The interconnection of the magnetic fields of the Earth and the Sun

Michael Lockwood

Magnetic reconnection facilitates the transfer of mass, energy, and momentum from the solar wind, through the Earth's magnetosphere and into the upper atmosphere. Recently, combined observations using both ground-based and satellite instruments have revealed much about how reconnection takes place. This new understanding has great significance for systems which exploit, or operate within, the Earth's plasma environment, as well as for a wide variety of scientific studies.

A plasma is an ionized gas. There are a variety of estimates of the amount of plasma in the Universe; however, most cosmologists agree that considerably more than 90 per cent of all known material is in the plasma state. We do not have a number for the mass in it, nor indeed do we know exactly how much of it there is. However, we do know that plasmas make up the vast majority of the matter we can see. We live in a highly unusual part of the cosmos in that the Earth's biosphere is a much larger region called the magnetosphere. We can study both these regions using satellites, but the ionosphere has been the focus of much recent attention. Nature supplies us with a large-scale plasma which surrounds the Earth. At altitudes between about 80 km and 1000 km lies the ionized upper atmosphere or ionosphere, which extends upward into a much larger region called the magnetosphere. We can study both these regions using ground-based remote sensing techniques, such as radar.

The free and charged particles of a plasma give it a very high electrical conductivity indeed. Hence currents can flow freely which cause the plasma to interact strongly with magnetic fields.

Magnetic reconnection is one process of plasma physics which has undoubted importance throughout the cosmos. It removes angular momentum and magnetic fields from condensing dust clouds, either of which would otherwise prevent stars from forming. It also allows the rapid liberation of large quantities of energy in stellar flare phenomena. The location closest to us where reconnection takes place is some 10 to 15 Earth radii (R_E) sunward of the Earth, where the magnetosphere meets the interplanetary medium at a boundary we call the magnetopause.

The dayside magnetopause. Like all main-sequence stars, the Sun ejects a continuous stream of plasma. Embedded within this 'solar wind' is a tenuous magnetic field of solar origin called the interplanetary magnetic field (IMF). By the time it reaches the Earth's orbit, the flow is highly supersonic (Mach number typically 6) and hence a bow shock forms at a stand-off distance of a few R_E from the obstacle presented by the Earth's geomagnetic field (figure 1). The solar wind flow compresses the geomagnetic field on the dayside and draws it out in a long (several thousand R_E) comet-like tail on the nightside.

The large dimensions of interplanetary space and the magnetosphere result in the magnetic fields being 'frozen' into the plasmas in these regions. The behaviour of the plasma and its embedded magnetic field is then described by a set of equations referred to as 'ideal magnetohydrodynamics' (MHD). It should be noted that this is an approximate description of general plasma behaviour, often referred to as the 'infinite conductivity limit', but one which can be applied here because the plasmas are so large. In the solar wind, the density of kinetic energy due to the bulk motion of the plasma is very much larger than the density of the magnetic field; consequently, the frozen-in condition results in the IMF being dragged along by the plasma flow.

However, in the magnetosphere the energy density in the geomagnetic field dominates over both the thermal and bulk-flow kinetic energy densities and here the charged particles are confined by the stronger magnetic field. The frozen-in condition would prevent solar wind plasma from crossing the magnetopause. In practice, some particles do enter but the magnetosphere is nevertheless a low-density cavity in the solar wind flow. Between the bow shock and the magnetopause lies the magnetosheath, where the plasma speed is lower than in the undisturbed solar wind. As a result, the frozen-in IMF field lines become draped over the nose of the magnetosphere, as shown in figure 1. Notice that the Earth always presents a northward-pointing magnetic field to the solar wind flow at the nose of the magnetosphere. Observations in interplanetary space and in the magnetosheath tell us that the magnetic fields in these regions point northward for exactly half the time and southward for the other half. Although ideal MHD applies throughout the vast majority of interplanetary space and the magnetosphere, there are localised by highly significant breakdowns of this condition. Figure 2 shows a segment of the dayside magnetopause at the nose of the magnetosphere. Figure 2(a) considers the boundary for a southward-pointing IMF in the ideal MHD limit. The shocked solar wind in the magnetosheath and the magnetospheric plasma are kept apart by an impermeable magnetopause, which carries a current associated with any discontinuity in the magnetic field across the boundary. All the geomagnetic field lines are then termed 'closed', which means that they connect the ionospheres of the two hemispheres (see figure 1). However, the solar wind flow would compress this current into a thin layer that ideal MHD would no longer apply. The equations show that the magnetic field would then diffuse from high to low values; that is, toward the centre of the plasma.
Consequences of reconnection

The interconnection of the magnetic fields of terrestrial and solar origin has far-reaching consequences for the Earth's magnetosphere, ionosphere, and neutral upper atmosphere. Solar wind particulates can stream down open field lines and impact upon the neutral upper atmosphere, causing it to emit the dusky cusp aurora and enhancing the ionospheric plasma by ionizing the neutral gas (principally atomic oxygen at the relevant altitudes, around 250 km). The optical auroral displays are dominated by the 630 nm ('red-line') oxygen emission which is readily stimulated by the cusp particles, with their magnetosheath energies of about 1 keV. In addition, in interplanetary space the open field lines are still frozen into the solar wind flow. In ideal MHD, which applies away from the reconnection X-line, magnetic field lines behave as if they were under tension and hence the segments of the open field lines inside the magnetosphere are also dragged anti-sunward over the polar caps. In this way, energy and momentum are also transferred across the magnetopause. These open field lines form the extended tail of the magnetosphere. Eventually, an open field line in one hemisphere will be pressed together with an anti-parallel partner from the other hemisphere and reconnection takes place at a second X-line in the current sheet at the centre of the geomagnetic tail (see figure 1). This yields a closed field line which will snap back sunward under the magnetic tension, releasing the energy extracted from the solar wind and stored in the magnetic field of the tail.

Consequently, reconnection sets up a large scale circulation of plasma in the outer magnetosphere, which is termed 'convection'. In the upper ionosphere, the plasma is dragged in a (usually) two-celled flow pattern at the higher latitudes of each hemisphere. In the lower ionosphere, currents can flow perpendicular to magnetic field lines (something which is not possible in the upper ionosphere or magnetosphere because the charged particles are frozen on to the magnetic field and can move only along it). The ionospheric field-perpendicular conductivity is caused by the charged particles colliding with the neutral gases of the upper atmosphere, causing it to emit the atmosphere, penetrate to lower altitudes than the dusky cusp particles and generate the characteristic current layer. Here anti-parallel fields reconfigure, as shown in figure 2(b). This is the process we call 'magnetic reconnection', and it is expected at the dayside magnetopause only if the IMF is southward. It produces 'open' field lines, which thread the magnetopause and connect the interplanetary and terrestrial magnetic fields. The reconnection takes place in a very small region, which is at the centre of the X field-line formation in figure 2(b) and which extends out of the plane of this diagram: it is therefore called the 'X-line'. The motion of magnetosheath and closed field lines towards the X-line and the motion of opened field lines away from it (as shown by the arrows in figure 2(b)) is consistent with an electric field (E) along the X-line. Hence a voltage appears across the X-line. From measurements in the ionosphere, we infer that this voltage can be as large as 150 kV, but is more often near 100 kV for southward IMF. We will return to figure 2(c) later.

Figure 1 The Earth's magnetic field is immersed in the solar wind. By deflecting this plasma flow, it generates a low-density cavity called the magnetosphere (unshaded). The Sun is to the left of the figure, which is for a southward orientation of the interplanetary magnetic field (IMF). The labels define various regions and boundaries discussed in the text. (Figure courtesy Rutherford Appleton Laboratory.)

Figure 2 Magnetic reconnection takes place at the dayside magnetopause, where IMF field lines are draped over the nose of the magnetosphere (see figure 1). (a) For only 'frozen-in' magnetic fields, the IMF in the magnetosheath and the geomagnetic field would be kept apart by an impermeable, current-carrying magnetopause. (b) Allowing for the very small width of the magnetopause, the frozen-in approximation breaks down and field lines diffuse towards the centre of the current layer. For southward magnetosheath field (as shown here) reconnection can occur with the northward-pointing geomagnetic field, giving 'open' field lines which thread the boundary. (c) For a short burst of enhanced reconnection, a pair of bubbles is formed in the magnetopause, containing looped field lines opened by the reconnection burst (from 5)).
557.7 nm (‘green-line’) oxygen emissions which dominate the nightside aurora. The trigger mechanisms for these substorms are not yet understood.

The description of the terrestrial ionosphere-magnetosphere system given here stems from a truly remarkable two-page paper by J. W. Dungey [1]. This was eventually published in 1961 after several years of debate and controversy which have continued to the present day. However, most solar-terrestrial physicists now believe reconnection does indeed act as described above. The evidence is as complex and varied as it is compelling, and this article does not attempt to review the arguments. Rather, it refers the reader to some reviews [2, 3, 4, 5] and concentrates on some recent studies of the variability of reconnection.

Observation techniques

Solar wind-magnetosphere-ionosphere coupling can be studied using in situ satellite measurements or by ground- and satellite-based remote sensing. In situ measurements can achieve very high resolution but suffer a great disadvantage because of the spacecraft motion. We call this the spatial/temporal ambiguity problem. In the ionosphere, satellite orbital periods can be as low as one hour; hence at any one position we can study temporal variations on time scales of order one hour and greater by comparing successive orbits. Furthermore, in one second the satellite moves only about 10 km. Hence, provided we are not concerned about spatial structure on this scale, it can be used to study temporal variations of period 1 second or less. However, a lone satellite cannot give us any information on variations on time scales of between about 1 second and 1 hour.

Remote sensing can generally provide only lower-resolution information but is ideal for looking at the fluctuations satellites cannot see. A ground-based radar can view the same volume of space for several hours (as the Earth rotates relatively slowly) and can achieve temporal resolutions approaching 1 second. One system which has recently exploited this possibility to great effect is the EISCAT (European Incoherent SCATter) radar in northern Scandinavia. This outstanding facility is developed and maintained by the EISCAT Scientific Association of six member nations – France, Germany, Norway, Sweden, Finland, and the U.K. Figure 3 shows two views of one of the three sites, near Tromsø in Norway. The 32-metre paraboloidal dish belongs to the UHF (930 MHz) system, while the huge 40 × 120-metre parabolic-cylinder antenna is part of the VHF (224 MHz) radar. These large, steerable antennae are required because the Incoherent Scatter technique employed yields only the faintest of echoes. Only $10^{-20}$ of the transmitted power is returned to the receiver, equivalent to studying a hard target the size of a small coin from a distance of 300 km. This presents a considerable technical challenge, but the rewards are great because the technique can yield the densities, temperatures, and flow velocities of the various ion species and of the electrons in the ionosphere. It is also interesting to note that the incoherent scatter technique is now also used with lasers as the principle diagnostic of laboratory plasmas, including those in fusion reactors.

Time-dependent reconnection.

Figure 4 provides a dramatic example of the effects of magnetic reconnection [6].
Figure 4 Simultaneous IMF and ionospheric flow observations on 27 October 1984. The top three panels show the sunward, duskward, and northward components of the IMF (B_s, B_d and B_n respectively), as observed by the AMPTE-UKS spacecraft. The bottom panel shows the plasma flow vectors and colour-coded ion temperatures, T_i, observed by EISCAT. The flow vectors have been rotated so that northward points to the right, to avoid congestion of the figure.

The AMPTE-UKS (Active Magnetospheric Particle Explorer-U.K. Satellite) spacecraft was able to make observations of the IMF immediately upstream of the bow shock (see figure 1). The importance of this position is that the undisturbed IMF was observed, but as close as possible to the nose of the magnetosphere. This allowed the propagation delay from the satellite to the magnetosphere to be determined with unprecedented accuracy. At the same time, the ionospheric flows in the dayside cusp region were observed by the EISCAT UHF radar. The top three panels of figure 4 show the IMF components: B_s is sunward and B_n is northward, coplanar with the Earth's magnetic axis, while B_d makes up the right-handed set and is roughly duskward. The third panel shows that at 1107 UT on this day, the IMF swung from northward to southward (that is, B_n turned negative). The vectors in the bottom panel are the flows observed in the cusp ionosphere. After the predicted propagation delay from the satellite to the radar, the flows increased dramatically. They then slowed, following a return of the IMF to weakly northward. Almost immediately, the IMF then swung to weakly southward and the ionospheric flows again responded, although much less strongly. The underlying coloured pixels give the ion temperatures observed by the radar: the enhancement is consistent with the flows and shows the heating caused as energy is extracted from the solar wind. The results show that energy and momentum cross the magnetopause at a much faster rate when the IMF is southward; that is, when there is the magnetic shear across the dayside magnetopause which can give reconnection.

The directness of the control of the dayside ionospheric flows by the IMF B_n component, as demonstrated by figure 4, has been confirmed by many other examples and by surveys of the entire EISCAT-AMPTE dataset. These results show us that convection is not a continuous quasi-steady circulation, as had been envisaged from averaged satellite data [7, 8]; rather, the pattern of flow is always changing with the cycle of energy storage in the geomagnetic tail and its subsequent release in substorms [9]. In addition, given that the IMF is rarely stable long enough to allow the undisturbed completion of a full cycle, the convection pattern does not usually even achieve steady oscillatory behaviour [10].

Evidence is now growing that even when the IMF is southward and stable, the rate of magnetopause reconnection, and hence the speed of the dayside ionospheric flows, is not steady. This
Figure 5 A sequence of false-colour auroral images taken near noon in the cusp region by an all-sky camera sensitive to the 557.7 nm green-line auroral light. For each image the outer circle is the horizon and the northward and westward directions are as marked in the first image. The images show two of a sequence of transient dayside aurorae.

Evidence comes from observations of the flow in the dayside cusp region by EISCAT, made in conjunction with ground-based optical observations of the cusp aurora on the Island of Spitzbergen [11, 5]. Figure 5 shows a sequence of nine false-colour images recorded by an all-sky TV camera which is sensitive only to the 557.7 nm, green auroral light. For each frame the outer circle is the horizon and the geographic northward and westward directions are as labelled on the first image. The images are one-second integrations and the sequence covers a 5-minute period showing a transient auroral intensification appearing on the eastern horizon in the second image. This subsequently moved westward at about 3 km s\(^{-1}\), reaching the meridian of the camera by the fourth image. The westward speed then decreased, and the event elongated and drifted poleward at about 1 km s\(^{-1}\), fading away in the last image. The first three images also show the

Figure 6 Photometer and EISCAT radar data as a function of time during the auroral transients shown in figure 5. The third image of figure 5 is mapped on to a geographic grid in (e), which also shows the meridian scanned every 18 seconds by the photometers and the simultaneous flow vectors determined by EISCAT. (a) shows that red-line transients, as observed by the 630 nm photometer, are superposed on a persistent cusp background, while (b) shows the 557.7 nm photometer observations of the green-line transients, also detected by the all-sky camera. (c) shows the flow vectors seen by EISCAT, while (d) is the corresponding voltage across the north-south extent of the radar field of view.
remnants of a preceding, weaker transient event. Figure 6 shows that these two events were also seen by photometers which scan the meridian of the camera every 18 seconds, and by the EISCAT radar. The fact that the transient events were seen in the green as well as the red oxygen emission line (panels b and a, respectively), tells us that some particles have been accelerated to above typical cusp energies. The vectors (panel c) are the flows seen by EISCAT, just to the south of the transient green-line aurora, as demonstrated by the mapped image in part (e). The flows are enhanced in a channel (in this case in the northern part of the radar field-of-view) which is found to be always co-located with the red-line auroral transient [12]. At all times, the motion of the transient aurora is the same as the local plasma flow seen by the radar. Panel (d) shows the voltage measured across the north-south extent of the radar field-of-view and which drives the observed plasma flows: it can be seen that the onset of the larger, second event is accompanied by a voltage of 50 kV. Given the event extends some distance north of the radar's viewing area, we estimate the voltage associated with this transient may be as large as 100 kV [11, 12]. Hence these events are associated with voltages which are large fractions of the typical voltage that reconnection generates across the magnetopause X-line when the IMF is southward.

There is now considerable evidence that these transient events are the effects of short (roughly 2-minute) bursts of reconnection at the dayside magnetopause. They are observed only when the IMF points southward and their motion is as expected for newly-opened field lines. The events shown in figures 5 and 6 move west initially, but others move east. The east/west motion is controlled by the B_y component of the IMF, in a way which is consistent with the magnetic tension on newly-opened field lines. The subsequent poleward motion is the effect of the solar wind flow. These bursts tend to recur every eight minutes when the IMF is continuously southward. This is very interesting, as satellites like AMPTE-UKS observe bumps in the magnetopause layer which for over a decade have been interpreted in terms of transient reconnection bursts and are called flux transfer events (FTEs) [13]. These are also observed only during southward IMF, when they too tend to recur every eight minutes. Recently, two examples of near-simultaneous observations of FTEs and dayside auroral and flow transients have been found [14]. Such observations are difficult because a satellite is not often close enough to both the magnetopause and to the same magnetic field lines observed in the ionosphere. Indeed, we cannot predict exactly where the satellite needs to be because we cannot accurately map the magnetic field lines from the cusp ionosphere to the magnetopause. In the cases observed, the delay between the magnetopause FTE and ionospheric signatures was as predicted. Some of these events are found to follow isolated swings of the IMF to southward; however, the others are found during periods when the IMF is continuously southward [11].

These observations show us that reconnection, at least sometimes, proceeds as a continual series of events when the IMF is constant and southward. They are also beginning to tell us about the way the reconnection takes place. Figure 7 illustrates three proposed models of FTEs, all invoking reconnection. A major difference between them is the length of the X-line. This cannot be defined by a lone satellite because, as the FTE moves past it, only the dimension in the direction of motion can be determined. The ionospheric events have dimensions of about 200 km north-south but over 1500 km east-west [12, 14]. As mentioned earlier, we do not know the topology of the magnetic field in the outer magnetosphere exactly; nevertheless, these dimensions imply a magnetopause reconnection X-line over 10 R_E in length. This rules out the original model of FTEs [13], shown schematically in figure 7(a), because the X-line invoked was much shorter than this and a near-circular ionospheric footprint was predicted with a twin vortical flow pattern (figure 7(d)). Similarly, the model shown in figure 7(c), with several parallel X-lines, has been predicted to give a single flow vortex (figure 7(f)), which is not consistent with the flow channels observed. However, the observations are consistent with the elongated signature in figure 7(e) and hence strongly support the model shown in 7(b).

To see how this model explains the FTE ‘bumps’ in the magnetopause, we return to figure 2. Conservation of energy and matter can be used to show that the angle of the X in figure 2(a) (marked α) increases with the rate of reconnection. If we have a burst of enhanced reconnection rate, the transient increment increases α produces a pair of bubbles in the magnetopause, one each side of the X-line, as shown in figure 2(c). The field lines opened by the burst of reconnection are looped through these bubbles. Allowing for any elongation of the X-line, the bubbles appear as cylinders in the magnetopause, as depicted in figure 7(b).

The search for ionospheric signatures of transient reconnection has been complicated by the ability of the other phenomena to mimic some of the predicted ionospheric signatures. In particular, variations in the dynamic pressure of the solar wind have been found to produce transient flows in the cusp ionosphere. There has been much debate about suggestion that they also produce some, or even all, of the magnetopause
PTEs. The reader is referred to references [5] and [12], which review this debate in greater detail and provide evidence that pressure pulses are not the cause of the events described here.

**Implications and the future**

These observations are teaching us much about how reconnection transfers energy, mass, and momentum from the solar wind into the magnetosphere. As well as being of vital importance to the study of solar-terrestrial coupling and to other disciplines of plasma physics, this is of importance to a wide variety of man-made systems. Most satellites have to operate within the magnetosphere or ionosphere and the energised particles cause charging problems, data errors, and even permanent damage. They also pose a health hazard to astronauts and high-altitude pilots. The heating of the upper atmospheric gases increases the frictional drag on satellites and causes them to spiral down to Earth. In addition, a great many systems either exploit or are disrupted by the ionosphere. These include navigation, radar, and communication systems and it is becoming clear that convection, and hence reconnection, influence such systems worldwide. The currents which flow in the ionosphere also cause destructive surges on long power lines and even facilitate the corrosion of oil pipelines. If we are to design systems which are more immune to these effects, or at least provide useful warnings of problems, we must understand the variations in the terrestrial plasma environment, many of which stem from the variable and dynamic process of reconnection.

Looking to the future, some very important observational goals will be achieved by ESA’s four Cluster spacecraft, due for launch in 1995. These will fly in close formation through the dayside magnetopause and the cusps and will be able to tell us a great deal about rapid temporal variations which we could not observe with a single spacecraft. To support this mission from the ground, the associate countries of EISCAT are planning to construct an advanced Polar Cap Radar on Spitsbergen [15] and there are other important proposals like the SUPERDARN network of bistatic HF radars. These advances are required because the techniques in use at present have been adequate to reveal the reconnection transients, but cannot resolve the important structure. In addition, these observations have been possible only under special circumstances, when the dayside aurora is in a southerly position and the radar echoes are exceptionally strong. It is probable that we have, thus far, detected only the largest of the events.

The future is an exciting one. The complexity, implications, applications, and controversy associated with reconnection have all grown in each decade since Dungey’s pioneering work in the 1950s. This trend is sure to continue.

**References**


