Looking through the oval window

The 2008 Rishbeth Prize for the best poster displayed at the National Astronomy Meeting in Belfast was awarded to "A superposed epoch analysis of auroral evolution during substorm growth, onset, and recovery" by S E Milan, A Grocott, C Forsyth, S M Imber, P D Boakes and B Hubert. Steve Milan summarizes.

Abstract

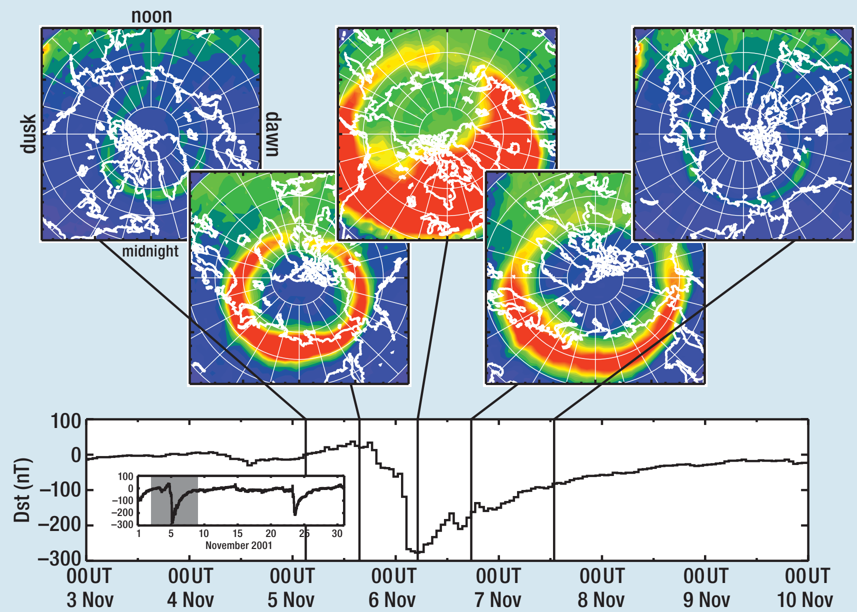
Magnetic reconnection between Earth's magnetosphere and the solar wind is a fundamental yet little-known process. Data from NASA's IMAGE spacecraft show changes in the auroral oval that reveal more about substorms, elements of the reconnection process. In particular, we found that the initiation of substorms depended on the amount of flux previously stored in the magnetosphere, but not exclusively. The strength of the ring current may also be a factor.

The magnetosphere can be a dynamic and hostile environment for humans and spacecraft, but almost all the activity within it is invisible. When we look up at the Sun at midday we are unaware that, 65 000 km above our heads, the solar wind is crashing into our magnetic shield in a region of space where twisting and knotting magnetic fields store vast quantities of magnetic energy. When looking at the night sky we don't see the huge reservoir of energy in our celestial backyard – the magnetotail – filling with this magnetic energy and solar-wind plasma until it reaches bursting point. Only aurora, visible at high latitudes, bear witness to the deposition of energy and particles in the atmosphere and the reconfiguration of near-Earth space when, every few hours, this stored energy is released suddenly and explosively in a process known as a substorm. And so it is to the aurora we look when we try to understand the complex series of physical processes that mediate the interaction between the solar wind and our planet.

Figure 1 shows five snapshots of the northern hemisphere UV aurora taken by the NASA IMAGE (Imager for Magnetopause-to-Aurora Global Exploration) spacecraft during a geomagnetic storm in November 2001, triggered by the impact of a solar-wind shock on the magnetosphere. The evolution of the storm can be traced by the variation in the Dst index, a measure of the deflection of the magnetic field at the equator due to the enhancement of a ring of current and the radiation belts that encircle the Earth. The range in intensity and location of the auroral oval is apparent, and it is this that provides key information on solar wind–magnetosphere–ionosphere coupling. However, although impressive, storms are not our primary consideration; rather it is substorms, which take place every few hours, that represent the quanta of solar wind–magnetosphere–ionosphere coupling.

The dynamic magnetosphere

The Sun blows off 10^9 kg of plasma into the heliosphere every second. Streaming outwards at 500 km s^{-1}, this solar wind carries the magnetic field of the Sun embedded within it to form the interplanetary magnetic field (IMF). The terrestrial magnetic field presents an almost impenetrable obstacle to the solar wind, which it deflects to leave a low-density region called the magnetosphere surrounding the Earth. The size and shape of the magnetosphere is determined by stress balance between the ram pressure of the solar wind and the magnetic field pressure on the inside, being compressed on the side of the Earth facing the Sun, the dayside, and forming a longer magnetotail on the nightside (figure 2). Inside, the magnetosphere is highly dynamic, tapping energy and momentum from the solar wind as it flows past. As first proposed by Dungey (1961), the exclusion of the solar wind by the magnetic field of the planet is not perfect: a plasma process known as magnetic reconnection can peel away the outer boundary of the dayside magnetosphere and interconnect it with the IMF, especially when the IMF is directed southwards. Interconnected field lines make the magnetopause leaky, allowing solar wind plasma to enter the magnetosphere. At the same time, these interconnected field lines are stretched by the flow of the solar wind to great lengths behind the planet to form the magnetotail, before being pinched off by magnetic reconnection on the nightside to dis-



1: Five false-colour snapshots of the northern auroral oval, taken by the SI12 camera of the IMAGE spacecraft in November 2001 during the evolution of a geomagnetic storm. The Dst index shows the deflection of the magnetic field at the equator of the Earth due to the presence of an enhanced ring current, characteristic of a geomagnetic storm.

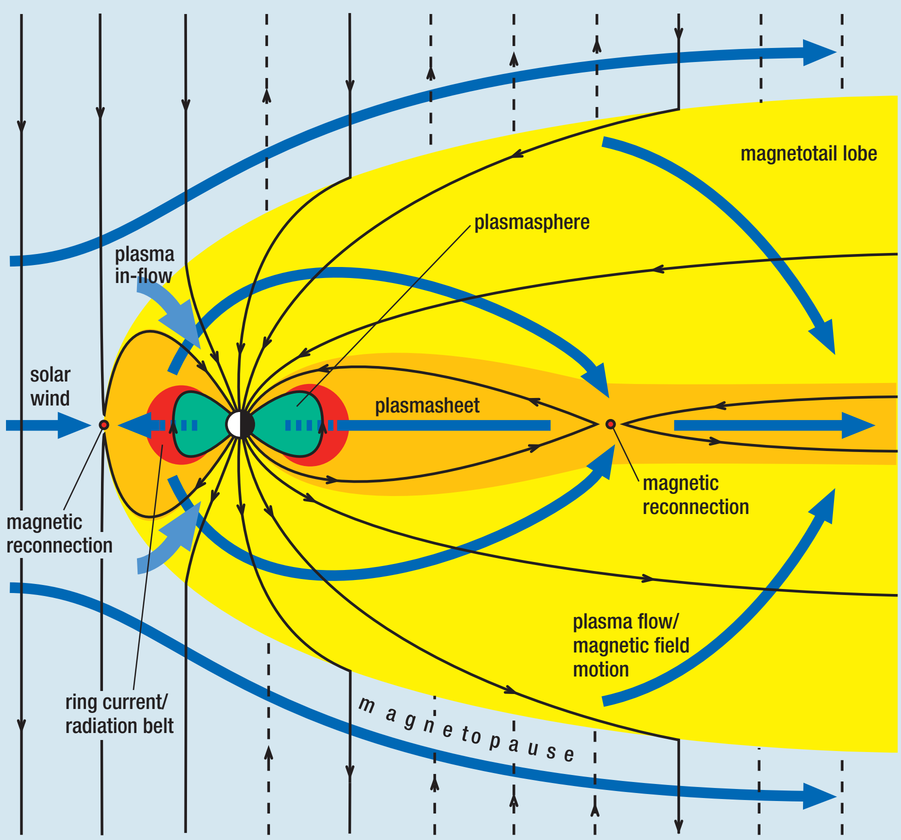
connect them from the solar wind. The resulting circulation of magnetic field and plasma within the magnetosphere is now known as the Dungey cycle, and it is this circulation that is responsible for creating structured magnetic field and plasma regions within the magnetosphere – for instance, the hot, dense plasmasheet and the evacuated magnetotail lobes.

The auroral ovals mark the locations where the plasmasheet reaches down into the atmosphere; they encircle the dim polar regions which are the footprints of the lobes. The magnetospheric circulation maps down into the polar atmosphere as well, exciting a large-scale circulation of the ionosphere that can be observed with ground-based radars, another key diagnostic of SW–M–I coupling.

There she blows!

Dungey envisaged the cycle of connecting and disconnecting solar-wind magnetic field lines, or opening and closing magnetospheric field lines, as a steady-state process, suggesting that reconnection at the dayside and reconnection in the tail occurred in step. It subsequently became apparent that this could not be the case, that the reconnection process on the night side could not anticipate the influence of the ever-changing solar wind and IMF conditions on the dayside reconnection rate. This inevitably means that the interconnection of the magnetosphere with the solar wind is variable and that the amount of magnetic flux forming the magnetotail lobes changes with time (e.g. Cowley and Lockwood 1992, Milan *et al.* 2003), being typically between 0.2 and 1 GWb, which is between 3% and 12% of the Earth's magnetic field (Milan *et al.* 2007). This in turn determines the size of the dim region inside the auroral oval and the latitude at which the auroral oval is seen. When the magnetosphere is only slightly open, the aurora is found at high latitude; when the level of interconnection is great, and the magnetotail is ripe with accumulated magnetic energy, the aurora migrates to lower latitudes.

The magnetotail inflates as it accumulates energy, flaring outwards and increasing the cross-section it presents to the solar wind. As the tail swells, the pressure within grows until it can no longer sustain itself and magnetic reconnection is triggered to pinch off magnetic flux and release stored energy, accelerating plasma into the nightside ionosphere to produce vivid auroral displays: this is the substorm.

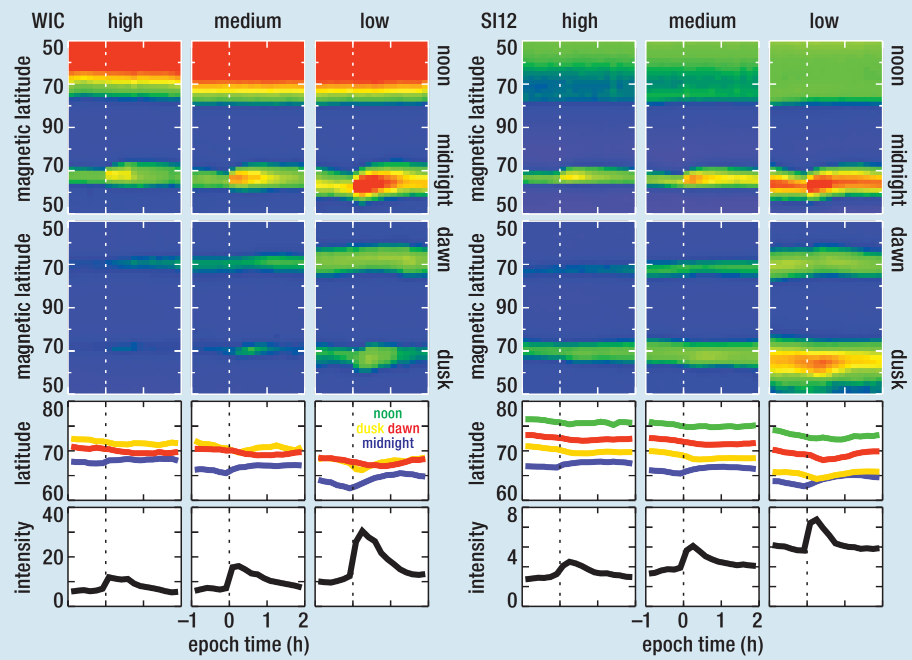


2: Schematic of the magnetosphere showing the configuration of the magnetic field (black arrowed lines) and the main plasma regions. Within the solar wind, full black lines show southward-directed interplanetary magnetic field (IMF), which interacts with the magnetosphere at the nose through magnetic reconnection, while dotted lines show northward IMF, which to first order does not and flows around the sides of the magnetosphere (not shown). Full blue arrows show the circulation of plasma and magnetic field in the Dungey cycle.

Life cycle of a substorm

Substorms were first identified in auroral observations in the 1960s (Akasofu 1964). They were found to occur in a relatively reproducible sequence of events lasting approximately an hour and repeating every few hours: a movement of the auroral ovals to lower latitudes as magnetic energy is accumulated in the tail (the substorm growth phase); a sudden and localized brightening of the nightside auroral oval as the substorm is triggered (substorm breakup); poleward, eastward and westward progression of the region of auroral disturbance, marking the pinch-off of magnetic flux (the expansion phase); and a return to more quiescent conditions (recovery phase). In the 40 or so years since these early studies there have been approximately 100 000 substorms, but they are still poorly understood. Many case studies of the auroral development during substorms have been documented over the years, showing that no two substorms are exactly alike but result from the complex interplay of different physical processes in several different regions of the magnetosphere. There is a danger of not seeing the wood for the trees, and so we decided to determine the average auroral evolution during a large sample of substorms. To this end we turned to observations from the IMAGE spacecraft, which spent five years mapping the dynamics of the Earth's auroral ovals with a suite of three far-ultraviolet cameras comprising the FUV instrument package.

We concentrated on the northern hemisphere aurora and, for orbital considerations, the period May 2000 to April 2002 was identified as most appropriate for study. We used the Wideband Imaging Camera (WIC), sensitive mostly to aurora produced by the impact of electrons on the atmosphere, and the Spectral Imaging (SI12) camera, sensitive to emissions from precipitating protons. SI12 has the advantage over WIC in that it is relatively insensitive to dayglow (sunlight scattered from the dayside of the planet); because of dayglow we had to restrict our WIC observations to the northern hemisphere winter months of January, November and December. Our analysis was aided considerably by a previous study that had determined the onset times and locations of all substorms observed within the first two years of the IMAGE dataset (Frey *et al.* 2004). As substorms can take place when the auroral oval is contracted or expanded, we subdivided the dataset by onset latitude (as recorded by Frey *et al.*) into three categories: high, low and medium latitude onsets, that is those where the onset brightening occurred at latitudes above 68°, below 65°, and those between; these limits were chosen to provide similar numbers of substorms in each category. It was than a relatively straightforward, though computationally intensive, task to map the auroral images into a suitable magnetic coordinate system and then perform a superposed epoch



3: Results of the superposed epoch analysis of auroral evolution during more than 2000 substorms.

analysis of the average auroral configuration in ten-minute intervals during the three-hour period spanning one hour before to two hours after substorm onset. More than 2000 substorms were included in the analysis (a subset of 500 for WIC), requiring the processing of some 250 000 individual images.

The bigger they come...

The results, shown in figure 3, are separated into two parts: those from WIC and those from SI12 (left and right, respectively). The upper row of panels show the average auroral evolution from one hour prior to substorm onset to two hours after, along the noon–midnight meridian of the auroral oval (see figure 1). At the top of each of these panels is a contribution from dayglow, which in the case of WIC entirely masks the noon-sector auroral evolution. The second row shows the auroral evolution along the dawn–dusk meridian. The third row shows the variation in latitude of the peak in brightness of the noon, midnight, dawn and dusk meridians of the oval. Finally, the bottom panel shows the variation in intensity of the midnight-sector oval. Several conclusions can be drawn immediately.

The oval brightness and expansion phase intensity increases dramatically from the *high* to the *low*-latitude substorms. Low-latitude substorms correspond to periods when the interconnection between the magnetosphere and the solar wind is greatest, when most magnetic energy has accumulated in the magnetotail, and which hence have the greatest capacity for energy deposition in the atmosphere. Indeed, a superposed epoch analysis of the concurrent IMF conditions for each substorm (not shown) demonstrates that the IMF is more southward prior to the low-latitude substorms.

The dawn-sector oval is brighter than the dusk sector in the WIC observations, and vice versa for SI12. This can be understood in terms of the behaviour of electrons and protons injected into the inner magnetosphere from the tail: the charged particles are trapped by the dipolar magnetic field, gyrating around the field lines, bouncing between the northern and southern hemispheres, and drifting eastwards in the case of electrons and westwards in the case of protons. Hence the dawn or dusk sector is favoured for deposition in the atmosphere.

There is an equatorward progression of the auroral oval during the growth phase (i.e. prior to onset) marking the increase in the interconnection of the magnetosphere and solar wind. This is followed by a brightening and poleward progression of the nightside auroral oval after onset, marking the pinch-off of open flux by the substorm reconnection process and the acceleration of charged particles into the atmosphere. In addition, a post-onset brightening of the dusk-sector oval is apparent in both WIC and SI12, indicating considerable asymmetry of the substorm “auroral bulge”.

Finally, the equatorward progression of the oval is most pronounced for low-latitude substorms, as is the post-onset poleward progression of the midnight sector oval. Along other meridians the oval can continue to progress equatorwards following onset for anywhere between 20 minutes and 1 hour, indicating that although open magnetic flux is being pinched off by reconnection in the tail, it can continue to accumulate through dayside reconnection if the IMF conditions are favourable.

Further questions

These results demonstrate that substorms are more intense when the magnetosphere has previously accumulated a greater amount of energy. However, this raises interesting questions about the mechanism by which the explosive energy release is initiated. Several mechanisms have been discussed in the past, usually small-scale plasma instability modes that can lead to the onset of reconnection in the tail, but clearly these must be triggered by changing conditions in the large-scale coupled solar wind–magnetosphere system. As previously noted, as the magnetotail inflates, the pressure inside increases, which is favourable for the onset of reconnection. This is speculated to be why most substorms occur during periods of southwards IMF, when the amount of open flux in the magnetosphere is increasing. This cannot be the full story, however, otherwise all substorms would be triggered close to a specific threshold level of open flux; that is, they would all have similar onset latitudes. That this is not the case argues that there are additional factors that stabilize or destabilize the tail to magnetic reconnection.

One suggestion is that the presence of an intensified ring current, as occurs during geomagnetic storms, modifies the magnetic configuration of the tail, making reconnection less likely and requiring greater pressures to trigger onset (Milan *et al.* 2008). This would explain why the auroral oval expands to low latitudes during storm conditions, as in figure 1. A superposed epoch analysis of the Dst index (not shown) does indeed demonstrate that the low-latitude substorm category is associated with an intensified ring current. This study indicates that a superposed epoch analysis of auroral images is a powerful technique for drawing out crucial information from a vast dataset. Our sample of more than 2000 substorms is sufficiently large that in future we will be able to subdivide it further to examine the magnetospheric response to differing solar wind inputs, and so shed further light on our invisible magnetic bubble. ●

Steve Milan, A Grocott, C Forsyth, S M Imber and P D Boakes, Department of Physics and Astronomy, University of Leicester, Leicester LE1 7RH, UK; B Hubert, University of Liège, Belgium. steve.milan@ion.le.ac.uk

Acknowledgments. We would like to thank the PI of the IMAGE FUV instrument, Stephen Mende of the University of California, Berkeley, for the use of WIC and SI12 data.

References

Akasofu S-I 1964 *Planet. Space Sci.* **12** 273.
Dungey J W 1961 *Phys. Rev. Let.* **6** 49.
Cowley S W H and Lockwood M 1992 *Ann. Geophysicae* **10** 103.
Frey H U *et al.* 2004 *JGR* **109** A10304.
Milan S *et al.* 2003 *Ann. Geophysicae* **21** 1121.
Milan S *et al.* 2007 *JGR* **112** A01209.
Milan S *et al.* 2008 *JGR* submitted.