

# Creation of the substorm current wedge through the perturbation of the directly driven current system: a new model for substorm expansion

G. Rostoker<sup>1</sup> and E. Friedrich<sup>1,\*</sup>

<sup>1</sup>Department of Physics, University of Alberta, Edmonton, Alberta, T6G 2J1, Canada

\* now at: NSERC Postdoctoral Fellow at University of Calgary, Calgary, Alberta, Canada

Received: 1 September 2004 – Revised: 21 March 2005 – Accepted: 23 May 2005 – Published: 15 September 2005

**Abstract.** The past four decades have seen a considerable amount of research on the study of magnetospheric substorms, and over most of these years the expansive phase of the substorm has been associated with the development of a three dimensional current system that has been termed the substorm current wedge. This current system has been thought to be a consequence of the short-circuiting of crosstail current through the ionosphere, and is viewed as a distinctive current system operating independently from the directly driven current with which it co-exists. The purpose of this paper is to show that the substorm current wedge should be viewed as an equivalent current system rather than a real current system. It will be shown that the magnetic perturbation pattern associated with the current wedge can be modeled as purely a perturbation of the directly driven current system in the midnight sector.

**Keywords.** Magnetospheric physics (Auroral phenomena; Current systems; Magnetotail; Storms and substorms)

## 1 Introduction

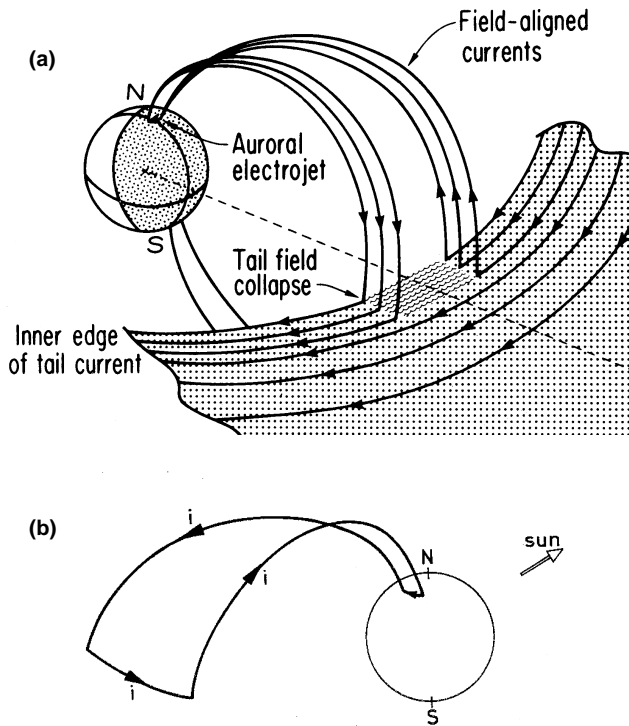
As far back as the beginning of the 20th century, the electric current system associated with the onset and early development of the substorm expansive phase (EP) was thought to be three dimensional, involving a downward field-aligned current (FAC) in the postmidnight sector connected to an upward FAC in the evening sector by an electrojet current flowing westward in the auroral ionosphere (cf. Birkeland, 1908).

The evidence for this presumption were the magnetic field perturbations observed at high latitudes during the strong auroral displays that were later termed substorms by Akasofu (1964). Until the end of the 1960's, most studies of the electric currents associated with substorms tended to refer to

the actual pattern as the equivalent current system of the disturbance. However, in the early 1970's both Akasofu and Chapman (1972) and McPherron et al. (1973) suggested that the current system was the result of the short-circuiting of crosstail current through the ionosphere (Fig. 1a). This led to a common view that the real current system for the substorm disturbance had been identified. It should, however, be recognized that there have been no measurements of electric current which have uniquely identified the FAC associated with the substorm EP. Any discussions of the three dimensional current system associated with that phenomenon are based on magnetometer measurements alone, and hence the substorm current wedge (as it is now called) is still only an equivalent current system, albeit three dimensional in character.

The three dimensional equivalent current system for the EP has been modeled by Bonnevier et al. (1970) (Fig. 1b) and compared with observations of the development of the expansive phase current system, that was later established by Kisabeth and Rostoker (1974) to occur in a series of stepwise intensifications, each one poleward of the preceding one. At that time, the equivalent nature of the model current system shown in Fig. 1b was confined to the closure current in the equatorial plane. That is, the eastward equivalent current flow in the equatorial plane could be viewed as a decrease in westward crosstail current flow at the onset of the substorm EP. It was this realization that led to the concept of the short-circuiting of some portion of the crosstail current to permit the substorm current wedge to form.

However, even in the early 1970's there was some question as to whether or not the short-circuit model was the only available explanation of the observations. Rostoker (1974) pointed out that it was possible to obtain the magnetic field perturbations associated with the substorm current wedge by simply stepping the upward FAC of the directly driven current system suddenly westward thereby extending the pre-existing westward electrojet further to the west (cf. Fig. 2). That is, the magnetic perturbation pattern associated with



**Fig. 1.** (a) Three-dimensional real current system proposed by McPherron et al. (1973) to explain the observed magnetic field perturbation pattern at the earth's surface during the development of the expansive phase of a magnetospheric substorm. In this framework, the expansive phase onset signals the short-circuiting of crosstail current through the auroral ionosphere (after McPherron et al., 1973). (b) Three dimensional equivalent current system used by Bonnevier et al. (1970) to represent the magnetic perturbation pattern associated with the EP of magnetospheric substorms. In this representation, dawnward equivalent current flow in the magnetic equatorial plane could be considered to represent a decrease in duskward real current flow. In a similar vein, downward equivalent current flow could be considered to represent a decrease in pre-existing real upward current flow (after Bonnevier et al., 1970).

such a westward shift of the upward FAC, and the concomitant westward extension of the directly driven westward electrojet, reproduce precisely the magnetic perturbation pattern attributed to the current wedge, although the downward FAC at the eastern edge of the wedge actually represents a decrease in the upward FAC present just prior to EP onset. In the short term, while this kind of real current system behavior could explain the ground magnetic field perturbation pattern during substorm expansive phase, the concept was abandoned. This was because, if such a large current was already flowing in the magnetosphere and ionosphere prior to EP onset, its buildup would have been identifiable as a perturbation from the quiet time baseline. Since expansive phase perturbations often seem to develop in situations where there is only a very weak ( $<50$  nT) deviation of the disturbance field from the background quiet time field, it seemed unlikely that the proposal by Rostoker (1974) was consistent with the observations. However, in the early 1990's a set of observations

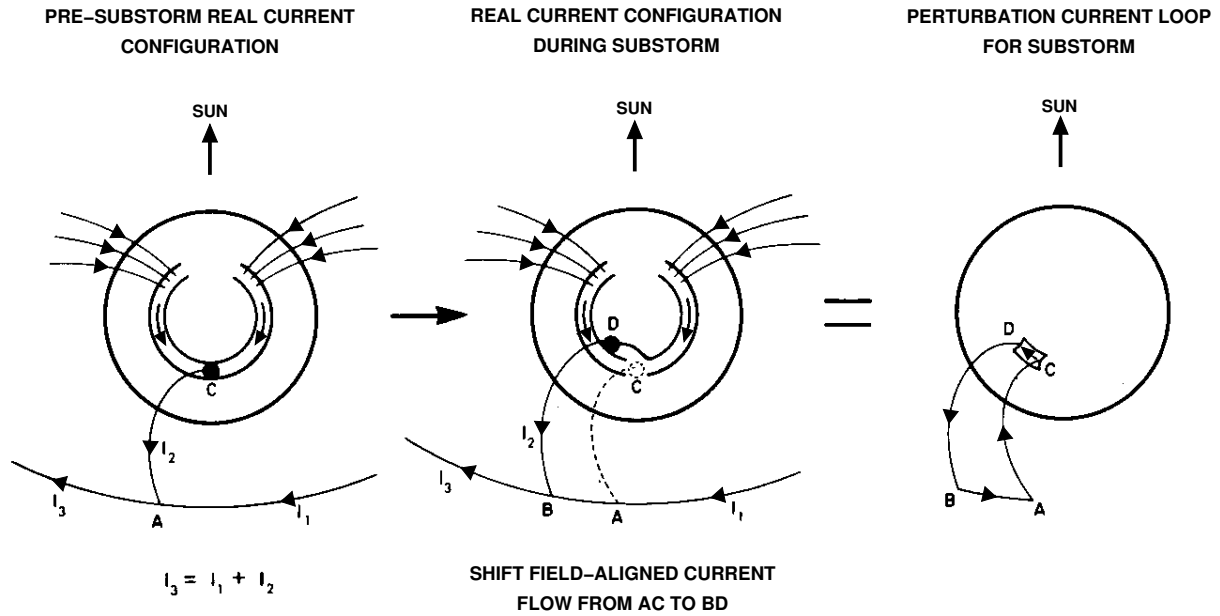
appeared which, together with proper modeling, caused a re-evaluation of the Rostoker (1974) suggestion. A preliminary version of this development was presented by Friedrich and Rostoker (1998). In the present paper we shall look at this idea in more detail and present additional modeling results that support the concept.

## 2 Motivation for the choice of the model current system

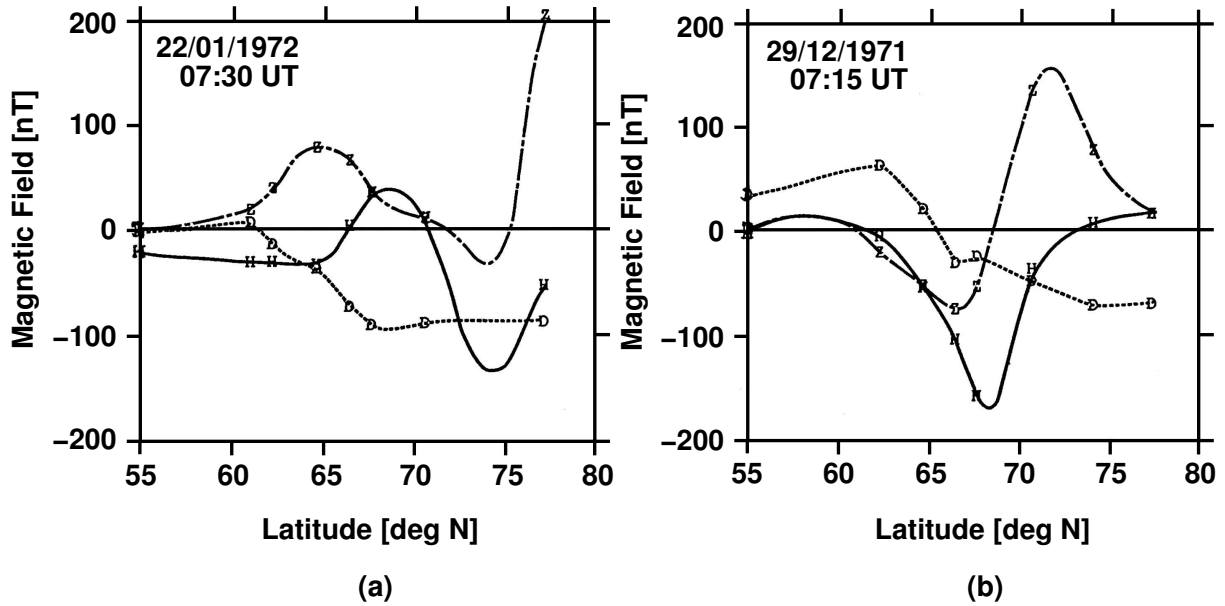
As mentioned in the Introduction, there appears to be more than one way in which to model the magnetic perturbation pattern associated with the substorm expansive phase. We begin our analysis of this contention by noting that the onset of a substorm EP is presaged by a period of enhanced energy input into the magnetosphere. The energy that enters from the solar wind may be thought of as being partitioned, with some of that energy being directly dissipated in the ionosphere and the rest being stored in the plasma and magnetic field of the magnetotail. How much is directly dissipated and how much is stored differs from event to event, depending on many preconditions, not the least of which is the conductivity of the ionosphere. The important point to note here is that, by the time of EP onset, there are already significant directly driven currents flowing in the magnetosphere-ionosphere system. These directly driven currents feature the two large-scale eastward and westward electrojets crossing the dusk and dawn meridians respectively, and the anti-parallel FAC sheets that flow into and out of the ionosphere in the electrojet regions together with their predominantly north-south ionospheric closure currents. Approaching the midnight meridian, the electrojet current gradually bleeds up the field lines leading to net upward current flow across the midnight sector from dusk to dawn. This net upward FAC is seen on the ground as a level shift in the east-west component of the perturbation magnetic field (cf. Hughes and Rostoker, 1977, 1979).

The Birkeland current system (involving azimuthally distributed anti-parallel current sheets connected by nearly meridional ionospheric closure currents) is essentially solenoidal in character. That is, ideally one does not detect the perturbation magnetic field associated with these currents unless one lies within the volume of space occupied by the FAC. Of course, these current configurations are not infinitely long solenoids and hence some magnetic flux leaks out the ends and can be detected on the ground. It is this leakage that provided the motivation to further explore the suggestion of Rostoker (1974) that a perturbation in the directly driven system could explain the magnetic perturbations attributed to the substorm current wedge as a real current system resulting from short-circuiting of the crosstail current.

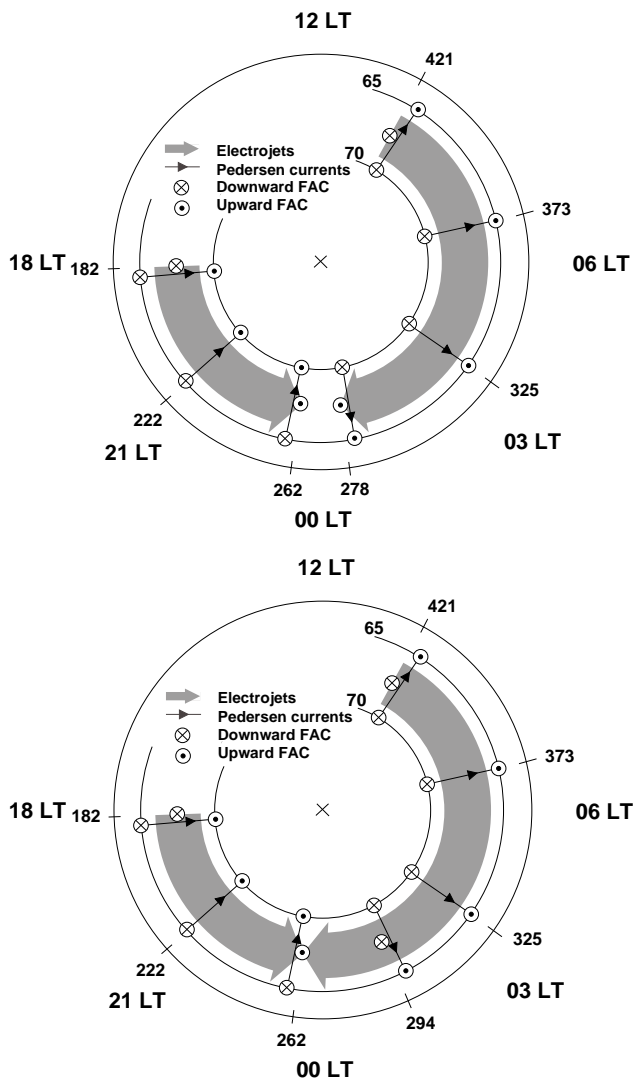
Figure 3a shows a latitude profile taken around midnight along a meridian line through the Alberta sector as monitored by a north-south line of magnetometers. Rather than featuring a bipolar signature in the Z-component across the oval, one sees a unipolar (positive) Z-component perturbation that is not typical of any electrojet signature. Figure 3b shows a



**Fig. 2.** Schematic showing the development of a substorm in terms of the change of current flow at substorm expansive phase onset. The unbalanced portion of the field-aligned currents is shown in this diagram together with the ionospheric eastward and westward directly driven electrojets. The magnetic perturbation pattern of the expansive phase can be explained by a westward jump of a pre-existing region of upward FAC flow and the associated westward expansion of the westward electrojet (after Rostoker, 1974).



**Fig. 3.** (a) Latitude profile taken on the Alberta line ( $\sim 115^\circ$  geographic) close to magnetic midnight. The three components of the perturbation magnetic field (H, D, Z) are in local magnetic coordinates. The positive Z-component across the average auroral oval latitude range is not a typical electrojet profile (which is bipolar in character) and is thought to be due to the combined edge effects of the meridional currents that are the ionospheric closure currents for the Birkeland current sheets in the morning and evening sectors. The large perturbation at the highest latitude is likely associated with the poleward border of the auroral oval, where poleward border intensifications take place. (b) Same as Fig. 3a except for a time close to midnight in which the latitude profile is representative of the morning sector westward electrojet. The asymmetric Z-component profile (large positive in the poleward region and small negative in the equatorward region) can be thought of as the combined effect of the electrojet itself (for which the positive and negative Z-component extrema should be approximately equal) and the edge effect of the morning side meridional closure current for the Birkeland currents.



**Fig. 4.** (a) Schematic of the directly driven currents in the auroral oval between dawn and dusk, used in the model to describe the growth phase of a substorm just before the EP onset. (b) Same as Fig. 4a but for conditions just after substorm EP onset.

latitude profile of the morning sector westward electrojet at a time close to midnight. A clear asymmetry is seen in the  $Z$ -component profile, the perturbation being more positive than negative across the electrojet region. This profile can be understood as the combined effect of the  $Z$ -component perturbation due to the electrojet (i.e. equal positive and negative maxima) and the  $Z$ -component of the leakage of magnetic flux out of the solenoidal Birkeland current configurations associated with the eastward and westward electrojet regions (i.e. a unipolar positive perturbation). This type of observation provided the motivation to view the magnetic perturbation pattern associated with substorm EP onset as a redistribution of the FAC in the midnight sector in such a way as to open the solenoid and allow its associated magnetic flux to be detected on the ground.

It should be emphasized that the profiles shown in Fig. 3 are taken at approximately the same local time, and yet they are quite different in structure. This is because the line separating the dusk cell from the dawn cell of the directly driven electrojet region is not always precisely located at local magnetic midnight. Sometimes the dividing line is slightly into the morning sector and sometimes it is slightly into the evening sector, although there is a tendency for it to be displaced towards the evening side at times of higher activity.

In the following section we shall show model results demonstrating that this kind of approach can adequately explain the observations, and in the Discussion section we shall provide a physically reasonable scenario in which the substorm EP onset is a natural consequence of the perturbation of the directly driven system of currents.

### 3 Model results

On the large scale, our model involves the two directly driven electrojets (eastward across dusk and westward across dawn) together with associated north-south current flow consistent with the known electric field and conductivity distribution in the auroral oval. The north-south currents connect to the Birkeland current sheets while the east-west current sheets are fed by downward field-aligned current in the dayside auroral region with the current gradually bleeding up the field lines in the nightside auroral region. The modeling was carried out in a dipole field geometry using the Biot-Savart law from programs described in Kisabeth and Rostoker (1977).

Figures 4a and b show the pre-onset and post onset configuration of electrojets and Birkeland currents associated with the directly driven system. The magnitudes of the currents in each local time sector are shown in Table 1 for the pre-onset configuration and in Table 2 for the post-onset configuration. From these Tables, it can be seen that we model the three dimensional current configuration by breaking the system down into cells of current in the ionosphere extending latitudinally between  $65^\circ$  and  $70^\circ$  and longitudinally over  $8^\circ$ . Within each cell, the ionospheric current is broken into its north-south and east-west components, with the east-west component being viewed as primarily Hall current and the north-south component primarily as Pedersen current. (An exception to this is found in the region of the substorm westward electrojet that marks the expansive phase onset, which we view as a Cowling current.) The north-south current in each cell is connected by field-aligned current sheets at the northern and southern edges of the cell to radial closure current in the equatorial plane, thus forming a current loop. The location of the closure current is determined by two L shells that mark the edges of the field-aligned current sheets, assuming dipole field geometry. The east-west current in each cell is connected by field-aligned current sheets at the eastern and western edges of the cell to azimuthal closure current, again forming a current loop. Thus, for each cell the condition that  $\nabla \cdot \vec{J} = 0$  is satisfied.

**Table 1.** Current magnitudes for Pre-onset model. \*

	Longitude (° East)	Hall Current (MA)	Pedersen Current (MA)	
Evening Sector	182–190	0.045	0.022	
	190–198	0.045	0.022	
(Eastward Hall/ Northward Pedersen current)	198–206	0.045	0.022	
	206–214	0.045	0.022	
	214–222	0.040	0.020	
	222–230	0.035	0.018	
	230–238	0.030	0.015	
	238–246	0.025	0.013	
	246–254	0.020	0.010	
	254–262	0.015	0.008	
	Morning Sector	278–286	0.056	0.030
		286–294	0.065	0.038
(Westward Hall/ Southward Pedersen current)	294–302	0.075	0.038	
	302–310	0.086	0.044	
	310–318	0.095	0.050	
	318–325	0.104	0.057	
	325–333	0.113	0.063	
	333–341	0.123	0.069	
	341–349	0.132	0.075	
	349–357	0.132	0.075	
	357–365	0.132	0.075	
	365–373	0.132	0.075	
373–381	0.132	0.075		
381–389	0.132	0.075		
389–397	0.123	0.069		
397–405	0.114	0.063		
405–413	0.105	0.057		
413–421	0.096	0.050		

\* All currents flow between latitudes 65° and 70° N latitude.

Considering the electrojet configuration as a whole, the east-west (Hall) current was allowed to bleed up the magnetic field lines as the electrojet current approached midnight, and the north-south Pedersen closure currents for the Birkeland current system decreased in intensity accordingly from the dusk and dawn meridians towards midnight. For the 15° of longitude centered on midnight, no electrojet or meridional current flows in the pre-onset model. The relative magnitudes of the Hall and Pedersen currents were chosen based on reported ionospheric conductivity ratios (cf. Wallis and Budzinski, 1981). During the growth phase, the  $\Sigma_H/\Sigma_P$  ratio is typically  $\sim 1$ – $2$  and the ratio of the morning sector to evening sector currents is  $\sim 3$ : $1$ .

The westward electrojet in the morning sector extends eastward past dawn into the pre-noon quadrant. The electrojet is fed by downward current most of which, to minimize the number of cells required in the model, flows into the ionosphere over a restricted range of longitudes (389° to 421°). (In reality, this downward current would be more spread out across the pre-noon quadrant.) The eastward electrojet, stretches westward to 10° past the dusk meridian and all the field-aligned current feeding this electrojet is concentrated at 190°.

The perturbation patterns are shown in the latitude and longitude profile formats so that they can be compared with observations that are often shown in similar formats. We begin by showing model pre-onset profiles, which the reader

**Table 2.** Current magnitudes for Post-onset model. \*

	Longitude (°East)	Hall Current (MA)	Pedersen Current (MA)	
Evening Sector	182–190	0.056	0.028	
	190–198	0.056	0.028	
(Eastward Hall/ Northward Pedersen current)	198–206	0.056	0.028	
	206–214	0.056	0.028	
	214–222	0.050	0.025	
	222–230	0.044	0.023	
	230–238	0.038	0.019	
	238–246	0.031	0.016	
	246–254	0.025	0.013	
	254–262	0.019	0.010	
	Morning Sector	262–270	0.162	0.000
		270–278	0.162	0.000
(Westward Hall/ Southward Pedersen current)	278–286	0.162	0.000	
	286–294	0.162	0.000	
	294–302	0.135	0.038	
	302–310	0.140	0.044	
	310–318	0.145	0.050	
	318–325	0.150	0.057	
	325–333	0.155	0.063	
	333–341	0.160	0.069	
	341–349	0.165	0.075	
	349–357	0.165	0.075	
	357–365	0.165	0.075	
	365–373	0.165	0.075	
	373–381	0.165	0.075	
381–389	0.165	0.075		
389–397	0.154	0.069		
397–405	0.143	0.063		
405–413	0.131	0.057		
413–421	0.120	0.050		

\* All currents flow between latitudes 65° and 70° N latitude.

should consider as typical of what might be observed at auroral oval latitudes. Figures 5a, b and c show latitude profiles across the auroral oval at evening sector, midnight sector and morning sector locations respectively. One sees peak magnetic perturbations due to the eastward and westward electrojets in the evening and morning sectors of  $\sim 30$  nT and  $\sim -100$  nT respectively. These perturbations are mostly caused by the Hall current electrojets. In the midnight sector, the latitude profile is not typical of that associated with either an eastward or a westward electrojet and the peak magnetic perturbations in any component are  $< 15$  nT. The horizon-

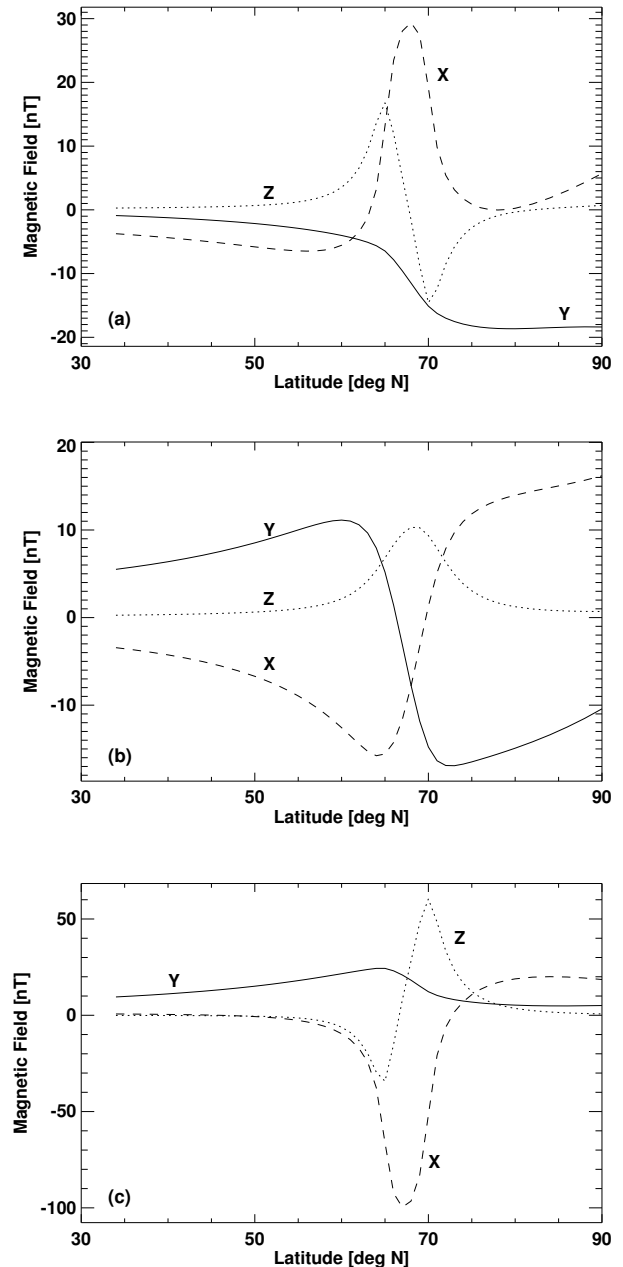
tal component perturbations are mostly due to field-aligned currents and the unipolar positive Z-component perturbation is due to the edge effects of the north-south Pedersen closure currents associated with the evening and morning sector Birkeland current systems. It is important to emphasize that perturbations of this level in the midnight sector would likely not be readily noticed as they would develop slowly over the time of the growth phase and are of the order of what one might expect from variations due to the quiet time current system.

We associate the onset of expansive phase activity with a sudden brightening of an auroral arc across the midnight meridian, signaling the onset of westward electrojet flow across that sector. We argue that this new westward current flow is an extension of the morning sector westward electrojet that is part of the directly driven system. We further argue that a longitudinally confined section of upward FAC that was at the western edge of the pre-onset westward electrojet shifts westward while the downward FAC to which it connects does not shift. The effect of this reconfiguration of the Birkeland currents is to “open the solenoid” permitting the ground observer to detect the magnetic flux from the westward portion of the Birkeland currents. As the downward FAC in the poleward portion of the morning sector Birkeland currents does not move in longitude, this effectively presents to the observer an up-down current pair connected by westward ionospheric current, i.e. a current wedge configuration. The EP westward electrojet, from its eastern edge of  $294^\circ$  to its western edge at  $262^\circ$ , is considered to be a Cowling current: that is, the sum of the enhanced westward electrojet across the dawn sector and the original (southward) Pedersen current from  $278^\circ$  to  $294^\circ$ . Finally, we note that consistent with a  $\sim 15^\circ$  shift of the westward electrojet into the evening sector, the eastward electrojet retracts by the same number of degrees. The final EP westward electrojet therefore has a longitudinal extent of  $32^\circ$ .

Along with the change of the spatial configuration of the Birkeland currents, we also note that conductivity changes occur that affect mainly the morning sector. Based on the contention of Kamide and Vickrey (1983) that the westward electrojet is conductivity dominated in the midnight and morning sector, and the observations by Brekke et al. (1974) and Opgenoorth et al. (1994) that the ratio  $\Sigma_H/\Sigma_P$  increases to  $>2:1$  during active times in those sectors, we effectively increased  $\Sigma_H/\Sigma_P$  in the morning hours while keeping the evening sector conductivity ratio almost unchanged.

This increased morning sector conductivity translates into increased directly driven westward electrojet current flowing towards midnight in the morning sector. Therefore our change from pre-onset to post-onset conditions involves both the sudden shift of the section of upward FAC between  $278^\circ$  and  $293^\circ$  to the western edge of the post-onset westward electrojet, as well as an increase in the strength of the directly driven westward electrojet current.

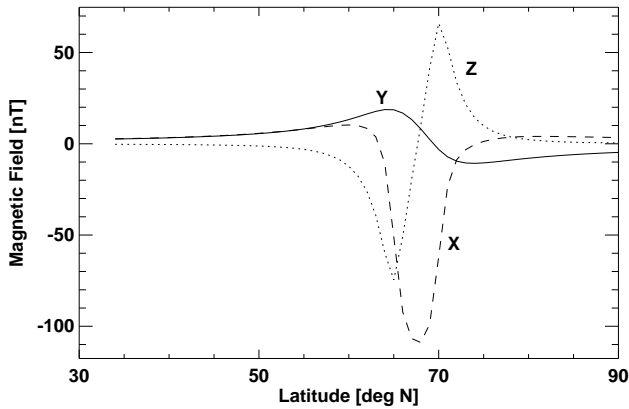
Figure 4b shows the new configuration of the modeled current system while Fig. 6 shows the change in the magnetic field on the ground as a latitude profile along the midnight meridian. The magnitudes of the Hall and Pedersen currents in each sector immediately after EP onset are shown in Table 2. If one compares the pre-onset and post-onset morning sector westward electrojets, it is immediately apparent that the latitude profile of the disturbance along the  $270^\circ$  meridian (i.e. the central meridian of the newly established westward current) has the same form, and features a peak magnitude of the X-component disturbance, that one sees as repeatable characteristics of the latitude profiles associated with observations of an EP onset in the region of that onset. The profiles



**Fig. 5.** Model latitude profiles for pre-substorm onset conditions in the evening (a), midnight (b) and morning (c) sectors of the auroral oval.

of the evening/morning sector eastward/westward electrojet (not shown) are also typical of those detected in those sectors immediately after EP onset.

Figure 7 shows a cartoon longitude profile of the components of the magnetic field disturbance typically detected at low latitudes at the time of EP onset (Panel a) and the model results for that same region on the earth’s surface (Panel b). It is evident that the modeled and observed disturbances closely resemble one another insofar as the horizontal components of the magnetic perturbations are concerned. The magnitudes of the magnetic perturbations at low latitude are only of the



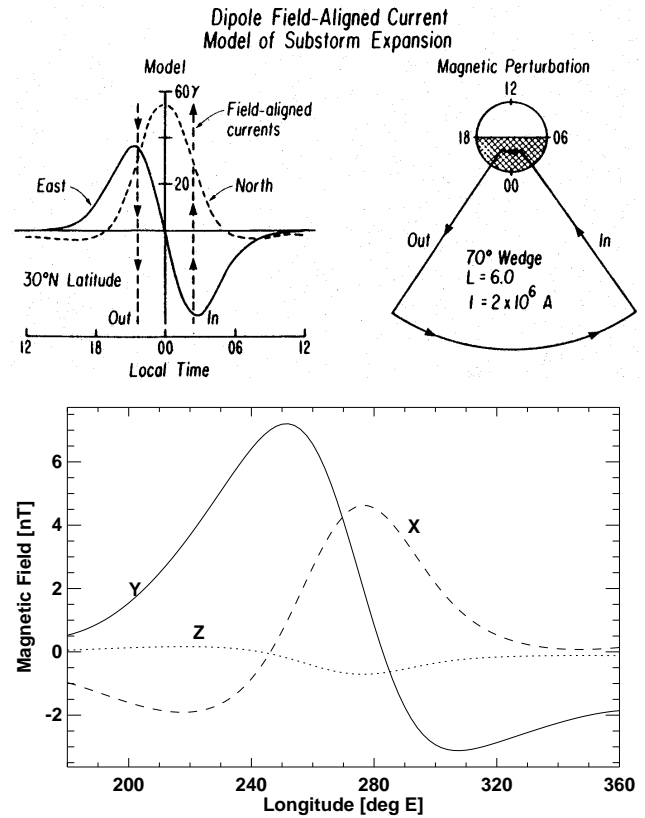
**Fig. 6.** Model latitude profile across the midnight sector auroral oval after substorm expansive phase. This is precisely the type of profile seen shortly after EP onset (cf. Bonnevier et al., 1970; Kisabeth and Rostoker, 1974).

order of a few nT, whereas observations feature some events that are an order of magnitude larger than what our model results provide. It is important to note, however, that our model only deals with the initial onset of the substorm EP. Normal expansive phases develop through discrete intensifications of auroral arc systems, each poleward of the previous one and each providing a component of disturbance of approximately the same size (Kisabeth and Rostoker, 1974). Typically, two to three such enhancements are found in a major substorm expansive phase before the activity reaches the poleward edge of the pre-existing diffuse aurora that approximately marks the open-closed field line boundary (Blanchard et al., 1994; Friedrich et al., 2001). Thus, our model allows for the normal range of magnitudes of EP disturbances observed at low latitudes.

In summary, we believe that the observed magnetic perturbations detected on the ground (i.e. Panel a) in association with EP onset can be produced simply by a perturbation in the pre-existing directly driven current system; that is, a short-circuiting of crosstail current may not be required at the time of the substorm EP onset. This does not prove that short-circuiting does not take place; it only demonstrates that the observations can be explained in a different fashion. In the Discussion section that follows, we shall demonstrate that our explanation is well founded in terms of conforming to the observations and to our theoretical understanding of what happens in the magnetosphere-ionosphere system when a substorm EP onset takes place.

**4 Discussion**

We have shown in the preceding section that the magnetic perturbation pattern at the earth’s surface associated with the onset of a substorm expansive phase can be explained by the sudden extension of a longitudinally limited (~ one time zone) portion of the directly driven westward electro-



**Fig. 7. (a)** Model longitude profile across the low latitude region observed during substorm expansive phase conditions (after McPherron et al., 1973). Note that 1 gamma ( $\gamma$ )=1 nT. The model profile for the line current system shown on the right side of the figure satisfies observations of the polarities of magnetic bays at middle latitudes documented by Rostoker (1966) among others. **(b)** Model longitude profile of the magnetic perturbation at latitude 45° for our change in the directly driven current system. Magnetic midnight is at 270° E. The high degree of similarity between these two profiles of the horizontal components of the magnetic field perturbation in Figure 7 strongly supports our model of the process in terms of representing a plausible explanation of the development of the substorm current system.

jet. This involves a sudden westward shift of an azimuthally limited section of upward region 2 current, while the region 1 downward current poleward of that shifted section remains in place. The consequence of this reconfiguration of region 2 current is that the solenoidal configuration of the Birkeland currents and their associated ionospheric and magnetospheric closure currents is broken, allowing the perturbation magnetic field to be detected at lower latitudes. We now discuss the observational and theoretical evidence supporting this new model of the current wedge.

We begin by looking at the observational evidence that has been used to support the concept of the short-circuiting of crosstail current through the ionosphere as a longitudinally limited westward electrojet. The signature of such a rerouting of current would be consistent with dipolarization of the magnetic field around geostationary orbit and beyond. While



this could be explained by a reduction in crosstail current, electric circuit theory would suggest that a portion of the current flowing from the dawn flank to the dusk flank finds a path through the auroral ionosphere to be less resistive than continuing to flow across the tail. However, Vasyliunas [private communication] has pointed out that if one can compare current flows in the ionosphere and magnetosphere (which is not at all certain), the conductivity encountered by the crosstail current is of the order  $\sim 300$  mho, far higher than could be expected for the region of the ionosphere through which the westward electrojet would be expected to flow. This casts some doubt on the basic premise for the existence of the current wedge as a real current system, in that there is no obvious reason why the crosstail current should divert to flow through the ionosphere. To our knowledge, all evidence for the short-circuit concept of the current wedge stems from magnetic field measurements at a limited number of observation points either on the ground or in space. These measurements cannot distinguish between the two possible ways of explaining the observations, and thus do not contribute to a solution of the uniqueness problem.

In contrast, there is a significant body of observations which support the concept espoused in this paper; namely, that the pre-existing westward electrojet is extended further into the evening sector at the time of onset of the substorm expansive phase. The primary question now is: What physical mechanism is responsible for such an extension? We shall examine both what happens in the region of ionosphere in which the auroral breakup takes place and the region of space in and around the magnetic equatorial plane/plane of the neutral sheet to try to provide an answer to this question.

However, before addressing the question of the physical mechanism, it is useful to describe the framework in which we attempt to understand the substorm process. We subscribe to the view that, during the growth phase of the substorm, energy is stored in the near-earth magnetotail as evidenced by the development of the weak field region in the vicinity of magnetic midnight. We accept the evidence of Samson et al. (1992) that onset occurs close to the inner edge of the midnight sector plasma sheet, where region 2 current is expected to flow. We associate onset with the breakdown of the shielding electric field that has built up during the course of the growth phase (cf. Rostoker, 1996) which involves an increase in region 2 field-aligned current. In response to the increase in region 2 field-aligned currents, the other currents that link to region 2 currents (i.e. ionospheric, region 1 and magnetospheric closure current) also increase.

This increase is generally brief, and is followed immediately by a general decrease in the strength of the directly driven currents during the substorm recovery phase. We further consider that the initial stage of the expansive phase involves the disappearance of the weak field region through the progressively tailward collapse of the crosstail current responsible for the weak field region. This collapse is in accord with the observations of Jacquy et al. (1991) and Ohtani et al. (1992) insofar as evidence for downtail evolution is concerned. When the collapse of the tail current reaches the neu-

tral line, expected to be somewhere in the vicinity of  $\sim 20$ – $40 R_E$  downtail, enhanced reconnection commences and tail lobe magnetic field energy is accessed to power the ensuing substorm activity. This ensuing activity takes the form of poleward border activations on the high latitude branch of the double oval (cf. Rostoker, 2002). The time from onset to the activation of the poleward border of the oval can be as small as a few minutes (cf. Friedrich et al., 2001), the delay being related to the time of propagation of fast mode Alfvén waves from the onset region near the earth to the locale of the neutral line. It should be noted that the Harang discontinuity in this model follows the scheme discussed in Rostoker (1996) in which the high latitude convection reversal and the region of electric field disturbance at the equatorward edge of the midnight sector auroral oval are not connected. The evening sector high latitude convection reversal is the site of the poleward border activations mentioned above that accompany the latter stages of the substorm after the closed field line region has become fully involved and enhanced tail lobe reconnection has commenced.

In this whole process, it may be said that the tail is “unloading” energy, which is dissipated in the auroral ionosphere or stored in the energetic trapped particle population that is left after the substorm expansive phase has concluded. We should like to emphasize that this “unloading” has two components. The first component involves processes that take place on the closed field lines of the plasma sheet between the onset region and the neutral line. The second component involves the open magnetic field lines of the tail lobe. As regards the energetics of the substorm process, our model pertains only to the first component and the large portion of any substorm may well be associated with the dissipation of the tail lobe energy after enhanced reconnection has commenced. Additionally, our model calculations in this paper pertain only to the onset of the substorm and hence the first component of “unloading” as described above. There may be two or more steps in the development of the expansive phase before enhanced tail lobe reconnection begins (cf. Kisabeth and Rostoker, 1974), but each of these steps would be modeled in the same manner as the initial onset. In keeping with the studies of Kisabeth and Rostoker (1974) and Wiens and Rostoker (1975), each step would involve the formation of a new arc poleward of the previous one and with its western edge westward of the previous one.

#### 4.1 The magnetic equatorial plane/neutral sheet

To begin, the upward FAC present prior to expansive phase onset comprises region 1 in the afternoon and evening sectors and region 2 in the morning and post-midnight sectors. The physics behind the production of these two regions of upward FAC is, in all probability, quite different and the regions of space in which they are generated are very different in character. Following the substorm framework of Rostoker (1996), we consider that region 2 upward FAC in the morning sector occupies a significant region of the morning sector plasma sheet outward from its inner edge. We

consider region 1 upward FAC in the afternoon and evening sector to map to the interface region between the low latitude boundary layer and central plasma sheet (CPS). Since the substorm onset is clearly a region 2 phenomenon in terms of the locale where it takes place (i.e., near the inner edge of the plasma sheet), there is no reason to believe that region 1 downward FAC in the morning sector (also mapping to the interface between the CPS) should change position during the early stages of the EP onset.

Secondly, since region 2 currents are most likely generated through pressure gradients in the near-earth plasma sheet, we will argue that the changes associated with substorm EP onset are caused by changes in the pressure gradient and, in particular, in the azimuthal component of the pressure gradient. Liu and Rostoker (1991) have shown that there is a significant azimuthal component of the pressure gradient from the cool morning sector plasma sheet to the warmer center, and that the direction of this gradient is consistent with upward FAC flow. The gradient itself is a consequence of the gradient/curvature drift of plasma sheet ions, the warmer ones moving more towards the evening side compared to their cooler counterparts. The upward FAC current in the presence of this azimuthal pressure gradient can be understood in the following way. Assuming

$$\nabla \cdot (\vec{J}_{par} + \vec{J}_{\perp}) = 0$$

it can be shown (cf. Hasagawa and Sato, 1979) that the FAC density associated with the pressure gradient is given by

$$J_{par} = 2B \int \frac{\vec{J}_{\perp} \cdot \nabla B}{B^2} dl$$

Noting that, for the assumption of isotropic pressure

$$\vec{J}_{\perp} = \frac{\vec{B} \times \nabla p}{B^2}$$

together with the vector identity

$$(\vec{B} \times \nabla p) \cdot \nabla B = -(\vec{B} \times \nabla B) \cdot \nabla p$$

and the expression for gradient/curvature drift

$$\vec{V}_{cg} = \frac{(\varepsilon_{\perp} + 2\varepsilon_{par}) \vec{B} \times \nabla B}{qB^3}$$

it can be shown that

$$J_{par} = -2qB \int \frac{\vec{V}_{cg} \cdot \nabla p}{f(\varepsilon) B} dl$$

where the total thermal energy of the current carrying particles is  $\varepsilon = \varepsilon_{perp} + \varepsilon_{par}$  and  $f(\varepsilon)$  is a function of the thermal energy.

For a pressure gradient directed duskward and the direction of the gradient curvature drift of the dominant current carrier (i.e. ions) duskward,  $J_{par} < 0$ , indicating flow of current out of the ionosphere into the equatorial plane. From this, it is clear that the character of the azimuthal component

of the pressure gradient in the inner plasma sheet will determine the distribution of upward FAC across the midnight sector auroral oval. We argue that the development of the azimuthally localized region of weak magnetic field across the midnight sector known to evolve during the substorm growth phase alters the azimuthal distribution of pressure, with a strengthening duskward component of the pressure gradient being a property of the eastern part of the weak field region. When the pressure gradient becomes sufficiently large,  $J_{par}$  becomes large enough to activate the field-aligned acceleration mechanism that produces the energetic electrons responsible for the brightening of the breakup arc.

Finally, we note that several mechanisms have been proposed in the literature to explain the development of the breakup arc and its subsequent distortion into a surge-like auroral form. Our model described above does not depend on which of these is valid; however, we tend to favor the hybrid shear flow ballooning mode discussed by Voronkov et al. (1997, 2000) which permits development of an auroral vortex structure from the discrete breakup arc. In this theory, the collapse of the weak field region is the consequence of a ballooning instability that becomes non-linearly unstable in the presence of a radial pressure gradient and velocity shear.

#### 4.2 The auroral ionosphere

Where FAC current flows into the equatorial plane, the current flow at the foot of the field line in the ionosphere must be outward (i.e.  $J_{par} > 0$  in a coordinate system with downward being positive). Assuming an infinitesimally thin ionospheric current sheet and  $\nabla \cdot \vec{V} = 0$ , the field-aligned current can be expressed in terms of the ionospheric electric field  $\vec{E}$  and height-integrated Hall and Pedersen conductivities by

$$J_{par} = \Sigma_p \nabla \cdot \vec{E} + \nabla \Sigma_p \cdot \vec{E} + \nabla \Sigma_H \cdot \frac{(\vec{B} \times \vec{E})}{B}$$

In terms of the divergence of ionospheric electrojet current flow into and out of the ionosphere, we concentrate on the third term of this expression assuming the electric field to be primarily meridional typical of the morning hours. We would expect the westward component of the electric field to be rather weak approaching the end of the growth phase due to the growth of the shielding electric field.

From the ionospheric conductivity distribution defined by Wallis and Budzinski (1981), we find that  $\nabla \Sigma_H$  points eastward in the post-midnight quadrant so that the third term in the expression for  $J_{par}$  above is negative, indicating upward field-aligned current flow. This is what one would expect intuitively, since diminishing Hall conductivity as one moves westward would be consistent with upward field-aligned current as the westward electrojet current is diverted to flow into the magnetosphere.

We now look at the changes one would expect in the magnetosphere as the substorm growth phase develops. The clearest development is that of a weak (magnetic) field region in the midnight sector. As this weak field region

develops rather slowly (on the order of tens of minutes), the assumption of approximate magnetohydrostatic equilibrium demands that the decrease in magnetic field energy density be balanced either by an increase in thermal pressure or an increase in magnetic tension, or both of these parameters. If the increase in thermal pressure is the dominant response to the weakened magnetic field, this would enhance the azimuthal thermal pressure gradient in the region just to the east of the midnight meridian. This, in turn, would be consistent with enhanced upward field-aligned current flow in this region. The collapse of the weak field region at substorm EP onset involves a release of stored energy in the midnight sector of the near-Earth plasma sheet. Some of this energy goes to the formation of the substorm expansive phase auroral arcs (i.e. degradation to heat through collisions of precipitating particles with ionospheric constituents) and part goes to providing energy consistent with Ohmic heating associated with the new current flow in the ionosphere. The evolution of the breakup arc system involves surge formation at its western end which is consistent with a sharp conductivity discontinuity developing that leads to upward FAC in a region where there was no significant upward FAC flow prior to EP onset. In contrast, in the eastern part of the weak field region the azimuthal dawn-to-dusk pressure gradient decreases suddenly, consistent with a reduction in upward FAC. It is this reduction in region 2 upward FAC in the dawn side part of the weak field region that permits the ground observer to “see” the magnetic effects of the downward region 1 FAC region poleward of the region 2 currents. We argue that this is the large scale physics behind the development of the initial arc systems in the substorm expansive phase up to the time of activation of the poleward border of the oval (i.e. the development of poleward border intensifications).

## 5 Conclusions

In this paper, we argue that the equivalent current system that is presently called the substorm current wedge is exactly that – an equivalent current system, rather than a real current system. While it may be true that the wedge consists of newly developed upward and downward FAC connected through the ionosphere by a westward electrojet, that is not the only possible interpretation of the observations. We have presented an alternative explanation that does not call on the creation of a new current system at EP onset, but only on a perturbation to the pre-existing currents (i.e. the directly driven system). Our model study shows that both the spatial character and level of strength of magnetic perturbations associated with EP onset can be explained by spatial changes in the locale of upward FAC, and we have discussed how that might take place as a consequence of changes in the thermal pressure distribution associated with the growth phase weak field region. We believe that this interpretation of the data should be seriously considered as we try to further understand the development of magnetospheric substorms.

In conclusion, we should like to emphasize that our model applies only to the perturbations associated with the onset and initial development of the expansive phase. After onset, a full expansive phase would involve a series of auroral activations, each one poleward of the previous ones (cf. Kisabeth and Rostoker, 1974). These all occur on closed field lines, mapping progressively further back in the tail. At some point, the complete closed field line region has become engaged, after which reconnection of lobe field lines provides the energy for the balance of the substorm. Disturbances due to lobe field line reconnection take the form of poleward border intensifications (cf. Lyons et al., 1999) and are not part of the substorm current wedge formation. The model presented in this paper does not deal with the poleward border intensifications, and is relevant only to pseudo-breakups and to the onset and subsequent activations that occur in closed field line geometry during the initial stages of an expansive phase.

*Acknowledgements.* The authors are grateful to M. Jankowska for her assistance in running the model and in preparation of some of the figures. This research was supported by the Natural Sciences and Engineering Research Council of Canada.

Topical Editor T. Pulkkinen thanks V. Papitashvili and another referee for their help in evaluating this paper.

## References

- Akasofu, S.-I.: The development of the auroral substorm, *Planet. Space Sci.*, 12, 273–282, 1964.
- Akasofu, S.-I. and Chapman, S.: *Solar Terrestrial Physics*, Oxford Univ. Press, London, 1972.
- Birkeland, K.: *The Norwegian Aurora Polaris Expedition 1902–1903*, 1, Sect. 1, Aschhoug and Co., Christiania, 1908.
- Blanchard, G. T., Lyons, L. R., Samson, J. C., and Rich, F. J.: Locating the polar cap boundary from observations of 6300 Å auroral emission, *J. Geophys. Res.*, 100, 7855–7862, 1994.
- Bonnevier, B., Boström, R., and Rostoker, G.: A three-dimensional model current system for polar magnetic substorms, *J. Geophys. Res.*, 75, 107–122, 1970.
- Brekke, A., Doupnik, J. R., and Banks, P. M.: Incoherent scatter measurements of E region conductivities and currents in the auroral zone, *J. Geophys. Res.*, 79, 3773–3790, 1974.
- Friedrich, E. and Rostoker, G.: Reconfiguration of the directly driven currents during a substorm expansive phase: Implications for the substorm current wedge, in *SUBSTORMS-4*, (Ed.) Kokubun, S. and Kamide, Y., 83–86, Terra Sci. Publ. Co., Kluwer Acad. Publ., Tokyo, 1998.
- Friedrich, E., Samson, J. C., Voronkov, I., and Rostoker, G.: Dynamics of the substorm expansive phase, *J. Geophys. Res.*, 106, 13 145–13 163, 2001.
- Hasegawa, A. and Sato, T.: Generation of field aligned current during substorm, in *Dynamics of the Magnetosphere*, (Ed.) Akasofu, S.-I., 529–541, D. Reidel Publ. Co., Dordrecht-Holland, 1980.
- Hughes, T. J. and Rostoker, G.: Current flow in the magnetosphere and ionosphere during periods of moderate activity, *J. Geophys. Res.*, 82, 2271–2282, 1977.
- Hughes, T. J. and Rostoker, G.: A comprehensive model current system for high latitude magnetic activity 1. The steady state system, *Geophys. J. Roy. Astron. Soc.*, 58, 525–569, 1979.

- Jacquey, C., Sauvaud, J., A. and Dandouras, J.: Location and propagation of the magnetotail current disruption during substorm expansion: analysis and simulation of an ISEE multi-onset event, *Geophys. Res. Lett.*, 18, 389–392, 1991.
- Kamide, Y. and Vickrey, J. F.: Relative contribution of ionospheric conductivity and electric field to the auroral electrojets, *J. Geophys. Res.*, 88, 7989–7996, 1983.
- Kisabeth, J. L. and Rostoker, G.: The expansive phase of magnetospheric substorms 1. Development of the auroral electrojets and auroral arc configuration during a substorm, *J. Geophys. Res.*, 79, 972–984, 1974.
- Kisabeth, J. L. and Rostoker, G.: Modelling of three-dimensional current systems associated with magnetospheric substorms, *Geophys. J. Roy. Astron. Soc.*, 49, 655–683, 1977.
- Liu, W. W. and Rostoker, G.: Effects of dawn-dusk pressure asymmetry on convection in the central plasma sheet, *J. Geophys. Res.*, 96, 11 501–11 512, 1991.
- Lyons, L. R., Nagai, T., Blanchard, G. T., Samson, J. C., Yamamoto, T., Mukai, T., Nishida, A., and Kokobun, S.: Association between Geotail plasma flows and auroral poleward boundary intensifications observed by CANOPUS photometers, *J. Geophys. Res.*, 104, 4485–4500, 1999.
- McPherron, R. L., Russell, C. T., and Aubry, M. P.: Satellite studies of magnetospheric substorms on 15 August 1968, 9. Phenomenological model for substorms, *J. Geophys. Res.*, 78, 3131–3149, 1973.
- Ohtani, S., Kokubun, S., and Russell, C. T.: Radial expansion of the tail current disruption during substorms: a new approach to the substorm onset region, *J. Geophys. Res.*, 97, 3129–3136, 1992.
- Opgenoorth, H., Persson, M. A., Pulkkinen, T. I., and Pellinen, R. J.: Recovery phase of magnetospheric substorms and its association with morning-sector auroras, *J. Geophys. Res.*, 99, 4115–4129, 1994.
- Rostoker, G.: Midlatitude transition bays and their relation to the spatial movement of overhead current systems, *J. Geophys. Res.*, 71, 79–95, 1966.
- Rostoker, G.: Current flow in the magnetosphere during magnetospheric substorms, *J. Geophys. Res.*, 79, 1994–1998, 1974.
- Rostoker, G.: Phenomenology and physics of magnetospheric substorms, *J. Geophys. Res.*, 101, 12 955–12 973, 1996.
- Rostoker, G.: Identification of substorm expansive phase onsets, *J. Geophys. Res.*, 107 (11), doi:1029/2001JA003504, 2002.
- Samson, J. C., Lyons, L. R., Newell, P. T., Creutzberg, F., and Xu, B.: Proton aurora and substorm intensifications, *Geophys. Res. Lett.*, 19, 2167–2170, 1992.
- Voronkov, I., Rankin, R., Frycz, P., Tikhonchuk, V. T., and Samson, J. C.: Coupling of shear flow and pressure gradient instabilities, *J. Geophys. Res.*, 102, 9639–9650, 1997.
- Voronkov, I., Donovan, E. F., Jackel, B. J., and Samson, J. C.: Large-scale vortex dynamics in the evening and midnight auroral zones: Observations and simulations, *J. Geophys. Res.*, 105, 18 505–18 518, 2000.
- Wallis, D. D. and Budzinski, E. E.: Empirical models of height integrated conductivities, *J. Geophys. Res.*, 86, 125–137, 1981.
- Wiens, R. G. and Rostoker, G.: Characteristics of the development of the westward electrojet during the expansive phase of magnetospheric substorms, *J. Geophys. Res.*, 80, 2109–2128, 1975.