The science of space weather

BY JONATHAN P. EASTWOOD*

Space Sciences Laboratory, University of California, Berkeley, 7 Gauss Way, Berkeley, CA 94720, USA

The basic physics underpinning space weather is reviewed, beginning with a brief overview of the main causes of variability in the near-Earth space environment. Although many plasma phenomena contribute to space weather, one of the most important is magnetic reconnection, and recent cutting edge research in this field is reviewed. We then place this research in context by discussing a number of specific types of space weather in more detail. As society inexorably increases its dependence on space, the necessity of predicting and mitigating space weather will become ever more acute. This requires a deep understanding of the complexities inherent in the plasmas that fill space and has prompted the development of a new generation of scientific space missions at the international level.

Keywords: Sun; solar–terrestrial physics; magnetic reconnection; space weather

1. Introduction

Fifty years ago, on 31 January 1958, the satellite Explorer 1 was launched, resulting in the discovery of belts of particle radiation about the Earth (van Allen & Frank 1959), trapped in what has become known as the magnetosphere. This directly revealed that ‘empty space’ was not in fact empty, but filled with particles and magnetic fields. Since then, it has become clear that the invisible link between the Sun and the Earth, mediated by the solar wind, results in a complex plasma environment in near-Earth space.

Understanding this complexity is a fundamental problem in physics. However, it is a problem which has a high-profile, significant and rapidly growing impact on society, because conditions in space can be extremely hazardous. The space business touches our everyday life in increasingly numerous ways, from the mundane, such as global positioning system units in cars, to the exotic, such as the trip of a lifetime on a suborbital rocket flight; it is estimated that in 2007 the revenues of the global space economy surpassed $250 billion (The Space Foundation 2008).

The term space weather is defined as the conditions on the Sun and in the solar wind, Earth’s magnetosphere, ionosphere, and thermosphere that can influence the performance and reliability of space-borne and ground-based technological systems and endanger human life or health. As described in this article, adverse...
space weather has caused power blackouts, the failure of radio communications, the loss of satellites, implementation of astronaut safety procedures on the International Space Station (ISS), and many other problems.

In §2 we begin by briefly reviewing the Sun–Earth system and some different sources of space weather. Although there are many different plasma processes at work, it will be seen that one of the most important is magnetic reconnection; in §3, recent cutting edge work in this field is reviewed. In §4, this work is placed in context by describing specific examples of adverse space weather in more detail. Finally, in §5, we look to the future and describe the new generation of international projects designed to study the complexity of collisionless space plasmas and to forecast space weather. We also speculate on the future exploitation of space, which will mean that weather forecasting will become as important in space as it is now on the ground.

2. The Sun–Earth connection

Coronal mass ejections (CMEs), massive transient eruptions of solar coronal material into space, are one of the most significant sources of space weather. Figure 1 shows an example of a CME captured by the LASCO instrument on board the SOHO spacecraft. The Sun’s magnetic dipole reverses every ~11 years and the number of sunspots rises and falls on a similar time scale. CMEs are more common during ‘solar maximum’ when there are more sunspots and the dipole is in the process of reversing. If a CME strikes the Earth’s magnetosphere, it can generate significant space weather effects, as described in more detail below.
Solar flares are another very important space weather phenomenon, where energy stored in the coronal magnetic field is rapidly released. It is widely thought that this rapid release is due to magnetic reconnection (e.g. Harra 2002) as shown in figure 2. Accelerated electrons travel down the magnetic field lines into the dense chromosphere resulting in the emission of hard X-rays. The heated chromospheric plasma subsequently evaporates, resulting in soft X-ray emission from the loop itself. Figure 2 also shows a flare observed by the extreme-ultraviolet imaging telescope (EIT) on SOHO.

Figure 2. (a) A cartoon of the basic physical mechanism at work in a solar flare. Magnetic reconnection in the corona converts two open magnetic field lines (purple) into closed (red) and open (blue) field lines and injects plasma back towards the surface of the Sun, resulting in X-ray and ultraviolet emission. (b(i)(ii)) Two ultraviolet 195 Å images of the Sun are shown before and, 12 minutes later, during a solar flare. Courtesy of SOHO/EIT consortium. SOHO is a project of international cooperation between the European Space Agency (ESA) and NASA.

Figure 3. During southward IMF, magnetic reconnection occurs at the subsolar magnetopause between the IMF (blue) and magnetospheric field lines (red). This results in ‘open’ field lines (purple), which are convected over the poles by the solar wind flow. Magnetic energy, stored in the lobes, is eventually released during a substorm.

Solar flares are another very important space weather phenomenon, where energy stored in the coronal magnetic field is rapidly released. It is widely thought that this rapid release is due to magnetic reconnection (e.g. Harra 2002) as shown in figure 2. Accelerated electrons travel down the magnetic field lines into the dense chromosphere resulting in the emission of hard X-rays. The heated chromospheric plasma subsequently evaporates, resulting in soft X-ray emission from the loop itself. Figure 2 also shows a flare observed by the extreme-ultraviolet imaging telescope (EIT) on SOHO.
Both CMEs and solar flares are sources of a third important space weather phenomenon: relativistic protons and electrons referred to as solar energetic particles (SEPs). ‘Impulsive’ SEP events are thought to be associated with solar flares and are produced in the solar corona directly whereas ‘gradual’ SEP events are produced by the shock waves driven by CMEs as they propagate into the heliosphere (Reames 2002).

Another important source of space weather does not tend to occur during solar maximum. During solar minimum the solar dipole is well defined and fast solar wind emanates from polar coronal holes with slow solar wind observed closer to the equator. However, because of dipole tilt effects and the irregular shape of coronal holes, fast solar wind can find itself running into slow solar wind further ahead. This leads to the formation of corotating interaction regions (CIRs), compressed regions of plasma following a spiral pattern in the heliosphere (Tsurutani et al. 2006), which can cause space weather effects due to their interaction with the magnetosphere.

The terrestrial magnetosphere itself is mainly controlled by the solar wind via magnetic reconnection (Dungey 1961; Fairfield & Cahill 1966). If the interplanetary magnetic field (IMF) points northward, parallel to the Earth’s magnetic field at the subsolar magnetopause there is very little geomagnetic activity, and the system appears to be quite stable (Kennel 1995). During southward IMF, oppositely directed field lines undergo reconnection at the subsolar magnetopause as shown in figure 3. Magnetic flux is steadily added to the tail lobes, increasing the stored magnetic energy. Eventually, this energy is explosively released in a substorm, resulting in dynamic auroral displays which significantly disturb the Earth’s magnetic field at ground level, together with the ejection of plasma out of the magnetotail (Kennel 1995). Although the exact sequence of events within the magnetotail that leads to a substorm being triggered is still unclear (this is the goal of the currently operating NASA THEMIS mission), it is widely acknowledged that magnetic reconnection in the magnetotail is the ultimate explanation for the release of the stored magnetic energy.

If the IMF points southward for a prolonged period of time, and if the solar wind speed is also high, substorm activity can develop into a geomagnetic storm, with intense auroras at low latitudes and a noticeable depression in the Earth’s low latitude magnetic field strength due to increased equatorial currents in space. In general, the largest geomagnetic disturbances are associated with fast CMEs, because they can often contain long intervals of strongly southward IMF (e.g. Richardson et al. 2001). Similarly, intervals of high-speed, southward IMF solar wind in the trailing part of CIRs are a significant source of magnetic storms, although they tend to be weaker than the ones generated by CMEs (Tsurutani et al. 2006).

3. An example of the physics underpinning space weather prediction: magnetic reconnection

Although there are many plasma phenomena that contribute to space weather, such as collisionless shocks, turbulence, wave–particle interactions, etc., the process of magnetic reconnection is considered crucial because it is thought to
enable the storage and subsequent release of energy in both the solar atmosphere and the terrestrial magnetosphere. Here we discuss the physics of magnetic reconnection and some recent cutting edge work in more detail.

In plasmas, individual particles orbit the magnetic field with a characteristic gyroradius, which depends among other things on the ambient magnetic field. If the magnetic field varies on scales much larger than the gyroradius, the particles gyrate on the ‘same’ field lines and the plasma is said to be ‘frozen in’ (e.g. Dendy 1990). The consequence of this is that two plasmas attached to ‘different’ magnetic fields (for example the solar and terrestrial dipole) cannot mix, and a current sheet (e.g. the magnetopause) forms between them (figure 4).

At this current sheet the magnetic field can vary on scales similar to the gyroradius, and the plasma particles may no longer be frozen to the field. If the fields on either side are oppositely directed, this may result in reconnection (figure 4), where the plasma becomes ‘demagnetized’ in a small central diffusion region, resulting in an X-line configuration and plasma jets. To ultimately understand the large scale consequences of reconnection, and predict where, when and how reconnection will occur, it is necessary to study the physics that occurs within the diffusion region.

Recent observational and theoretical results have shed new light on the properties of the diffusion region. In particular, is now clear that it has a multi-scale structure. Ions and electrons have different gyroradii because of their different mass, and in fact the ions demagnetize on scales comparable to the ion inertial length (defined by the gyroradius of an ion moving at the Alfvén speed) in the ion diffusion region, whereas the electrons demagnetize on the much smaller electron inertial length scale. This means that, except in the case of a positron–electron plasma, collisionless reconnection is fundamentally a two scale process. It has been suggested that this plays a crucial role in controlling the dynamics of reconnection, and in particular enables ‘fast’ reconnection (Birn et al. 2001) with energy release rates comparable to those observed in the solar corona and magnetosphere.

The ion diffusion region is small (a few hundred kilometres wide at the magnetopause), and only within the last decade have magnetospheric spacecraft made in situ observations confirming its existence and the predictions of this two scale model (Fujimoto et al. 1997; Nagai et al. 2001; Øieroset et al. 2001; Mozer et al. 2002). More recently, a great deal of progress has stemmed from observations made by the European Space Agency (ESA)/NASA Cluster mission (Escoubet et al. 2001). Cluster was launched in 2000 and consists of four identical spacecraft flying in formation in polar orbits around the Earth. The orbits are tuned so that the spacecraft form a tetrahedron in regions of interest. Four spacecraft allow spatial structure to be decoupled from temporal evolution (Paschmann & Daly 1998). One can determine the motion and thickness of boundaries, measure the current within the tetrahedron volume and determine the dispersion relation of the local plasma waves. These techniques have been used to study the diffusion region at the magnetopause (Vaivads et al. 2004), in the magnetotail (Runov et al. 2003; Borg et al. 2005; Wygant et al. 2005; Nakamura et al. 2006; Xiao et al. 2006; Eastwood et al. 2007) and in the magnetosheath (Phan et al. 2007).

Another important topic in reconnection research is understanding the role it plays in energetic particle production. As mentioned above, reconnection is thought to lie at the heart of solar flares, which deposit a significant fraction of

Phil. Trans. R. Soc. A (2008)
their energy in energetic electrons (Lin et al. 2003). Spacecraft have made in situ observations of energetic (hundreds of keV) electrons near reconnection sites in the magnetotail, helping to establish their exact non-thermal properties (Øieroset et al. 2002; Asano et al. 2008). Although the specific source of these energetic particles is still uncertain, one proposed mechanism suggests that they arise from the interaction of particles with ‘secondary magnetic islands’ in the outflow jets (Drake et al. 2006), which have been observed in association with reconnection (Eastwood et al. 2007).

The importance of reconnection is underlined by the development of the NASA Magnetospheric Multi-Scale mission scheduled to launch in 2014. This mission, designed to specifically study the magnetic reconnection process, was ranked as the highest priority medium-scale mission in the 2003 National Research Council decadal survey of solar and space physics research.

4. Examples of space weather

(a) Power grids

Geomagnetic storms drive large variations in the magnetic field that can induce geoelectric fields in the Earth’s surface over large areas (Kappenman 2005). These electric fields drive geomagnetically induced currents (GICs) through conducting networks such as power grids. GICs can cause core saturation in power transformers eventually leading to their failure. One of the most well-documented blackouts occurred on 13 March 1989 when the Hydro-Quebec power system failed, as a result of the GICs caused by a CME-driven geomagnetic storm (Bolduc 2002). More recently, on 30 October 2003, a geomagnetic storm caused a blackout in southern Sweden (Pulkkinen et al. 2005) and also caused noticeable anomalies in the Scottish power grid (Thomson et al. 2005). Modern networks will tend to collect GICs in the highly conducting high-voltage lines and a large storm could cause a cascading failure of the grid, in a manner similar to the 14 August 2003 North American blackout which was caused by a power plant tripping offline (Kappenman 2003).
(b) Pipeline corrosion

Pipelines are also susceptible to GICs that introduce a potential difference between the pipe and the soil it is buried in. As well as directly affecting the corrosion of pipes, GICs also interfere with operation of corrosion protection systems. Long pipes, with modern coatings and which lie at high latitude are most exposed to the effects of GICs; GICs in the Finnish natural gas pipeline have been measured since 1998, leading to the development of a nowcasting service, GIC Now!, implemented by the Finnish Meteorological Institute within the ESA Space Weather Applications Pilot Project (Viljanen et al. 2006).

(c) The global positioning system

Global positioning system (GPS) receivers use signals from multiple satellites in orbit around the Earth to triangulate their position. As the radio signals travel through the ionosphere between ground receivers and the satellites, ionospheric irregularities can cause both diffraction and refraction (Kintner et al. 2007). Both effects can degrade GPS signals, causing receivers to lose lock or generate ranging errors, and effects are more severe at solar maximum and during geomagnetic storms. GPS is also vulnerable to the radio emission that accompanies solar flares and which can effectively drown out the signal. Solar flare radio emission will affect all the GPS satellites simultaneously, and it is thought that the radio bursts at solar maximum from large flares are capable of adversely affecting GPS receivers. A solar radio burst can last a few hours, during which time the receiver may not be able to produce a navigation solution (Cerruti et al. 2006).

(d) Air travel

Polar routes, such as those used to connect the continental USA and Asia, are most susceptible to space weather. The Federal Aviation Authority (FAA) has established a number of requirements for polar flights, including reliable communication over the entire flight route. Above 82°N, geosynchronous satellites cannot be used to communicate, and HF radio is used. During periods of solar activity, HF radio is often disrupted meaning that flights must be re-routed, costing in the region of $100 000 per flight (Balch 2004). During the major solar activity of October–November 2003, NOAA’s Space Weather Prediction Center (SWPC) participated in three to five teleconferences per day with the major US airlines to help plan polar operations (Balch 2004). Polar flights may also be more exposed to SEP events since energetic particles can more easily access the polar regions. The FAA issued its first, and so far only, advisory during the October 2003 storms. SEPs could affect avionics when charged particles become embedded in the system electronics, causing memory corruption or other anomalies that might impact the operation of the vehicle (e.g. Jones et al. 2005).

(e) Satellite anomalies

Space weather can cause a variety of satellite anomalies such as surface charging, surface damage, deep dielectric charging, solar panel degradation, and background counts in sensors (Iucci et al. 2005). The varying magnetic
field can cause satellite orientation problems, and the expansion of the Earth’s atmosphere from SEP-related heating increases satellite drag and changes their orbit. During the October–November 2003 storms, 33 Earth orbiting satellite anomalies were reported (Webb & Allen 2004). Accurate predictions of geomagnetic activity can save money. During the 2003 storms, the SWPC was able to assist a commercial satellite company by correctly predicting that the activity would not exceed the company’s threshold beyond which they would have to put their satellites into safe mode. Such an action would have cost several million dollars per day, and so good knowledge of the space environment enabled this cost to be avoided (Balch 2004).

(f) Spaceflight

On 28 October 2003 during the October–November 2003 storms, solar flare activity caused flight controllers to ask astronauts on the ISS to remain in the most shielded part for five separate 20 min periods (Balch 2004). Solar activity in December 2006 also forced mission controllers to advise astronauts on the ISS and the shuttle to remain in the most shielded parts of their vehicles. While astronauts in Earth orbit can take evasive action, astronauts on the Moon are potentially much more exposed. During the Apollo era, a large solar eruption occurred on 4 August 1972, between Apollo 16 (April 1972) and 17 (December 1972). If this solar eruption had occurred during these missions, and the astronauts were on the surface of the Moon protected only by their spacesuits, they would have suffered possibly fatal radiation sickness (Miroshnichenko 2003).

5. Future developments

The science of space weather is particularly challenging because it requires us to understand the behaviour of large systems such as the solar corona and the magnetosphere which are irreducibly complex. As such, there is a vigorous programme of research in this area and table 1 lists some future satellite missions planned by the international community to study different aspects of solar–terrestrial physics.

Table 1. Future missions.

<table>
<thead>
<tr>
<th>mission name (agency)</th>
<th>target</th>
<th>launch date</th>
</tr>
</thead>
<tbody>
<tr>
<td>C/NOFS (USAF/NASA)</td>
<td>ionospheric scintillation</td>
<td>16 Apr 2008</td>
</tr>
<tr>
<td>Solar Dynamics Observatory (NASA)</td>
<td>solar magnetic field</td>
<td>2009</td>
</tr>
<tr>
<td>Radiation Belt Storm Probes (NASA)</td>
<td>radiation belts</td>
<td>2011</td>
</tr>
<tr>
<td>Magnetospheric Multi-Scale (NASA)</td>
<td>magnetic reconnection</td>
<td>2014</td>
</tr>
<tr>
<td>Solar Orbiter (ESA/NASA)</td>
<td>solar corona</td>
<td>2015</td>
</tr>
<tr>
<td>Solar Probe Plus(^a) (NASA)</td>
<td>solar corona ((in situ))</td>
<td>2015</td>
</tr>
<tr>
<td>Cross Scale(^b) (ESA/JAXA)</td>
<td>multi-scale plasma coupling</td>
<td>2017</td>
</tr>
<tr>
<td>Solar Sentinels(^a) (NASA)</td>
<td>inner heliosphere</td>
<td>2018</td>
</tr>
</tbody>
</table>

\(^a\)Planned.

\(^b\)Under study.
Scientific investigations into the physics of space weather are also driven by very demanding and strong external pressures. Firstly, our everyday life has an ever-increasing reliance on complex technological systems in space and these systems are becoming increasingly sensitive to the vagaries of space weather, in part because of their very complexity. Secondly, it is likely that in the near future, access to low Earth orbit, by people and satellites, will become dominated by the demands of private industry rather than governmental space programmes. The most high profile examples of this trend are new companies such as Virgin Galactic who are exploiting popular demand for space tourism by offering sub-orbital rocket flights. Space travel is a popular dream, and even though the current cost of a seat is several hundred thousand dollars, many people are willing to pay for a ticket. The price will certainly fall in the future, opening up access further. Thirdly, the major national and international space agencies are expected to continue to expend their human exploration effort on the Moon and Mars. However, it is incredibly difficult, and consequently expensive, to reach Mars. Therefore, although it is possible that other targets such as asteroids will be identified, it is most likely that some time will be spent establishing an essentially continuous presence on the Moon.

All of these diverse activities will require knowledge of the space weather conditions to protect both people and assets in space, leading to a significant societal demand for reliable space weather prediction and mitigation. As a result, the market for commercial space weather services is likely to be highly competitive.

The first 50 years of the space age has seen a dramatic jump in our understanding of the near-Earth space environment and the first steps into space. Within the next few years, it will be possible for an average member of the public to make a once of a lifetime trip into space, and this is likely to open the floodgates for a vast expansion in access to and commercial exploitation of space. As such, it is important that research into the physics that underpins space weather is properly supported, since without proper scientific foundations, the exploration and exploitation of space could be a risky business indeed.

For their comments on the manuscript, and many useful discussions about space weather and reconnection, J.P.E. would like to thank Andre Balogh, Laura Peticolas, Tai Phan and Michael Shay. This work was supported by NASA grant NNX07AF32G.

References


Jonathan P. Eastwood studied at Imperial College London, UK, where he graduated in 2000 with an MSci degree in physics (1st class with honours), having spent his third year studying abroad at l’Ecole Superieure de Physique et Chimie Industrielles de la ville de Paris (ESPCI, Paris), France. He received his PhD in physics in 2003 from Imperial College London, UK, where he worked on the ESA Cluster mission and studied the behaviour and properties of the Earth's bow shock. Subsequently, he was awarded a National Research Council Resident Research Associateship at NASA Goddard Space Flight Center, Maryland, USA, and in 2005 moved to the University of California at Berkeley, USA, where he is an assistant research physicist in the Space Sciences Laboratory. In his research, he aims to understand the basic science that governs space weather, in particular the physics of shocks and magnetic reconnection.