Magnetospheric Substorms—Definition and Signatures


INTRODUCTION

The study of episodes of energy dissipation in the earth’s ionosphere and magnetosphere developed historically through the measurements of perturbations in various parameters detected by ground-based equipment. In the early stages of this research, impulsive increases at auroral zone latitudes in the levels of auroral luminosity, magnetic fluctuations in the frequency range $f > 0.3$ mHz, electromagnetic noise in the VLF and ELF frequency range, and absorption of cosmic noise were isolated as being common features of the midnight sector. The pioneering work of Akasofu and Chapman [1961] led to the concept of a violent magnetospheric process that they termed the substorm and that featured all the perturbations measured by ground-based equipment described above.

It has now been a decade since Akasofu [1968] published a treatise linking high-latitude processes associated with auroral substorms in a framework that was called the magnetospheric substorm. The time sequence of events for each process was linked to that of the auroral substorm as defined by Akasofu [1964]. Since that time a vast amount of research has been carried out that has greatly enhanced our understanding of the substorm phenomenon. However, it has become clear over this decade of research that the substorm phenomenon is more complex than it was originally envisaged. The scattered arrays of monitoring equipment have provided a complicated combination of spatial and temporal variations that are just now beginning to be appreciated. Furthermore, different groups have used different signatures to define the occurrence of substorms. Unfortunately, this has resulted in substorm researchers talking about different phenomena but giving them the common label of substorm. In order to alleviate these problems, nine magnetospheric physicists active in the area of substorm research met in Victoria, British Columbia, Canada, for three days in August 1978 to attempt to reach a consensus on what the definition of a substorm should be, based on current understanding of the phenomenon, and to back this definition with a set of usable signatures that could be employed by space researchers working on substorm-related problems.

The purpose of this paper is to present the results of the Victoria workshop, which represents a consensus reached by all nine participants. We hope that the definition of the substorm presented herein will provide a common framework in which researchers can couch their results. The reader is referred to Akasofu [1968] and Akasofu [1977] for definitions of the various terms describing substorm phenomenology used in the following text.

THE EVENTS LEADING UP TO THE ONSET OF SUBSTORM ACTIVITY

A magnetospheric substorm is a transient process initiated on the nightside of the earth in which a significant amount of energy derived from the solar wind–magnetosphere interaction is deposited in the auroral ionosphere and in the magnetosphere. In order to address the question of the beginning of a substorm event, it is first necessary to ask whether or not one can find periods of time when substorm activity does not take place. This is not a trivial question, because the present substorm index $AE$ and other indices are not accurate enough to rule out the possibility of small localized substorms. Since the authors do not believe that there is any inherent difference between the physical processes responsible for small substorms and their larger-magnitude counterparts, the question of whether or not there are sustained periods when no substorm activity is in progress should be treated carefully. On the basis of the experience of the authors, it was felt that past studies [e.g., Kamide et al., 1977] strongly suggest that if the interplanetary magnetic field (IMF) has a significant northward component ($B_z = |B|$) for an interval of time greater than several substorm time constants, the magnetosphere becomes quiet and asymptotically approaches a state of lowest activity that we term the ground state. In this interval of time, magnetospheric convection continues at a low level as does energetic particle precipitation, but the probability of substorm occurrence approaches zero. The magnetotail continues to exist...
throughout such quiet intervals and will continue to exist when the magnetosphere is in the ground state.

As the IMF turns away from northward, the region of sunward convection in the auroral oval expands equatorward. In this time period, magnetometer data suggest that there can be an interval of enhanced convection prior to the first substorm expansion [e.g., Rostoker, 1974]. The actual length of this interval probably depends on how closely the magnetosphere approached the ground state. After this initial period a substorm may be triggered by processes internal or external to the magnetosphere. The total energy released in the solar wind—magnetosphere interaction depends both on the solar wind velocity and on the magnitude and direction of the IMF. Finally, it should be noted that not all substorms need be preceded by the sequence of events described above (viz., a southward turning of the IMF). Substorms can be triggered during periods in which the IMF has a northward component or during sustained intervals where the IMF has a southward component. The occurrence of substorms is therefore clearly influenced by the internal dynamics of the magnetosphere, with the general occurrence frequency being modulated by the level of the solar wind—magnetosphere interaction.

**Definition and Structure of Magnetospheric Substorms**

After some considerable discussion the authors agreed on an operational definition of a magnetospheric substorm in terms of measured parameters commonly available to researchers. The definition is as follows: the term magnetospheric substorm describes an interval of increased energy dissipation confined, for the most part, to the region of the auroral oval. The onset of this process is signaled by explosive increases in auroral luminosity in the midnight sector, and the entire process encompasses an interval during which the strength of the current in the auroral electrojets increases from and returns to the background level from which the substorm arose. During this interval there may be a sequence of intensifications of the westward electrojet, each associated with a Pi 2 micropulsation burst and a westward travelling surge. As the substorm develops, the region of discrete auroras in the midnight sector expands poleward and westward (the poleward bulge). Eventually, the region of disturbed aurora reaches a maximum latitude and begins to recover toward its presubstorm location. The interval of time between the first Pi 2 burst and the time the aurora reaches a maximum latitude has been called the expansion phase. The interval during which the aurora in the midnight sector returns to lower latitudes is called the recovery phase.

This working definition differs considerably from the original definition of the auroral substorm by Akasofu [1964], in whose framework the magnetospheric substorm phenomenon has been described to date. Most noteworthy is the removal of time scales that were assigned to the auroral substorm. The authors, on the basis of their experience, feel that the magnetospheric substorm as defined above has a time scale of ~1-3 hours. However, there are no statistical studies to date demonstrating the average substorm length or the limits of the substorm interval, namely, how long or how short a disturbance must be for it to be termed a substorm.

A second noteworthy facet of this definition is that multiple onsets and multiple surges may occur within the body of a single magnetospheric substorm. (In the original auroral substorm definition there was but one surge permitted.) This stems from the studies of substorm sequences by Wiens and Rostoker [1975] and multiple-onset substorms by Pytte et al. [1976], both of which show that the substorm westward electrojet develops as a series of discrete expansions of the disturbed region of the ionosphere. Each discrete onset has its own Pi 2 micropulsation burst and is associated with a surge form in the auroral structure.

The question then arises of whether each onset should be defined as a substorm or whether the ensemble of onsets should carry that definition. The authors agreed that each discrete expansion should be termed a substorm intensification, while the ensemble of the onset plus all intensifications up to the time of maximum poleward expansion of the substorm—disturbed region should be termed the expansion phase of the magnetospheric substorm.

Although the following fact does not enter into the actual substorm definition above, it should be noted that the authors are in agreement that the substorm onset is signaled by the brightening of an arc located near the equatorward boundary of the oval of discrete auroras. In the original definition of the auroral substorm as laid out by Akasofu [1964], the auroral observations on which the morphology was defined were obtained by using all-sky cameras, which were often unable to detect diffuse auroral glow. This inadvertently led to the impression that the substorm onset occurs near the equatorward edge of the auroral oval and thus near the inner edge of the plasma sheet. Later, Akasofu pointed out that the substorm onset occurs near the equatorward edge of the discrete aurora, far removed, in many cases, from the equatorward edge of the diffuse auroras [Akasofu, 1977]. In the context of the low-altitude plasma sheet established by Winningham et al. [1975], the substorm onset appears to occur near the equatorward edge of the boundary plasma sheet (BPS), as was indicated by the Isis 2 observations reported by Lui and Burrows [1978].

**Signatures of the Magnetospheric Substorm**

**The Auroral Signature**

The original time history of an auroral substorm proposed by Akasofu [1964] involved the brightening of an arc in the equatorward portion of the midnight sector auroral oval as the signature of onset. This was followed by rapid poleward expansion and the formation of a single westward travelling surge and the development of eastward drifting patches and K bands in the morning sector. The above signatures constituted the expansion phase, which ended at the time when the auroral disturbed region reached its maximum poleward position. The recovery phase encompassed the return of the disturbed region to its presubstorm configuration.

In the picture developed by the present authors, this general pattern must be interpreted and qualified in the following fashion. First, as was pointed out earlier, the arc whose brightening signals the onset of the breakup lies not in the equatorward portion of the auroral oval but rather near the equatorward edge of the oval of discrete auroras. In the late evening sector there may then be diffuse auroras well equatorward of the site of the brightened arc. Second, it is observed [e.g., Kawasaki and Rostoker, 1979] that a discrete arc brightening can occur in the region immediately to the west of a westward propagating surge. Since the surge may appear suddenly and explosively shortly thereafter, the sequence of an arc brightening and the arrival of the surge (which itself involved poleward expansion) may well satisfy the requirements of the definition of onset as originally proposed by Akasofu.
LOCATION OF OBSERVATORY

<table>
<thead>
<tr>
<th>LOCATION OF OBSERVATORY</th>
<th>ONSET TO EAST OF OBSERVATORY</th>
<th>ONSET ON MERIDIAN OF OBSERVATORY</th>
<th>ONSET TO WEST OF OBSERVATORY</th>
</tr>
</thead>
<tbody>
<tr>
<td>NORTH OF CENTER OF WESTWARD ELECTROJET</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOUTH OF CENTER OF WESTWARD ELECTROJET</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOW LATITUDES</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CODE

O - ONSET OF SUBSTORM
A - ARRIVAL OF SURGE AT OBSERVATORY MERIDIAN
R - ONSET OF RECOVERY

Fig. 1. Schematic representation of magnetograms from various sites placed at different positions with respect to the high-latitude substorm current systems. These synthetic records can be compared with real data to establish the position and time sequence of events for an interval of substorm activity. Typical deflections at high-latitude observatories range from ~100 to 1000 nT, while typical deflections at low-latitude stations range from a few nanoteslas to a few tens of nanoteslas.

[1964], and yet it may only represent an intensification of the substorm and the movement of the disturbance into the field of view of the observer (all-sky camera). Thus the brightening of an arc followed by the formation of a westward travelling surge must, for a single observer, be interpreted as being either an onset or an intensification of a substorm initiated earlier that has finally entered the observer's field of view. Data from other sources are necessary to resolve the potential ambiguity in the use of arc brightenings to define a substorm onset; this can be achieved by using either an adequate array of all-sky cameras/photometers on the ground or satellite coverage effected by imagers flown on polar-orbiting satellites (such as DMSP, Isis 2, or KYOKKO).

To summarize, we first emphasize that our definition of the magnetospheric substorm permits the development of multiple surges during the course of the expansion phase. Accordingly, the use of an auroral arc brightening as the indicator of substorm onset must be accompanied by evidence that such a brightening was the first in the sequence of the many brightenings that may take place during the substorm expansion phase. At the very least this may be achieved by demonstrating that no brightening occurred to the east of the station in question in the time frame of a few minutes prior to the brightening that is being proposed as the onset. A second point of importance is that the position of maximum poleward expansion may not occur in the same longitudinal region as the original onset. This is because the disturbance region expands, on the average, toward the northwest, and thus the maximum geomagnetic latitude reached is likely to be well to the west of the region of onset. Again, adequate ground array data or high-altitude imager data are needed to establish the time at which maximum poleward expansion is reached; a single all-sky camera or photometer is unlikely to be able to detect both the onset and the end of the expansion phase of a magnetospheric substorm.

The Magnetic Signature

For the present time, magnetograms from either auroral oval or low-latitude (A < 50°) magnetic observatories can be used with caution to define substorm onsets and recovery. Use of subauroral (60° > A > 50°) and polar cap (A > 75°) magnetograms is not recommended at present because the former is subject to mixed low-latitude and high-latitude responses, while the latter involves identification of perturbations whose morphology is not yet well defined and that look different from station to station owing to the proximity of polar cap stations to the geomagnetic pole.

Figure 1 illustrates schematically the magnetic signatures
found on auroral oval and low-latitude magnetograms, which can be used to define (1) the earliest detectable time of onset, (2) the location of the observatory with respect to the substorm-disturbed region, and (3) the recovery time of the substorm. The figure shows the various magnetogram signatures in \((H, D, Z)\) for auroral oval stations and \((H, D)\) for low-latitude stations depending on the position of the observatory with respect to the substorm westward electrojet and its associated field-aligned currents. It should be noted that during the time between onset and the beginning of the recovery phase several quasi-periodic intensifications are shown. This emphasizes that a substorm, by our definition, can encompass many electrojet intensifications (with their associated auroral surge forms) that occur over a time span of several minutes to over an hour. The following perturbation trends should be noted and used as indicators of the position of the substorm-disturbed region.

1. At auroral oval stations, positive (negative) \(Z\) indicates an observatory to the north (south) of the center of the westward electrojet. A positive spike in \(D\) indicates the presence of surge forms close to the observatory. Observation of significant negative \(H\) and negative \(D\) with no discernible positive \(D\) spike indicates that the substorm onset took place just to the west of the observatory. A significant negative \(H\) with little response in \(D\) indicates that the substorm onset took place well to the west of the observatory.

2. At low latitudes a positive (negative) excursion in \(H\) indicates that the observatory is near, but to the east (west) of, the field-aligned currents associated with the western edge of the substorm-disturbed region. Positive (negative) \(D\) indicates that the observatory is to the west (east) of the western edge of the substorm-disturbed region. As regards the determination of the onset time of the substorm from auroral zone magnetograms, it should only be pointed out that at stations to the west of the initial position of the surges, the substorm onset cannot usually be unambiguously defined given the present array of observatories. However, one can be reasonably sure that an auroral oval observatory is observing the true substorm onset if the sharp positive \(D\) spike is not preceded by a slowly developing negative \(H\) bay (see Figure 1). Otherwise great care must be exercised to ensure that sharp negative \(H\) bays (presently used to define substorm onsets) do not simply reflect local intensifications of the substorm westward electrojet rather than the onset of the substorm. Insofar as identification of substorm onsets from low-latitude magnetograms is concerned, the \(D\) component generally provides the best indicator of the earliest detectable onset time. The \(H\) component signature depends on the position of the observatory with respect to the meridian to which the westward edge of the westward electrojet extends; one can detect the passage of the western edge of the westward electrojet from east to west across the meridian of the observatory by the transition from a negative-going to a positive-going \(H\) component signature.

Finally, we note that it is presently customary to use the \(AE\) index (synthesized from auroral zone magnetograms) to determine onsets and identify recovery of substorm activity. While \(AE\) is useful to establish the gross level of magnetospheric activity, it should not be used for the study of the time development of substorms. This is because the distribution of stations used to derive \(AE\) has too many gaps to permit accurate identification of the substorm onset and the end of the expansion phase and because much useful information is lost through ignoring the \(D\) and \(Z\) components and considering the \(H\) component only. We recommend that researchers studying the time development of substorms and substorm-related phenomena utilize the original magnetograms from as many relevant observatories as possible for obtaining information on the development of the substorm. Magnetograms from many observatories distributed around the globe may be obtained on request from the World Data Centers.

**The Pi Micropulsation Signature**

Our definition of a substorm has, as a requisite feature, the onset of at least one Pi 2 pulsation burst. The appearance of the Pi 2 signifies that the magnetospheric event is of an impulsive nature, which is a key characteristic of a substorm. Pi 2 micropulsations are conventionally thought of as damped wave trains whose initial pulse is easily detectable both in the auroral oval and at middle and low latitudes on the nightside. Saito [1961] first emphasized that there was a one-to-one correspondence between Pi 2 bursts and substorms, but Rostoker [1968] showed that the situation was more complicated in that there was generally a minimum of two identifiable Pi 2 bursts associated with every substorm. At present, it is recognized that each substorm intensification has an associated Pi 2 burst, so that a magnetospheric substorm, by our definition, can have several associated Pi 2 bursts, only one of which actually coincides with the onset of the substorm.

At the present time, Pi 2's can be identified both on standard magnetograms and on recordings from induction coil magnetographs. Pi 2 amplitudes vary from \(\sim 100\) nT in the source region to fractions of a nanotesla at low latitudes, which calls for recorders with large amplitude ranges. Standard magnetograms do not satisfy this requisite, and hence only the large Pi 2's are detectable unambiguously at lower latitudes. Since there exists, at the present, no amplitude criterion for Pi 2's, great care must be exercised in ruling out the presence of Pi 2 activity on the basis of standard magnetograph and induction coil magnetograph recordings. For example, the failure to detect a Pi 2 at a given time does not rule out the existence of substorm activity, as the perturbation may simply be below the threshold of conventional recorders.

As a rule of thumb, a well-defined Pi 2 onset gives an accurate (\(\pm 1\) min) identification of the onset time of a magnetospheric substorm onset or intensification. However, low-amplitude micropulsation activity in the Pi 2 frequency band can also indicate substorm onsets or intensifications even though the beginning of the wave train may not be easily identified. It is important to recognize that a Pi 2 wave train does not always start with a large amplitude pulse as observed at all field points. Often the initial pulse may fall off faster in amplitude than the rest of the wave train as one moves away from the source region, leading to wave trains that appear to build up in amplitude. Such wave trains are legitimate Pi 2's and are as good an indicator of substorm onsets and intensifications as the more conventional damped wave train forms.

**The Electric Field Signature**

In comparison with ground-based magnetic and optical observations the electric field phenomenology of substorms is still in a state of rapid development, and a comprehensive picture of the spatial and temporal development of the ionospheric electric field during a substorm has not yet been fully achieved. For this reason the electric field behavior is not included to specifically in our definition of the magnetospheric substorm. However, in light of recent major achievements by
incoherent backscatter facilities (such as Chatanika) and auroral radar systems (such as Stare), some brief comments on substorm electric field signatures are warranted.

Perhaps the only reliable substorm electric field signature available at the present time is the increase in the westward electric field in the region behind a westward travelling surge. There are several studies that also suggest that there is a southward turning of the electric field at the onset of a substorm expansion phase. However, it is not known whether or not this effect is spatial (involving the expansion of a pre-existing region of southward electric field) or temporal (involving the appearance of a southward field in the substorm-disturbed region). While one would like to use high-altitude electric field measurements to complement those obtained by using balloon-borne and rocket-borne probes and ground-based techniques, it has become increasingly apparent that parallel electric fields exist at high altitudes, decoupling the magnetosphere from the ionosphere over regions of limited spatial extent. Thus it is not possible to infer unequivocally ionosphere electric fields from high-altitude electric field measurements, particularly on auroral oval field lines.

Finally, we note that the southward electric fields associated with substorms have been observed to have magnitudes ranging from a few millivolts per meter to ~200 mV/m. However, up to this point it has been impossible to relate quantitatively the strength of the electric field to the strength of the substorm as measured by more conventional means.

**Intensity Criteria for Substorms**

It will be noted that in our definition of the substorm we have made no attempt to define some threshold of a given parameter below which it can be said with certainty that no substorm has occurred. The reason for this is that despite the many discussions regarding how strong a characteristic signature must be before it can be accepted as a substorm effect, no agreement has yet been reached on a lower limit for each effect above which a substorm can be affirmed and below which it can be denied. The various substorm signatures vary in intensity from event to event, and the degree of spatial localization and complex auroral structure of the substorm-disturbed region make it difficult to assign some 'integrated intensity' to any single event.

Many researchers have chosen the magnetic index $AE$ (or its component index $AL$) as a quasi-quantitative indicator of substorm strength. Unfortunately, this index is not perfect, as it registers a lower limit to the peak current rather than an integrated current. Examination of several years of values of either $AE$ or $AL$ yields a clear impression that stronger currents flow during some substorms than during others. However, this impression has no firm basis in fact, owing to the inadequate nature of the data base used. Therefore it is not, at present, possible to use the magnitude of the $AE$ or the $AL$ index alone to confirm or deny the presence of a substorm and to quantify its intensity.

Similar problems affect any attempt to quantify a substorm by its level of auroral luminosity. This is due to the fact that it is the combination of auroral luminosity and electric field (both primary and polarization) that determines the level of energy dissipation in an auroral-disturbed region, and without knowledge of the global electric field distribution it is not possible to evaluate quantitatively the intensity of a given substorm.

Finally, we note that the energy release associated with a substorm event may lead to injection of part of the energy into the ring current, where it is subsequently dissipated over a much longer period of time than one normally associates with the substorm process. The proportion of the energy dissipated in the ionosphere during the substorm expansion phase vis-à-vis the amount injected in the ring current appears to vary in an as yet undetermined fashion. Once again, the problem of assigning an intensity criterion to substorms appears impossible given our present state of knowledge.

We can summarize the question of intensity criteria for substorms simply by saying that the area is one that requires extensive study. It is not possible at this time to look at some localized measurements of any parameters (i.e., magnetic field, auroral luminosity, etc.) and rule out the presence of substorm activity. Perhaps future high-altitude satellites carrying auroral imaging devices may be able to determine the timing of substorm phases. Only then will it be possible to assign intensity criteria to substorms.

**Recommendations**

In this paper we have provided what we believe to be an operational definition of a magnetospheric substorm, and we have indicated the signatures one should expect to observe for some of the more commonly used ground-based detectors. One need only use one of the specific detectors whose signatures have been presented above in order to be entitled to discuss substorm morphology, as long as one heeds the limitations of that observational technique.

It is important, however, to emphasize that present arrays of ground-based instrumentation are often inadequate to identify unambiguously the true onset of a magnetospheric substorm and follow its development in space and time. The substorm is a dynamical process, and the key to understanding it is monitoring its temporal behavior. To this end, we have established networks of all-sky cameras and magnetometers. Satellite photography has all too clearly demonstrated the imperfections in our present network coverage. Our understanding of auroral morphology took a great step forward when satellite pictures became available. However, these photographs are merely snapshots. To understand the relationship between the auroral motions and the development of the auroral electrojet during substorms requires a satellite monitor of the aurora capable of following the auroral oval over the whole night sky during the course of a substorm. This monitor should be able to distinguish discrete and diffuse auroral forms, to resolve features to better than 10 km, and to relate these forms to positions on the earth to such a resolution. It should be able to return such images at least every minute for a period of several hours. We realize that such a monitor would put severe demands on spacecraft altitude, telemetry rate, and stability, but we view such a system as being necessary for achieving a better understanding of substorm phenomena.

We also draw attention to the utility of radar measurements for providing a monitor of substorm dynamical effects in the ionosphere, albeit over a limited band of longitudes. The approach used in the Stare radar system located in Scandinavia has proven to be particularly useful and is relatively inexpensive. We would encourage the establishment of a number of such systems so that the ionospheric electric field configuration can be established simultaneously in several local time regimes.

Another difficulty in studying substorms is the analogue na-
ture of much of the ground-based data. Considerable effort is expended in digitizing records that, with present techniques, could easily have been obtained by using digital equipment. We recommend that wherever possible data be gathered in digital form and quickly archived in central facilities. The International Magnetospheric Study has gone some considerable way in achieving this latter aim insofar as some ground-based facilities are concerned; however, this approach will have to be taken globally before an adequate, accessible data base will be available with which the substorm problem can be attached.

Acknowledgments. This study was carried out at the University of Victoria in Victoria, British Columbia, Canada. The authors are greatly indebted to R. E. Horita for his help in arranging the facilities at the university and to the university for the hospitality extended to them during the course of the workshop. The authors are grateful to the various funding agencies in the United States, Canada, and the Federal Republic of Germany for providing travel funds in support of the Victoria workshop.

REFERENCES


(Received October 1, 1979; revised November 2, 1979; accepted November 20, 1979.)