EARTH’S MAGNETOSPHERIC CUSPS

M. F. Smith
Laboratory for Extraterrestrial Physics
NASA Goddard Space Flight Center
Greenbelt, Maryland

M. Lockwood
Rutherford Appleton Laboratory
Chilton, Didcot, Oxon, England

Abstract. Earth’s cusps are magnetic field features in the magnetosphere associated with regions through which plasma from the Sun can have direct access to the upper atmosphere. Recently, new ground-based observations, combined with in situ satellite measurements, have led the way in reinterpreting cusp signatures. These observations, combined with theoretical advances, have stimulated new interest in the solar wind–magnetosphere–ionosphere coupling chain. This coupling process is important because it causes both momentum and energy from the solar wind to enter into the near-Earth region. Here we describe the current ideas concerning the cusps and the supporting observational evidence which have evolved over the past 30 years. We include discussion on the plasma entry process, particle motion between the magnetopause and ionosphere, ground optical and radar measurements, and transient events. We also review the important questions that remain to be answered.

1. WHAT IS THE CUSP?

The concept of cusps in the Earth’s exterior magnetic field first appeared in the papers by Chapman and Ferraro [1931a, b], which postulated that Earth’s magnetic field would generate a low-density cavity in a stream of ejected solar matter. This cavity we now call the magnetosphere. Their work was aimed at explaining the sudden commencement of geomagnetic storms. We now know that these are associated with a compression of Earth’s magnetic field produced by increases in the continuous ionized particle flux emanating from the Sun (the solar wind). In the Chapman and Ferraro model, the oncoming edge of the stream from the Sun was assumed to be planar (Figure 1a). Since the stream is a fully ionized gas (a plasma), the edge appears to Earth as an electrically conducting plane. The problem then reduces to superposing the dipole magnetic field of Earth with its image in the conducting plane. A set of planar currents, called Chapman-Ferraro currents, is required to cancel Earth’s dipole field. The Chapman-Ferraro paper was the first to show how the Earth’s magnetic field is confined by the solar wind. The paper implicitly introduced the concept of the magnetosphere, the region of space in which Earth’s field dominates the behavior of the plasma. The magnetopause current flow is centered on the two magnetic null points, one in each hemisphere (Figure 1b). These null points were termed “horns” by Chapman and Ferraro, who suggested that these nulls would be points of entry for external plasma into the magnetosphere. The horns later became known as the magnetic cusps. Chapman and Ferraro imagined that the solar wind only impinged upon the Earth in isolated events, each producing a sudden commencement of a geomagnetic storm. We now know that the flow is continuous but highly variable [e.g., Hapgood et al., 1991]. In reality, the current-carrying plane is a roughly bullet-shaped surface (the magnetopause) that surrounds Earth (Figure 1c).

When the solar wind impinges on Earth, and for all planets that have intrinsic magnetic fields, it causes the downstream side of the magnetic field to be elongated into a long tail, while the dayside becomes compressed. The existence of a tail requires a cross-tail current carried by magnetospheric plasma, splitting the magnetic flux into two parts or lobes. The topology of Earth’s magnetic field can be derived by adding these tail currents to the Chapman-Ferraro system (Figure 1c). For further discussion of the history and development of models of the magnetospheric field and the magnetopause, see the review by Siscoe [1987].

Following development of more realistic magnetospheric field models [e.g., Spreiter and Briggs, 1962; Midgley and Davis, 1963; Mead, 1964], many authors speculated on the access that solar wind plasma might have, through the cusps, to low altitudes [Piddington, 1965; Spreiter and Summers, 1967; Willis, 1969]. The speculation was initially confirmed by the first in situ observations made by the ISIS 2 spacecraft inside the magnetosphere of particles with characteristics very similar to the shocked solar wind (magnetosheath) plasma [Heikila and Winningham, 1971]. These data, along with those of Frank [1971], showed that the solar wind has almost unimpeded access to the cusp regions. The first indication that the entry process was more complex came from noting that although the magnetic null point of the cusp...
been used interchangeably for many years with little attempt to quantify the differences until the recent work of Newell and Meng [1988, p. 14,550]. The conceptual definition used by them is

The low-altitude cusp is the dayside region in which the entry of magnetosheath plasma to low altitudes is most direct. Entry into a region is considered more direct if more particles make it in (the number flux is higher) and if such particles maintain more of their original energy spectral characteristics.

Newell and Meng [1988, 1992] and Newell et al. [1991] used over 60,000 satellite passes through the high-latitude ionosphere to provide definitions of the cusp and cleft precipitations (as well as other, lower density, precipitations such as the “mantle” and “polar cap”) and to produce a statistical map of the average ionospheric precipitation pattern (Figure 2).

2. THE OPEN MAGNETOSPHERE

Magnetic fields and plasma interact in a complex manner. From the pioneering work of Nobel Laureate Hannes Alfven was developed a system of equations called magnetohydrodynamics (MHD). The large spatial scales of space plasmas allow the application of these equations in the “high conductivity limit” also known as “ideal MHD.” In this approximation the magnetic field is “frozen-in” to the plasma, such that plasma can flow freely along the magnetic field lines. The plasma, however, can only move perpendicular to the magnetic field lines by carrying them with it. In the solar wind the energy density of the plasma flow dominates the energy stored in the magnetic field, and, as a result, the frozen-in condition causes the solar wind to drag a weak magnetic field of solar origin, the interplanetary mag-

Figure 2. A schematic map showing regions where the various types of precipitation are observed in the upper atmosphere based on a statistical survey of 60,000 low-altitude satellite passes [from Newell and Meng, 1992]. The magnetosheath-like precipitations are the cleft low-latitude boundary layer (LLBL), cusp, mantle, and polar cap. CPS, central plasma sheet; BPS, boundary plasma layer.
nentic field (IMF), from the Sun to Earth. Within the Earth's magnetosphere, however, the density of energy stored in the magnetic field is much greater than that in the plasma, and the frozen-in condition generally results in the magnetic field constraining the plasma, although the field does undergo considerable distortions. At the lower boundary of the magnetosphere is the ionized upper atmosphere or "ionosphere." Here Earth's intrinsic magnetic field is much greater than that produced by the magnetospheric-ionospheric currents and is, to a good approximation, constant. The ionosphere is thus considered to be incompressible.

If the frozen-in condition were to strictly apply at all times and places, the magnetic fields and plasma of interplanetary space and of the magnetosphere would not mix nor interact. Hence the observations of solar-wind-like plasma at low altitudes in Earth's magnetosphere and ionosphere would not occur. Even if plasma entry into the magnetosphere occurred through the magnetic null points of the cusps, that plasma would only reach the ionosphere at a singular point. Cusp precipitation into the ionosphere over extended regions (Figure 2) therefore reveals some form of breakdown in ideal MHD and in the frozen-in condition.

The most important process that breaks the ideal MHD approximation is magnetic reconnection, which allows solar wind plasma to enter the magnetosphere. In reconnection, fluid elements connected to one another via a magnetic field line may become disconnected from each other and reconnected to a different set of fluid elements [after Axford, 1984]. The reconnection mechanism was suggested as a source of power for solar flares by Giovanelli [1946] and for the terrestrial aura by Hoyle [1949]. The concept of magnetic reconnection as we now understand it however, was put forward by Dungey [1958, 1961]. As the reconnection mechanism allows mass, energy, and momentum to enter Earth's (and other planet's) magnetosphere, it is one of the most important in space physics and has considerable implications in other areas of science, such as fusion research and astrophysics.

The orientation of the IMF with respect to Earth's magnetic field is variable and highly influential. In Figure 3 the IMF points southward. Since the solar wind flow is supersonic, a bow shock (BS) forms upstream of Earth's magnetosphere. This slows the flow, which stagnates at the nose of the magnetosphere, such that the frozen-in IMF field lines become draped over the magnetosphere [see, e.g., Crooker et al., 1985]. The region between the bow shock and the magnetopause (MP) is the magnetosheath (MS). Without the process of magnetic reconnection, all Earth's field lines would be "closed" (labeled c in Figure 3). In other words, the magnetic field lines would connect the ionospheres of opposite hemispheres and never pass through the magnetopause. Reconnection, however, can take place at a so-called X or neutral line (X) such that a closed field line (c) and an IMF field line reconfigure to produce an open field line (o), which threads the magnetopause and connects the terrestrial and interplanetary regions. Outside the magnetosphere, the open field lines are frozen-in to the magnetosheath and solar wind flow and hence continue to flow away from the Sun.

The newly opened field lines are highly kinked (Figure 3), and the equations of MHD show that such field lines are subject to a curvature force, often referred to as "magnetic tension," which acts to straighten them. This tension force moves the newly opened field lines away from the stagnation region at the nose of the magnetosphere. They are subsequently dragged antisunward by the magnetosheath flow through the magnetic cusps (C in Figure 3). Such open field lines therefore extend the Earth's field line into the tail lobes described in section 1. The open flux tends to accumulate in the tail until enhanced reconnection takes place in the current sheet separating the two lobes. This reconnection in the tail converts open field lines back to closed ones and returns magnetic flux to the interplanetary medium. The closed field lines move sunward, back toward the dayside magnetopause, completing a circulation of flux in the magnetosphere and ionosphere which is termed "convection." The rate of transfer of open flux into the tail is rarely matched to the rate of return of closed flux at any one instant, although they have the same long-term average. This imbalance produces cycles of growth and
destruction of the open magnetic flux in the tail called substorm cycles.

Dungey [1968] introduced the concept of reconnection between the IMF and the geomagnetic field and showed that the breakdown of ideal MHD was localized (at X in Figure 3). Elsewhere, ideal MHD applies to the open field lines, such that the solar wind plasma is free to flow along them into the magnetosphere and down into the ionosphere. In practice, there is some scattering of particles, particularly ions, as they pass through the magnetopause current sheet, resulting in some ions being reflected back into the magnetosheath. Observations, however, show that the distribution of ions transmitted through the boundary is otherwise largely as predicted if these scattering effects are neglected (see section 3). For southward IMF the antisunward convection of newly opened field lines corresponds to an electric field \( \mathbf{E} \) perpendicular to the magnetic field in Earth's reference frame which is directed roughly from dawn to dusk. The magnetopause currents \( (\mathbf{J}) \) (Figure 1c) are in the same direction as the dawn-to-dusk electric field at latitudes between the two cusps so that there is ohmic heating \( (\mathbf{J} \cdot \mathbf{E} > 0) \). Elsewhere, however, the currents and electric field are opposed, and hence \( \mathbf{J} \cdot \mathbf{E} < 0 \). From Poynting's theorem [e.g., Cowley, 1991], this means that plasma will be accelerated at the dayside magnetopause [Hill and Reiff, 1977; Cowley, 1991] but decelerated near the tail lobe boundary. In steady state these regions represent sinks and sources of Poynting flux, respectively. In nonsteady situations the storage or extraction of energy from the magnetic field must also be considered. The mechanism by which the acceleration takes place was first described in detail by Levy et al. [1964] and is discussed in section 3.

Although the IMF points southward in Figure 3, in general, it can have any orientation. The IMF is usually described in a right-hand Cartesian coordinate system called GSM, in which the \( X \) axis points from Earth to Sun, the \( Z \) axis lies in the plane containing both the Earth's dipole axis and the \( X \) axis, and the \( Y \) axis makes up the right-handed set and points roughly from dawn to dusk. Hence, in Figure 3, \( B_z \) is negative and \( B_x \) is zero. Generally, the IMF has a large \( B_z \) component (Figure 4). An IMF \( B_z \) produces a curvature of newly opened field lines such that they are subject to a tension force which not only has a component in the \( Z \) direction but also the \( Y \) direction. For positive \( B_z \), the field lines in the northern hemisphere move toward dawn, whereas in the southern hemisphere they move toward dusk. The sense of these flows is reversed for negative \( B_z \). In the ionosphere this maps to eastward or westward flows which are oppositely directed in the two hemispheres and whose direction depends upon the IMF \( B_z \) [Iversen et al., 1972; Atkinson, 1972; Cowley, 1981; Greenwald et al., 1990]. This is called the Svalgaard-Mansurov effect after the scientists who first noted the associated effect on currents in the lower ionosphere. The east-west motion is superposed on a poleward drift, which corresponds to the average antisunward evolution of the field lines from the \( X \) line into the tail lobe. The \( B_z \)-dependent east-west flows persist while the field line is curved, becoming more poleward as the field lines straighten out (Figure 4). The momentum is transferred down into the ionosphere by a pair of oppositely directed field-aligned currents [Southwood and Hughes, 1983; Cowley, 1991] (Figure 4). The field-aligned currents poleward of the region of east-west flow are called the "cusp" currents [Iijima and Potemra, 1976; Saunders, 1989], whereas those equatorward of it are part of a more widespread current (associated with the antisunward convection) called "region 1." Plasma crossing the magnetopause will precipitate down the newly opened field lines and into the ionosphere in roughly the same region as where the tension force causes the \( B_z \)-dependent east-west flow (see section 4).

Much of the detailed knowledge of the effects of reconnection is for intervals of southward IMF, although reconnection still occurs when the IMF points strongly northward. The reconfiguration then no longer takes place at the low-latitude magnetopause (between the magnetic cusps) because there the magnetic fields no longer have antiparallel components. This condition, however, can occur poleward of the magnetic cusps for \( B_z > 0 \) [Dungey, 1963; Russell, 1972], and reconnection signatures at these locations have indeed been observed when \( B_z > 0 \) [Gosling et al., 1991; Paschmann et al., 1990; Kessel et al., 1996]. The newly opened field lines produced by reconnection poleward of the cusps have more complex curvature geometries, and the combined action of magnetic "tension" and magnetosheath flow produces sunward or east-west flow. It is probable that the main effect is to take previously opened field lines which thread the magnetopause far down the tail and to reconfigure them such that they thread the dayside magnetopause.

The "open magnetosphere" model describes how magnetic reconnection at the magnetopause facilitates mass, energy, and momentum transfer from the solar wind into the magnetosphere and ionosphere (Figures 3 and 4). Mass transfer across the magnetopause occurs along open field lines from the time it is opened until it is closed by reconnection in the tail current sheet that separates the lobes. Experimental evidence for this model is extensive and now generally accepted but is not within the scope of this review. Here we describe only results which relate to solar wind particle entry across the magnetopause and into the cusp ionosphere.

3. SOLAR WIND PARTICLE ENTRY INTO THE MAGNETOSPHERE

Magnetic reconnection facilitates the entry of plasma into the magnetosphere and modifies incoming particles when they cross the magnetopause boundary (Figure 5). By applying ideal MHD and conserving momentum
across the rotational discontinuity at the magnetopause, the following equation can be obtained [Landau and Lifshitz, 1960]:

\[ \mathbf{V}_2 - \mathbf{V}_1 = \pm \left( \mathbf{B}_2 - \mathbf{B}_1 \right) / \left( \mu_0 \rho_0 \right)^{1/2} \]  \hspace{1cm} (1)

where the subscripts 1 and 2 refer to observations on either side of the boundary, \( \mathbf{V} \) is the bulk plasma velocity in Earth's frame of reference, \( \mathbf{B} \) is the magnetic field, \( \rho_0 \) is the density, and \( \mu_0 \) is the permeability of free space. Equation (1) is for the simplified case of an isotropic plasma but can readily be generalized for the real magnetopause where the field-perpendicular pressure exceeds the field-parallel value. The sign of the right-hand side of the equation denotes the polarity of the boundary normal field (\( \mathbf{B}_n \)) and therefore an observation point either north or south of the X line. Equation (1) allows for the plasma to be either accelerated or deaccelerated as it crosses from the magnetosheath into the magnetosphere along open field lines produced by reconnection. Both ions and electrons are accelerated, but the change in velocity for electrons is negligible. Accelerated magnetosheath ions were first reported in the magnetosphere by Paschmann et al. [1979].

An alternative approach to describing particle entry uses the de Hoffman-Teller (dHT) reference frame, which slides along the magnetopause at a speed \( V_{HT} \) such that the electric field goes to zero [de Hoffman and Teller, 1950]. From (1) it can be shown that plasma flows through this ideal-MHD rotational discontinuity at the local Alfvén speed, \( \mathbf{B} / (\mu_0 \rho_0)^{1/2} \) in this frame of reference [Hudson, 1970]. Transformation back into Earth's rest frame then gives

\[ \mathbf{V} = \mathbf{V}_{HT} + \mathbf{R} / (\mu_0 \rho_0)^{1/2} \]  \hspace{1cm} (2)

\( V_{HT} \) can be thought of as the velocity with which the field lines move along the boundary [Cowley, 1982] and is given by \( \mathbf{E}_i / \mathbf{B}_n \), where \( \mathbf{E}_i \) is the tangential electric field in Earth's frame and \( \mathbf{B}_n \) is the boundary normal magnetic field (\( \mathbf{B}_n \) is only nonzero where there are open field lines which thread the boundary). Equation (2) is sometimes called the Walen relation and predicts a linear relationship between the flow velocity and the magnetic field. It can be used to test for convecting open field lines and hence for reconnection (the so-called momentum or stress balances). This calculation has been performed by a large number of authors and has confirmed that stress...
balance often holds at the magnetopause [e.g., Aggson et al., 1983; Paschmann, 1985; Paschmann et al., 1979, 1986, 1990; Sonnerup et al., 1981, 1990; Gosling et al., 1990c; Maynard et al., 1991; Smith and Rodgers, 1991]. Sanchez et al. [1990] and Sanchez and Siscoe [1990] have confirmed stress balance at the tail lobe boundary, 30 Earth radii \( (R_E) \) downtail. Thus once the field line has been opened, it stays open until reconnection in the tail closes it. Both the magnetosheath and magnetospheric plasma will continue to stream across the magnetopause as long as the field line remains open.

As particles move along the newly opened field lines, the field lines are themselves convecting under the tangential electric field \( E_T \). The transmitted magnetosheath electron population has parallel velocities greatly in excess of the transmitted magnetosheath ion population. The perpendicular (convection) speeds of the two populations are identical, being that of the field line onto which they are frozen. The locus of the most energetic electrons/ions, injected from the magnetosheath into the magnetosphere at the X line, is called the electron/ion edge (E2/I2 in Figure 5), respectively. These edges are the limits where magnetosheath electrons and ions can be seen. The edge of the electron population is considerably closer to the open/closed field line boundary because of higher parallel speed. Corresponding edges exist for the escaping magnetospheric particles. Both the electron and ion edges have been observed by the ISEE spacecraft [Gosling et al., 1990b] (Figure 6). The existence of the electron and ion edges is further support for reconnection and, as we shall show later, has consequences for the ionospheric cusp.

Cowley [1982] predicted the particle distributions, both in the magnetosphere and magnetosheath, that arise from the motion of ions across the magnetopause along open field lines (Figure 7). The distribution, shown in Figure 7, arises because the sheath plasma is a drifting Maxwellian. In the field line rest frame, however, only plasma which has a velocity directed inward across the magnetopause can enter the magnetosphere. This produces a cutoff in the distribution at \( V_T = 0 \) in the dHT frame. In Earth’s rest frame, all injected particles are accelerated by a fraction of the de Hoffmann-Teller frame velocity (Figure 8), so the cutoff appears at \( V_T > 0 \).

In the Earth’s frame of reference both the dHT frame and the field lines in the open LLBL move along the magnetopause in the \( Z \) direction at speed \( V_{HT} \), typically 100–300 km s\(^{-1}\), corresponding to a field-perpendicular convection velocity of about 10–30 km s\(^{-1}\) [Aggson et al., 1983; Sonnerup et al., 1990]. Cowley’s [1982] predictions were made using a geometric construction like that shown in Figure 8b. For the simplest case, the Alfvén speed is the same on the two sides of the boundary. As discussed above, the theory of a time-stationary, one-dimensional ideal-MHD RD reveals that the plasma will flow through the KD at the local Alfvén speed \( V_A \) in the dHT frame (equation (2)): sheath ions approaching the magnetopause have \( V_x > 0 \) and \( V_z < 0 \) (s in Figure 8b), but once injected they have \( V_x < 0 \) but \( V_z > 0 \) (i in Figure 8b). About the bulk flow speed is the thermal spread, and only ions with \( V_z < 0 \) will cross the magnetopause. The minimum injected field-aligned velocity in the dHT frame is zero, which gives a minimum field-aligned velocity in the Earth’s frame of \( V_{HT} \cos \theta \) (Figure 8b). Close to the X line \( \cos \theta \) is nearly unity, and thus the low-velocity cutoff of the injected population is predicted to be \( V_{HT} \). Hence the cutoff appears at \( V_{HT} \), the de Hoffmann-Teller velocity.

The incoming sheath distribution is thus not only accelerated but truncated as it crosses the magnetopause. The characteristic “D” distributions that arise from the Cowley [1982] theory have been observed by many spacecraft [e.g., Gosling et al., 1990c; Fuselier et al., 1991] and reproduced by both MHD and hybrid numerical simulations [Lin and Lee, 1993]. Smith and Rodgers [1991] also showed that the distribution cutoff is at \( V_{HT} \).
Figure 6. Contours of electron and ion distribution function in the GSE XY plane seen by the ISEE 1 spacecraft on August 12, 1978 [Gosling et al., 1990b]: (top) ions; (bottom) electrons. The arrow corresponds to the projection of the magnetic field in the GSE XY plane. The spacecraft starts in the (left) magnetosphere and traverses deeper into the boundary layer, toward the magnetopause current layer. In the magnetosphere, only the hot magnetospheric components are seen for both ions and electrons. As the spacecraft enters the (middle) edge of the boundary layer, cold electrons are observed because the satellite has crossed the electron edge (E2) in Figure 5. One electron beam (in the direction of the magnetic field) is directly injected across the magnetopause, the other contains electrons injected earlier which have mirrored near Earth and returned to the satellite to form a beam flowing antiparallel to the field. At the same time, there is a reduction in the higher-energy field-aligned electron component as magnetospheric electrons start to escape into the magnetosheath. However, at this time the ions remain magnetosphere-like. As the (right) boundary layer is penetrated further, the hot magnetospheric electrons completely disappear and the electron distribution becomes sheath-like. In addition, the first cold ions from the sheath are observed as well as a reduction in the ion component. Thus the satellite has crossed the ion edge (I2). Interestingly, electron fluxes increase significantly when the ion edge is traversed.

Calculated independently from stress balance across the magnetopause (Figure 7). Although we have only dealt with the D distributions, all the distributions on both sides of the magnetopause have been observed as predicted by Cowley [see Fuselier et al., 1991].

Evidence of heating at the magnetopause [Paschmann, 1984] and problems in mass conservation across the boundary [Paschmann et al., 1986; Fuselier et al., 1993] however, suggest that the magnetopause rotational discontinuity is not quite ideal. In addition, the current sheet is thin compared to ion gyroradii, giving pitch angle scattering which is not considered in Cowley’s theory other than allowing only a proportion of the incident ions to be transmitted through the magnetopause. These nonlinear ion dynamic effects will scatter individual ions but do not appear to greatly distort the distribution function. A further complication is that the magnetopause rotational discontinuity has a complex structure [Biernat et al., 1989]. The great success of the ideal MHD theory in predicting the D-shaped distributions means that these departures can be regarded as second order.

Combined with the other observations, these data provide unambiguous evidence that reconnection plays the major role in plasma entry into the magnetosphere. The simple theory of Cowley [1982] is remarkably successful in explaining magnetopause particle populations and impressive because it was published a decade before observers confirmed its accuracy.

4. PARTICLE MOTION BETWEEN THE MAGNETOPAUSE AND IONOSPHERE

Reconnection allows particles to stream continuously into the magnetosphere from the magnetosheath when the field line is open. The characteristics of the injected population, however, change with time as the field line is transported into the tail lobe. For a field line which threads the dayside magnetopause (at latitudes below the magnetic cusp C in Figure 3), the ions are accelerated and have a bulk flow directed toward the Earth. Later, however, the same field line will thread the magnetopause tailward of C. Here the injected ions will be
Figure 7. Predicted and observed ion distribution functions in field-parallel and field-perpendicular coordinates (top). Predicted ion distributions between the magnetopause and the ion edge [after Cowley, 1982]. Populations of various origins are shown. The shaded population is the transmitted magnetosheath ion distribution with a cutoff at the de Hoffmann-Teller (dHT) frame velocity $V_{HT}$. This population can be thought of as a characteristic of open field lines. Other ions are from the ionosphere and ring current and are incident on and reflected from the magnetopause. (bottom) Observed ion distribution function observed by AMPTE UKS within an accelerated ion flow region [Smith and Rodgers, 1991]. The contour levels have been chosen to highlight the transmitted magnetosheath population. Note the existence of the D-shaped distribution corresponding to that predicted by Cowley (another example can be seen in the upper right-hand panel of Figure 6).
Figure 8. (a) Shows a rotational discontinuity (RD) formed by a newly opened field line which evolves along the magnetopause at speed $V_f$ in the Earth’s frame of reference. The figure is in the dHT frame of reference in which the newly-opened field lines are at rest and so moves at the de Hoffman-Teller velocity $V_{HT} = V_f$. The Z-axis points along the magnetopause in the direction of motion of the newly-opened field lines and the X-axis is the outward normal to the boundary. The open field line makes angles $\Phi$ and $\theta$ with the magnetopause on its magnetosheath and magnetospheric sides, respectively. These angles increase as the field line evolves away from the reconnection site. (b) A velocity space plot showing field-aligned bulk flow at the Alfvén speed $V_A$ in the de Hoffman-Teller frame $(V_x, V_y$ with origin at O) both into the RD in the magnetosheath (at s) and away from the RD in the magnetosphere (at i). This geometric construction allows the calculation of the minimum and bulk flow field-aligned speed of injected particles in the Earth’s frame $(V_x', V_y'$ with origin at O’). (Adapted from Cowley [1982].)

deaccelerated on crossing the boundary and will also be flowing super-Alfvénically tailward, away from the Earth. As a result, little of the injected plasma from tailward of the cusps will reach the Earth and that which does forms a very low density, low-energy precipitation in the Earth’s frame of reference [Reiff et al., 1977; Newell and Meng, 1987]. Thus plasma with densities comparable to the magnetosheath are seen on each field line for a limited time after it reconnects. As the open field lines move poleward, this causes “cusp” precipitation to be seen over a limited range of latitudes in the ionosphere. Poleward of the cusp the precipitating flux decreases to low values, despite the field line remaining open. Consequently, the precipitation characteristics evolve from “cusp” to “mantle” to “polar cap” as the field lines move poleward away from the open-closed boundary (see Figure 2) [Cowley et al., 1991a].

A further complication in the spectra of particles seen at low latitudes is the fact that ions have a variety of flight times from the magnetopause to an observing satellite in the midaltitude magnetosphere or in the ionosphere. This gives rise to the “velocity filter effect” as first discussed by Rosenbauer et al. [1975], Shelley et al. [1976], Reiff et al. [1977], and Hill and Reiff [1977]. Ions of different field aligned velocity, injected simultaneously across the magnetopause onto any one field line, have different flight times along that field line. Hence they have different arrival times in the ionosphere and, as the field line is convecting, are spatially dispersed along the locus of the field line. In other words, the trajectories of ions precipitating down field lines are only field-aligned if there is no convection and the magnetosphere is stagnant [Lockwood and Smith, 1993]. The magnetosphere is generally not stagnant, and, as a result, the spectrum of field-parallel velocities seen simultaneously at low altitudes implies a spread of injection locations [Lockwood and Smith, 1993; Onsager et al., 1993; Onsager, 1995].

Each newly opened field line passing through any one point $P_n$ on the magnetopause reaches that point at a time $t_n$ after it reconnects. A D-shaped distribution of ions is injected into the magnetosphere and accelerated at each $P_n$ (see previous section). Considering field-aligned ions and scatter-free adiabatic motion, the satellite sees all ions with the same energy of field-aligned motion $E_n$ and (by Liouville’s theorem) distribution function $f$ as they had at $P_n$. Each ion has a time of flight between injection and observation of $d_n (m/2E_n)^{1/2}$, where $d_n$ is the field aligned distance from $P_n$ to the satellite and $m$ is the ion mass. If at time $t$ we observe a field line reconnected at $t_0$, the energy and observation time are related by the time-of-flight equation:

$$d_n (m/2E_n)^{1/2} = t_s - (t_0 + t_n)$$

Equation (3) shows that each injection location on the magnetopause can only provide one energy of zero pitch angle ion to the spectrum seen by the satellite [Onsager et al., 1995]. In addition, it shows that the energy supplied by each injection point (for constant $t_s$ and $d_n$) falls with increasing time elapsed since reconnection ($t_0 - t_0$). Carlson and Torbert [1980] applied the time-of-flight equation to a series of injections observed on a sounding rocket flight. By assuming all the ions were injected at one time they determined an injection distance of between 7 and 19 $R_E$, consistent with injection from the dayside magnetopause. Recently, Lockwood [1995a] has shown how the time-of-flight (3) can be used to predict the dispersion characteristics of the cusp ions for a variety of situations.

The velocity filter effect was initially thought of as resulting from poleward convection across a static open/closed separatrix [e.g., Reiff et al., 1977]. In these steady state cases, $(t_0 - t_0)$ increases with distance as the field line moves away from the open-closed separatrix. Thus a satellite moving away from the open-closed boundary at speeds faster than the convection sees field lines whose elapsed time since reconnection is increasingly greater (i.e., $t_0 - t_0$ is increasing), and according to (3) the ion energies $E_n$ seen from all points on the magnetopause decrease with observation time $t_s$. Conversely, a satellite moving toward the boundary sees field lines whose elapsed time since reconnection is progressively smaller ($t_0 - t_0$, decreasing) and all ion energies rise with $t_s$. For
southward IMF (antisunward convection) this leads to ion energies falling with latitude, whereas for northward IMF (sunward convection) it leads to ion energies rising with latitude [Woch and Lundin, 1992a]. Lockwood and Smith [1994] have considered how time elapsed since reconnection and therefore how ion energies behave in other circumstances, for example, when the satellite moves away from the open-closed boundary at speeds lower than the convection speeds. This applies for higher-altitude spacecraft and for any spacecraft orbits that are nearly tangential to the open-closed boundary.

Equation (3) is more general than the described steady state cases and allows for motions of the open-closed boundary because it employs time elapsed since reconnection \( t_c - t_0 \). Observations have shown the dayside magnetopause to lie closer to Earth for a given compressing solar wind dynamic pressure when the IMF is more southward [Maehara, 1974]. This is explained by rapid reconnection removing dayside magnetospheric flux by opening flux tubes so that they can be transferred into the tail lobe by the tension force and the solar wind flow. If this occurs faster than flux can be replenished by sunward convection of closed field lines, the magnetopause moves inward [Aubry et al., 1970; Fairfield, 1971; Holzer and Slavin, 1979]. This “erosion” would correspond to an equatorward motion of the open-closed boundary in the inner magnetosphere and ionosphere [Freeman and Southwood, 1988]. Because cusp precipitation is seen on newly opened field lines for southward IMF, that is, just poleward of the open-closed boundary we would expect erosion to produce an equatorward motion of the equatorward edge of the cusp precipitation. This was observed with in situ particle measurements [Burch, 1973]. The effect is also seen in red line aurorae and radar echoes caused by the cusp precipitation [Vorobyev et al., 1975; Horwitz and Akasofu, 1977; Foster et al., 1980]. When the boundary migrates equatorward, the most recently opened field line is still at the equatorward edge of the cusp. The time elapsed since reconnection will still increase with latitude giving the same poleward ion dispersion as for steady state (a static boundary) and poleward convection [Lockwood and Smith, 1989]. No convection in Earth’s frame is required to cause the velocity filter effect in these nonsteady erosion cases; only convection across the open-closed separatrix in its own rest frame which, by definition, is the process of magnetic reconnection, will produce the velocity filter effect.

5. OBSERVATIONS OF CUSP PRECIPITATION

Since the original ISIS 2 measurements many spacecraft have observed the cusp at both high and low altitudes [e.g., Russell et al., 1971; Shelley et al., 1976; Reiff et al., 1980; Burch et al., 1982; Lundin and Dubinin, 1984; Newell and Meng, 1988; Aparicio et al., 1991; Woch and Lundin, 1992b; Yamauchi et al., 1993; Norberg et al., 1994]. Cusp precipitation is characterized by the energy dispersion of the ions. Ion energy falls with increasing latitudes as expected from the velocity filter effect for southward IMF conditions (Plates 1 and 2).

In addition to the energy-latitude dispersion, the pitch angle (PA)/time spectrogram shows characteristic “V”-shaped signatures (Plate 1). This effect is also due to the relative times of flight of the ions. An ion that has a velocity vector completely along the magnetic field (a particle with 0° PA) will have a shorter time of flight than one that has the same total energy but spirals around the field line (a finite PA). For example, to observe a large PA ion at the same time and location as a 0° PA particle injected at the same point on the magnetopause (and thus at the same time) requires the large particle to have a higher total energy [Burch et al., 1982; Menietti and Burch, 1988]. This leads to the characteristic “V” shape in the midlatitude data (Plate 1). The convergence of the dipolar-like magnetic field, however, allows only those particles injected with PAs near 0° to reach lower altitudes, and hence no “V” signatures are found in the lower-altitude DMSP data (Plate 2).

Electrons behave somewhat differently from ions. At the edge of the cusp closest to the magnetic separatrix, electron V signatures are sometimes observed [Burch et al., 1982] (for example, around 0717:20 in Plate 1). The equatorward boundaries of the electron and ion precipitations are the edges in the ionosphere corresponding to those seen at the magnetopause (see section 3). The majority of electrons in the cusp, although sometimes highly structured, do not show energy-PA or energy-latitude dispersion. Burch [1985] showed that these electrons are largely responsible for maintaining quasi-neutrality in the cusp and that the electron density throughout the cusp remains equal to the ion density, as predicted by Reiff et al. [1977]. Thus, as mentioned earlier, electron fluxes must increase considerably across the ion edges (Plates 1 and 2), as observed at the magnetopause (Figure 6).

6. EVIDENCE THAT THE CUSP IS ON NEWLY OPENED FIELD LINES

Evidence from several sources shows that the precipitation termed “cusp” seen at lower altitudes is the same magnetosheath plasma that crosses the magnetopause along newly opened field lines and produces the D-shaped ion distribution function. The first set of evidence comes from observations that some cusp ions are accelerated to energies above those in the magnetosheath [Hill and Reiff, 1977]. This is consistent with particle entry having occurred on newly opened field lines which accelerate particles to a fraction of \( V_{HI} \), as discussed in section 3. Because of the velocity filter effect (see section 4), accelerated ions will be seen on the edge of the cusp nearest the open closed separatrix, that is, on the equatorward edge of the cusp for antisunward con-
Plate 1. An example spectrogram of cusp (top) electrons and (bottom) ions from a poleward pass of the DE 1 spacecraft at middle (19,000 km) altitudes. Color contours of differential energy flux are shown as a function of energy and observing time. (middle) The pitch angle variation of the observations as the satellite spins. The dispersion is such that ion energy falls with increasing latitude [from Burch et al., 1986].

vection when $B_z < 0$, and on the poleward edge for sunward convection when $B_z > 0$. Such accelerated ions have been reported by Newell and Meng [1992] and Woch and Lundin [1992b].

Further evidence that the cusp is on newly opened field lines comes from observations that the cusp/cleft plasma is found on eastward/westward convecting flux tubes for strongly negative/positive IMF $B_z$ in the northern hemisphere (and the opposite for the southern hemisphere) [Burch et al., 1985; Lockwood and Smith, 1989], as at the dayside magnetopause [Gosling et al., 1990a] (see section 2). The tension force giving the east-west flow only occurs on newly opened field lines, and hence the cusp precipitation must also be on such field lines. Taguchi et al. [1993] found that the cusp and much of the cleft precipitation always occurred in the region bounded by the cusp/region 1 field-aligned current pair which transfer the east-west flow to the ionosphere. Hence we can infer that not only the cusp but also much of the cleft precipitation must be on open field lines.

The velocity filter effect, described in section 4 provides yet further evidence. The dispersion, seen in Plates 1 and 2, is also seen in the regions bordering the cusp, in particular in the region poleward of the cusp called the mantle. Two consequences arise: first, the magnetic field lines must be moving through, as opposed to around, the cusp region, a feature predicted only by the open field line model. Second, the variation in the ion energies precipitating into the ionosphere shows that the plasma was injected over an extended region of the dayside magnetopause [Onsager et al., 1993; Lockwood and Smith, 1993; Phillips et al., 1993]. This is also consistent with reconnection, as the plasma must stream through the magnetopause continuously once the field line has been opened [Cowley et al., 1991b]. The necessarily broad entry region eliminates the original idea of local plasma entry around a magnetic null as a mechanism for populating the cusp.

The final piece of evidence concerning the open nature of the particle cusp concerns the variation in the cusp location with the IMF orientation. Burch [1973] and Carbery and Meng [1988] showed that the equatorward edge of the cusp migrates to lower magnetic latitudes when the IMF turns southward, consistent with magnetic reconnection eroding the magnetosphere. In addition, Cundill et al. [1989] and Newell et al. [1989] showed that the cusp movement in local time is dependent on the IMF $B_z$ component. Again, this is predicted by the reconnection model as a result of the effects of
Plate 1. An example spectrogram of cusp (top) electrons and (bottom) ions from a poleward pass of the DE 1 spacecraft at middle (19,000 km) altitudes. Color contours of differential energy flux are shown as a function of energy and observing time. (middle) The pitch angle variation of the observations as the satellite spins. The dispersion is such that ion energy falls with increasing latitude [from Burch et al., 1986].
Plate 2. An example of cusp (top) electrons and (bottom) ions from an equatorward pass of the DMSP F7 satellite at low (800 km) altitudes. This plot is similar to Plate 1, except the satellite only observes ions with pitch angles close to zero and the energy scale on the ion spectrogram has been inverted. As in Plate 1 the ion energy falls with increasing latitude [after Newell et al., 1991].

7. QUASI-STEADY AND TRANSIENT DAYSIDE RECONNECTION

For many years, bursts of magnetic reconnection have been invoked as a cause of transient phenomena near the dayside magnetopause. These events were termed flux transfer events (FTEs) by Russell and Elphic [1978] and have been widely observed [e.g., Haerendel et al., 1978; Paschmann et al., 1982]. The main characteristic of these events is a bipolar signature in the magnetic field component normal to the magnetopause surface. Originally, this signature alone defined an event, but it has become clear that many FTEs also have a characteristic plasma signature at their center, consisting of a mixture of magnetosheath and magnetospheric plasma [e.g., Thomsen et al., 1987; Farrugia et al., 1988; Rijnbeek et al., 1984; Smith and Owen, 1992]. Many FTEs are associated with accelerated ion flows, as predicted by reconnection-based models [e.g., Paschmann et al., 1982].

Other mechanisms that have been proposed to explain FTE signatures include impulsive plasma penetration [Heikkila, 1982], which has been shown to be incorrect by Owen and Cowley [1991], and the plasma penetration model of Lemaire [1977], which has been challenged by Smith and Curran [1990]. Sibeck [1992, 1993] has proposed that FTE signatures are responses of the magnetosphere to dynamic pressure changes in the solar wind. Sibeck and Smith [1992] showed that, in theory, ion flow measurements could distinguish between the pressure pulse and reconnection mechanisms.
The event studied, however, did not fit either model. This is not a problem for the reconnection model as boundary motions can be added to the passage of a region of enhanced newly opened flux. On the other hand, the pressure pulse model describes what the boundary motion must be. Lockwood [1991] has argued against the pressure mechanism, as it would tend to produce tripolar rather than bipolar signatures and because it is inconsistent with the observed dependence on sheath field orientation. A recent correction to this statistical argument by Sibeck and Newell [1995] does not alter this conclusion. Smith and Owen [1992] argued that the plasma in an FTE is not the same as in the sheath. Eipch et al. [1994] found no correlation between solar wind dynamic pressure changes and magnetopause FTEs. Heapgood and Lockwood [1995] used two spacecraft to show that an FTE signature was a thickening of the LBL, as predicted by reconnection theory, as opposed to the thinning expected from the pressure pulse. The evidence thus strongly favors the reconnection model. Sibeck’s work, however, has highlighted the need for care in interpreting magnetopause signatures because some so-called FTEs may be surface waves due to pressure pulses.

The effects of an enhancement in the reconnection rate on the magnetopause magnetic field configuration can be understood by applying conservation of mass and energy to the inflow and outflow of plasma at the reconnection site (Figure 9). As the outflow speed \( V_{HT} = E_r/B_n \) is defined by the local Alfvén speed in the inflow region, an increase in the reconnection rate \( E_r \) must increase \( B_n \) and hence widen the region through which the plasma flows out of the reconnection site (the expansion fan). A pulse of reconnection thus produces a bulge in the reconnection layer (Figure 9) which yields the classic FTE bipolar magnetic field signature as the structure moves over the spacecraft. This simplification of the explanation of the reconnection layer was suggested by Southwood et al. [1988] and is supported by MHD simulations [Scholer, 1988], and by analytic theory of time-dependent Petschek reconnection [Semenov et al., 1992]. The bulge and the newly reconnected flux which threads it, move away from the reconnection site at the speed \( V_{HT} \). The ion edges (11 and 12 in Figure 5) bound the region of accelerated flows [e.g., Paschmann et al., 1979; Gosling et al., 1990a].

It is important to note that the de Hoffman-Teller velocity \( V_{HT} \) is independent of the reconnection rate. As \( V_{HT} = E_r/B_n \), an increase in \( E_r \) causes a proportional increase in the boundary normal magnetic field \( B_n \), but the field line velocity \( V_{HT} \) is constant if the sheath flow and magnetic field remain constant. Observations of so-called “quasi-steady” reconnection tell us about the existence of accelerated flows and about the velocity \( V_{HT} \) but provide no information on the reconnection rate \( E_r \). Since they indicate that reconnection is ongoing the reconnection should be called “quasi-continuous.”

8. GROUND-BASED MEASUREMENTS AND IONOSPHERIC TRANSIENTS

Satellite observations yield a detailed view of cusp precipitation within the magnetosphere and ionosphere. An important limitation, however, to such in situ observations of the cusp at middle and low altitudes is that they provide a sequence of measurements as a time series. Since the satellite must be moving through the cusp, this sequence will, in general, result from a complex mixture of spatial gradients and temporal changes. This spatial/temporal ambiguity cannot be resolved by a lone satellite. In general, the simplest interpretation is in terms of cusp characteristics which vary along the satellite orbit but do not change with time.

On the other hand, ground-based or satellite remote sensing of the cusp region can resolve temporal changes and spatial gradients because, unlike a satellite, the same region can be viewed for extended periods. The problem is that the precision and resolution of such measurements is very low compared to the in situ data. It is instructive to compare the timescales of phenomena which can be resolved in data from in situ satellite and ground-based remote sensing experiments. Consider a satellite in the topside ionosphere, which will move at about 5 km s\(^{-1}\) and have an orbital period of approximately 90 min. If spatial gradients over a scale of 5 km or smaller were neglected then temporal fluctuations of period up to 1 s could be observed. Furthermore, by comparing one orbit with the next, variations of period 90 min and greater could be observed, but the in situ observations can give no information on fluctuations of the cusp in the interval between about 1 s and 90 min. A ground-based remote sensing can observe the cusp for a period of approximately 2 hours while the Earth rotates under it, depending on the size of the field of view. The resolution that can be achieved varies with the technique, but both optical instruments and radars can achieve subsecond sampling intervals. Hence ground-based observations can be used to study exactly those variations in the cusp which are hidden from in situ low-altitude satellite data. However, care must be taken with ground-based measurements which scan the ionosphere to produce an image, as they suffer the same kind of limitations as the satellite data.

Ground-based observations of the cusp/cleft region with high time resolution have provided a radically different view from that derived from satellite data. One major reason for this may be the importance of temporal variations of the cusp in the 1 s to 90 min band. In recent years much interest has focused on such variability with considerable progress made in terms of accommodating the ground-based and satellite observations.

8.1. The Dayside Aurora

The aurora that takes place on the nightside of Earth is caused by energetic (several keV) electrons which precipitate into the upper atmosphere. The high energy
of these electrons allows them to penetrate down to low altitudes (about 100 km) where they stimulate the atomic oxygen gas to emit mainly green (557.7 nm) and UV (295.8 and 297.2 nm) emissions.

On the other hand, the magnetosheath-like electron precipitation in the dayside cusp and cleft is of much lower energy (characteristic energies of order 100 eV) and preferentially excites red (630 and 636.4 nm) emissions at greater altitudes (250–500 km). The first observation of a localized patch of predominantly red emission on the dayside was made by Sanford [1964]. From the ratios of emission intensities of spectral lines, Eather and Mende [1971] showed that this emission was caused by “soft” electron precipitation, which Heikkila [1972] associated with the recently discovered cusp/cleft precipitation. Whalen et al. [1971] showed that the red aurora was poleward of the dayside auroral oval, as later viewed from space in the UV lines [Murphree et al., 1980; Elphinstone et al., 1992]. The cusp/cleft precipitation is known to cause considerable heating of the thermal ionospheric electron gas [Titov et al., 1976; Burch et al., 1982] which thermally excites much of the red aurora [Mantas and Walker, 1976; Roble and Rees, 1977; Wickwar et al., 1974; Wickwar and Kofman, 1984; Lockwood et al., 1995a]. Because of the complexities of the excitation and emission processes giving the various spectral lines, it is generally impossible to distinguish cleft and cusp precipitations from optical observations. Consequently, we describe the dayside red-dominant aurora as being caused by “cusp/cleft” precipitation.

The dayside aurora was shown to migrate equatorward following southward turnings of the IMF and hence, during the early phases of substorms as open flux is accumulated in the tail lobe [Vorobyev et al., 1975; Horwitz and Akasofu, 1977]. This is consistent with the motion of the cusp precipitation region [Burch, 1973]. Considerable debate arose about the cusp/cleft movement in response to southward turnings of the IMF and substorm activity [Eather et al., 1979; Sandholt et al., 1983; Meng, 1983; Eather, 1985].

Feldstein and Starkov [1967] reported that the dayside aurora displayed structured forms, as seen by white-light cameras (most sensitive to 557.7-nm emissions). Vorobyev et al. [1975] and Horwitz and Akasofu [1977] noted that as the band of emission drifted equatorward, these discrete forms broke away from the band and generally drifted poleward at speeds of typically 0.5–1 km s⁻¹. In these structured events, green emissions can dominate over the red, revealing electron precipitation of several kilovolts characteristic energy. Hence, if these electrons originate from solar wind, they have undergone considerable acceleration and would not be classed as “cusp” from their spectral characteristics [Newell and Meng, 1988]. These poleward moving structures have been extensively studied by Sandholt and coworkers using instruments on the Svalbard islands (one of the few locations in the northern hemisphere where the cusp/cleft aurora can be observed and then only at new moon near winter solstice) [Sandholt, 1988; Sandholt et al., 1985]. These authors defined a class of poleward moving events which they termed “dayside auroral breakups,” with transient forms observed in both red and green light. Such events are seen throughout the dayside and an example is shown in Plate 3. Their work has clearly
demonstrated these events to be associated with reconnection at the dayside magnetopause. Two pieces of evidence relate to their behavior for different IMF orientations. Like magnetopause FTE events, southward IMF appears almost to be both a necessary and sufficient condition for their occurrence and the distribution of repeat periods is very similar to that of FTEs, with a mean value of about 7 min [Fasol, 1995]. Also, they show longitudinal motion which depends on the sense of IMF $B_z$: for positive $B_z$, the events move westward (for northern hemisphere observations) [Sandholt et al., 1989a; Sandholt, 1988], whereas for negative $B_z$, they move eastward [Sandholt et al., 1992, 1993; Lockwood et al., 1994a]. The only known explanation for this is that they are subject to the azimuthal tension force of the Svalgaard-Mansurov effect and that they comprise newly opened field lines produced by reconnection (see Figure 4). Furthermore, the longitudinal motion of these events is greatest early in their lifetime and slows as their poleward drift increases, consistent with the pattern of motion of newly opened flux under the influence of tension and then magnetosheath flow. Therefore the red dayside breakup events must be on the same (newly opened) field lines as the cusp precipitation.

Lockwood et al. [1993a] studied one westward traveling and one eastward traveling event at the moment when the green transient (revealing keV electron precipitation) passed through zenith for the observing site and showed that the region of green-dominant emission was much narrower than the red-dominant transient. Furthermore, the most intense green light came from the poleward edge of the red transient for the westward traveling event but from the equatorward edge of the eastward traveling event. These authors concluded that the red-dominant transient delineates a patch of newly opened flux and that the green transient is the region of upward field aligned current needed to transfer the Svalgaard-Mansurov flow to the ionosphere. This interpretation is significantly different from that by Fasol et al. [1992], who associate the green transient with the reconnection X line and multiple brightenings of such transients as multiple X lines. In either case, the mechanisms which accelerate the solar wind electrons to keV energies in these thin (1–10 km) regions are not understood. Note that such a thin band of accelerated electrons would not stop the broader region being classified as cusp. Thus the presence of 557.7-nm transients does not indicate that the events are outside the cusp region. The dimensions of some events are up to 500 km north-south and 1500 km east-west [Leonov et al., 1992; Lockwood et al., 1993b; Pudovkin, 1991], while other events are smaller [Sandholt et al., 1992; Lockwood et al., 1994a].

8.2. Radar Observations

The cusp/cleft region was identified from its ionospheric heating effects in topside ionosonde data by Nishida [1967] and Titheridge [1976]. Using ground based ionosonde data, Stiles et al. [1977] defined the cusp/cleft region to be a highly structured and irregular region. Such irregularities have recently allowed Baker et al. [1990] to identify the cusp/cleft as broad echoes received by HF backscatter radars. The ionospheric heating in the cusp region has been observed by the Sondrestromfjord incoherent scatter (IS) radar [Wickwar and Kofman, 1984; Watermann et al., 1992, 1994; Nilsson et al., 1994].

The equatorward motion of the cusp/cleft region, as identified by the Chatanika IS radar, was reported during southward IMF by Foster et al. [1980], with peak equatorward excursion occurring near a substorm onset, consistent with the auroral motions and magnetospheric erosion. Observations by the EISCAT radar [Lockwood et al., 1989a, b, 1990b, 1993b, c, 1994a; Sandholt et al., 1989b; J. Moen et al., Variability of dayside high-latitude convection associated with a sequence of auroral transients, submitted to Geophysical Research Letters, 1994] have shown that within each auroral red transient event for large IMF $B_z$, there is a strong burst of longitudinal flow, consistent with the Svalgaard-Mansurov effect, and with ion-neutral frictional heating of the ion gas. The plasma flow within the auroral (630 nm dominant) events is, at all times, found to be equal to the phase motion of the events as a whole. This is an important finding, as it eliminates any explanations involving motions of sources across field lines [e.g., Lui and Sibeck, 1991] and provides a clear distinction between these events and large-scale traveling convection vortices generated by solar wind dynamic pressure pulses [Frits-Christensen et al., 1988; Sibeck, 1990; Lühr et al., 1993; Glassmeier et al., 1989; Glassmeier, 1992; Potemra et al., 1992]. Pincock et al. [1993] combined HF radar and satellite observations to show that cusp precipitation and a longitudinal flow burst were coincident. As the flow burst defined the cusp, the Pincock et al. [1993] data imply that the transient events reflect variations of the cusp as a whole, and not structure within the cusp.

The question then arises as to why flows are pulsed and why transients form and propagate. Sandholt et al. [1992] and Lockwood et al. [1993b] invoked pulsed reconnection as an explanation, that is, each pulse generates a region of newly opened flux. In this case, such events should be seen in association with magnetopause flux transfer events. Unfortunately, only one opportunity to date has allowed simultaneous observations of the cusp/cleft aurora and the magnetopause. In this case, some correspondence between magnetopause events and dayside breakup/flow burst events was observed [Elphic et al., 1990]. In addition, Fasol [1995] has recently shown that the distribution of event repeat periods is very similar to that for magnetopause FTEs, as reported by Lockwood and Wilde [1993]. The only alternative explanation, recently proposed by Newell and Sibeck [1993], suggests that the enhanced flows are due to enhanced magnetosheath $B_z$ (possibly caused by enhanced compression of the magnetosheath). For steady reconnection, however, this would produce a continuous
Plate 3. An example of a dayside auroral breakup event (sometimes also called a poleward moving auroral form (PMAF)), as seen at winter solstice from Ny Ålesund, Svalbard. The images are taken in 630-nm (red line) emissions from oxygen atoms which are excited by the precipitation of magnetosheath-like plasma in the cusp/cleft region. A sequence of false-color images is shown. For each, the circles are lines of constant zenith angle and the magnetic meridian is the diagonal line running from the top left corner (north/poleward) to near the bottom right corner (south), east/west being to the right/left of this line. The images shown are 1-s integrations, and 1 min apart, the time being given in the top left corner. For this station the local magnetic time is roughly UT + 3 hours, so these images were taken just after noon. Contours of the 630-nm intensity are color coded between the largest (white) and the lowest (blue/black). In the first observation (0930 UT) an event is already in progress, with a bulge in the poleward edge of the aurora seen to the west of the station. In subsequent images this grows poleward and migrates to the east, until it fades by 0940–0941 UT. By this time a second event is already present, being first seen as an equatorward motion of the equatorward boundary which spreads eastward in the interval 0934–0937 UT. This also migrates poleward, producing a poleward bulge of the poleward edge of the aurora by about 0944 UT, which moves eastward and fades by 0949 UT, as did the first event. The poleward motion of these overlapping events is most clearly seen in the vertical panel to the right of the plot which shows the intensities seen along the magnetic meridian with north to the left and time running from top to bottom [from Lockwood et al., 1995a].
region of red emission which moves back and forth in the east-west direction. Consequently, $B_{y}$ changes are not consistent with the observed patches of red aurora, which move repetitively in one direction [Lockwood et al., 1995a, b].

9. THE CUSP FOR TIME-VARYING RECONNECTION

As flow transients reveal newly opened field lines (subject to the tension force) on which cusp precipitation is expected, and because this precipitation is expected to cause red auroral transients, a link between FTEs and the cusp is strongly suggested. The first suggestion that the cusp precipitation region was the ionospheric signature of an FTE came from Menietti and Burch [1988]. They derived an estimate of the field-perpendicular width of the injection region from the pitch angle–energy dispersion of ions. Since this was of the same order (approximately $1 R_{E}$) as typical FTE dimensions at the magnetopause [Saunders et al., 1984], they associated the precipitation with an FTE. The evidence from ground-based instrumentation for ionospheric transients, with dimensions and repetition periods consistent with FTEs, led Lockwood and Smith [1989] to associate cusp precipitation with FTE signatures. The subsequent development of the “pulsating cusp” model [Smith and Lockwood, 1990; Lockwood and Smith, 1994] has paralleled advances in understanding flows and evolution of FTE signatures in the ionosphere (section 9.1) and led to the successful prediction of cusp ion steps (section 9.2).

9.1. FTE Signatures in the Ionosphere

The ionospheric signatures of flux transfer events were first discussed by Goertz et al. [1985] and Southwood [1985, 1987]. In these papers, the patch of newly opened flux produced by a reconnection burst was envisaged as moving poleward under the influence of magnetic tension and magnetosheath flow. Since the ionosphere is incompressible, this movement generates a twin vortex flow pattern in the ionosphere. To illustrate the flows and filamentary field-aligned currents produced, Southwood [1985, 1987] employed a circular newly opened flux tube. Extensions of this idea allowed for flux tube untwisting [Lee, 1986] and elliptical cross sections [Wei and Lee, 1990; Lockwood et al., 1990c].

Southwood noted that when the newly opened flux was appended to the older open flux in the tail lobe, it would no longer be moving faster than any surrounding flux tubes, and the pattern would disappear. His simple illustration of the model using a circular flux tube has been employed, often rather literally, in attempts to explain transient ionospheric flow signatures in the cusp/cleft region as resulting from magnetopause FTEs [Lockwood et al., 1989b; Bosqued et al., 1991; Basinska et al., 1989]. Many of the predicted features, however, arise from the assumed circular nature of the flux tube. Questions arise as to what shape the region of newly opened flux would initially have in reality and how the shape would evolve with time. One particular problem with the rigid circular flux tube interpretation is that it left the open-closed field line boundary in a highly distorted shape [Lockwood et al., 1988]. These problems were answered with studies of the effects of magnetopause erosion [Freeman and Southwood, 1988] and of the response of ionospheric flows to southward turnings of the IMF [Lockwood et al., 1990a]. Along with the theoretical work of Siscoe and Huang [1985], these studies lead to general concepts of ionospheric flow excitation [Cowley and Lockwood, 1992] which could be used to understand how patches of newly opened flux and their associated flows evolve with time following a burst of magnetopause reconnection [Cowley et al., 1991a; Smith et al., 1992].

As an example, Lockwood [1994] considered a burst of reconnection lasting 2 min, followed by a 7-min period during which the reconnection is switched completely off (Figure 10). The reconnection is envisaged as commencing at noon and spreading tailward. From observations, the speed of tailward expansion in the ionosphere is initially about 10 km s$^{-1}$ but falls to about 1 km s$^{-1}$ near dawn and dusk [Lockwood et al., 1986; Todd et al., 1986; Etemadi et al., 1988; Saunders et al., 1992; Lockwood et al., 1993b]. Tailward expansion may have one of a number of causes but will depend on the mapping of LLBL field lines to the ionosphere and, hence, upon the amount of dayside open flux and the degree to which the cusp is “opened” as a result [Crooker et al., 1991].

A region of rotational discontinuity, produced by the reconnection pulse, emerges from the (still active) X line (Figure 10, diagrams 1a and 1b). This newly opened flux forms an extrusion of the polar cap boundary (Figure 10, diagram 1c). The main effect of the reconnection is to move the open-closed boundary equatorward [Lockwood et al., 1993b], although some weak poleward flow begins one Alfvén wave travel time after reconnection onset.

At time $t = 2$ min the reconnection at noon switches off, and this switch off will expand tailward in the same manner as did the onset of reconnection. Therefore, at any one MLT the reconnection lasts for 2 min. At $t = 4$ min the reconnection has ceased at noon but is still ongoing at MLT on either side of noon (Figure 10, diagram 2b). At noon, the ionospheric open/closed boundary is equatorward of its zero flow equilibrium location, which depends upon the amount of open flux that exists at this time. From Cowley and Lockwood [1992], flow of the kind shown in Figure 10, diagram 2c will be excited as the system tends toward the new equilibrium. At noon the polar cap boundary is now said to be adiabatic and moving poleward with the plasma flow; at MLT beyond the active merging gaps, the flow is weakly equatorward.
For the assumed conditions the reconnection pulse does not expand beyond the midmorning and midafternoon sectors. At time $t = 7$ min, reconnection has ceased at all MLT (*Figures 10, diagrams 3a and 3b). Note that if the traveling reconnection enhancement persists for longer, it is possible for the next reconnection pulse to begin before the prior pulse has disappeared. The patch of RD produced by the burst of reconnection is now evolving over the dayside magnetopause. Each newly opened field line will evolve according to the magnetic tension and magnetosheath flow it experiences [Cowley and Owen, 1989], and both, as well as the pattern of motion, will depend upon the MLT. Therefore the patch of RD (and any associated FTE structure) will change in shape as it evolves. Inductive smoothing of ionospheric convection means that only at this stage have the flows increased to their peak speeds, even though there is now no reconnection present at any MLT (*Figure 10, diagram 3c). The decay time for this flow is of the order of 15 min, which is the timescale for the transfer of flux into the tail by the magnetosheath flow. Consequently, even without a second reconnection pulse, there would be decaying ionospheric flow for a further 15 min.

At times 9–11 min in Figure 10 the second reconnection pulse takes place at noon, and the subsequent behavior is as for the previous pulse. The second patch of RD, however, forms and evolves over the dayside while the first is still evolving into the tail lobe (*cf. Figure 10, diagrams 4a–6a and 4b–6b with 1a–3a and 1b–3b). The magnetopause has become antisunward moving patches of RD, separated by extensive regions of TD, with the ratio of the areas of RD and TD in the high-latitude boundary layer being the mark-space ratio of the reconnection pulses (in this example 2:7). Likewise, the flows caused by the second burst of reconnection are superposed on the residual flows due to the prior pulse of reconnection (*cf. Figures 10, diagrams 4c–6c and 1c–3c).

When the IMF $|B_y|$ is large, the concepts discussed above still apply. However, there is one additional consideration; namely, the newly opened flux experiences an initial azimuthal (east/west) motion due to the magnetic tension force, as discussed in section 2 (the Svalgaard Mansurov effect). This motion commences just one Alfvén travel time after reconnection but will decay as the field line straightens. As a result, the motion is east/west initially and evolves to poleward as the magnetosphere-ionosphere system tends towards the new equilibrium. The predictions in Figure 10 can easily be generalized to allow for this pattern of motion [see Cowley et al., 1991b; Lockwood et al., 1993b].

9.2. Cusp Ion Dispersion Steps

An important prediction of the above model is poleward propagating steps in the cusp ion dispersion signatures arising from reconnection which varies in time. A low-altitude satellite flying meridionally poleward along the noon meridian through the cusp at time $t = 13$ min in Figure 10, diagram 5c, for example, moves poleward much faster than the convection. Consequently, it will encounter both patches of newly opened flux produced by the two reconnection pulses. Across any one patch, the satellite will first encounter the field lines which were reconnected at the end of that reconnection pulse, whereas at the poleward edge of the patch, it will encounter field lines which were opened at the start of the reconnection pulse. The time elapsed since reconnection therefore increases as the satellite flies poleward, and the cusp ion energies decay with altitude. Hence, across each patch produced by a burst of transient reconnection the satellite observes ion energy-latitude dispersion. On crossing the boundary between the two patches (dashed line in Figure 13, diagram 5c), however, the satellite moves from a field line which was reconnected at time $t = 9$ min to one that was reconnected at time $t = 2$ min. The time elapsed since reconnection shows a discontinuous increase, and all cusp ion energies will step downward. The steps are only observed because each region of newly opened flux is appended immediately equatorward of the patch produced by the previous burst. This is a prediction unique to the Cowley et al. [1991b] model of FTE signatures, which allows for the change of shape of each patch as it evolves. Rigid, circular, or elliptical flux tube models only produce steps at the singular points where event signatures touch, and therefore the probability of observing steps would be negligibly small. Independent of these theoretical predictions by Cowley et al. [1991b], examples of such cusp ion steps were found in data from the DMSP satellites [Newell and Meng, 1991]. Additional examples have been presented by Escoubet et al. [1992] and Lockwood et al. [1993b].

Steps may also be produced in other ways. For example, if a satellite passes from flow streamlines emanating from one X line to those from a second X line, a step would generally be seen [Lockwood, 1995b]. Onsager et al. [1995] presented observations of steps by the DE 1 and DE 2 satellites in roughly the same location but at two different times, implying a spatial origin. A steady state spatial cause would not, however, produce the poleward moving events predicted in Figure 10. Temporally produced steps will be on the boundary between two poleward moving patches of newly opened flux. Closely conjugate EISCAT radar and DMSP-F10 satellite measurements [Lockwood et al., 1993b] have shown that cusp ion steps, observed by the satellite, are seen on the boundaries of the poleward moving events seen by the radar (Plates 4 and 5), as predicted by the reconnection burst theory (Figure 10). The ionospheric flows provide additional evidence that these events result from temporal and not spatial structure of the reconnection [Lockwood, 1995a] (Plate 5).

The events shown in Plate 5 were observed in March when sunlight prevents observations of the 630-nm aurora. The poleward moving events seen by the EISCAT radar in
electron temperature, however, are almost identical in nature to the poleward moving 630-nm transients discussed in section 9.1. Similarly is expected because the 630-nm dominant cusp aurora is generally caused by thermal excitation due to enhanced ionospheric electron temperatures produced by precipitating cusp particles. Lockwood et al. [1993b] show that the cleft, cusp, mantle, and polar cap precipitations were seen when the satellite intersected dif-
different poleward moving events. The clear implication is that the precipitation evolves within each event through these classifications. The regions are thus events seen at different phases of their evolution.

Lockwood and Smith [1994] noted that steps were only discontinuous changes in the cusp ion energy dispersion curves if the reconnection rate went to zero between pulses. If the reconnection continued, but at a much lower rate the steps would become less abrupt. They estimated that fluctuations in the reconnection rate of a factor of less than 2 would be indistinguishable from steady state dispersion plumes on particle spectrograms (e.g., Plate 2). They also noted that the form of the cusp ion steps would critically depend on the ratio of the satellite and convection speeds with respect to the open-closed boundary. Specifically, if the satellite moved away from the boundary more slowly than the convecting plasma, the steps would be upward. In such cases, a patch of more recently opened flux is convected over the satellite giving a sudden decrease in time elapsed since reconnection and hence an upward step. This has recently been verified by Pinnock et al. [1995], who observed the form of cusp ion steps predicted by Lockwood and Smith in data from the DMSP-F8 satellite while it moved longitudinally across the cusp. These events are also temporal in nature because poleward moving transient events were seen using the Halley Bay PACE radar, which was in close conjunction with the satellite.

9.3. The Pulsating Cusp Model

The pulsating cusp model [Smith and Lockwood, 1990] was initially based on an analysis of a DE 2 pass through the cusp in which the direction of convection and the cusp ion dispersion were inconsistent with a steady state interpretation [Lockwood and Smith, 1989]. The authors stress that an entirely pulsed dayside magnetopause reconnection rate is just one limit of more general behavior at the magnetopause. The other limit is that for which the reconnection rate is continuous and at a constant rate. A priori there is no apparent reason why the reconnection rate should favor any particular limit. The usual argument put forward to defend steady state scenarios is that the cusp precipitation often appears continuous. Given that it is reconnection which gives cusp particles access to the magnetosphere, it is argued that it too must be continuous [Newell and Sibeck, 1993]. Cusp precipitation on any newly opened field line, however, persists for an extended period of time. If that period of time exceeds the intervals without reconnection, then continuous cusp precipitation is observed even if the reconnection is entirely pulsed. Precipitation, with densities sufficient to class it as “cusp,” persists on any one field line for an extended period partly because of the extended time for which particles cross the dayside boundary, but mainly because of the large range of ion flight times from the magnetopause to the ionosphere. This extended time gives the cusp a latitudinal width as the field lines convect poleward and/or the open/closed boundary erodes equatorward. Lockwood and Davis [1995] have used a cusp ion pass (Plate 2) to estimate that precipitation persists for between 8 and 20 min, depending on the reconnection site.

Even for southward IMF, the cusp is only observed for about 70% of the time by low-altitude spacecraft at 12 MLT and less often at other local times [Newell et al., 1989]. Smith et al. [1992] used a simple model of completely pulsed ionospheric reconnection signatures to demonstrate how the local time distribution of the cusp occurrence rate could arise for a low-altitude spacecraft. Even in this extreme limit they were able to match the data of Newell et al. [1989]. Lockwood and Davis [1995] have also shown that the level to which the reconnection is pulsed does not effect the distribution of the cusp widths or occurrence probability.

An inversion of the theory of cusp ion steps [Cowley et al., 1991] allows the variation of reconnection rate from the cusp ion dispersion characteristics to be determined [Lockwood and Smith, 1992]. The theory only applies to low altitudes as it relies on the concept of ionospheric incompressibility. Pulses giving the stepped appearance of the cusp ion spectrumpogram in Plate 4 can be seen in Figure 11. The distance to the reconnection site has been calculated by comparison with the corresponding events seen by the EISCAT radars (Plate 5b) [Lockwood et al., 1995b]. A key part of this theory is that the lowest-energy ions seen at any one observation time have the longest time of flight; thus they were the first to be injected. In the reconnection model this means that they crossed the magnetopause close to the X line. Although this is always true for observations relatively close to the open/closed separatrix, it may not be valid for further away [Lockwood, 1995b].

Lockwood et al. [1994b] applied the theory to the cusp ion plume shown in Plate 4. The results show that the reconnection is highly pulsed, even in this seemingly steady state example. Newell and Meng [1995] also applied the theory to 22 cusp ion passes. They found that the intervals with no reconnection were rare and short, indicating that reconnection is generally continuous but variable in rate. Current theories which explain FTE signatures as transient thickenings of the open LLBL [Southwood et al., 1988; Scholer, 1988; Semenov et al., 1992] are consistent with this idea (Figure 9). The examples of cusp ion steps reported in the literature, however, show that in some cases the reconnection rate does fall to undetectable levels between pulses [Lockwood et al., 1994b].

Recently, the theory of ion injection and acceleration at the magnetopause (as discussed in section 3) has been used to compute the characteristics of the magnetosheath at the reconnection site [Lockwood et al., 1994b]. In particular, the Alfvén and bulk flow speeds of the sheath can be estimated at the X line, as can the velocity with which field lines emerge from the X line. Lockwood et al. [1994b, 1995b] have applied these techniques to the data shown in Plates 2 and 4, respectively,
and have found the results consistent with reconnection occurring near the subsolar magnetopause.

10. THE FUTURE

Earth's magnetic cusps are an important area for study in space physics because they provide information on the magnetopause boundary and its variations. This review has, in general, chronicled the development of our understanding of the behavior of magnetosheath ions injected into the magnetosphere. While much has been learned concerning the character and behavior of the cusps, many questions remain. For example, what is the relative importance of transient reconnection pulses in comparison to the more quasi-continuous form? The behavior of the electron population has generally been ignored and needs further study. The relationship of cusp transients to dayside acceleration (inverted V) events is only now being studied [Menietti and Smith, 1993]. Burch [1985] clearly demonstrated that quasi-neutrality of the cusp is maintained, but how it is achieved is largely unknown. In addition, the behavior of the cusp is best understood for southward IMF conditions. Relatively little is known about the cusp during the remaining 50% of the time when the field is northward. There are indications that lobe reconnection occurs preferentially in the summer hemisphere [Crooker and Rich, 1993]; therefore the very existence and behavior of the winter, northward IMF cusp is of great interest and not easily understood.

An additional important area of study concerns how much of the LLBL is an open field line. Observations of accelerated ion flows and D-shaped distributions clearly reveal that there is an open field line LLBL. There is debate, however, as to whether, and how much, the particle populations used as evidence for an LLBL on closed field lines may in fact be on open field lines (see review by Smith [1991]).

The cusp precipitation is known to show dipole tilt effects [Newell et al., 1989]. This also raises interesting questions about the effects of season and time of day on magnetic field topology. In turn, this will influence field line evolution velocities, the ion acceleration, and the time-of-flight dispersion characteristics. In addition, the cusps may show strong asymmetries in ionospheric conductivities at the solstices, influencing the Joule heating rates and the inductive time constants with which ionospheric flow can react to changes and transient events.

Evidence is growing that reconnection rate variations not only produce the characteristic FTE signatures on the dayside magnetopause but are also responsible for the auroral and flow transients seen in the cusp region. The origins of the narrow green-line transients, however, are still unknown. Links to other phenomena, such as polar cap patches, are likely but not yet established. An interesting phenomena called “double injections,” in which cusp plumes appear to overlap [Woch and Lundin, 1991, 1992b; Yamauchi, 1993; Norberg et al., 1994] deserves further attention. Lockwood and Smith [1994] showed that at least some of these double injections, at midaltitudes, are finite ion gyroradius effects around upward cusp ion steps produced by pulsed reconnection. On the other hand, Norberg et al. [1994] have shown examples at low altitudes and point out that these do not easily fit into the reconnection model. A possible solution has recently been suggested by Lockwood [1995c],

Plate 4. Cusp particles seen by the DMSP-F10 satellite during an equatorward pass over the field of view of the EISCAT radar. The format is as for Plate 2, other than the ion energy scale has not been inverted. The large arrows mark the times of major cusp ion steps [from Lockwood et al., 1993b].
Plate 5. Observations of the electron temperature by the EISCAT UHF radar during the pass of DMSP-F10 shown in Figure 10. The black line and large arrows show the satellite pass and locations of the cusp ion steps. Plate 5b shows a shorter time period at higher time resolution around the DMSP pass and also shows the flow vectors derived from the radar data.
but this is not completely satisfactory and further work remains.

The future for such studies is an exciting one, with a number of new facilities and missions. These include the four-spacecraft Cluster mission, the new ESR-EISCAT incoherent scatter radar combination, and the network of SuperDARN HF backscatter radars. In the longer term, initiatives such as magnetospheric imaging will be vital for understanding the complexities of the temporal and spatial structure of the cusps and magnetopause.

This review has covered what we believe to be the important theories and observations that led to our present understanding of the cusp. For further reading on work in the cusp/cleft region we recommend the International Union of Geodesy and Geophysics review by Crooker and Burke [1991].

ACKNOWLEDGMENTS. T. E. Cravens was the editor responsible for this paper. He thanks the interdisciplinary reviewer, B. Dutrow; a technical reviewer, J. L. Burch; and another anonymous technical reviewer.

REFERENCES


Burch, J. L., J. D. Menietti, and J. N. Barfield, DE-1 Observations of solar wind-magnetospheric coupling processes in the polar cusp, in Solar Wind-Magnetosphere Coupling, ed-


Formisano, V., HEOS-2 observations of the boundary layer from the magnetopause to the ionosphere, Planet. Space Sci., 28, 245, 1980.


Glassmeier, K.-H., Traveling magnetospheric convection twin-


Paschmann, G., Comment on “Electric field measurements at


Saunders, M. A., Recent ISEE observations of the magnetopause and low-altitude boundary layer: A review, J. Geophys., 52, 190, 1983.


Sibeck, D. G., A model for the transient magnetospheric


Southwood, D. J., and W. J. Hughes, Theory of hydromagnetic waves in the magnetosphere, Space Sci. Rev., 35, 301, 1983.


