flow. As a consequence, magnetic flux is added to the magnetotail. This flux is removed by a similar reconnection process at a neutral line in the distant tail. The rate of flux addition and removal is variable. Typically, the system accumulates flux in the tail faster than it can be removed. The system will release the flux surplus suddenly through the formation of a near-Earth reconnection site during a substorm; this may be triggered by a change in the IMF, but if the surplus is large enough such a trigger may even not be apparent. One immediate consequence of near-Earth reconnection is the tailward expulsion of a plasmoid that carries the excess flux away.

Another consequence of reconnection in the near-Earth tail is the injection of accelerated plasma into the inner magnetosphere. This leads to an enhancement of the magnetospheric electric field and modifies the convection patterns, which in turn has an impact on the plasmasphere: its outer layers are stripped off, often in the form of plasmaspheric plumes, swirls of plasma of ionospheric origin that may escape out of the magnetosphere. Reconnection at a neutral line creates electric potential differences across the magnetic field lines surrounding the reconnection site. Such perpendicular potential differences must convert into parallel potential variations when the field lines are rooted in the electrically conducting upper atmosphere because of current conservation. Such parallel potentials can accelerate particles from the magnetosphere and catapult them into the upper atmosphere where they can excite or ionize atoms and molecules; upon their return to the ground state these emit photons and thus create the auroras [see Paschmann et al., 2003, and references therein]. Polar lights therefore appear continuously in the auroral oval, the projection of the plasma sheet boundary in the atmosphere. During an actual substorm the powerful near-Earth reconnection process will create even stronger, intermittent auroras. From the geometry one can understand that the auroral oval on the night side is much broader during a substorm, with auroras occurring also at lower geomagnetic latitudes.

**Magnetospheric storms**

Sometimes the Earth is exposed to major perturbations in the solar wind; these are typically interplanetary coronal mass ejections or corotating interaction regions (see Janssens and Vanlommel, this issue). Both types of events are characterized by shocks as fast solar wind overtakes slower plasma, by an increase in solar wind ram pressure, and by abrupt changes in interplanetary magnetic field direction. During such periods a number of phenomena are triggered that are similar to those observed during substorms, except that they are much more intense. These events are called magnetospheric storms.

An increase in solar wind pressure leads to a compression of the magnetosphere. The subsolar magnetopause may be pushed Earthward even inward of geostationary orbit (at 42000 km from the Earth’s centre), sometimes with dire consequences for spacecraft there. Storms lead to strong near-Earth reconnection, large polar caps, very intense auroras down to low latitudes, injection of energetic particles into the inner magnetosphere, changes in the magnetospheric electric field, stripping of the plasmasphere, and corresponding ionization signatures in the ionosphere. Storms imply various particle acceleration processes, such as adiabatic acceleration during magnetospheric compression and acceleration by induced electric fields. As a consequence, a reservoir of energetic particles is created that is trapped in the Earth’s dipole field; the radiation belts that were discovered by Van Allen at the dawn of the Space Age and that are still of major technological interest because of the harmful effects of this ionizing radiation on man and machine.

Storms imply a drastic and sudden change of the magnetospheric magnetic field. A changing magnetic field implies induced electric fields, currents, and secondary magnetic fields, also in the ionosphere and at the Earth’s surface. The storm-time auroral electrojet is a prime example. Such geomagnetically induced currents may be disruptive to electric power utility grids and pipelines, two of the more dramatic space weather effects.

**Ionospheric storms**

Ionospheric storms [Buonsanto, 1999] are defined as strong ionospheric disturbances resulting, ultimately, from the highly dynamic processes occurring on the Sun and in the magnetosphere. Two major sources of such disturbances can be distinguished: solar flares-outbursts in X-rays that cause an almost immediate increase in ion production, and disturbances in the solar wind such as corotating interaction regions and coronal mass ejections (Janssens and Vanlommel, this issue). In the latter case, an ionospheric storm occurs in parallel with a magnetospheric storm.

The ionospheric perturbations during storms exceed the median variability and manifest complex behaviour depending on spatial location and season. The F-region response is easily detected from ionosonde observations of the ionospheric peak density and height. Similarly to geomagnetic storms, an ionospheric storm goes through an initial phase (lasting a few hours but occasionally brief or even missing) when the electron density and the electron content are greater than the ‘normal’ median values, followed by a main phase when the above quantities are below pre-event values. The main phase often lasts for several days and during this period the ionosphere gradually returns to its normal behaviour.

The multiple instruments available at the RMI Geophysical Centre in Dourbes [Jodogne and Stankov, 2002] are able to assess the geomagnetic and ionospheric storm conditions in Belgium in real time. For example, at 06:00 UT on March 17, 2013, Earth was hit by a coronal mass ejection, which resulted in an ionospheric storm as can be seen in Figure 3. The most obvious effect is the increase in the total electron content (black line in the middle panel) in day time on March 17, followed by a decrease during the night to a level well below that of the unperturbed nights before. After the storm, it takes a few days before the ionization level returns to quiet-time values.
Conclusion

The magnetosphere is a complex environment that shields the Earth from perturbations in the interplanetary medium. The ionosphere, as the partially ionized top layer of the atmosphere, plays an important role in these shielding processes as it determines the conductivity in the electromagnetic circuit that controls the transient response of the coupled magnetosphere – ionosphere system. A proper understanding of this system is of paramount importance if we want to protect ourselves from the harmful consequences of space weather phenomena. Thanks to the Federal Scientific Institutions and the Solar-Terrestrial Centre of Excellence initiative, Belgium currently assumes a leading role in Europe in this field, both in fundamental magnetospheric and ionospheric research, e.g. through its involvement with ESA’s Cluster mission, the EU’s 7-th Framework Programme, and in the area of scientific services, with an active involvement in ESA’s Space Situational Awareness – Space Weather Element programme.

Acknowledgements

This work was supported by the Belgian Science Policy Office (BELSPO) through PRODEX in the context of the Cluster mission and through the Solar-Terrestrial Centre of Excellence (STCE).

References