Optical characterization of the growth and spatial structure of a substorm onset arc

I. J. Rae,1 C. E. J. Watt,1 I. R. Mann,1 K. R. Murphy,1 J. C. Samson,1 K. Kabin,1 and V. Angelopoulos2

Received 18 February 2010; revised 29 April 2010; accepted 26 May 2010; published 12 October 2010.

[1] We present a detailed case study of the characteristics of auroral forms that constitute the first ionospheric signatures of substorm expansion phase onset. Analysis of the optical frequency and along-arc (azimuthal) wave number spectra provides the strongest constraint to date on the potential mechanisms and instabilities in the near-Earth magnetosphere that accompany auroral onset and which precede poleward arc expansion and auroral breakup. We evaluate the frequency and growth rates of the auroral forms as a function of azimuthal wave number to determine whether these wave characteristics are consistent with current models of the substorm onset mechanism. We find that the frequency, spatial scales, and growth rates of the auroral forms are most consistent with the cross-field current instability or a ballooning instability, most likely triggered close to the inner edge of the ion plasma sheet. This result is supportive of a near-Earth plasma sheet initiation of the substorm expansion phase. We also present evidence that the frequency and phase characteristics of the auroral undulations may be generated via resonant processes operating along the geomagnetic field. Our observations provide the most powerful constraint to date on the ionospheric manifestation of the physical processes operating during the first few minutes around auroral substorm onset.


1. Introduction

[2] In the mid-1950s, Heppner [1958] identified that the brightening and breakup of the most equatorward premidnight auroral arc was strongly associated with the formation of a sharp magnetic bay in ground magnetometers. Subsequently, the surrounding sequence of events was categorized and termed a “substorm” by Akasofu [1964]. Since this initial identification and characterization, the physical process or processes responsible for the initiation of magnetic energy release during the initiation of substorm expansion phase onset have since remained elusive [see, e.g., Lui, 2004, Angelopoulos, 2008; Angelopoulos et al., 2008a]. One of the primary reasons for the ambiguity in determining the causal sequence of events leading to the magnetic energy release during substorms is the inability to temporally and spatially resolve the responses of different processes during the first 2 min of onset [e.g., Lui, 1991]. Substorm initiation via magnetic reconnection at the near-Earth neutral line (NENL) [e.g., Hones, 1976] and via current disruption (CD) in the nearer magnetotail [e.g., Roux et al., 1991; Lui et al., 1991] have both been proposed as likely mechanisms for substorm triggering. There are a number of other phenomenological models that have also been proposed but have received less attention in the literature (e.g., boundary layer dynamics model [Rostoker and Eastman, 1987], near-geophysical onset [e.g., Maynard et al., 1996a], and global Alfvénic interaction [Song and Lysak, 2001; Lin et al., 2009]). The primary objective of the Time History Events and Macroscale Interactions during Substorms (THEMIS) mission is to resolve the causal sequence of energy release during substorm expansion phase onset and thus provide the required crucial information on the relative timing of the physical phenomena surrounding onset [Sibeck and Angelopoulos, 2008; Angelopoulos, 2008]. Interestingly, Angelopoulos et al. [2008b, 2009] have shown evidence of either a new or modified paradigm: These authors observing a clear tail reconnection-initiated auroral brightening before CD, a scenario that neither of NENL or CD paradigms can explain fully.

[3] In this paper we use the classical Akasofu definition of auroral substorm onset that “The first indication of a substorm is a sudden brightening of one of the quiet arcs lying in the midnight sector of the oval (or a sudden formation of an arc)” [Akasofu, 1977]. Using this definition, the magnetospheric source of auroral expansion phase onset clearly...
occurs at an earlier time than the first observable signature in the ionosphere. Friedrich et al. [2001] presented evidence of small azimuthal scale arc ripples during the onset process that evolved into vortical structures over time scales of approximately minutes. With the new capability of high time resolution optical data over an extended spatial region offered by the THEMIS All-Sky Imagers (ASIs) [Mende et al., 2008], finer temporal and spatial scale auroral features are now being discovered during the onset process. Most notably, Donovan et al. [2007, 2008], Liang et al. [2008], Sakaguchi et al. [2009], and Rae et al. [2009a, 2009b] presented evidence of small-scale arc beadings during the initiation of substorms and pseudobreakups that were characterized to have azimuthal wave numbers, m, with values $>100$. Sakaguchi et al. [2009] found that in some cases, the increases in auroral luminosity may start in a localized longitudinal region of $\sim 30$–60 km width and have longitudinal wavelengths of $\sim 100$ km. However, it is difficult to establish solely using optical data with a limited field of view (FOV) that these arc brightenings represent the first ionospheric onset of this global phenomenon since brightenings outside the FOV cannot be excluded. Indeed, not all wave-like auroral fluctuations necessarily associate with or lead to auroral breakup. Uritsky et al. [2009] showed clear evidence of periodic auroral arc intensity fluctuations observed 10–20 min before auroral onset. These fluctuations had longer wavelengths than those reported above in conjunction with onset [250–420 km compared to 70–100 km]. Uritsky et al. [2009] demonstrated that these fluctuations were consistent with in situ observations of current sheet flapping, suggesting that these pre-onset arc undulations are the ionospheric manifestation of a drift mode in a thinning magnetotail rather than being directly linked to the onset process itself.

While auroral dynamics have historically been used to characterize substorm onset on the ground, ULF waves observed by ground-based magnetometers have been shown to be intimately linked to substorms onset and have become increasingly vital in characterizing the initial seconds of substorm expansion phase onset. Traditionally, the presence of Pi2 (40–150 s period) [Jacobs et al., 1964] waves coincident with the formation of a ground magnetic bay are commonly used as evidence of substorm onset. However, it is difficult to resolve the “2 min” substorm problem with a Pi2 pulsation since the period range of this ULF wave is comparable to the uncertainty surrounding onset. Pi1B ($\sim 1$–10 s period) pulsations have been identified as localized pulsations close to the region of substorm onset [e.g., Bosinger and Yahnin, 1987; Arnoldy et al., 1987; Bosinger, 1989; Lessard et al., 2006] though typically these pulsations require a dense network of search coil magnetometers situated on the ground to be able to accurately resolve the onset timing and location of magnetic substorms. Both the Pi1B and Pi2 frequency ranges have traditionally been the most widely used for substorm timing, leaving the longer-period Pi1 (1–40 s period) range largely ignored. Recently, however, ULF waves in the band spanning from the longer-period Pi1 range (10 s period) to the shorter-period Pi2 range have been identified as the first observable ionospheric signal of ULF wave onset [Milling et al., 2008; Mann et al., 2008; Murphy et al., 2009a, 2009b; Rae et al., 2009a, 2009b]. Milling et al. [2008] were the first to demonstrate that the onset of long-period Pi1/short-period Pi2 ULF waves (hereafter termed Pi1-2 band) occur first in a localized epicenter on the ground, expanding at later times to reach larger radial distances away from this epicenter. Subsequently, Murphy et al. [2009a, 2009b] demonstrated that the onset of Pi1-2 ULF waves was spatially colocated with, but occur minutes before, the time of global auroral intensification as identified in the Frey substorm database [Frey and Mende, 2007]. Furthermore, Rae et al. [2009a, 2009b] demonstrated that the onset epicenter of Pi1-2 waves was spatially and temporally coincident with small-scale auroral fluctuations along the substorm onset arc before auroral breakup. Corroboratively, Sakaguchi et al. [2009] found in one case that there was also evidence of similar frequency ULF wave activity during a period that displayed $\sim 15$ s period arc intensity fluctuations, although this substorm onset occurred outside the field of view of the ASI.

In this paper we use the interval first discussed by Rae et al. [2009b] to investigate the possible wave modes responsible for the auroral arc structure and fluctuations observed during a clear, isolated substorm on 7 March 2007 observed by the THEMIS ASIs and ground-based observers (GBOs) [Russell et al., 2008, Peticolas et al., 2008] and the Canadian Array for Realtime Investigations of Magnetic Activity (CARISMA) [Mann et al., 2008] magnetometer network. We concentrate on the first 3 min of auroral fluctuations observed during this onset and characterize the auroral intensity as a function of both time and along-arc (approximately azimuthal) wave number and calculate their corresponding growth rates. Combined with the study by Rae et al. [2009b], this study provides the strongest constraint to date on the relationship between frequency and wave number of these auroral fluctuations. The inferred dispersion relation characteristics can be used to probe the type of instability or disturbances in the equatorial plane of closed field lines that are responsible for the initiation and generation of the structures accompanying auroral substorm expansion phase onset.

2. Auroral Instrumentation and Large-Scale Arc Morphology

In this study we use optical data from the THEMIS ASI at SNKQ (Sanikiluaq, 56.5°, 280.8° geographic latitude and longitude, respectively) in geographic coordinates projected to an assumed altitude of 110 km. Figure 1a shows the auroral fluctuations that occur in the SNKQ ASI that define the initiation of auroral expansion phase onset, together with the Gillam ASI for context as well as the onset contours of Pi1-2 ULF waves as identified by Rae et al. [2009b]. Figure 1b shows the relevant southwest portion of the SNKQ FOV that contains the auroral fluctuations. The location of any given pixel within the FOV is rotated into corrected geomagnetic coordinates (CGMs), and the emission intensities are interpolated onto a regularly spaced grid in CGM latitude and longitude. As shown in Figure 1b, the onset arc, and the periodic bead structuring within it, is closely aligned along a constant geomagnetic latitude of 63.7°. The horizontal lines in Figure 1b denote slices at constant magnetic latitude that are taken through the ASI FOV to investigate the characteristics of the auroral emission intensity as a function of longitude and time. The
auroral intensity as a function of longitude and time through the middle (dot-dashed) line is shown in Figure 1c. Clear from Figure 1c are the periodic longitudinal structures that signify both auroral and magnetic onset just before 0552 UT (the optical onset was identified to be within the 0551:54 UT SNKQ ASI frame by Rae et al. [2009b] through the visual inspection of native temporal resolution (3 s) difference THEMIS ASI images and using the Automated Wavelet Esti-
mation of Substorm Onset and Magnetic Events (AWESOME) algorithm outlined by Murphy et al. [2009a] to determine the ground magnetic onset).

3. Optical Analysis

[7] Figure 2 shows the dynamic power spectral density (PSD) as a function of longitudinal wave number ($k_{\text{lon}}$) and time computed from the data shown in Figure 1c. These data have been detrended in the spatial domain and a Hanning window has been applied in the longitudinal direction. The vertical pixel length in Figure 1c corresponds to $\sim 1.6$ km in longitude at 110 km altitude at 63.7° CGM latitude. Figure 2a shows the full range of wave numbers that is covered by the data shown in Figure 1c. The greatest values of PSD lie at low values of $k_{\text{lon}}$ throughout this interval, although it should be noted that the PSD grows for all $k_{\text{lon}}$ following the initial arc brightening and the onset of arc structuring and bead formation at $\sim 0552$ UT. Moreover, increased PSD in discrete wave numbers can clearly be seen at lower $k_{\text{lon}}$ following expansion phase onset at $\sim 0552$ UT. We show PSD for low $k_{\text{lon}}$ in more detail in Figure 2b, where we focus on the low wave number region below the white dashed line in Figure 2a. The corresponding dimensionless azimuthal wave numbers ($m$ values) for this range of wave numbers are indicated on the right-hand axis, calculated from the number of wavelengths that would fit within a circumference at this constant latitude. Figure 2b shows a peak in wave PSD at $k_{\text{lon}} \sim 10^{-4}$ m$^{-1}$ at 0552:12 UT (corresponding to $\sim 70$ km spatial scales and an $m$ number of $\sim 250$). The peak progresses to lower wave numbers around 0552:24 UT, which can be interpreted as evidence of the inverse energy cascade, consistent with the signature of an instability, identified visually by Rae et al. [2009b]. After 0552:30 UT, the wave PSD increases for all $k_{\text{lon}}$, and the amplitude of the fluctuations at these wave numbers remains high for $\sim 30$ s.

[8] Further analysis of the PSDs in Figure 2b reveals that the auroral perturbations with low wave numbers grow approximately exponentially between $\sim 0552:12$ and $\sim 0552:40$ UT, with higher wave numbers exhibiting growth before lower wave numbers. Figure 3 shows the growth rates as a function of $k_{\text{lon}}$ for each of the 5 latitudinal slices shown in Figure 1b. The time series of the amplitude of each wave number is analyzed to determine whether it contains a period of exponential growth. If the exponential growth lasts less than $\sim 30$ s duration, then that wave mode is ignored since the growth occurs for a duration that is less than the wave period. Only wave numbers below $1.6 \times 10^{-4}$ m$^{-1}$ were identified to fit this criteria. Although the growth rates

![Figure 2.](image-url)
some evidence of wave activity at higher $k_{\text{lon}} \sim 4-6 \times 10^{-3} \text{ m}^{-1}$; the PSDs are \sim 3 orders of magnitude less than those of the dominant wave modes in this interval. The discrete spectral peak between approximately 50–150 mrad s\(^{-1}\) corresponds to the bead structure that is clearly visible in Figure 1.

4. Discussion

[10] In this paper we present a detailed analysis of the spectral content and the dispersion relation of bead-like optical emissions seen to develop along the onset arc during substorm expansion phase onset. We use an interval where these bead structures are clear and where the onset of 24–96 s Pi1–2 waves has been identified at the same time and in the same location (see Rae et al. [2009b] and Figure 1b). This combined magnetic and optical signature clearly characterizes important features related to the initiation of the substorm expansion phase onset as seen in the ionosphere. The combination of ground-based magnetometer and THEMIS ASI measurements provide unprecedented temporal and spatial accuracy that clearly demonstrates that the magnetic and auroral bead fluctuations are unequivocally linked during the onset process.

[11] Figure 1c shows a very clear wave-like structure of auroral brightening along the onset arc with a frequency that lies in the Pi1–2 ULF wave band. The longitudinal auroral intensifications have a \sim 30–50 s periodicity for approximately three wave periods before auroral breakup, corresponding to a frequency in the center of the same Pi1–2 ULF wave band in which ionospheric magnetic perturbations at substorm onset are first seen [Milling et al., 2008; Mann et al., 2008; Murphy et al., 2009a, 2009b; Rae et al., 2009a, 2009b]. Further analysis of the white light auroral data shown in this paper demonstrate that there are enhancements in the amplitude of optical fluctuations with longitudinal (azimuthal) wave numbers $k_{\text{lon}} \sim 1 \times 10^{-3} \text{ m}^{-1}$ that are observed closely coincident with the Pi1–2 magnetic ULF wave onset. The optical display is subsequently dominated by larger spatial scale ($k_{\text{lon}} \sim 5 \times 10^{-3} \text{ m}^{-1}$) auroral fluctuations approximately tens of seconds later. This is consistent with the visual identification made by Rae et al. [2009b] of the auroral fluctuations occurring first at \sim 70 km spatial scales and

![Figure 3](image3.png)

**Figure 3.** Estimated growth rates $\gamma$ as a function of $k_{\text{lon}}$ for the 5 latitudinal slices shown in Figure 1b. See text for detail of the growth rate calculations.

![Figure 4](image4.png)

**Figure 4.** Auroral fluctuations as a function of angular frequency $\omega$ and azimuthal wave number, $k_{\text{lon}}$. Fluctuation amplitudes have been multiplied by $\omega$ to whiten the spectrum and reveal the details of higher frequency structure more easily.
undergoing an inverse spatial cascade to auroral vortices of \(\sim 100\) km scales. Furthermore, the exponential growth of the auroral fluctuations indicates that the ASI may be measuring an optical manifestation of a plasma instability in the equatorial plane of the magnetosphere. In this case, the ionospheric observations would be the imprint of the magnetospheric instability, whose wave properties are shown in Figures 3 and 4. Finally, analysis of the \(\omega-k_{\text{lon}}\) dispersion relation during this interval reveals a number of interesting characteristics, most notably that at low \(k_{\text{lon}}\) there is a discrete long period disturbance with \(\omega\sim50\sim150\) mrad s\(^{-1}\) that tends to dominate the spectrum, with a secondary spectral peak at \(\sim300\sim450\) mrad s\(^{-1}\), also at low \(k_{\text{lon}}\), as well as some weaker features at higher \(k_{\text{lon}}\). Our observations place a number of important constraints on the disturbances that are taking place along the auroral onset arc at the equatorward edge of the oval and in particular on the instabilities that may be responsible for initiating them.

[12] A number of plasma instabilities have been proposed to play pivotal roles in initiating magnetospheric substorm expansion phase onset (see Lui, 2004, and references therein for a comprehensive review). These include the cross-field current instability [e.g., Lui et al., 1991], shear flow and kinetic ballooning instabilities [e.g., Voronkov et al., 1997; Horton et al., 2001; Cheng, 2004], lower-hybrid drift [e.g., Yoon et al., 1994], drift kink/sausage [e.g., Zhu and Wingle, 1996], current driven Alfvénic [e.g., Perraut et al., 2000], Kelvin-Helmholtz [e.g., Yoon et al., 1996], tearing [e.g., Coppi et al., 1966], and entropy antidiffusion [e.g., Lee et al., 1998]. Note that most of these estimates of the growth rates and wave numbers have been derived using linear theory, whereas the optical undulations may be the manifestation of a nonlinear instability. Nevertheless, it is still edifying to make comparisons between the growth rates estimated from the optical data and linear theoretical predictions.

[13] To deduce information regarding the magnetospheric source of the auroral filaments, we assume that the perpendicular wave numbers of the disturbance scale along the magnetic field according to a Tsyganenko (T96) mapping and use this to estimate the likely azimuthal wave numbers of instabilities or disturbances in the equatorial magnetosphere. We also assume that the wave frequencies observed in the ionosphere are the same as those present in the magnetosphere, effectively neglecting any possible impact of dispersion or phase mixing effects as the disturbances propagate from the equatorial magnetosphere to the ionosphere. This is a reasonable assumption given that the difference between the calculated Alfvén transit times along field lines bracketing the latitudinal extent of the auroral undulations is \(\sim 3\) s, using the self-consistent model detailed in the study by Kabin et al. [2007]. We then compare the mapped wave numbers and growth rates to predictions made by the studies of various onset instability models outlined above. Using the T96 magnetic field model, Figure 5 shows the estimated source location of the auroral undulations in the equatorial plane of the magnetosphere. The azimuthal wavelength in the equatorial plane is then estimated from the average distance between the center of each feature, giving \(k_{\text{lon,m}}\sim 2.9 \times 10^{-6}\) m\(^{-1}\). From the in situ magnetic field strength of \(\sim 90\) nT observed at the approximately conjugate GOES east (GOES-12) satellite, we predict the equatorial proton gyrofrequency \(\Omega_i\sim 7\) rad s\(^{-1}\). We assume a

**Figure 5.** Estimated source location of the auroral undulations in the equatorial plane (\(Z_{\text{GSM}} = 0\)) of the magnetosphere (red). Black lines indicate contours of constant magnetic field strength in the T96 magnetic field model. Inset shows an enlarged view of the region.
number density of \( n_i = 10^6 \text{ m}^{-3} \) and an ion temperature of 10 keV in the equatorial plane, based upon general plasma sheet characteristics around geosynchronous orbit in the magnetotail [see for example Wang et al., 2006]. The radial extent of the disturbance in the magnetosphere is estimated from the mapping shown in Figure 5 assuming that the auroral emissions are at the constant altitude discussed in section 2 and in the study by Rae et al. [2009b] and spans the range \( L = 5.6 \) to \( L = 6.3 \) in the nightside equatorial plane. Thus, the azimuthal structuring has a wavelength of \( \sim 0.34 R_E \), and the radial extent is \( \sim 0.7 R_E \) in the magnetosphere.

[14] We can immediately rule out several of the proposed instabilities in the following way. The tearing instability [e.g., Coppi et al., 1966] has a slow growth rate whose e-folding time is \( \sim 1 \text{ h} \) and moreover has a radial \( k \) structure, not the longitudinal structuring observed in our case study. Similarly, the drift kink/sausage instability [e.g., Zhu and Winglee, 1996] also predicts radial structuring and a growth rate that is too low. The current-driven Alfvénic instability [e.g., Perraut et al., 2000] excites ULF waves with frequencies of the order of the proton gyrofrequency and has growth rates that peak at low \( k_{\text{lon,m}} \). However, both the predicted growth rates and frequencies are larger by an order of magnitude than those observed during this case study. Similar considerations rule out the lower-Hybrid drift instability [e.g., Yoon et al., 1994].

[15] The cross-field current instability [Lui et al., 1991] predicts growth rates that peak at low azimuthal wave numbers with remarkably similar \( k_{\text{lon,m}} \) and growth rates \( \gamma \) to those observed during this interval. Moreover, the frequencies of the waves with maximum growth rates are also extremely similar to those we observed in the ionospheric auroral forms. We must therefore regard this instability as being a strong candidate for explaining our observations, for initiating auroral substorm onset, and perhaps even magnetospheric substorm expansion phase onset. The shear flow ballooning instability (SFBI) outlined by Voronkov et al. [1997] also generally has properties consistent with those revealed by the analysis presented in this paper; the predicted perpendicular wave numbers are in the azimuthal direction with \( k_{\text{lon,m}} \sim 3.7 \times 10^{-6} \text{ m}^{-1} \). This compares favorably with the dominant growing mode \( k_{\text{lon}} \sim 2.9 \times 10^{-6} \text{ m}^{-1} \). The growth rates predicted for the SFBI in Voronkov et al. [1997] are peaked at low \( k_{\text{lon,m}} \) and are similar in magnitude to those observed in the ionosphere \( (\gamma \sim 0.07 \text{ s}^{-1}) \) theoretically as compared to \( \gamma \sim 0.08\text{–}0.12 \text{ s}^{-1} \) estimated from the auroral observations. Crabtree et al. [2003] predicted similar growth rates for the kinetic ballooning instability, although the growth rates are dependent on the size of the analytic domain and so cannot be tested here without further information. We therefore also regard the ballooning instability as being another potentially strong candidate for the initiation of auroral expansion phase onset and perhaps also substorm expansion phase onset.

[16] Interestingly, this ionospheric data set provides a number of other potential constraints on the possible physical processes that occurred during this substorm onset. The monochromatic nature of the optical beading is suggestive of a resonant process that selects a combination of wavelength and wave number according to the field line and plasma configuration in the inner magnetosphere. Although the general motion of the auroral fluctuations along the onset arc is ultimately poleward, we can interpret the individual periodic stripes as phase fronts of a wave. The time series shown in Animation S1 of Rae et al. [2009b] demonstrates that each phase front moves eastward and equatorward at the beginning of the interval. This behavior can be described by a model of a high-\( m \) field line resonance that includes an internal energy source [e.g., Mann, 1998]. A local energy source in the resonant region could be a stretched field line topology, a large pressure gradient, or a region containing shear flow that could couple to a ballooning mode [e.g., Voronkov et al., 1997]. We use a self-consistent ULF eigenmode model in a realistic magnetic field geometry [Kabin et al., 2007] to compute the fundamental toroidal and poloidal eigenfrequencies of field lines that span the geomagnetic latitudes of the auroral undulations. We find that the fundamental toroidal modes have periods of between 31 and 39 s and the fundamental poloidal modes of these field lines have periods of between 66 and 86 s (for an equatorial number density of \( 10^6 \text{ m}^{-3} \) and a density variation along the field that \( \propto r^3 \)). Since the period of the auroral undulations shown in Figure 1b are \( \sim 30\text{–}60 \text{ s} \) and the first observable magnetic signatures are in the 24–96 s period band, this raises the intriguing possibility that the frequency and wave number of these auroral undulations may be selected via resonant processes along the geomagnetic field.

[17] Given the paucity of in situ measurements during this interval, several instabilities cannot be ruled in or out as being responsible for features seen in the onset arc for this particular auroral substorm expansion phase onset because their frequencies or growth rates strongly depend upon magnetospheric parameters that are unknown for this case study. The Kelvin–Helmholtz instability [e.g., Yoon et al., 1996] is predicted to have growth rates that peak at low \( k_{\text{lon}} \) and are comparable to those observed during this interval, although this prediction relies on the scale length of the velocity shear corresponding to \( \sim 20\% \) of the scale length of the current sheet before expansion phase onset. The frequencies predicted by Yoon et al. [1996] are significantly larger than those observed during this case study, although again these frequencies will vary with magnetospheric parameters. Entropy antidiffusion also predicts an e-folding time of \( \sim 15 \text{ s} \) [e.g., Lee et al., 1998] that is consistent with the growth rates of the auroral fluctuations observed in this case but provides no information on the frequency content of the excited ULF waves. However, it must be noted that candidate magnetospheric instabilities do not just have to predict the correct structure and growth rate as deduced from the ground-based observations, but they would also have to account for the monochromatic nature of the auroral undulations along the onset arc as discussed above.

[18] With regard to the likely mechanism that provides the communication between the equatorial magnetosphere and ionosphere, Rae et al. [2009b] postulated that this rapid communication could be explained by the acceleration of electrons via shear Alfvén waves [e.g., Watt et al., 2005; Watt and Rankin, 2009]. Observationally, both Semeter and Blixt [2006] and Semeter et al. [2008] presented clear evidence that some auroral displays during substorms are intimately associated with dispersive Alfvén waves. However, in this interval the perpendicular wave number are much lower than those associated with dispersive Alfvén waves. When the perpendicular wave numbers are mapped
along the geomagnetic field to the auroral acceleration region and the equatorial magnetosphere, we find that the dominant perpendicular scale lengths are too large to account for the formation of parallel electric fields and subsequent electron acceleration [Watt and Rankin, 2007]. Although the spatial and temporal resolution of the THEMIS ASI may not be able to observe the finer scale structuring that could be contained within a larger scale arc morphology, other processes may be at work that can explain the rapid connectivity between the equatorial plane and the ionosphere. Electron time scale connectivity could still be the key, but the acceleration of the auroral electrons is not necessarily linked to dispersive shear Alfvén waves. Clearly, more work is required to probe this connectivity.

The evidence presented in this paper and in that of Rae et al. [2009b] demonstrates clearly that there is a plasma instability operating at onset in a region that is closely conjugate to the inner edge of the plasma sheet that manifests itself as both optical and magnetic fluctuations in the Pi1-2 ULF wave band. In the framework of the CD paradigm, this near-Earth plasma sheet instability causes current disruption, diverting the cross-tail current into the ionosphere, forming the Substorm Current Wedge (SCW) and energizing the aura [e.g., Lui, 2004]. Unfortunately, this is not a testable hypothesis in this case study.

It has been proposed that instabilities in the near-Earth magnetosphere are instigated by the braking of plasma sheet flows generated by reconnection at the NENL (e.g., the NENL substorm onset paradigm). Kepko et al. [2009] presented a case study whereby a faint equatorward propagating localized (2° in latitude) auroral arc was observed in the red line auroral emissions that was only detected with the THEMIS ASI cameras in retrospect. This N–S arc was interpreted as evidence of earthward moving fast flows in the tail. Nishimura et al. [2010] and Lyons et al. [2010] presented statistical evidence of the presence of N–S equatorward moving arc structures during a large number of auroral onset events several degrees poleward of the onset arc and some 2–8 min before onset. In their statistical model, the N–S arc is diverted westward or eastward in the general direction of convection, and once the new plasma has reached sufficiently far into the inner magnetosphere, the onset arc brightens and poleward auroral expansion ensues. This case study implies that the earthward moving fast flow would result in a plasma instability that would cause the initiation of the auroral substorm with characteristics as detailed in this paper. The source of the free energy for this plasma instability could be the magnetic shear and flow diversion ahead of the fast flow region suggested by Kepko et al. [2009]. However, we have good optical coverage poleward of the onset arc that verifies that no such visible equatorward moving N–S arc is evident in the THEMIS ASI data as studied by Nishimura et al. [2010] and Lyons et al. [2010]. Further, Rae et al. [2009b] demonstrated using 3 s difference images from this interval that the region poleward of the onset arc contained no significant auroral movement before the initial auroral brightening. This therefore leads to the conclusion that this substorm onset was not initiated via NENL-like mechanisms.

It is important to note that the observations shown in this paper and that of Rae et al. [2009b] are entirely consistent with the near-geophysical onset (NGO) model proposed and outlined by Maynard et al. [1996a, 1996b, 1997, 1998] and Erickson et al. [2000] using the CRRES satellite. In the proposed NGO event sequence, drift waves grow in amplitude within the background dawn-dusk convection electric fields. Once the wave amplitude grows sufficiently large, significant field-aligned Poynting flux is launched toward the ionosphere, reflecting and either constructively or destructively interfering in the equatorial plane of the magnetosphere. If constructive interference occurs, a much larger burst of wave energy is launched along the field toward the ionosphere and a substorm occurs. Conversely, if destructive interference occurs, a pseudobreakup results. Subsequent to this set of events (Erickson et al. [2000], stage 1) is the formation of the substorm current wedge, electrojet intensification and dipolarization, and harnessing the energy stored in the lobe (see Erickson et al. [2000], and references therein, stage 2). This is consistent with the observations shown in this paper and in that of Rae et al. [2009b] that demonstrated that the onset time of Pi1-2 ULF waves occurred approximately minutes before the formation of the SCW and dipolarization in space. Perhaps more intriguingly, Maynard et al. [1996a] and Erickson et al. [2000] both identified the 30–100 s period range (corresponding to the Pi1-2 range defined here) in the near-Earth magnetosphere as being the critical ULF wave frequency range for substorm onset to occur. Specifically, in the one event that there was sufficient temporal resolution to resolve all frequency bands throughout the Pi1 and Pi2 ULF wave bands, Maynard et al. [1996b] found that there was only significant parallel Poynting vector in the 30–100 s period band, which is remarkably similar to the 24–96 s ULF wave band that has recently been shown to be the first observable magnetic signature observed on the ground. It is also interesting to note that the characteristics of the geomagnetic field play a pivotal role in the NGO onset mechanism, which may provide evidence that the mechanism responsible for the magnetic and auroral fluctuations may be selected via resonant processes along the geomagnetic field. Clearly, more work is required in order to verify a number of these points.

5. Conclusions

In this paper we present detailed analysis of ground-based auroral white light THEMIS All-Sky Imager data to examine the spatiotemporal and spectral characteristics of the auroral disturbances that develop along substorm onset arcs. By assuming a direct connection of these disturbances to the equatorial plane, we further examine the plasma instabilities that might have been responsible for triggering the auroral substorm expansion phase onset. We analyze the auroral intensity as a function of longitude and time and time to demonstrate that the auroral fluctuation amplitudes grow between ~0552 and 0553 UT. Analysis of the separate $k_{||}$ modes reveals that the growth is approximately exponential across low wave numbers aligned along the arc, suggesting that the auroral display is the optical manifestation in the ionosphere of the growth of a magnetospheric instability.

Our results have clear ramifications for the reconnection versus current disruption debate and more and combined with the results shown in the study by Rae et al. [2009b], we can assert the following:
[24] (1) The first signatures in the ionosphere of substorm expansion phase onset are marked by the contemporaneous and colocated signature of Pi1-2 magnetic and optical ULF waves (this paper and that of Rae et al., 2009b), which is remarkably consistent with the in situ observations shown by Maynard et al. [1996a] and Erickson et al. [2000].

[25] (2) If the arc structure is directly linked to a magnetospheric instability, then (assuming negligible dispersion or phase mixing) it will have a frequency \( \omega \sim 0.1 \text{ rad s}^{-1} \) (\( \sim 60 \text{ s period} \)), azimuthal wave number \( k_{\text{az},\text{m}} \sim 3 \times 10^{-6} \text{ m}^{-1} \) in the equatorial magnetosphere and very fast growth rates \( \gamma \sim 0.1 \text{ s}^{-1} \).

[26] (3) The most likely plasma instability that could be responsible for the initiation of this auroral substorm expansion phase onset is either the cross-field current instability or the shear flow or kinetic ballooning instability, although we cannot clearly rule out Kelvin-Helmholtz or entropy antidiffusion based upon theoretical predictions in the literature.

[27] (4) Any model of magnetospheric and auroral substorm expansion phase onset must be able to explain the specific structuring and morphology of the onset arc observed in the ionosphere. In any near-Earth paradigm (CD or NGO), these instabilities would represent magnetospheric substorm expansion phase onset. In the case of the NENL paradigm, these plasma instabilities would need to be triggered by the braking of earthward propagating plasma sheet flows. Indeed, if this is the case it could be that the instabilities required by the CD model might still occur within the paradigm of the NENL model following flow braking in the near-Earth region, their features being as shown here.

[28] (5) The monochromatic nature of the optical undulations, their equatorward phase propagation and the close similarity between the frequencies of the optical and magnetic perturbations and modeled fundamental field line eigenfrequencies may be explained via a field line resonant process.

[29] In conclusion, we have analyzed ground magnetic and optical measurements of a clear, isolated auroral substorm onset in detail to determine the likely plasma instability responsible for its initiation. There are several approximations within this study, most notably the assumption of a ~10 keV ion temperature in the plasma sheet, which cannot be verified with in situ measurements in this case study. It is the synthesis of ground-based and plasma sheet measurements during an isolated substorm using the THEMIS satellite constellation that will provide important further refinements to these estimates, fully establishing the likely instability responsible for auroral substorm expansion phase onset. The combination of measurements is vital for fully determining the time sequence of events during the expansion phase of the magnetospheric substorm.

References


Heppner, J. P. (1958), A study of the relationships between the auroral borealis and the geomagnetic disturbances caused by electric currents in the ionosphere, DR135, Defence Research Board of Canada, Ottawa, Ontario, Canada.


