Tail Reconnection Triggering Substorm Onset

Vassilis Angelopoulos,1* James P. McFadden,2 Davin Larson,2 Charles W. Carlson,2 Stephen B. Mende,2 Harald Frey,2 Tai Phan,2 David G. Sibeck,4 Karl-Heinz Glassmeier,4 Uli Auster,1 Eric Donovan,1 Ian R. Mann,6 I. Jonathan Rae,6 Christopher T. Russell,1 Andrei Runov,1 Xu-Zhi Zhou,1 Larry Kepko7

Magnetospheric substorms explosively release solar wind energy previously stored in Earth’s magnetotail, encompassing the entire magnetosphere and producing spectacular auroral displays. It has been unclear whether a substorm is triggered by a disruption of the electrical current flowing across the near-Earth magnetotail, at ~10 RE (Earth radius, or 6374 kilometers), or by the process of magnetic reconnection typically seen farther out in the magnetotail, at ~20 to 30 RE. We report on simultaneous measurements in the magnetotail at multiple distances, at the time of substorm onset. Reconnection was observed at 20 RE, at least 1.5 minutes before auroral intensification, at least 2 minutes before substorm expansion, and about 3 minutes before near-Earth current disruption. These results demonstrate that substorms are likely initiated by tail reconnection.

Substorms are global reconfigurations of the magnetosphere involving storage of solar wind energy in Earth’s magnetotail and its abrupt conversion to particle heating and kinetic energy (I, 2). Because phenomena related to the onset of substorms are initially localized (within 1 to 2 RE) in space but expand quickly to engulf a large portion of the magnetosphere (3, 4), fortuitous (and thus unoptimized) conjunctions between single satellite missions have been unable to pinpoint the exact location of the substorm trigger in space. This has led to diverging theoretical efforts to explain the onset mechanism (3, 6). The key question is whether substorm phenomena are triggered by a near-Earth dipolarization (current disruption) process, at ~10 RE, or by the process of magnetic reconnection at ~20 to 30 RE. Both processes operate during substorms, but attempts (7, 8) to delineate the causal relation between them and substorm onset were limited due to temporal resolution, spatial propagation effects, and/or lack of simultaneous observations of the above key regions. The THEMIS mission (9, 10) was designed to address this question. The mission uses five identical satellites (hereafter termed “probes”) on orbits, enabling recurrent probe alignments parallel to the Sun-Earth line (probes within 8XGSM ≤ 2 RE from each other; GSM: Geocentric Solar Magnetospheric coordinate system). The probes can thus monitor tail phenomena simultaneously at ~10 RE and at ~20 to 30 RE downtail, while mapping magnetically over a network of ground-based observatories (GBOs), which can determine the meridian and time of substorm onset on the ground (11). Here we present timing results from THEMIS for isolated substorms, which demonstrate that the substorm trigger mechanism is magnetic reconnection.

Substorm Signatures

The ground signatures of substorms consist of a rapid auroral intensification, a breakup of auroral forms into smaller filaments, and a poleward expansion and a westward surge of the most intense auroral arcs. Those are within 1 to 2 min of each other and have often been used synonymously with substorm onset. Ground magnetic signatures of currents associated with auroral arc intensification include abrupt increases in the auroral electrojet (AE) index (12) and irregular pulsations in the 40- to 150-s range called Pi2s (13), or at lower periods (14), observed at high-latitude (auroral) and mid-latitude (subauroral) magnetic stations. Such magnetic signatures are known to coincide with auroral intensification typically within 1 to 2 min (15) and are also used for substorm onset identification.

Before substorm onset, during the substorm growth phase, stable arcs intensify and move equatorward while the magnetotail plasma sheet thins and the cross-tail current increases (16–19). During the expansion phase, auroral arcs typically advance toward the poleward edge of the auroral oval and a current wedge develops in space, at ~10 RE, composed of field-aligned currents into and out of the ionosphere (17). Current wedge formation is also referred to as dipolarization, because the field becomes more dipole-like, or as current disruption, because it is consistent with a disruption (reduction) of the duskward cross-

Fig. 1. Projections of THEMIS probes in X-ZGSM plane along with representative field lines and neutral sheet location in GSM coordinates at 04:45 UT on 26 February 2008. Times refer to the time delays in Table 1.
Fig. 2. (left). (A) THEMIS index, $AE_{TH}$, computed from THEMIS GBOs (11) and select CARISMA stations. (B and C) Band-pass–filtered (3.0 to 120 s) magnetograms from Gillam (east-west component; station magnetic coordinates: 66.5°N, 28.0°W) and Carson City, NV (north-south component; 45.1°N, 57.0°W), in nT. The filter includes the Pi2 band but extends to higher frequencies to reduce aliasing. (D) Integrated auroral intensity from the northern half of Gillam ASI’s field of view (arbitrary units). (E) Latitude of poleward-most extent of auroral luminosity at station Gillam. (F and G) Composite images (mosaics) from THEMIS ASIs around the location of substorm onset, over a continental outline. Stations used in both stills are Gillam (bottom left), Rankin Inlet (top left; station magnetic coordinates: 72.7°N, 25.1°W), and Sanikiluaq (bottom right; 66.9°N, 3.5°W). Gillam had partially cloudy skies, reflecting the Moon from low in the horizon and obscuring the north-east and south-west views (the latter is reflection on the dome), though this does not affect our findings. The red line is the midnight meridian. Color circles indicate ionospheric footpoints of THEMIS probes (same symbols as in Fig. 1) based on the T96 mapping model (24). The arrow indicates optical onset location. Fig. 3 (right). (A to S) Overview of magnetic field and particle data for 30 min, at 3-s resolution, around the 04:54:00 UT substorm. There are six panels per probe, arranged from top to bottom for probes P1, P2, and P3, plus one (bottom) panel showing only 30- to 200-keV ion spectra on P5, from the SST instrument. The six panels per probe are (from top to bottom) as follows: magnetic field measured by the FGM instrument (35); ion-flow velocity measured by the ESA instrument; energy spectra of 0.005- to 2000-keV ions from the SST and ESA instruments (next two panels); and energy spectra of 0.005- to 2000-keV electrons from the same instruments. All energy spectrograms show omnidirectional differential energy flux (eflux) in units of $eV/(cm^2\cdot s\cdot str\cdot eV)$. The abrupt change in eflux for each species exists because the data below 25 keV and above 30 keV were obtained by two instruments (ESA and SST, respectively) with different instrument geometric factors. Magnetic field and velocity are in GSM coordinates; $X$, $Y$, and $Z$ components are shown in blue, green, and red, respectively.
tail current (5). Further downtail, fast tailward flows threaded by southward magnetic fields, or Earthward flows threaded by northward fields, are observed near substorm expansion onset and have been interpreted as evidence for magnetic reconnection (20, 21). Substorm expansion is followed by substorm recovery, during which auroral forms remain active at the poleward boundary of the auroral oval for hours, until they eventually reduce in intensity and move equatorward, often starting another substorm sequence. Arc intensification alone does not necessarily constitute a substorm, even though it may involve the same underlying physics as substorms (22). The sequence of growth phase, expansion, and recovery of the aurora constitutes a bona fide substorm process. Due to the gradual intensification of auroral arcs at growth phase, it is typically easier to identify a substorm onset by its poleward expansion. The high cadence of the THEMIS ground measurements allows us to explicitly differentiate between the aforementioned observational determinations of onset, and we intentionally retain this differentiation as these phenomena manifest different magnetospheric-ionosphere coupling processes.

Substorm Timing on 26 February 2008
At 4:50 UT, the THEMIS probes were aligned along the Sun-Earth line, less than 1 \( R_E \) from the nominal neutral sheet (Fig. 1). A sudden increase of the THEMIS Auroral Electrojet Index (AE\(_{11}\)) to 200 nT was observed at 04:54:00 UT, indicating an isolated substorm onset by its poleward expansion. The high cadence of the THEMIS ground measurements allows us to explicitly differentiate between the aforementioned observational determinations of onset, and we intentionally retain this differentiation as these phenomena manifest different magnetospheric-ionosphere coupling processes.

Table 1. Summary of timing results during the 26 Feb 2008 04:53:45 UT substorm onset, in order of time sequence. The last column is the time delay, assuming reconnection onset at 04:50:03 UT, at 20 \( R_E \), which was interpreted at base on our interpretation of data and an estimate of an average Alfvén speed in the plasma sheet of 500 km/s.

<table>
<thead>
<tr>
<th>Event</th>
<th>Observed time (UT)</th>
<th>Inferred delay (seconds since 04:50:03 UT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reconnection onset</td>
<td>04:50:03 (inferred)</td>
<td>( T_{Rx} = 0 )</td>
</tr>
<tr>
<td>Reconnection effects at P1</td>
<td>04:50:28</td>
<td>25</td>
</tr>
<tr>
<td>Reconnection effects at P2</td>
<td>04:50:38</td>
<td>35</td>
</tr>
<tr>
<td>Auroral intensification</td>
<td>04:51:39</td>
<td>( T_{Al} = 96 )</td>
</tr>
<tr>
<td>High-latitude Pi2 onset</td>
<td>04:52:00</td>
<td>117</td>
</tr>
<tr>
<td>Substorm expansion onset</td>
<td>04:52:21</td>
<td>( T_{Rx} = 138 )</td>
</tr>
<tr>
<td>Earthward flow onset at P3</td>
<td>04:52:27</td>
<td>144</td>
</tr>
<tr>
<td>Mid-latitude Pi2 onset</td>
<td>04:53:05</td>
<td>182</td>
</tr>
<tr>
<td>Dipolarization at P3</td>
<td>04:53:05</td>
<td>( T_{Ed} = 182 )</td>
</tr>
<tr>
<td>Auroral electrojet increase</td>
<td>04:54:00</td>
<td>237</td>
</tr>
</tbody>
</table>

Overview of tail signatures. Probes P1 and P2 recorded a decreasing magnitude of Earthward component, \( B_\parallel \), between 04:45 and 05:01 UT (Fig. 3, A and G), consistent with a decreasing current sheet thickness and an increasing current sheet density, as expected at substorm growth phase. The plasma sheet ion density and average energy, obtained from the data shown in Fig. 3, D and J, were about one particle per cubic centimeter and ~1 keV, respectively. These are attributes of a cold-dense plasma sheet after prolonged intervals of northward interplanetary field (25). Probe P1 observed at ~05:44 UT tailward flows \((V_\parallel < 0)\) accompanied by southward \((B_\parallel < 0)\) and duskward \((B_\parallel > 0)\) excursions of the magnetic field, followed at ~05:02 UT by Earthward flows and opposite-polarity magnetic field perturbations. Probe P2 observed Earthward flows of the same nature and at about the same time as P1 (05:02 UT). The flow and field signatures at P1 are expected from a reconnection site first located Earthward, then retreating (or reappearing) tailward of P1, starting at ~05:01 UT. The observed \( B_\parallel \) component variations \((B_\parallel > 0\) tailward of the reconnection site; \( B_\parallel < 0\) Earthward of it) are also classical Hall signatures of reconnection (26–28).

Before the onset of the fast tailward flows, both probes observed two ion components: a 500-eV component, commensurate with the cold plasma sheet preceding the event, and a 10-kV component that appeared gradually. Both probes also observed relatively low-temperature electrons (100 to 200 eV) of decreasing flux. Despite the plasma sheet thinning, P1 and P2 remained within the plasma sheet. As the fast tailward flows were observed at P1, the average energy of the ions and electrons increased to 10 keV and 1 keV, respectively. After the tailward retreat of the reconnection site at 05:01 UT, both probes observed even hotter plasma (20-kV ions, 2-keV electrons) and crossed the neutral sheet as evidenced by the near-zero transitions of \( B_\parallel \) at around 05:02 UT. This is evidence of plasma heating at the reconnection outflow and plasma sheet dipolarization at 05:02 UT at 22 \( R_E \); the distance of P1.

If the flows are due to reconnection, the current sheet should resemble a slingshot-like, standing Alfvén wave (27, 28). To evaluate the shear stress balance, we examined the correlation between the measured ion flows, \( \Delta V_i \), and the flows predicted from reconnection outflow, \( \Delta V_A \propto \pm \Delta B \cdot N_i^{-1/2} \), on P1 and P2 \((V_A:\) Alfvén speed, \( B: \) magnetic field, \( N_i: \) the ion density). The correlation coefficients for the tailward and Earthward flows on P1 \((04:53:30 \text{ to } 04:58:30; 05:01:00 \text{ to } 05:13:00)\) and for the Earthward flows on P2 \((05:01:30 \text{ to } 05:06:00)\) were 0.79, 0.61, and 0.86, whereas the slopes were ~0.53, 0.32, and 0.51, respectively. The slopes are likely underestimates because of temporal variations in the reconnection process and because we have not yet included energetic
particles in the velocity determination. Alternatively, the stress balance at the peak velocity near the center of the plasma sheet gives ratios: \(|\Delta V_p/|\Delta V_x| \sim 90\% \) \((V_p,\text{ ion velocity})\), 67\% and 88\% for the tailward and Earthward flows on P1 (04:54:30 UT, 05:02:00 UT) and the Earthward flows on P2 (05:02:30 UT), respectively. Under the caveat of the need to include the superthermal ion corrections to the ion flow velocity, mostly required for the hotter, Earthward flows, the observations are consistent with the Alfvénic acceleration expected from reconnection.

P3 was near the neutral sheet. It observed fast (>400 km/s) Earthward flows starting at about 04:52:27 UT, followed by a transient increase in the northward component of the magnetic field (transient dipolarization) at 04:53:05 UT. The onset of fast flows was followed by a more permanent dipolarization at 04:54:40 UT, signifying the development of a substorm current wedge in near-Earth space. The transient dipolarization at ~04:53:05 UT is interpreted as the first indication of a substorm current wedge at P3.

P5, near geosynchronous altitude, saw an energy-dispersed ion injection of the 50- to 200-keV ions (Fig. 3S). The flux increase at the lowest-energy (highest-flux) channel is an exception, because it responds to the local plasma and is correlated with convective velocity changes measured at the same time. The energetic-particle dispersion (more energetic particles drifting faster than lower-energy particles) is consistent with a duskward drift of those particles to the location of P5 after an injection near midnight. Such dispersed injections are classic signatures of substorms observed by geosynchronous orbit satellites (29).

**Observations around the time of onset.**

The fast tailward reconnection flows \((V_z < -100 \text{ km/s})\) on P1 started at ~04:52:30 UT. They were preceded by a northward convective flow \((V_y > 50 \text{ km/s})\) and an accompanying southward deflection of the magnetic field \((\delta B_z < 0)\), which we interpret as evidence of onset of reconnection inflow toward the neutral sheet at 04:50:28 UT. Ion-velocity distributions (Fig. 4C) show two components: a relatively isotropic component below ~500 km/s and a duskward and tailward streaming component above 1000 km/s. These are the cold and hot ions seen earlier in the spectra of Fig. 3D. Similar behavior was found on probe P2. The anisotropy of the few-keV ions intensified by 04:51:14 UT (see distribution function in Fig. 4C); even 1-keV ions (~310 km/s) exhibited pronounced duskward drift. This is consistent with a diamagnetic ion current. The inferred gradient scale is approximately the gyroradius of a 1-keV proton in the local field (~20 nT), i.e., ~600 km. This is further evidence that the current sheet was thin and the current density high.

Electron velocity space distribution functions (Fig. 4D) exhibit a bidirectional anisotropy before 04:50:54 UT, which intensified in the ensuing minutes. As seen in Fig. 4E, 50- to 300-eV electrons were streaming toward the reconnection site (180\° pitch angle, i.e., approximately Earthward), while 400- to 2000-eV electrons were streaming away from the reconnection site (0\° pitch angle, i.e., ~tailward). Such electron streaming is a signature of reconnection due to the Hall current system (21), suggesting that reconnection had started near the location of P1 by 04:50:28 UT.

Probe P2 was farther away from the neutral sheet than P1, as evidenced by the enhanced magnetic field (Fig. 3G) and ion energy spectra (Fig. 3, I and J). No direct connection of field lines at P2 to the reconnection site was evident until after onset. However, similar to P1 at 04:50:28 UT, observations at P2 at 04:50:38 UT show the beginning of inflow toward the reconnection site \((V_y > 0)\,\text{Fig. 4G}) and the start of a positive deflection of \(B_z\) along with a bipolar \(B_y\). These are signatures of an Earthward flux transfer event (30, 31), signifying tail reconnection somewhere tailward of P2. The simultaneous deflection of \(B_y\) northward at P1 and southward at P2 suggests that a reconnection topology was established at that time between the two probes, i.e., between 17 and 22 \(R_E\) (Fig. 1).

Probe P3 observed a slow ramp-up of the Earthward flow velocity \((V_y > 50 \text{ km/s})\) at 04:52:27 UT (Figs. 3N and 4I), followed by a

---

**Fig. 4.** Data from P1, P2, and P3 during the first few minutes before substorm onset. (A, F, and H) Magnetic field as in Figs. 2 and 3, except that on P1 the resolution is four samples per second. On P1 and P2 the 
\(X_{PSM}\) component (but not the others) was detrended (high-pass filtered) by subtracting a 6-min running average, to reveal details. (B, G, and D) Ion velocity as in Figs. 2 and 4. For dotted lines, see text. (C and D) Ion- and electron-velocity distribution functions near the spin plane in despun spacecraft coordinates (+X is Earthward and is to the right of the page; +Y is dawnward and is positive to the top of the page); units are in particles/(cm\(^3\) km\(^3\) s\(^3\)). (E) Energy flux spectra of electrons along (0\°), opposite (18\°), and perpendicular (90\°) to the magnetic field for the times indicated.
deflection of Bz northward (dipolarization) and high-speed flows. The slow flows seen at P3 before the dipolarization did not necessarily emanate directly from the reconnection region; rather, they may have been nearby plasma that accelerated Earthward due to the forces from the establishment of a reconnection topology further downtail. The first signatures of dipolarization were timed at 04:53:05 UT.

Together, the observations at P1, P2, and P3 make a compelling case for onset of tail reconnection at or before t0 = 04:50:28 UT, between P1 and P2 (Table 1). Approximating the Alfvén speed near the reconnection site as 500 km/s (based on a local density measurement of one particle per cubic centimeter and a magnetic field of 20 nT), we can infer the downtail location of the source, x0, and the time of reconnection onset, t0. Noting that the reconnection pulse can travel a distance of 5 Re (the P1-P2 interprobe separation) in 60 s, we obtain (t1 - t0) + (t2 - t0) = 60 s, resulting in t0 = 04:50:03 UT and x0 = 20 Re.

Time history of events. The inferred reconnection onset at 04:50:03 UT preceded the onset of auroral intensification by 96 s (Table 1). Arc intensification was followed by high-latitude P12 onset, 21 s later. The high-latitude P12 onset may signify the arrival of the field-aligned current pulse generated by the reconnection flows in the tail. It is unlikely that a shear Alfvén wave, starting at ~500 km/s, can travel from 20 Re to the ionosphere in 96 s, due to both the high density in the plasma sheet and the large distance to the source.

Conversely, kinetic Alfvén waves (32) of an ion acoustic gyroradius scale may exceed the local Alfvén speed by a factor of \(\sqrt{2}\) and arrive faster. Those waves can also accelerate electrons (33), which may result in visible aurorae ahead of the wave. This would explain the observations of arc intensification 20 s earlier than the P12 onset and would be consistent with reported observations of Alfvénic aurorae (34).

Twenty-one seconds after the high-latitude P12 onset, the aurorae started to expand poleward. This expansion started 6 s before the arrival of an Earthward flow perturbation at P3 and about 40 s before the dipolarization at P3. Therefore, the initial poleward motion of the aurora cannot be caused by the near-Earth flux pile-up of reconnection flows. It is likely associated with the change in magnetic field mapping, because reconnection at 20 Re results in the engulfment of higher magnetic latitude flux in the reconnection process. Flux pileup and current wedge formation may, however, be responsible for later stages of poleward expansion; careful modeling of the current wedge currents in realistic magnetotail fields is needed to properly address this question.

Across field lines, the reconnection process started to affect the inner magnetosphere at 11 Re (P3’s location) 144 s after onset. This time delay relative to reconnection onset is commensurate with the magnetosonic speed of the 1- to 4 keV plasma at the neutral sheet between P2 and P3, i.e., ~500 km/s. The first evidence of intense dipolarization, interpreted as the reconnected flux arrival at that same location, was seen 30 s after the first indication of Earthward flow and 3 min after reconnection onset at 20 Re. The latter is also commensurate with the simultaneously observed plasma flow speed, ~400 km/s. The onset of mid-latitude P12 pulsations was observed simultaneously with the dipolarization at 11 Re. Mid-latitude P12 onset has been interpreted previously as an integrated response to the field-aligned currents from the flows contributing to current wedge formation, but not necessarily due to the current disruption process in the near-Earth region. The observed P12 onset time at 04:53:05 UT is consistent with such an interpretation.

Other Substorm Events
In the aforementioned substorm event, the plasma sheet was atypically cold and dense, suggesting that the slower Alfvén and magnetosonic speeds may result in longer-than-usual communication times between the various regions within the plasma sheet, as well as between the tail and the ionosphere. This may have been responsible for the easy temporal differentiation of the substorm signatures on the ground and in space observed in this event. To demonstrate that our findings are typical of other events, we examined two additional isolated substorms on 16 and 22 February, both between 04:30 and 05:00 UT, and reached similar conclusions (23).

In the events analyzed, the earliest indication of substorm onset followed the first evidence of tail reconnection by <96 s and preceded the earliest indication of current disruption by >1 min. It is surprising how quickly the aurora intensifies in response to reconnection onset (<96 s). Electron acceleration by reconnection-generated kinetic Alfvén waves may explain this tight coupling between the ionosphere and the reconnecting plasma sheet. Our observations, however, raise another question: What growth-phase process preconditions and destabilizes tail reconnection during spontaneous and externally driven substorms? Because substorm arcs intensify gradually and the magnetotail thins slowly before onset (over several minutes), the entire magnetotail from geosynchronous altitude to 30 Re would have enough time to partake in that process.

References and Notes
23. Materials and methods are available as supporting material on Science Online.
36. This work was supported by NASA contract NASS-02099. The work of K.-H.G. was financially supported by the German Ministerium für Wirtschaft und Technologie and the German Zentrum für Luft- und Raumfahrt under grant 50QP0402. Logistical support for fielding and retrieval of the THEMIS-GBO and CARISMA (Canadian Array for Realtime Investigations of Magnetic Activity) data is provided by the Canadian Space Agency. CARISMA is operated by the University of Alberta. Special thanks to A. Prentice for careful editing of this paper.