Magnetopause Coupling Processes and Ionospheric Responses: A Theoretical Perspective [and Discussion]

D. J. Southwood, S. Quegan and P. J. Christiansen

Phil. Trans. R. Soc. Lond. A 1989 **328**, 79-92
doi: 10.1098/rsta.1989.0025

Phil. Trans. R. Soc. Lond. A **328**, 79–92 (1989)
Printed in Great Britain

Magnetopause coupling processes and ionospheric responses: a theoretical perspective

D. J. SOUTHWOOD

Department of Physics, Imperial College, London, SW7 2AZ, U.K.

In recent years much progress has been made in establishing the mechanisms for mass, momentum and energy transfer from the solar wind into the terrestrial magnetosphere; in particular, the importance of reconnection, at least at disturbed times, is generally agreed. In the simplest circumstances, where dayside and night-side reconnection rates are balanced and steady, the simple open magnetospheric model would pertain. In general, however, reconnection is unsteady, day–night flux transfer occurs in an irregular way and the full complexity of the solar-wind–magnetosphere–ionosphere system becomes apparent. A hierarchy of coupling times, for each of which different processes dominate, needs to be considered.

INTRODUCTION

In recent years a large measure of agreement on what are the major solar-wind–magnetosphere coupling mechanisms has emerged. There is general acceptance that the direction of the interplanetary magnetic field (IMF) is critical in establishing the level and efficacy of coupling and thus that it is the crucial solar input parameter in geomagnetic activity. Magnetic reconnection between solar and terrestrial magnetic fields gives a natural qualitative explanation of the southward IMF influence. Unlike in laboratory and solar plasma physics the use of the term 'reconnection' in solar terrestrial physics remains a little controversial (in part out of habit). As we indicate below, its use does not imply there are not a lot of open questions concerning the manner in which it occurs, but it seems perverse to claim that the phenomenon does not occur in some way or that it is not important.

The reconnection model of the magnetosphere provides a natural way of not only describing the circulation system set up in the interior of the magnetosphere but also the high- and mid-latitude circulation in the ionosphere. Furthermore, as we illustrate later, it also explains well the global current systems including the high-latitude Birkeland field-aligned currents (region I currents, Iijima & Potemra 1978) associated with geomagnetic activity. The model also provides a simple rationale for the location and distribution of the many different plasma régimes found in the vicinity of Earth (e.g. plasmasphere, ring current, plasma sheet, mantle, lobe, etc.) (Cowley 1980).

Happily for the theorist, the overall order imposed by the open model still leaves plenty of questions. For example, there is much to understand concerning the nature of linkages within the coupled solar-wind–magnetosphere–ionosphere system. Such questions provide the theme of this review.

STEADY MAGNETOSPHERIC CIRCULATION

The magnetosphere has a large-scale internal circulation system driven by a combination of viscous drag (Axford & Hines 1961) and reconnection (Dungey 1961) with the latter process being dominant (Cowley 1984), at least when the interplanetary field has a southward component.

Magnetohydrodynamics (MHD) is valid through most of the convection system although it fails to describe either reconnection or the viscous transfer process. The latter processes both take place in thin regions with scale lengths small enough that MHD is not valid. In figure 1 we show Dungey's original sketch of the open-field model. Circulation of plasma and magnetic flux in the Dungey model is described using the MHD notion of frozen-in field everywhere, save in the vicinity of the magnetic neutral line (A) on the dayside and in the centre of magnetotail (C). At A and C, MHD breaks down and 'reconnects' field lines.

The convection pattern can be outlined by following the motion of individual magnetic flux tubes through the system. A solar tube frozen into the solar wind flow and a terrestrial tube first reconnect on the dayside (at A in the figure). Two new 'open' tubes then result. These extend from the polar ionosphere to the solar wind in each hemisphere. Both move anti-sunward and become stretched out to become part of the tail. Both eventually sink to the centre of the tail where each reconnects with a tube from the opposite polar cap. The two newly joined field lines then separate, one moving sunwards, the other rejoining the solar wind flow. The sunward moving tube is now closed (both feet on the terrestrial ionosphere). The sunward flow eventually returns it to the dayside where it can reconnect with an interplanetary line again for the cycle to then repeat.

Note that, as the circulation proceeds, most of the plasma does not go near points A or C and can be thought of as frozen to the field, but reconnection of solar and terrestrial fields is essential to the setting up of the internal magnetospheric motion. In fact, in the simplest magnetohydrodynamic model, reconnection need be included only as a boundary condition specifying the speed of the internal magnetospheric flow. (Exactly similar considerations apply

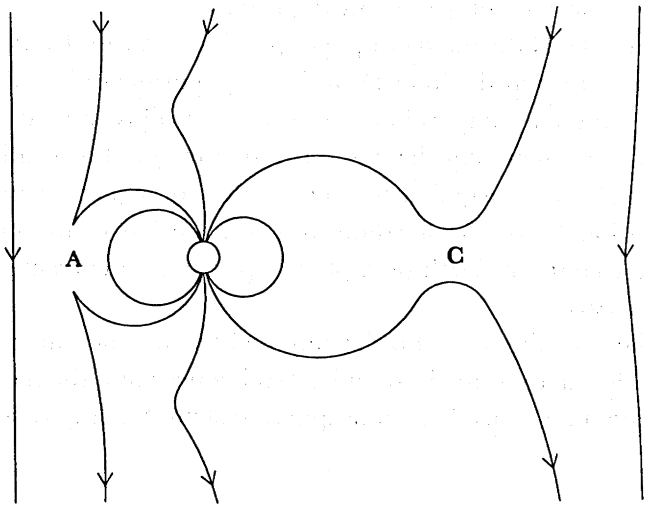

FIGURE 1. A sketch illustrating the open magnetosphere field topology in the noon–midnight meridian. (After Dungey 1961.)

to any other non-MHD momentum transfer process such as anomalous diffusion at the magnetopause.)

Figure 2 shows a dawn–dusk view. In steady state the electric field is derived from an electrostatic potential. The magnetohydrodynamic 'frozen in' field condition

$$\boldsymbol{E} = -\boldsymbol{v} \times \boldsymbol{B} \tag{1}$$

has the useful consequence that both the magnetic field line and the flow lines lie on electric field equipotentials. Figure 2 shows how the solar wind (dawn–dusk) electric field thus can be directly mapped down to ionospheric heights.

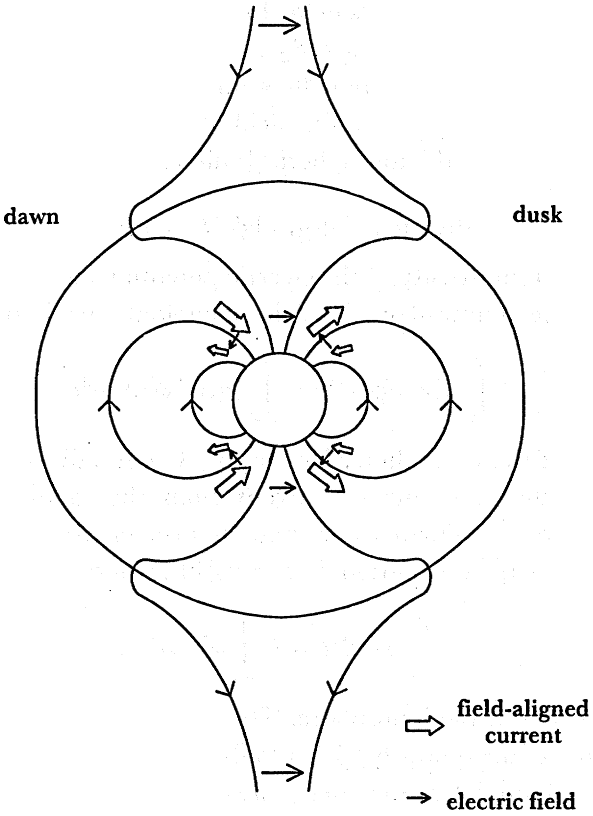

FIGURE 2. A dawn–dusk view of the electric field and field-aligned current system imposed on the ionosphere by reconnection with a southward interplanetary field.

The polar caps are defined as the regions of the ionosphere that are magnetically connected to the IMF. The reconnection rate determines how much solar wind magnetic flux becomes connected to the terrestrial polar caps per unit time. In steady state, the reconnection rates by day and night must match and, with fixed solar wind conditions given, the circumference of each polar cap is proportional to the reconnection rate.

In the simplest description one assumes that there is no significant feedback on the solar wind flow; i.e. the flow is not slowed by the connection to the terrestrial plasma. In steady state the solar wind provides a fixed EMF (electromotive force) which sets the total potential across each polar cap (see, for example, Vasyliunas 1975 and references therein). The actual local time distribution of the potential around the polar cap boundary is determined by the location of reconnection. In the ionosphere, the potential imposed around the polar cap will give rise to a flow at latitudes below the polar cap; this corresponds to the return flow of flux from the magnetotail on closed field lines mentioned earlier.

Consider the polar cap ionosphere further. It is resistive and horizontal Pedersen currents are driven from dawn to dusk across the cap by the solar-wind-induced electric field (cf. figure 2). Similarly, in the mid- and low-latitude ionosphere, the electric field associated with the return flow of flux drives Pedersen currents which also have a net divergence at the polar cap boundary. Both sets of Pedersen currents close by downward (upward) Birkeland current sheets on the dawn (dusk) edges of the polar cap (illustrated by the broad arrows in figure 2). Ultimately the currents close by current flow across the field in interplanetary space. The sense of the $\boldsymbol{j} \times \boldsymbol{B}$ force associated with the cross field closure currents is such as to slow the solar wind (see, for example, Southwood & Hughes 1983).

The link between the Birkeland current flow and resistive (Pedersen) current flow in the ionosphere can be made more precise; the field-aligned current flux into and out of the ionosphere is directly related to the ionospheric Joule energy dissipation. Consider the vector identity

$$\operatorname{div}(\boldsymbol{J}\phi) = +\operatorname{grad}\phi \cdot \boldsymbol{J} + \phi \operatorname{div} \boldsymbol{J}.$$

Let $\boldsymbol{J}$ be the electrical current density, ϕ the electric potential and integrate the expression over the entire volume of the ionosphere. As $\boldsymbol{J}$ is solenoidal one finds the following integral relation,

$$\int_V \operatorname{div}(\boldsymbol{J}\phi)\,\mathrm{d}^3\boldsymbol{r} = \int_V [\operatorname{grad}\phi \cdot \boldsymbol{J}]\,\mathrm{d}^3\boldsymbol{r},$$

where V is the volume of the ionosphere. Now apply Gauss's divergence theorem to the left-hand side and noting that the flux of current from the ionosphere into the insulating atmosphere is everywhere zero. Furthermore note that the integrand of the right-hand side the volumetric rate of Joule energy dissipation $\boldsymbol{J}\cdot\boldsymbol{E} = \sigma_{\mathrm{P}}|\boldsymbol{E}|^2$, where σ_{P} is the Pedersen conductivity of the ionosphere. One finds

$$\int_S \phi \boldsymbol{J}\cdot \mathrm{d}^2\boldsymbol{r} = -\int_V \boldsymbol{J}\cdot\boldsymbol{E}\,\mathrm{d}^3\boldsymbol{r}, \tag{2}$$

where S is the surface area of the ionosphere. The Joule dissipation in the ionosphere is thus equal to the vertical flux of the quantity, $\phi\boldsymbol{J}$, into the ionosphere from the magnetosphere. The latter quantity can be identified as the energy flux into the ionosphere. The vertical current into and out of the ionosphere is provided by Birkeland currents and we can rewrite the integrand on the left-hand side of (2) as $\phi I \sin\chi$, where I is the Birkeland current density and χ is the local magnetic dip angle.

Equation (2) shows that

(i) field-aligned current flow into and out of the ionosphere is a necessary concomitant of ionospheric Joule dissipation;

(ii) field-aligned Birkeland current between ionosphere and solar wind flows as a result of the EMF required to maintain ionospheric currents.

The region II Birkeland currents (smaller broad arrows sketched in figure 2) on closed field lines also transmit energy (Southwood 1977), but we shall not discuss this system in detail here.

As long as the solar wind is a good voltage source, the amount and indeed location of the field-aligned current flow in the polar cap is determined entirely by the ionosphere. If the conductivity is uniformly distributed, the only divergence of ionospheric current is at the dawn and dusk flanks of the polar cap, as described above. The resulting current system is sketched for the Northern Hemisphere in figure 3. The dawn and dusk current pair can be identified with the region I currents detected above the auroral zones by polar orbiting spacecraft (Iijima & Potemra 1978). The quantity $\phi I \sin\chi$ is almost directly observable from appropriately instrumented polar orbiting spacecraft and, although it is not feasible to derive a global value, values derived in a single meridian would be informative.

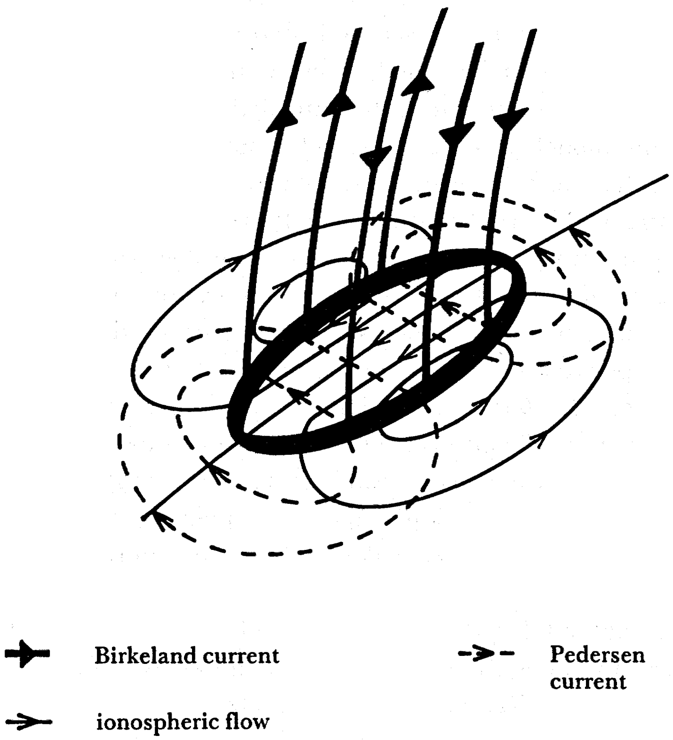

FIGURE 3. Sketch of the basic field-aligned current system associated with polar cap flow and return at lower latitudes in the open magnetosphere model.

Subtle complications are introduced by non-uniformity in the ionospheric conductivity; equation (2) can be used to show that subsidary field-aligned current systems are set up wherever there is a gradient in Pedersen conductance. Further complication occurs if the Hall conductivity distribution is non-uniform. In steady state, the ionospheric flow must be incompressible; the Hall current flow is antiparallel to the plasma flow and where the conductivity is uniform the Hall current flow is divergence free. Non-uniformity creates barriers to the flow and, in steady state, the flow will twist to avoid flowing across regions of Hall conductivity gradients.

DRIVEN OR ENERGY STORAGE–RELEASE MODEL?

The steady-state picture provided in this section hardly, if ever, holds. It provides no more than a framework which can be successively modified to examine the effect of departures from the steady situation. Two well-known and controversial models of terrestrial response to solar wind input can be thought of as direct developments.

The driven model (Perrault & Akasofu 1978; Akasofu 1980) is based on the contention that the solar wind input and output are directly proportional. Akasofu and his colleagues introduced an input parameter, $\epsilon \propto v_{sw} B^2 \sin^4 \frac{1}{2}\theta$ where v_{sw} is the solar wind velocity, B is the IMF, and θ is the angle between the yz IMF component and z GSM axis (geocentric solar magnetospheric). The main modulation of ϵ comes from the B_z component. Evidently the driven model supports the idea of a steady-state response by the magnetosphere–ionosphere system if the input is steady.

The energy storage–release model was proposed first by McPherron (1970). In this model an essential feature of the magnetospheric response is that energy is stored in the magnetotail and subsequently impulsively released. The model views of the magnetospheric substorm as the release of prestored stress from the tail (rather like an earthquake). Even with a steady-state input the proponents of the model would still predict a non-steady convection system in the magnetosphere and ionosphere as substorms continued to take place.

In fact, both Akasofu and McPherron camps predict lags of between about 15 min and more than 1 h between input and output (Akasofu 1980; Bargatze *et al.* 1985). The relative dearth of occasions where the solar wind input is constant over timescales of many hours has meant that incontrovertible evidence to sustain either model has been hard to find.

Matters are further complicated by the use of data derived from magnetometers from the auroral zone as a measure of solar terrestrial activity. Magnetometers respond to the fields of Hall electrojet currents flowing in the auroral zone but the zone moves with activity level and planetary observatory coverage is not even (Allen & Kroehl 1979). Moreover, as Cowley (S. W. H. Cowley, personal communication 1988) has pointed out, the nightside Hall conducting region is sustained only when sufficiently hard (more than 1 keV) electron precipitation is occurring; F-region flow measurements using the EISCAT (European incoherent scatter) radar show that not all large ionospheric (and thus, magnetospheric) flows are accompanied by such precipitation. Flows occurring in the absence of an E-region are not recorded in magnetically based indices such as *AE*, *AL*, etc.

Issues raised by the models will emerge in the later discussion. For the present we note that there is no disagreement that where the solar wind input varies on timescales of less than 1 h the coupling between solar wind and ionosphere and magnetosphere is unsteady.

Unmatched day and nightside reconnection rates

The continual variation in size of the polar cap is a major consequence of the unsteadiness of solar-wind–ionosphere coupling on scales of less than 1 h. In a truly driven system (i.e. with no lag), day and night reconnection rates would match. In actuality, the system is weakly enough coupled that day and night reconnection rates are usually not matched. A day–night difference in reconnection rate corresponds to an increase or decrease in polar cap magnetic flux and as the ionospheric magnetic field does not change the polar cap area must vary. When dayside reconnection is dominant the polar cap area grows, the location of the auroral zone should generally move equatorward; dominant nightside reconnection leads to a decrease of polar cap area and poleward motion of the boundary.

The magnetospheric manifestation of day–night reconnection rate imbalance is the expansion phase of magnetospheric substorms where the aurora retreats polewards (see, for example, Akasofu 1977) and the magnetotail collapses (see, for example, McPherron *et al.*

1973). In the expansion phase the nightside reconnection is dominant and the polar cap area is shrinking.

The most dramatic direct observations of circumstances where dayside reconnection rate is dominant are the reports of magnetopause erosion (Aubry *et al.* 1970; Paschmann *et al.* 1986). Only recently has attention turned to examine what ionospheric signature there may be of dayside erosion. Siscoe & Huang (1985) have modelled the case of uniform polar cap expansion. Studies have been done to correlate polar cap boundary position with interplanetary and geomagnetic parameters during disturbed times (see, for example, Meng 1983; Rodger & Broom 1986). Much of this work assumed that the polar cap boundary moves equatorward uniformly over all local times. Freeman & Southwood (1988) have examined the consequences at ionospheric heights of reconnection taking place over a localized region of the dayside magnetopause. In the vicinity of the longitude where reconnection is greatest, equatorward bulges in the polar cap boundary are set up; elsewhere the ionospheric flow adjusts to the reshaped boundary. One of the important results that Freeman & Southwood derive is that apparently anomalous flow components (e.g. southward flows on the dayside) can be introduced during longitudinally dependent erosion.

Just as the expansion phase of a substorm corresponds to a shrinking of the polar cap, so it is natural to associate erosion and polar cap growth with the growth phase of the geomagnetic substorm in the storage–release model (McPherron *et al.* 1973). As erosion progresses the polar cap flux increases. The associated magnetic energy is eventually released once a threshold is passed in the substorm expansion phase.

DAYSIDE EROSION

The timescale on which erosion is observed (30 min) is considerably less than the time of several hours that it takes a tube to move from day to night (assuming an ionospheric velocity of a few times 10^{-1} km s^{-1}). (The solar wind itself moves a distance of order 100 R_E† in half an hour.) The rate of flux addition to the polar cap is of order 5×10^4 Wb s^{-1} (i.e. corresponding to a cross polar cap potential of 50 kV). As the total polar cap flux is of order 5×10^9 Wb, it changes only by a small fraction in the time involved. What then are the processes that give rise to erosion?

First note that reconnection on the dayside magnetopause at latitudes below the polar cusp only occurs when the external field is southward. Reconnection will still occur when the external field is northward, but it only drives a polar cap circulation (see, for example, Reiff & Burch 1985). Negative B_z acts as a switch or gate for circulation through the whole system. Once dayside reconnection has started, the solar wind will eventually become connected by (region I) field-aligned currents to the polar cap. Ultimately the wind itself can be assumed to have enough momentum to maintain the current system without disruption but the new system will take a finite time to set up.

Can the delay associated with setting up the new stress balance between newly connected flux and the ionosphere account for the erosion timescale? The immediate answer seems to be 'No'. As we discuss later, there is indeed a finite time during which ionospheric flow and magnetosphere are not matched, but it is a matter of minutes as is borne out by the correlations between interplanetary fields and ionospheric flows (Rishbeth *et al.* 1985; Etemadi *et al.* 1988).

† $R_E = 6.37 \times 10^6$ m.

However, there must be some impediment to flow over or out of the nightside polar cap on the erosion timescale. It has long been known that the magnetotail field does increase by a significant factor in the growth phase of substorms (Fairfield & Ness 1970). The polar cap flow would be slowed by the corresponding build up of magnetic pressure. Dynamically the compression is a result of the solar wind dynamic pressure; Coronoti & Kennel (1972) attribute the effect to increased flaring of the tail. The increasing field gives rise to increased tail current; the electromagnetic effect is that part of the solar wind voltage is diverted to build up the energy stored in the inductance formed by the magnetotail.

In the driven model the erosion time is attributed to the delay associated with the build-up of tail flux. Akasofu (1980) refers to the time lag associated with the magnetotail inductance, i.e. the build-up of tail field and thus current; up to this point the description is very similar to what proponents of the storage–release model would say. Beyond this point the argument is difficult to sustain theoretically. An inductance discharging into a resistor (Akasofu's model of what is occurring) does not lead to a simple time lag. The current through the resistor is related to the source voltage $V(t)$ by a relation of the form

$$L\,I(t) = \int_0^t V(t')\exp\left[\frac{R}{L}(t-t')\right]\mathrm{d}t',$$

thus the result is a convolution. A simple lag only results if the system is sinusoidally excited. Otherwise the effect of the inductance is to smooth out variations in the source voltage on timescales short compared to L/R. In particular, there seems nothing that could provide the dramatic magnetospheric changes on a timescale of minutes that are detected at the expansion onset of a substorm.

In the McPherron model, the inductive build-up persists until a sudden change in tail current triggers the tail collapse across a localized region of the tail (the substorm wedge). The tail current change is a result of the onset of reconnection and thus is associated with an increase in the current sheet resistivity. The tail collapse is associated with the discharge of an inductance with relatively high internal resistivity while the build-up of flux has occurred while the resistivity is low.

The author finds it hard not to side with the proponents of the storage–release model on this issue. The magnetospheric response to substorm onset is sudden (McPherron 1970) and has sudden effects throughout the magnetosphere and ionosphere. The smoothing of the coupled magnetospheric inductance and ionospheric resistance proposed by Akasofu (1980) in the driven model works strongly against the model's credibility.

There seem no other candidates to explain substorm onset. Anomalous field-aligned current processes are undoubtedly important in auroral particle processes, but a sudden change in parallel anomalous resistivity does not work as a switch for substorm onset if the source of EMF is a voltage source. A sudden change in tail resistivity associated with impulsive reconnection seems the best explanation of expansion onset (see, for example, Birn 1980; Schindler 1984). In the expansion itself the near tail field undergoes a very substantial reconfiguration of the order of 5 min, followed by a sustained period of convection on the closed field lines for a matter of 1 h or more. During the collapse the nightside polar cap area is reduced. The response in the ionospheric electric field is, of course, smoothed by the absence of induction fields there and it is in the later 'electrostatic' phase that the asymmetry in the day–night polar cap boundary position which develops during erosion is reduced.

PATCHY OR SPORADIC DAYSIDE RECONNECTION

The nightside reconnection rate seems subject to sudden changes. The rapid response to changes in interplanetary B_z in dayside ionospheric flows remarked upon earlier (Etemadi *et al.* 1988; see also Lockwood & Freeman's paper in this Symposium) suggests that dayside reconnection responds rapidly to B_z changes.

Although there is no sign of a delay resembling the erosion/tail flux build-up time described in the previous section, the occurrence of flux transfer events (FTES) at the dayside magnetopause suggests that reconnection is subject to much shorter time variations; indeed, the bulk of flux transfer from day to night takes place in a manner that is unsteady on the time-scale of minutes (see, for example, Rijnbeek *et al.* 1984; Southwood 1987). Since their discovery in spacecraft magnetometer data from near the magnetopause, FTES have been attributed to localized flux tubes connecting solar and terrestrial field (Russell & Elphic 1978).

Russell & Elphic proposed that FTES were due to the passage past the spacecraft of a reconnected flux tube of limited perpendicular scale size. Southwood (1987) analysed the subsequent evolution of such a tube. The localized connected tube is a little like a miniaturization of the polar cap and Southwood proposed that a localized current system like figure 3 would be set up around the tube. However, there will be a finite time during which the newly connected field reconfigures (i.e. unfolds). During this time it might not be able to provide adequate stress to drive flows in the ionosphere (see, for example, Southwood 1987) as new flow conditions are communicated rapidly at ionospheric heights (the ionospheric flow is close to incompressible and a compressional MHD wave can travel from pole to equator in much less than a minute). Once a flow surge associated with the onset of reconnection has propagated down to the ionosphere on a connected flux tube a large drag is rapidly brought into play. In electrical terms, a large resistance is coupled suddenly into the voltage on the (small) connected tube. The flow is slowed and, as the tube unwraps, its behaviour is capacitative.

Southwood (1987) has speculated that the sudden coupling in of the ionospheric resistance to the newly connected tube could itself cause a modification of the reconnection rate and thus be involved in the mechanism of event formation. The travel time for Alfvén waves from the magnetopause to the ionosphere lies somewhere between the characteristic timescale of individual events (a minute or so) and the repetition rate for events of *ca.* 7 min (Rijnbeek *et al.* 1984).

Once the flux tube is unwrapped and is being stretched as it connects into the solar wind proper it will become a voltage source as in the earlier description. If the polar cap flux is increasing the tube will gradually slow as it drapes over the pre-existing flux by which time the isolated tube has been effectively assimilated into the polar cap.

It would be surprising if there were no identifiable FTE signature in ionospheric phenomena. Cowley (1984) first discussed the matter. Cowley (1984) and Southwood (1985, 1987) pointed out that the immediate effect of FTES would be to set up bulges in the ionospheric polar cap. Using Russell & Elphic's (1978) original model, several workers (e.g. Saunders *et al.* 1983; Rijnbeek *et al.* 1984) suggested that the typical cross section was of order 1 R_E^2 giving rise to an ionospheric footprint with a scale size of order a few hundred kilometres (Southwood 1987). Other FTE models predict larger footprints (Southwood *et al.* 1988; Lee & Fu 1985). As the companion paper in this Symposium by Lockwood & Freeman points out there are a variety of currently active lines of research.

MICROSCOPIC BOUNDARY PROCESSES

In the case of the nightside reconnection we proposed the timescale for variations in rate was likely to lie in the microscopic plasma behaviour in the thin current sheet where MHD breaks down.

Although the occurrence of magnetopause reconnection was finally and unambiguously established experimentally by Paschmann *et al.* (1979) by using field and plasma data from the *ISEE* spacecraft from the dayside magnetopause, many questions remain about how it takes place. Paschmann *et al.* (1979) detected a layer of accelerated flow at the magnetopause. In particular, the acceleration was found to be consistent with the jump in velocity imparted as material passes through a planar tangential discontinuity. In an isotropic plasma the condition is that the jump in field and velocity satisfy

$$(\mu_0 \rho)^{\frac{1}{2}} \boldsymbol{v}_{\mathrm{T}} = \boldsymbol{B}_{\mathrm{T}}.$$

Subsequent work (see, for example, the review by Baumjohann & Paschmann 1987) led to the discovery of many more instances of reconnection flow at the magnetopause and refined analysis has also allowed analysis of the energy balance across the thin reconnection layer. Paschmann *et al.* (1986) have shown that there is substantial heating within the reconnection layer and so it follows that the layer is not a simple standing dissipation-free MHD Alfvén wave.

The absence of a complete theory of current sheet microscopic processes is fairly critical, particularly if a threshold effect exists. It may well be that one can characterize the behaviour within the dissipative region by an anomalous resistivity or similar ansatz, but it is not necessarily so. In current sheets the mass difference between ions and electrons can give rise to radically different orbits and sheet residence times, but charge neutrality needs to be maintained (see Dungey 1988 for a recent discussion). Even if wave driven scattering is present, the particle orbits may not be simple. The observational situation in regard to turbulence within the magnetopause current sheet is not clear, but it has long been known that on the nightside the current sheet is much quieter than regions in strong field where emergent particle fluxes are detected on the boundaries of the plasma sheet (see, for example, Gurnett *et al.* 1976).

The evidence thus shows that reconnection occurs, and it should now be accepted rather as the existence of collisionless shocks is agreed by all. However, just as the microscopic processes that account for the heating and redistribution of energy within the shock are not fully agreed, so the microscopic transfer processes in reconnection are not agreed.

In fact, similar uncertainties surround the detail of any viscous process occurring at the magnetopause in addition to reconnection. Some level of viscous interaction process at the magnetopause is allowed by most authors. It is generally agreed that the process would be sustained by wave–particle scattering (see, for example, Eviatar & Wolf 1968). Intense bursts of low-frequency noise are indeed seen in the magnetopause vicinity and these could provide enough spatial scattering to transport sheath ions across the magnetopause (Tsurutani & Thorne 1982). How charge balance is maintained, as it must be, is not clear (Baumjohann & Paschmann 1987) as there is no evidence that there is sufficient noise to scatter comparable numbers of electrons. In the absence of scattering some quasi-steady field and current system must be set up to maintain charge neutrality in the boundary layer. The exact manner in which ions interact with the wave fields may need further investigation. The scattering is a resonant process. The orbits of particles through the wave fields are thus critical to determining the level of scattering and the motion of ions parallel to the field may be important in establishing the interaction time and thus the effectiveness of the resonant process. Note that in FTES, for which there is ample evidence of an origin in reconnection rather than diffusion, there are very high levels of wave noise (LaBelle *et al.* 1987). In these events, it is hard to argue that noise induced is other than a secondary effect.

MAGNETOHYDRODYNAMIC WAVES ON CLOSED FIELD LINES

The timescale determining the formation of FTES is short enough that one can be sure that the system as a whole is not involved and Alfvén wave propagation plays a necessary part in their evolution. On closed field lines standing waves can occur and rapid changes can produce oscillatory responses. Indeed, FTES may excite oscillatory MHD waves (see, for example, Glassmeier *et al.* 1985; Gillis *et al.* 1987). One possibility is that they produce a localized surge in magnetospheric flow on closed field lines as flux is sucked into the reconnection region near the Equator. The delay before the ionosphere is coupled in gives rise to oscillations as stress is redistributed along the field (Southwood 1987). Alternatively, the pressure perturbation created as the FTE moves over the magnetopause can pump hydromagnetic signals deep within the magnetospheric cavity just as the Kelvin–Helmholtz instability is thought to do (Southwood 1974; Chen & Hasegawa 1974).

In fact MHD waves are expected in association with any change in flow conditions on timescales of less than *ca.* 10 min. Each closed field line is a resonator for the Alfvén (transverse) MHD mode because the mode is field guided. The impulsive change in the applied electric field gives rise to an electric field of the form

$$E(t) = E_0(1-\sin(\omega_A(L)\,t)\,\mathrm{e}^{-\gamma t}), \tag{3}$$

where $\omega_A(L)$ is the local field-line resonant frequency. Higher harmonics of the field-line resonance frequency, $\omega_A(L)$, may be introduced depending on the spatial distribution of the impulse along the field. One can derive an estimate of $\omega_A(L)$ by using the WKB approximation

$$\omega_A(L)\int\frac{\mathrm{d}s}{A} = \pi,$$

where A is the Alfvén velocity and the integral is taken from ionosphere to ionosphere, although the precise expression is a function of the polarization (Singer *et al.* 1981). The damping time, γ^{-1}, is given approximately by $LR_E\Sigma_P\mu_0$ (Newton *et al.* 1978).

The Alfvén mode is the only MHD mode that carries field-aligned current (Southwood & Hughes 1983) and thus is likely to be excited in the vicinity of field-aligned currents whenever magnetosphere–ionosphere coupling conditions change on closed field lines. The region II system (Iijima & Potemra 1978) is the major system of field-aligned currents in the closed field region (cf. figure 2) and is due to the ring-current shielding effect (Vasyliunas 1972; Jaggi & Wolf 1973; Southwood 1977). There is a particular timescale associated with the temporal changes in the region II system, the shielding time. Southwood & Kivelson (1989) give a recent derivation of the timescale, T_S. T_S is proportional to the local outward gradient of the ring-current pressure and inversely proportional to the local height-integrated conductivity.

Southwood & Kivelson (1989) estimate the typical dayside shielding time to be *ca.* 2 h, but it could be less than 1 h if the ring-current inner edge is steep.

Any impulsive increase in convection will evoke a modification of the region II current system which will evolve with timescale T_S. In the initial stages of the perturbation there will be oscillations as the field lines establish a new stress balance between magnetosphere and ionosphere; on a longer scale the shielding effect comes into play to reduce the applied field. Thus equation (3) is modified and the electric field response to a sudden change in the applied conditions then takes the form

$$E(t) = [\alpha + (1-\alpha)\exp(-\gamma/T_S)]\,[1-\sin(\omega_A(L)\,t)\exp(-\gamma t)], \tag{4}$$

where the parameter α is the amplitude of the final shielding. The situation is further complicated by the coupling that inhomogeneity introduces between MHD waves. Expressions such as (3) or (4) are only strictly valid for the north–south electric field component unless the signals are confined in L shell and/or longitude. When the source is large-scale coupling may be very important and a global mode excited (Kivelson & Southwood 1986). Global mode frequencies are determined by the (discrete) eigenfrequencies of the compressional fast mode in the magnetospheric cavity although much of the energy deposited in the global mode is fed into those field lines with Alfvén resonance frequencies matching the global mode (see, for example, Kivelson & Southwood 1986; Allan *et al.* 1886*a*, *b*; Zhu & Kivelson 1988).

Concluding remarks

Solar terrestrial physics is now a sophisticated discipline. Solar-wind–magnetosphere–ionosphere coupling is one of its central topics and not surprisingly has progressed greatly in the last twenty years. This review has been written in the confidence that reconnection is important, that, at least when the external field is southward and geomagnetic activity moderate or high, the magnetosphere is open. Some of the excitement now centres on the shorter-term responses of the system. Nevertheless, crucial questions remain even on basic issues. Neither the system nor the subject can be said to be yet predictable and it will be some time before we run out of things to find or phenomena to surprise us.

The author thanks his colleague, S. W. H. Cowley, for several critical suggestions made during the course of preparing this paper.

References

Akasofu, S.-I. 1977 *Physics of magnegospheric substorms*. Hingham, Massachusetts: Reidel.

Akasofu, S.-I. 1980 The solar-wind–magnetosphere coupling and magnetospheric disturbances. *Planet. Space Sci.* **28**, 495.

Allan, W., White, S. P. & Poulter, E. M. 1986*a* Impulse-excited hydromagnetic cavity and field line resonances in the magnetosphere. *Planet. Space Sci.* **34**, 371.

Allan, W., White, S. P. & Poulter, E. M. 1986*b* Hydromagnetic wave coupling in the magnetosphere-plasmapause effects on impulse-excited resonances. *Planet. Space Sci.* **34**, 1189.

Allen, J. H. & Kroehl, H. W. 1979 Spatial and temporal distributions of magnetic effects of auroral electrojets as derived from AE indices. *J. geophys. Res.* **84**, 7138.

Aubry, M. P., Russell, C. T. & Kivelson, M. G, 1970 Inward motion of the magnetopause before a substorm. *J. geophys. Res.* **75**, 34.

Axford, W. I. & Hines, C. O. 1961 A unifying theory of high-latitude geophysical phenomena and geomagnetic storms. *Can. J. Phys.* **39**, 1433.

Bargatze, L. F., Baker, D. N., McPherron, R. L. & Hones, E. W. Jr 1985 Magnetosphere impulse response for many levels of geomagnetic activity. *J. geophys. Res.* **90**, 6387.

Baumjohann, W. & Paschmann, G. 1987 Solar-wind–magnetosphere coupling: processes and observations. *Physica Scr.* **T18**, 61.

Birn, J. 1980 Computer studies of the dynamic evolution of the geomagnetic tail. *J. geophys. Res.* **85**, 1214.

Chen, L. & Hasegawa, A. 1974 A theory of long period magnetic pulsations 1. Steady state excitation of field line resonances. *J. geophys. Res.* **79**, 1024–1032.

Coroniti, F. V. & Kennel, C. F. 1972 Changes in magnetospheric configuration during the substorm growth phase. *J. geophys. Res.* **77**, 3361.

Cowley, S. W. H. 1980 Plasma populations in a simple open model magnetosphere. *Space Sci. Rev.* **26**, 217.

Cowley, S. W. H. 1984 Evidence for the occurrence and importance of reconnection between the Earth's magnetic field and the interplanetary field. In *Magnetic Reconnection in Space and Laboratory Plasmas. Geophys. Monograph* no. 30, p. 375. Washington, D.C.: American Geophysical Union.

Dungey, J. W. 1961 Interplanetary magnetic field and the auroral zones. *Phys. Rev. Lett.* **6**, 47–48.

Dungey, J. W. 1988 Electron viscosity. Preprint, Imperial College, London.

Etemadi, A., Cowley, S. W. H., Lockwood, M., Bromage, B. J. I., Willis, D. M. & Lühr, H. 1988 The dependence of high latitude dayside ionospheric flows on the north-south component of the IMF: A high time correlation analysis using EISCAT Polar and *AMPTE-UKS* and *-IRM* data. *Planet. Space Sci.* **36**, 471.

Eviatar, A. & Wolf, R. A. 1968 Transfer processes in the magnetopause. *J. geophys. Res.* **73**, 5561.

Fairfield, D. H. & Ness, N. F. 1970 Configuration of the geomagnetic tail during substorms. *J. geophys. Res.* **75**, 7032.

Freeman, M. P. & Southwood, D. J. 1988 The effect of magnetospheric erosion on mid- and high-latitude ionospheric flows. *Planet. Space Sci.* **36**, 509.

Gillis, E. J., Rijnbeek, R. P., Kling, R., Speiser, T. W. & Fritz, T. A. 1987 Do flux transfer events cause long-period pulsations in the dayside magnetosphere? *J. geophys. Res.* **92**, 5820.

Glassmeier, K.-H., Lester, M., Mier-Jedrejowicz, W. A. C., Green, C. A., Rostoker, G., Orr, D., Wedeken, U., Junginger, H. & Amata, E. 1985 Pc5 pulsations and their possible source mechanisms: a case study. *J. geophys.* **55**, 108.

Gurnett, D. A., Frank, L. A. & Lepping, R. P. 1976 Plasma waves in the distant magnetotail. *J. geophys. Res.* **81**, 6059.

Iijima, T. & Potemra, T. A. 1978 Large-scale characteristics of field aligned currents associated with substorms. *J. geophys. Res.* **83**, 599.

Jaggi, R. K. & Wolf, R. A. 1973 Self-consistent calculation of the motion of a sheet of ions in the magnetosphere. *J. geophys. Res.* **78**, 2852.

Kivelson, M. G. & Southwood, D. J. 1986 Coupling of global magnetospheric MHD eigenmodes to field-line resonances. *J. geophys. Res.* **91**, 4345.

LaBelle, J., Treumann, R. A., Haerendel, G., Bauer, O. H., Paschmann, G., Baumjohann, W., Lühr, H., Anderson, R. R., Koons, H. C. & Holzworth, R. H. 1987 *AMPTE/IRM* observations of waves associated with flux transfer events in the magnetosphere. *J. geophys. Res.* **92**, 5827.

Lee, L. C. & Fu, Z. F. 1985 A theory of magnetic flux transfer at the Earth's magnetopause. *Geophys. Res. Lett.* **12**, 105.

McPherron, R. L. 1970 Growth phase of magnetospheric substorms. *J. geophys. Res.* **28**, 5592.

McPherron, R. L., Russell, C. T. & Aubry, M. P. 1973 Satellite studies of magnetospheric substorms on August 15, 1968, 9. Phenomenological model for substorms. *J. geophys. Res.* **78**, 3131–3149.

Meng, C.-I. 1983 Case studies of the storm time variation of the polar cusp. *J. geophys. Res.* **88**, 137.

Newton, R. S., Southwood, D. J. & Hughes, W. J. 1978 The damping of pulsations by the ionosphere. *Planet. Space Sci.* **26**, 201.

Paschmann, G., Papamastorakis, I., Baumjohann, W., Sckopke, N., Carlson, C. W., Sonnerup, B. U. O. & Lühr, H. 1986 The magnetopause for large magnetic shear. *AMPTE/IRM* observations. *J. geophys. Res.* **91**, 11099.

Paschmann, G., Sonnerup, B. U. O., Papamastorakis, I., Sckopke, N., Haerendel, G., Bame, S. J., Asbridge, J. R., Gosling, J. T., Russell, C. T. & Elphic, R. C. 1979 Plasma acceleration at the Earth's magnetopause: evidence for reconnection. *Nature, Lond.* **282**, 243.

Perrault, P. & Akasofu, S.-I. 1978 A study of geomagnetic storms. *Geophys. Jl R. astr. Soc.* **54**, 547.

Reiff, P. H. & Burch, J. L. 1985 IMF B_y-dependent plasma flow and Birkeland currents in the dayside magnetosphere. 2. A global model for northward and southward IMF. *J. geophys. Res.* **90**, 1595.

Rijnbeek, R. P., Cowley, S. W. H., Southwood, D. J. & Russell, C. T. 1984 A survey of dayside flux transfer events observed by *ISEE* 1 and 2 magnetometers. *J. geophys. Res.* **89**, 786.

Rishbeth, H., Smith, P. R., Cowley, S. W. H., Willis, D. M., van Eyken, A. P., Bromage, B. J. I. & Crothers, S. 1985 Ionospheric response to changes in the interplanetary field observed by EISCAT and *AMPTE-UKS. Nature, Lond.* **318**, 451.

Rodger, A. S. & Broom, S. M. 1986 The ionospheric signature of the polar cleft over Halley, Antarctica. *Bull. Br. Antarct. Surv.* **72**, 1.

Russell, C. T. & Elphic, R. C. 1978 Initial *ISEE* magnetometer results: magnetopause observations. *Space Sci. Rev.* **22**, 681.
Saunders, M. A., Russell, C. T. & Sckopke, N. 1984 Flux transfer events: scale size and interior structure. *Geophys. Res. Lett.* **11**, 131.
Schindler, K. 1984 Spontaneous reconnection. In *Magnetic reconnection in Space and Laboratory Plasmas* (ed. E. W. Hones Jr), p. 9. *Geophys. Monograph* no. 30. Washington, D.C.: American Geophysical Union.
Singer, H. J., Southwood, D. J., Walker, R. J. & Kivelson, M. G. 1981 Alfvén wave resonance in a realistic magnetic field geometry. *J. geophys. Res.* **86**, 4589.
Siscoe, G. L. & Huang, T. S. 1985 Polar cap inflation and deflation. *J. geophys. Res.* **90**, 543.
Southwood, D. J. 1974 Some features of field line resonance in the magnetosphere. *Planet. Space Sci.* **22**, 483–491.
Southwood, D. J. 1977 The role of hot plasma in magnetospheric convection. *J. geophys. Res.* **82**, 5512.
Southwood, D. J. 1985 Theoretical aspects of ionosphere–magnetosphere–solar-wind coupling. In *Physics of Ionosphere–Magnetosphere*. *Adv. Space Res.* **5**, 4–7.
Southwood, D. J. 1987 The ionospheric signature of flux transfer events. *J. geophys. Res.* **92**, 307.
Southwood, D. J., Farrugia, C. J. & Saunders, M. A. 1988 What are flux transfer events? *Planet. Space Sci.* **36**, 503.
Southwood, D. J. & Hughes, W. J. 1983 Theory of hydromagnetic waves in the magnetosphere. *Space Sci. Rev.* **35**, 301.
Southwood, D. J. & Kivelson, M. G. 1989 Magnetospheric interchange motion. *J. geophys. Res.* (In the Press.)
Tsurutani, B. T. & Thorne, R. M. 1982 Diffusion processes in the magnetopause boundary layer. *Geophys. Res. Lett.* **9**, 1247.
Vasyliunas, V. M. 1972 The interrelationship of magnetospheric processes. In *Earth's Magnetospheric Processes* (ed. B. M. McCormac), p. 29. Hingham, Massachusetts: Reidel.
Vasyliunas, V. M. 1975 Concepts of magnetospheric convection. In *The Magnetospheres of the Earth and Jupiter*, p. 179. Hingham, Massachusetts: Reidel.
Zhu, X. & Kivelson, M. G. 1988 Analytic formulation and quantitative solutions of coupled ULF wave problem. *J. geophys. Res.* **93**, 8602.

Discussion

S. Quegan (*University of Sheffield, U.K.*). Professor Southwood used the idea of an impulse response function extensively in his talk, i.e. ideas which are relevant to linear stable (time-invariant) systems. To what extent is linearity and time-invariance appropriate for the variety of phenomena he described? At what scale does he cease to believe in impulse response functions?

D. J. Southwood. I think nonlinearity is important in the system, but in order to isolate where it is important one needs to analyse (and thus be capable of removing) all potential linear effects. Workers such as Bargatze *et al.* (1985), by using linear prediction filter analysis have accepted this implicitly by providing different filters at different levels of geomagnetic activity.

P. J. Christiansen (*Space and Plasma Physics Group, University of Sussex, U.K.*). Professor Southwood talked about decoupling of the ionosphere from magnetospheric connection in short timescales. Could he clarify what he means by short?

D. J. Southwood. I meant on timescales short compared with the decay time of an Alfvén wave; a typical time would be 5–10 min in the outer-ring current regions.