Cluster's last stand?

Article

Published Version

Open Access

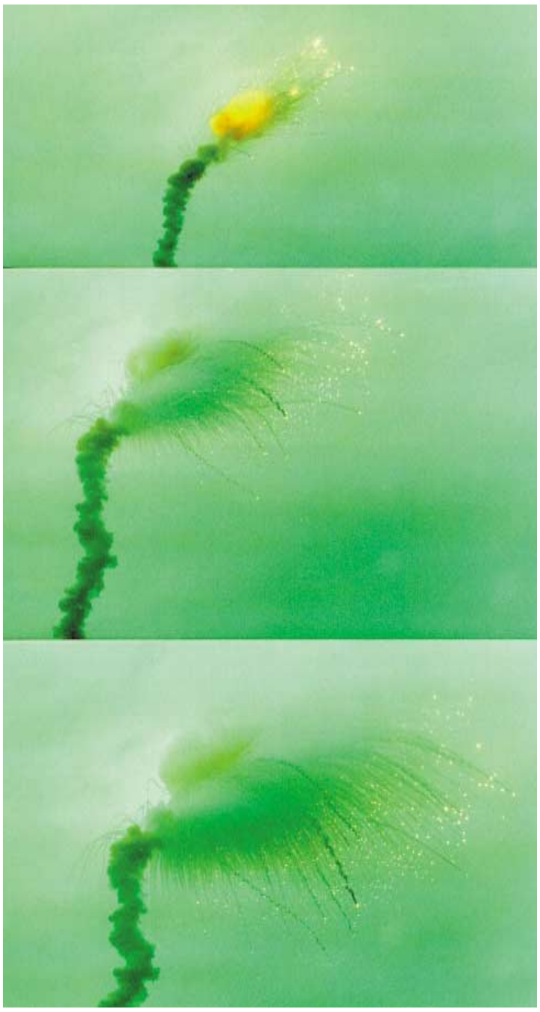
Lockwood, M. ORCID: https://orcid.org/0000-0002-7397-2172 (1997) Cluster's last stand? Astronomy & Geophysics, 38 (1). pp. 21-25. ISSN 1366-8781 doi: https://doi.org/10.1093/astrog/38.1.21 Available at https://centaur.reading.ac.uk/38780/

Cluster's last stand?

When the Ariane 5 rocket exploded on its first launch, it blew apart a cornerstone of European space research. **Mike Lockwood** tells what made Cluster so special, and reports on plans to replace it.

On 4 June last year the first attempt to make three-dimensional measurements in space was lost when the Ariane 5 rocket veered off course and self-destructed, 39 s into its maiden flight. On board were four identical spacecraft which made up Cluster, a mission that the European Space Agency called a "cornerstone" of its Horizon 2000 scientific programme. A full description of the Cluster satellites is given in a special issue of *Space Science Reviews* (Escoubet *et al.* 1997). Their loss dealt a devastating blow to the Cluster scientists and to those working on other missions and projects planned to interact with Cluster. Many discoveries have been made during the 15 years in which Cluster progressed from an idea to the state-of-the-art satellites that were on top of Ariane 501 on 4 June. However, these discoveries invariably underline rather than undermine the importance of Cluster. Now plans to recover the unique and exciting research that was to be done using Cluster are well advanced.

Cluster was a mission with two roles. First, it was part of an armada of spacecraft to study the large-scale variations of the magnetosphere (see p22). Different parts of this complex, coupled system interact with each other but respond differently to variations in the interplanetary medium. The International Solar-Terrestrial Programme (ISTP) is one of several initiatives to make use of widely spaced satellites and ground-based instruments to study these interactions and responses. (Distances in magnetospheric physics are usually measured in units of a mean Earth radius, 1 R_E = 6370 km, and the ISTP satellites are typically 10–100 R_E apart.) The second aim of Cluster was to make the first three-dimensional measurements in space by comparing the data from the four identical craft when separated by about 1 R_E or less. The interpretation of data from lone satellites is ambiguous, and this is also true for those that fly in pairs. These ambiguities lead to a number of fundamental controversies, and Cluster was unique among the ISTP missions in addressing them. They will remain controversial until a replacement mission is flown. It is this which gives such urgency and force to the plans to replace the lost mission with Cluster II.



At a height of 3500 m, a software error pitched Ariane 501 off course, making it self-destruct with the loss of the four Cluster spacecraft. The error was exposed because the rocket's trajectory, until then faultless, gave it greater horizontal velocities than Ariane 4, source of some of the software. Pictures by Prof. André Balogh, Imperial College.

Almost every aspect of magnetospheric research highlights the importance of Cluster. For example, a key phenomenon is magnetic reconnection, which we know takes place in the current sheets at the magnetopause and at the centre of the geomagnetic tail, as originally proposed by Dungey (1961). However, we do not know how the rate of reconnection varies in space and time, nor are the local conditions which control this rate understood. Observations to date have detected the "smoking gun" of accelerated ion flows away from reconnection sites, and have confirmed that reconnection is present from one-dimensional tests of the balance of stresses near the current sheet (Paschmann *et al.* 1979, Sonnerup *et al.* 1990). However, the direct products of reconnection, namely a

The four identical Cluster satellites were designed to orbit in close formation, measuring the charged particles, fields and waves of the near-Earth space environment. Each would have provided such high resolution data that researchers would, for the first time, have been able to discern how important structures and boundaries of the Earth's magnetosphere distort, evolve and move. The loss of this unique and exciting research capability has been a devastating blow to the worldwide community of scientists studying how particles and energy from the Sun are extracted from the solar wind and deposited in the Earth's upper atmosphere. Now there are plans to salvage as much as possible from the investment in Cluster through Cluster II. These make use of Phoenix, the one remaining Cluster satellite, which could fly around the turn of the century.

The magnetosphere

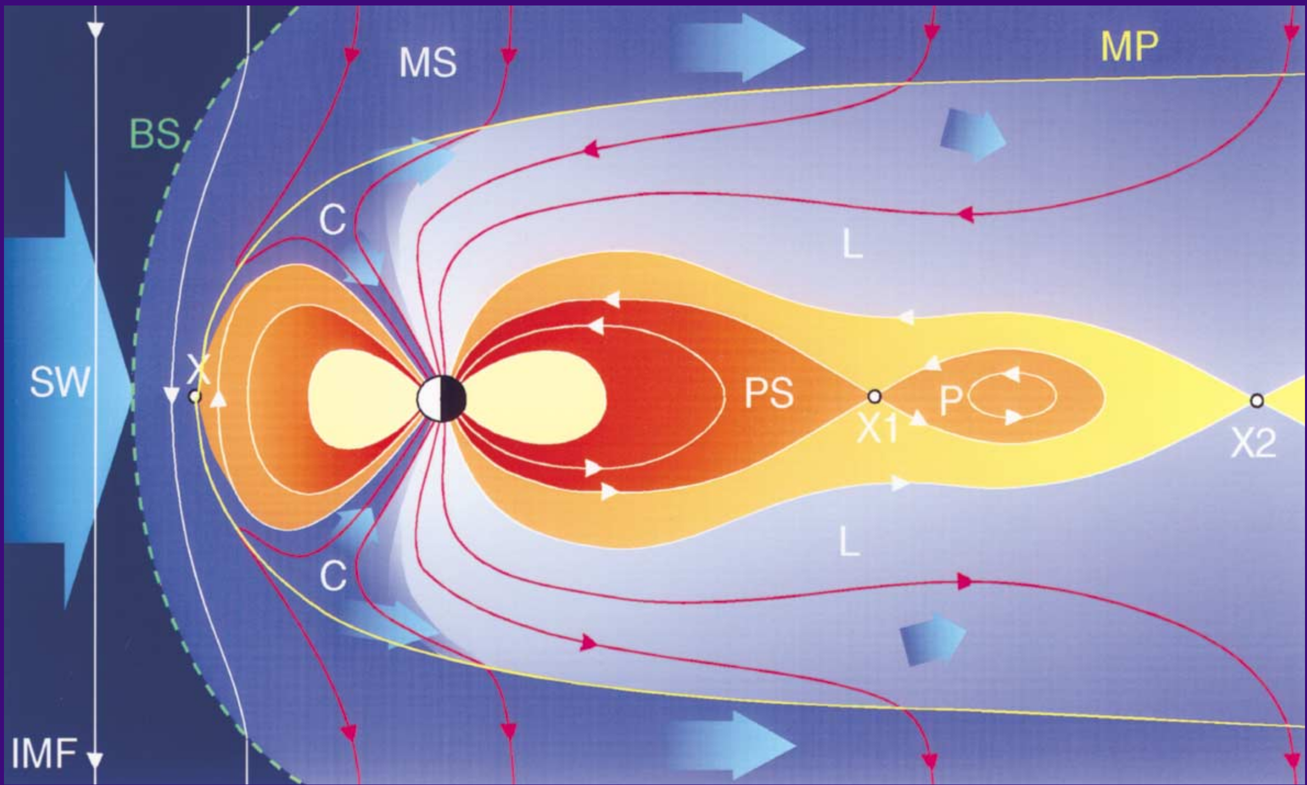
The area of space surrounding the Earth is called the magnetosphere and the behaviour of the ionized gases (plasma) within it is dominated by the Earth's magnetic field. The diagram shows a noon-to-midnight cross-section of the magnetosphere, with the Sun to the left and the North Pole at the top of the Earth. The geomagnetic field presents an obstacle to the continuous flow of solar wind plasma, generating the low-density magnetospheric cavity. It is bounded by a current layer called the magnetopause (shown in yellow and labelled MP). Because the solar wind is supersonic, a collisionless bow shock (BS) forms upstream, and between it and the magnetopause is a region of slowed, heated and generally rather turbulent plasma called the magnetosheath (MS). The high electrical conductivity of the plasmas, and their large spatial scales, result in the charged particles and magnetic fields being frozen together throughout most of this system: the plasmas can move along magnetic field lines but not across them. If this applied strictly, it would stop the particles and energy of the solar wind from entering the magnetosphere. That this is not the case is due to one of the most interesting phenomena in solar-terrestrial physics, a localized breakdown of the frozen-in theorem, known as magnetic reconnection.

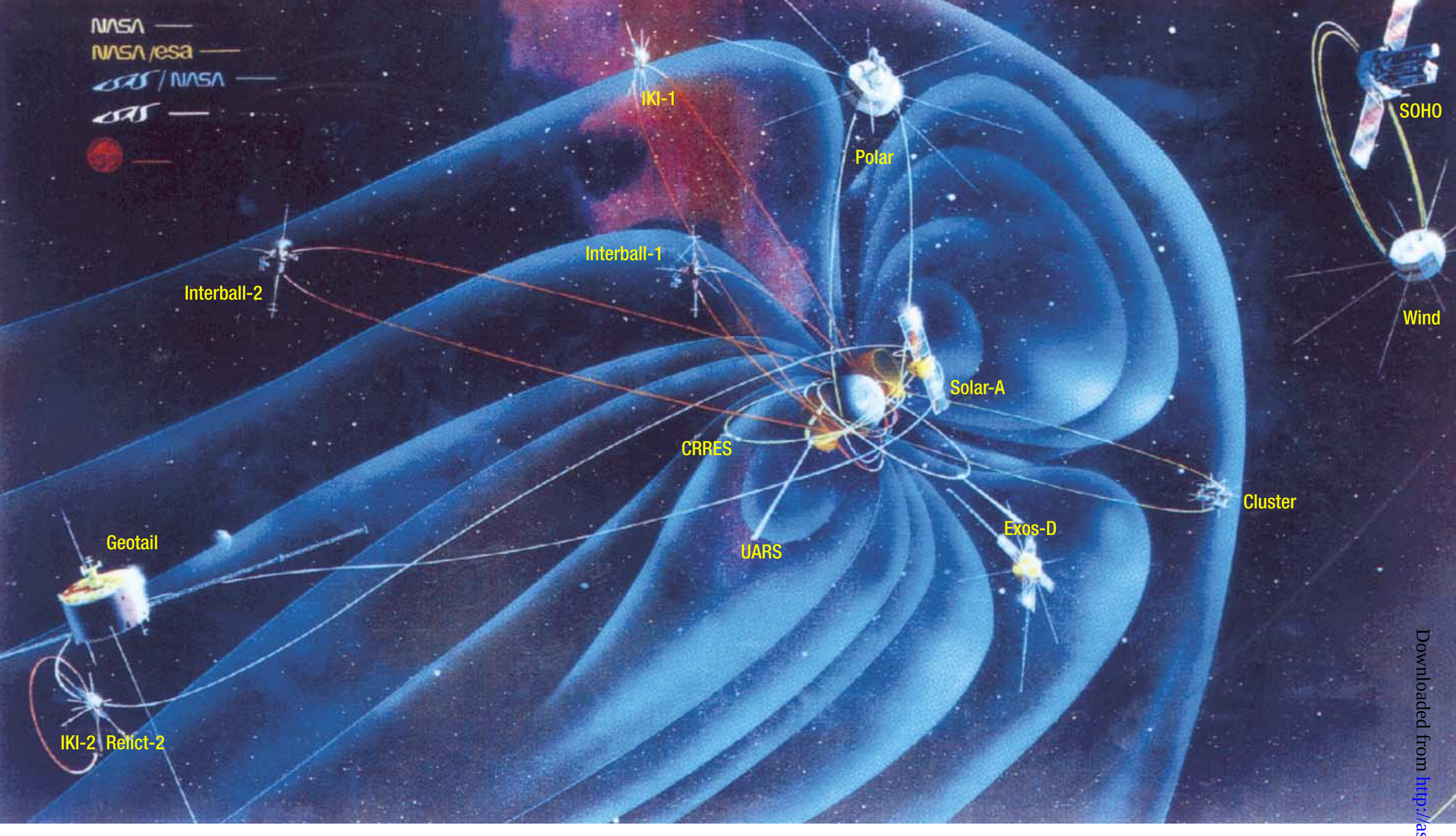
Embedded in the solar wind is a weak magnetic field of solar origin, the interplanetary magnetic field (IMF). When this has a southward component (as it has for half the time, as in the diagram) it can reconfigure with the north-pointing geomagnetic field at a reconnection site such as X and generate "open" field lines which thread the magnetopause: here shown in purple. Plasma is free to flow along these open field lines across the boundary, allowing magnetospheric plasma to escape into the magnetosheath and solar wind plasma to enter. In the magnetosheath and the interplanetary medium, open field lines are still frozen into the flow and this sweeps them antisunward where they accumulate in the tail lobes, L. The solar wind plasma which enters through the dayside magnetopause is hotter and denser, because of the bow shock. It is accelerated towards the Earth on crossing the boundary and can precipitate directly into the Earth's upper atmosphere through two funnel-shaped regions called the cusps, C (Smith and Lockwood 1996). These are narrow in latitude because the solar wind plasma that crosses the magnetopause away from the reconnection site is less dense, cooler and is slowed on crossing the boundary: it is mainly swept into the tail with the field lines and very little reaches the polar atmosphere, poleward of the cusp. The solar wind electrons heat the ionized upper atmosphere (the ionosphere) in the cusps and give characteristic red auroras. In the northern hemisphere, Svalbard is uniquely good for observing this aurora.

The open magnetic flux in the tail cannot accumulate for ever. The magnetic energy density stored there increases, as does the current that flows across the central region of the tail (called the plasma sheet; PS). After about 45–60 min of rapid reconnection at X, this current becomes unstable near Earth and is diverted into the ionosphere. This is called a substorm (Elphinstone *et al.* 1996). During quiet times, open field lines may be closed again (slowly) by reconnection at a distant site, such as X2 in the cross-tail current sheet. In a substorm, reconnection begins much more rapidly at a site like X1, called a Near-Earth Neutral Line. Between X1 and X2 a magnetic island or "plasmoid" (P) forms and is released when X1 starts to reconnect open field lines. The lobe field lines have been stretched out by the solar wind flow and when they are reconnected at X1 they snap elastically earthward. This makes the energy stored in the tail lobe energize the plasma sheet. It also generates greatly enhanced green auroras and currents that deposit much energy in the upper atmosphere.

Magnetic reconnection allows the Earth's magnetosphere to extract in the order of 2% of the incident energy of the solar wind flow. Two-thirds of this returns to the interplanetary medium (much of it as plasmoids). The remainder has significant effects on the magnetosphere, the ionosphere and the neutral upper atmosphere as well as on a whole variety of man-made operational systems. These effects are global and highly variable on timescales from minutes up to the 11-year solar cycle. The variability arises from fluctuations in the solar wind flow, variations in the direction of the IMF and intrinsic time constants of the system. A review of energy flow through this system has been made by Cowley (1991). Processes such as reconnection, collisionless shocks and particle acceleration have applications in disciplines ranging from astrophysics to the development of fusion reactors. The magnetosphere is an excellent natural laboratory where they can be studied on a variety of scales, both by remote sensing and *in situ*, and where the plasma is not disturbed by boundaries or diagnostic probes.

For further reading on space plasma physics in general, see Kivelson and Russell (1995).





The International Solar-Terrestrial Programme (ISTP) involves the coordinated deployment of a large number of satellites to study the magnetosphere and its interaction with the solar wind. Some are shown here. They range from the Solar and Heliospheric Observatory (SOHO), making observations of the Sun and the solar wind, to Geotail, in the far geomagnetic tail. Polar, Wind and the paired Interball craft are also already returning data. Others, like the German Equator-S, are yet to be launched, while some missions of the original concept, like CRRES, have already ended.

Complementing systems on mainland Scandinavia, the new EISCAT Svalbard radar (right) has just commenced operations on the island of Spitsbergen, a northerly location where the cusp aurora can be studied. Making measurements in co-ordination with Cluster was one of the major objectives of this, and several other, new ground-based observatories.
Photograph by Tony van Eyken (EISCAT Scientific Association).

magnetic field threading the sheet and an electric field tangential to it (associated with the motion of the field lines away from the reconnection site), cannot be measured because the sheet's orientation and motion are not known. To do this requires four craft in close formation (see 'Why four?' p24).

Since the ISEE mission in the late 1970s, bumps in the magnetopause current sheet have been interpreted as resulting from pulses of enhanced reconnection rate (called flux transfer events; Russell and Elphic 1978). Although recent observations in the cusp ionosphere by radars and low-altitude satellites have shown that these pulses do occur, the interpretation of the bumps has remained ambiguous: as a result, their dimensions and importance are still controversial (Lockwood 1996). The EISCAT radars can be used to view the cusp in ways that avoid the temporal-spatial ambiguities that plague the interpretation of data from low-altitude satellites. If the reconnection were continuous and at a constant rate, we would expect the cusp to be a steady hot region. The data shown reveal a very different situation, with the cusp comprising a series of poleward-moving events (Lockwood *et al.* 1993). These are very similar to auroral transients that have been seen at winter solstice by optical instruments on Svalbard, when the cusp ionosphere is in the dark. The events are well explained by reconnection pulses, each generating a patch of newly opened field lines which migrates poleward as the field lines are moved into the tail lobe by the solar wind flow. A satellite which flew through the events saw the solar wind electrons responsible for the heating, plus characteristic "step" signatures in the precipitating ions. These are also predicted for pulsed reconnection. Both the magnetopause bumps and the transient events in the cusp ionosphere are quasi-periodic with a mean repeat time of 8 min; the origin and significance of this time constant remain a mystery.

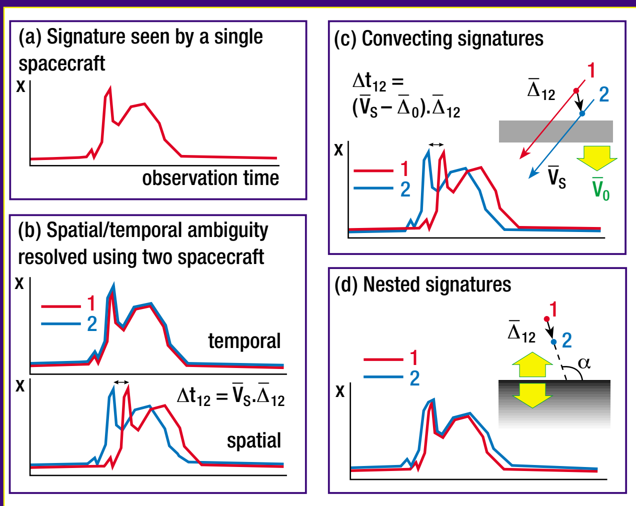
Our inability to define exactly when and where reconnection occurs is also at the heart of another fundamental debate, namely how substorms are triggered. The smoking-gun signatures of reconnection imply that the Near-Earth Neutral Line (NENL; X1 in diagram p22) is usually more than 20 R_E away from the Earth down the tail. However, substorms appear to be initiated much closer to the Earth than this, near 10 R_E. This has led to the development of two distinct, and mutually exclusive, families of substorm model (see review by Elphinstone *et al.* 1996). In one, reconnection at the NENL drives the current in the near-Earth tail unstable; in the other the current instability forms first and later develops into reconnection as it spreads down the tail. Differentiating between the two requires an understanding of the spatial extent and occurrence of the rapid flows along the current sheet that are produced by reconnection, and of the growth of the instability in the cross-tail current. This, like magnetopause reconnection, is just one of a myriad of fundamental problems on which Cluster would have had a huge impact.

Cluster in context

One measure of how important Cluster was to the development of solar-terrestrial physics is the number of supporting facilities that grew around it. For example, a number of large ground-based observatories have been developed which would have been used to place the Cluster observations in context. Chief among these was a new incoherent scatter radar, constructed on the Svalbard archipelago by the EISCAT Scientific Association of six European nations that already operate two such radars in northern Scandinavia. The new radar is in a uniquely good location for making observations of the dayside cusp ionosphere, and these would have been compared with Cluster data from the dayside magnetopause; on the nightside it observes the polar cap and would have been used to study substorms in concert with the other EISCAT radars and Cluster, when in the tail. The network of EISCAT radars forms an excellent facility in its own right, and will become yet more powerful with the addition of the second antenna on Svalbard, made possible by Japan joining EISCAT. It will now be used in conjunction with the other ISTP satellites.

Why four?

When we receive a sequence of data from a lone satellite, we cannot tell if the variations it has seen were caused by temporal changes of the entire region surrounding the craft or by spatial changes through which it has flown. This applies to any of the important measurements of solar-terrestrial physics: in (a) x could be, among many other things, the temperature of one of the charged particle species; the magnitude of a vector such as the plasma flow or the electric or magnetic field; or the amplitude of a plasma wave at a given frequency. Because of this spatial-temporal ambiguity, some missions such as ISEE (International Sun–Earth Explorer), AMPTE (Active Magnetospheric Particle Tracer Experiment) and recently Interball, have deployed craft in pairs. In (b) we show the same variation of x seen by both of such a pair of satellites. When the two craft always see the same conditions at the same time (in the upper plot), the variation is temporal over a region at least as big as the separation of the two craft, Δ_{12}. The lower plot shows that both craft see the same signature, but 1 sees it after 2 with a delay of Δt_{12}. If this lag is what we expect for the satellite velocity, V_s, then the variation was caused by the craft flying through a spatial structure.



Such simple cases do not often apply: spatial structures are usually changing and in motion.

One complex signature, classed as "convecting", is shown in (c). Both craft fly through a spatial structure which is moving at velocity V_o, influencing the delay Δt_{12}: the relative velocity of the structure and the satellite is now (V_s – V_o). Two satellites can provide the component of this relative velocity along their separation, giving just one component of V_o. For all three components of a structure's velocity, we need three independent separations and hence four satellites.

Another type of signature often seen by ISEE and AMPTE is called "nested". This is shown in (d), and occurs when structures only partially move over the satellites and then retreat. In the figure, the boundary of a structure moves over satellite 2 and then over 1, but its motion subsequently reverses and it passes back over 1, then 2. In examples like this, the structure often has some temporal evolution: the sequence seen leaving the structure is not the reverse of that seen on entry. The lag between a point on the structure passing over the two craft depends on the boundary velocity and its orientation with respect to the craft (Δt_{12} is maximum if the angle α is 90° but falls to 0 if α is 0). Again, we need four craft in order to understand the boundary in three dimensions.

Four-craft observations will show how spatial structures move, the orientation of important boundaries and current sheets, and will resolve spatial and temporal ambiguities in three dimensions for the first time. This should be as significant an advance in solar-terrestrial physics as the first *in situ* satellite observations more than 30 years ago.

However, there are other advantages to operating four craft in close formation. Comparing the phase of a wave detected at each can determine its direction of motion for the first time: in effect, we will have the first telescope for plasma waves in space. In addition, one cannot measure currents flowing in the plasma directly, but four craft in a tetrahedral formation will provide the curl of the magnetic field vector. From this, the current density can be calculated using Ampère's Law.

Many studies of the magnetosphere concern the balance of stresses between the magnetic field and the thermal and dynamic pressures of the plasma. With four-craft observations we can include gradients in the pressure properly for the first time, likewise the acceleration and other temporal variations. Because it makes such novel observations possible, many methods of analysis have been developed specifically for Cluster (Mattock 1995). With all of these applications, it is vital to have identical, high-resolution and well calibrated instrumentation so that comparisons reflect the true gradients, velocities and changes.

Poleward-moving hot spots in the cusp ionosphere, shown by the temperature of the ionospheric electron gas measured by the EISCAT UHF radar over Svalbard. The heating is caused by solar wind electrons precipitating down newly-opened field lines. It generates a series of poleward-moving events, consistent with pulsed reconnection. Eight major events are seen in 90 minutes, close to the overall average repeat period of 8 m when the IMF points south. Such events raise a host of new questions about magnetopause reconnection, which only Cluster can address.

Nevertheless, the destruction of Cluster was a serious loss. The same applies to a pair of HF radars called CUTLASS, which are now determining the ionospheric flows in an extended region around the EISCAT radars. They make up part of two chains in both hemispheres called SuperDARN. With the ISTP satellites, the new radars and a host of other novel ground-based instruments, the next few years will be a time of unparalleled activity in solar-terrestrial physics. It is therefore somewhat ironic that the most novel and ambitious project of them all, Cluster, should be absent.

2001: A space odyssey?

The problems facing efforts to fly a replacement mission have been legion, but a way ahead is now emerging. There is a fifth Cluster craft, a spare with a nearly-complete set of instrumentation. Soon after the disaster, ESA took the decision to make it ready for launch. Called Phoenix, this could be launched by late 1997 and so could act as a partial substitute for Cluster I in ISTP. However, this does not meet the goal of three-dimensional observations. Funding and a launch for four more craft was too expensive for both ESA and the national agencies and, in addition, there were potential disruptions to other missions.

A variety of options have been considered, including using Phoenix as a "mother" satellite with three smaller craft around it. However, the costs of designing and building these "daughters" was high and some capabilities were lost. The problem with launching Phoenix in 1997 was that three new craft, of any design, could not be built and launched to join it within its guaranteed lifetime of 2.5 years. At its meeting in November, ESA's Science Programme Committee accepted in principle a compromise plan to delay the launch of Phoenix to about 2001, by when it could still play some role in the last ISTP operations, even though several of the other missions will have ended. Three new craft will be made ready and launched, either with it or soon after, to make up Cluster II. One reason why this decision makes good sense is that the research done since Cluster was first conceived has already provided us with a considerable understanding of how effects and signatures in different parts of the magnetosphere–ionosphere system are related. This, and the early rewards of ISTP, will enable us to interpret other observations better, such as those from the EISCAT and CUTLASS radars, and to use them to place the Cluster II observations in context.

It has been estimated that the investment in Cluster I was 490 million accounting units (MAU) at today's prices (an accounting unit being roughly the same as a US dollar), with a further 100 MAU for the instruments, contributed directly by the various national agencies. A surprising amount of this has survived in the form of designs, operations planning, systems and data-handling facilities. The plans for Cluster II use much of this enormous initial investment to produce and fly three new craft (with all operations support) for 120 MAU. An additional 30 MAU is already being spent to make ready Phoenix. The remaining cost to ESA is for the launch (or launches) and the best of several options for 60 MAU will be chosen. In addition, the impact on other forthcoming missions is small, the chief effect being a delay to M4, which has yet to be selected. Now the national agencies need to find funds to build instruments for the three new craft, and initial indications are generally promising.

I, like many scientists in solar-terrestrial physics, firmly believe that the concept embodied by Cluster is so revolutionary, and so important, that it would eventually be carried through, even if Cluster II were not possible. Now it seems that the first three-dimensional measurements in space may be made before the year 2001 and the Cluster space odyssey may happen after all.

Mike Lockwood is based at the Rutherford Appleton Laboratory, Chilton, Didcot, Oxfordshire, UK. *He is grateful to A Balogh, S W H Cowley, S Schwartz and M A Hapgood for many discussions relevant to this article.*

References

Cowley S W H 1991 *Annales Geophys.* **9** 176–187.
Dungey J W 1961 *Phys. Rev. Lett.* **6** 47–48.
Elphinstone R D *et al.* 1996 *Rev. Geophys.* **34** 169–232.
Escoubet C P *et al.* (in press) *Space Sci. Rev.*
Kivelson M and **Russell C T** (eds) 1995 *Introduction to Space Physics* (Cambridge University Press).
Lockwood M 1996 EOS, *Trans. Am. Geophys. Union* **77(26)** 246 – 250.
Lockwood M *et al.* 1993 *Nature* **361** 424–428.
Mattock C (ed) 1995 ESA-SP-371, ESA Publications Division, ESTEC, Noordwijk, the Netherlands.
Paschmann G *et al.* 1979, Nature **282** 243–246.
Russell C T and **Elphic R C** 1978 *Space Sci. Rev.* **22** 681–715.
Smith M F and **Lockwood M** 1996 *Rev. Geophys.* **34**(2) 233–260.
Sonnerup B Ü *et al.* 1990 *J. Geophys. Res.* **95** 10 541–10 557.