



OPEN ACCESS

EDITED BY
Xu-Zhi Zhou,
Peking University, China

REVIEWED BY
Joseph E. Borovsky,
Space Science Institute (SSI), United States
George K. Parks,
University of California, United States

*CORRESPONDENCE
Yan Song,
✉ yan@umn.edu

RECEIVED 18 September 2023
ACCEPTED 31 October 2023
PUBLISHED 15 November 2023

CITATION
Song Y and Lysak RL (2023), A
mechanism of the auroral substorm
expansion onset: electric discharge in
the double layer.
Front. Astron. Space Sci. 10:1296626.
doi: 10.3389/fspas.2023.1296626

COPYRIGHT
© 2023 Song and Lysak. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which does
not comply with these terms.

A mechanism of the auroral substorm expansion onset: electric discharge in the double layer

Yan Song* and R. L. Lysak

School of Physics and Astronomy, University of Minnesota, Minneapolis, MN, United States

Magnetospheric substorms often occur when a significant amount of energy from the solar wind is deposited and stored in the magnetotail during the growth phase and release explosively in the expansion phase, which accelerates charged particles causing the rapid auroral intensification. A physical mechanism is needed to release the energy explosively. The formation of double layers is a likely mechanism for the energy release and the acceleration of particles and triggers the onset of the expansion phase. We suggest that the localized parallel electric field that forms the double layer results from the displacement current complying with Ampere's law for the dynamic case. The double layers are embedded in low density cavities surrounded by enhanced magnetic stresses. Positive feedback in the electric field generation may cause rapid release of the accumulated energy. The Poynting flux carried by Alfvén waves continuously supplies the energy to maintain strong electric fields during the rapid development of auroral substorms.

KEYWORDS

substorm onset, auroral substorm, double layer, displacement current, parallel electric field

1 Introduction

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 (Rostoker et al., 1980). Based on observations of auroras from All-Sky camera data Akasofu (1964) introduced a concept of the auroral substorm which is the manifestation of the magnetospheric substorm process and is one of the spectacular displays of the explosive energy release and charged particle acceleration phenomena that frequently occur in cosmic plasmas. The substorm onset is defined by the onset of initial brightening of the breakup arc. Understanding the formation of substorm aurorae and their evolution is very important in understanding the mechanism of the substorm process.

The substorm process consists of the growth phase, the expansion phase and the recovery phase. The basic physical processes that lead to substorm onset include the following:

The growth phase is a period of enhanced magnetic energy storage in the plasma sheet. During the growth phase (e.g., McPherron, 1972), energy and momentum transfer from the solar wind into the magnetosphere through the interaction between the solar wind and magnetosphere, which occurs throughout the magnetopause current sheet, stretches the magnetotail and accumulates energy in the tail. Accumulation

energy is a necessary condition to drive the substorm process. Also, changes of tail configuration (thinning of the plasma sheet) will drive the magnetosphere toward an unstable state (Hsu and McPherron, 2012).

Before the expansion onset, a decrease in the momentum input from the solar wind and other interactions may break the force balance on the stressed magnetotail producing a strong earthward body force acting in the whole magnetotail in a short time (Lin et al., 2009). As a consequence, the whole magnetotail tends to return to a more dipolar configuration releasing the previously stored free magnetic energy and initiating fast mode waves and shear plasma flows. This process provides favorable conditions for the excitation and intensification of magnetotail disturbances, which may develop into magnetospheric substorms. The braking of the earthward flows that leads to the pileup of northward magnetic flux carried by the flow in the near-Earth region (Shiokawa et al., 1998) may enhance the accumulation of magnetic energy in this region. The energy required for the auroral substorm before the onset is mainly stored in the magnetotail within $X_{\text{GSM}} \sim -10R_E$ (Akasofu, 2015; 2017).

The expansion phase is a period of large-scale energy release in the magnetosphere-ionosphere coupling system, which accelerates auroral particles causing the substorm auroral intensification and poleward expansion. In the process the stored electromagnetic energy is converted into the kinetic energy of charged particles in the acceleration region.

The onset mechanism of magnetospheric substorms has been explored extensively. Several substorm onset models have been proposed, which include the near-Earth neutral line (NENL) or reconnection model (e.g., Hones, 1984; Baker et al., 1996), the cross-field current disruption model (Lui, 1996, 2004), the ballooning mode instability model (Roux et al., 1991; Cheng, 2004), the convection reduction model (e.g., Lyons, 1995), and magnetosphere-ionosphere coupling models (e.g., Haerendel, 1992; Kan and Sun, 1996; Haerendel, 2015). In a review Lui (2004) classified the various substorm onset mechanism into two categories: one category with near Earth initiation of substorm process, in which the expansion onset is related to the plasma instabilities, such as the cross-field current instability and kinetic ballooning instability (Cheng, 2004), initiated on magnetic field lines linked typically to the most equatorward auroral arc (from near the geosynchronous orbit to $\sim 10 R_E$). The second category, represented by NENL model, with mid-tail initiation of substorm onset, link the mid-tail activity to near Earth disturbances through fast plasma flows resulting from magnetic reconnection at $\sim 20\text{--}30 R_E$ down the tail. In this model (see e.g., Baker et al., 1996) reconnection (presumably due to the growth of a tearing instability) initiates during the late growth phase in the region of thinned plasma sheet. The strong earthward flows caused by the reconnection drive the field aligned current and disturbances on the field lines that related to the auroral sudden brightening and power the substorm onset. A comparison between the two categories of substorm onset mechanisms was discussed in Lui (2004); Cheng (2004), as well as Baker et al. (1996).

More recent studies of substorms have suggested various processes of how the energy is accumulated prior to the auroral breakup developed near the equatorward border of the auroral oval. Sergeev et al. (Sergeev et al., 2011; Sergeev et al., 2012) suggested that prior to breakup, the bursty bulk flow activity initiated by magnetotail reconnection approaching the inner magnetosphere

may influence parameters in the outer portion of the dipole-like region, causing auroral breakup. It has also been suggested that there is a precursor auroral sequence prior to substorm auroral onset (e.g., Nishimura et al., 2010; Ohtani et al., 2022). The sequence is initiated by a poleward boundary intensification and followed by a north-south arc (auroral streamer) moving equatorward toward the onset latitude. The onset is preceded by enhanced earthward plasma flows, and the Earthward-transport of the new plasma leads to a near-Earth instability and auroral breakup. Akasofu (2015) suggested that much of the Poynting flux produced on the magnetopause is directed toward the dipole-like main body of the magnetosphere (within $10 R_E$). When the stored energy reaches 2×10^{22} erg in the main body of the magnetosphere, the magnetosphere becomes unstable, and unloads it in the process of stabilizing itself, which is manifested as the expansion phase.

In all of the above scenarios, energy accumulation during the growth phase is a necessary condition to develop substorms. A physical mechanism is needed to release the energy explosively at the onset latitude. Since the auroral sudden brightening at the substorm onset indicates the acceleration of charge particles, the electric field generation is an important element in onset mechanisms. For example, Cheng (2004) applied the kinetic effect in the ballooning instability to show the formation of enhanced perturbed parallel electric field associated with a charge separation due to electrons and ions move differently in the direction perpendicular to the ambient magnetic field. This paper will focus on the mechanism of the expansion onset by emphasizing the generation of field aligned currents in auroral regions in the form of double layers, and show how the explosive energy release occurs and accelerates auroral particles to high energy.

2 The acceleration of charged particles and the role of the displacement current

Charged particles can only be accelerated by electric fields. Electrostatic fields parallel to the magnetic field, E_{\parallel} , which are related to the charge-separation are the most powerful means to efficiently accelerate charged particles.

In previous studies [e.g., Song and Lysak, 2001, 2006] we have developed a dynamical theory which explains how the parallel electric field can be generated under certain conditions, especially in low density plasmas with varying fields. We pointed out that the displacement current, $\partial E/\partial t$, in the Ampere's law, which has often been ignored in the discussion of global magnetospheric dynamics, is the key to the generation of the electric field in low density plasma. It is the parallel displacement current that represents the establishment of the parallel electric field. Maxwell's addition of the displacement current to the steady state Ampere's law is of crucial importance under dynamic plasma conditions where rapidly changing fields and explosive energy releases occur. Without the displacement current there would be no electromagnetic radiation, and there would be no electric fields generated. Actually, in the discussion of time-varying fields in electrodynamics (Feynman et al., 1963) it has been pointed out that we must give up the idea that E is zero in conductors: "When the fields are changing, the charges in conductors do not, in general, have time to rearrange

themselves to make the field zero. They are set in motion, but never reach equilibrium." So, in rapidly varying fields, a magnetic field line is not an equipotential. A localized potential drop may form in plasmas that are inhomogeneous in space and vary rapidly in time. Based on the above theoretical considerations, we will describe in the following sections how parallel electric fields and associated potential drops are generated from enhanced displacement currents under certain conditions within plasma structures called "double layers" that accelerate charged particles in the auroral region, which may cause the explosive energy release and trigger the onset of auroral substorms. The generation of the quasistatic electric field by the displacement current provide a possible mechanism for efficiently accelerating charged particles. There are also other mechanisms to generate electric fields without involving the displacement current (e.g., Block, 1978; Fälthammar, 1983; Lysak et al., 2020), which may also play important roles for charged particle acceleration and auroral formation under various conditions.

3 The double layer as a mechanism for explosive energy release and auroral particle acceleration

The displacement current which represents the generation of the electric field in the parallel direction can be expressed from Ampere's law as

$$\partial E_{\parallel} / \partial t = c(\nabla \times \mathbf{B})_{\parallel} - 4\pi J_{\parallel} \quad (1)$$

where the field aligned current $J_{\parallel} = nev$, while n , and v are the electron density and drift velocity, respectively. A magnetic shear, $\nabla \times \mathbf{B}$, that is not balanced by the field aligned current (FAC) may establish the parallel electric field before the electrons respond. Song and Lysak [2006] have derived that the ratio between the displacement current and the particle current, $\Gamma = |(1/4\pi) \partial E_{\parallel} / \partial t / J_{\parallel}| \propto \gamma^2 / n$, where the electric field $E_{\parallel} \approx (4\pi / \omega_{pe}^2) (\partial J_{\parallel} / \partial t)$ is assumed, $\omega_{pe} = (4\pi ne^2 / m_e)^{1/2}$ is the electron plasma frequency, and γ is used to denote the time derivative. This implies that at low density, for a certain $(\nabla \times \mathbf{B})_{\parallel}$, when the plasma is too tenuous to carry the J_{\parallel} , $\partial E_{\parallel} / \partial t$ becomes more significant, although it may look relatively small. For example, a displacement current of $10^{-6} \mu\text{A}/\text{m}^2$, which is much smaller than typical auroral currents of $\sim 1 \mu\text{A}/\text{m}^2$, may represent the establishment of $E_{\parallel} = 100 \text{ mV}/\text{m}$ in 1 s, which is significant in accelerating charged particles.

With the generation of parallel electric fields, localized potential drops along magnetic fields can be formed. For example, in a piece of magnetic flux tube in the auroral acceleration region with radius R and height h , the potential drop $V_{\parallel} \Delta S = - \int dS \int_0^h E_{\parallel} dz$, where ΔS is the cross section of the flux tube and z is in the direction of the magnetic field. Such a special structure with E_{\parallel} and V_{\parallel} established in it has been called a **double layer** (e.g., Block, 1978; Alfvén, 1986). It is within such structures that charge particles are accelerated, and the accumulated magnetic energy is converted to the kinetic energy of charged particles. Note that while the formation process of the double layers involves the inductive contribution due to the energy input from the Poynting flux of Alfvén waves, the parallel acceleration of electrons and ions is primarily electrostatic. This

means that the inductive field is important in setting up the E_{\parallel} ; if a steady input of energy is provided then the system will settle into an electrostatic state.

We also notice that the displacement current (or more precisely, its divergence) is related to the formation of charge separation, since from Gauss's law we have $d\rho/dt = (1/4\pi) \nabla \cdot (d\mathbf{E}/dt)$, where ρ is the charge density. Thus, the displacement current in the structure may change the distribution of the charge density and play an important role in the formation of charge separation, which causes the quasi-neutrality of the plasma in the structure to be locally violated. Since the plasma tends to be quasi-neutral, the quasi-neutrality condition should hold over most of the field line and quasi-neutrality can be violated only in localized regions. This implies that in the M-I coupling system, the existence of double layers (DLs) is localized in areas which satisfy certain plasma conditions. As summarized in the previous work (e.g., Block, 1978), certain conditions must be fulfilled for the formation of double layers, such as that the electric field is much stronger inside the DL than outside, so the integrated positive and negative charges nearly cancel each other; a minimum current is necessary to form a DL (the Bohm condition). Such conditions can only be satisfied in confined areas.

As we suggested above, the displacement current play a crucial role in the generation of the parallel electric field that is associated with the charge separation in the low density plasma. The displacement current would become significant when the magnetic shear cannot be balanced by the charged particle flow and enhances the parallel electric field, forming the double layer. It implies E_{\parallel} forms for large B/n , favoring low densities. The above conditions are often met in the auroral density cavity at altitudes of $1 R_E$ or less where the magnetic field is strong and the density is low (Chaston et al., 2006). Space observations of parallel electric fields and potential drops at several thousand km above the ionosphere (e.g., Ergun et al., 2004) and the estimated thickness of the double layer (see for example, Table 1 in Block, 1978, and review by Lysak et al., 2020) have shown that the formation of double layers is localized in regions where certain conditions are met.

The fact that parallel electric fields that associated with field aligned currents form double layers rather than extended parallel electric fields has been shown in numerical simulations (e.g., Borovsky and Joyce, 1983; Génot et al., 2001; Newman et al., 2001). Considering that in the auroral region it is likely that the plasma initially carries a current due to larger-scale processes, and that this current drives the formation of a DL at the site of some inhomogeneity, Newman et al. (2001) did simulations by imposing a current, and showed that the parallel electric field develops as a double layer. What that simulation did not show is how the current is formed. We show that the current is required by the twisting of the magnetic field that is discussed in the following section 4.

4 The required large magnetic stress (twisted flux tube) and its relation with the vorticity in the plasma flow

The required large $\nabla \times \mathbf{B}$ represents a localized magnetic stress concentration. From Ampere's law in cylindrical coordinates

$(\nabla \times \mathbf{B})_{\parallel} = (1/r)\partial(rB_{\phi})/\partial r = (4\pi/c)J'_{\parallel}$, where $J'_{\parallel} = J_{\parallel} + (1/4\pi)\partial E_{\parallel}/\partial t$, we know that the total current in the system, including the displacement current as well as the particle current, is closely related to the azimuthal component of the magnetic field B_{ϕ} , i.e., the twist of the magnetic flux tube. The relationship can be expressed as (Song and Lysak, 2001; 2006) the volume integration of J'_{\parallel} over the flux tube

$$\langle J'_{\parallel} \rangle \cdot h = \int_0^h dz \int_0^R J'_{\parallel} 2\pi r dr \approx (c/2) \int_0^h dz \int_0^R B_{\phi} dr \approx (c/2)\Phi_B \quad (2)$$

where $\Phi_B = \int_0^h dz \int_0^R dr B_{\phi}$ is the azimuthal magnetic flux through the vertical area hR of the flux tube. Eq. 2 can be rewritten as $\partial \langle V_{\parallel} \rangle_S / \partial t = \langle I_{\parallel} \rangle (4\pi \cdot h / \Delta S - (2\pi c / \Delta S)\Phi_B)$, where $\langle V_{\parallel} \rangle_S$ is the average of the potential drop over the cross section of the flux tube ΔS , which shows that the localized azimuthal magnetic flux is needed to support the voltage drop in the flux tube.

We have further pointed out that the magnetic shear that leads to large $\nabla \times \mathbf{B}$ can be formed by the vorticity in the plasma flow: $\partial \Phi_B / \partial t = (1/2\pi) B_{\parallel} O_{\parallel}|_0^h$, where

$$O_{\parallel} = \int_0^h dz \int_0^R \Omega_{\parallel} 2\pi r dr \approx \int_0^h dz \int_0^R 2\pi u_{\phi} dr \approx 2\pi \Phi_u \quad (3)$$

is the volume integration of the vorticity $\Omega_{\parallel} = (\nabla \times \mathbf{u})_{\parallel}$, and Φ_u is the azimuthal velocity flux. For auroral flux tubes, the vorticity that causes the azimuthal magnetic flux can be produced by shear flows at the equatorial region, together with the drag force at the ionosphere in the magnetosphere-ionosphere coupling system. Thus, the storage and release of the free magnetic energy and the kinetic energy of plasma bulk flow are closely related to the local generation of E_{\parallel} and V_{\parallel} .

5 The main physical processes and conditions in generating electric field and potential drop in the acceleration region

We have shown that the space plasma under certain conditions can generate and maintain localized electric fields and potential drops along magnetic field lines, which will accelerate charged particles. The main physical processes in such special plasma regions and their occurrence conditions can be summarized as the following:

(1) Shear flows in the plasma produce vorticity and thus enhance the azimuthal magnetic flux (magnetic shear) as expressed in Eq. 3. In the Earth's magnetosphere-ionosphere coupling system, the shear flows can be caused by earthward flows accumulating magnetic energy in the near-Earth region before the expansion onset as described in Introduction section. The enhanced azimuthal magnetic flux and the drag force at the ionosphere cause a twist of the auroral magnetic flux tube and generate a localized magnetic shear ($\nabla \times \mathbf{B}$) associated with the propagation of Alfvén waves. (2) The generation of parallel electric fields as well as the associated FACs, which comply with the Ampere's law in dynamically varying magnetized plasmas, and the formation of the associated charge separation and potential drop along magnetic field lines as described in Section 3. In a flux tube with a field-aligned current, if there is a spot in the flux tube where the plasma particles have difficulty

carrying that current a charge separation will occur. This also means that if the current carriers go above the Bohm criterion (the drift velocity greater than the thermal speed) (Block, 1978) there is an instability that will give rise to a charge separation forming double layers. (3) Subsequent acceleration of charged particles due to the rapid establishment of electric fields and potential drops, which can be explosive. The formation of density cavities due to the acceleration of particles away from the acceleration region may produce positive feedback causing rapid development of the electric field.

Since low density favors the generation of the electric field, the structures are often nested in a low-density cavity and surrounded by enhanced magnetic stresses. Such density cavities have been associated with the acceleration of particles by kinetic Alfvén waves (Chaston et al., 2006). During this process, the initially generated localized electric fields deepen the seed low density cavity by accelerating charged particles out of the structure and lowering the density, which can in turn further enhance the generation of electric fields. Such a positive feedback process may rapidly develop a powerful parallel electric field, and the structure will act as an efficient high energy particle accelerator. The process can be explosive, during which a large amount of electromagnetic energy is rapidly converted into the kinetic energy of charged particles. The explosive energy release that accelerates charged particles to high energy is a form of electric discharge associated with the charge separation in the double layer. In the auroral acceleration region, such explosive energy release and associated accelerations of keV electrons may lead to auroral substorms. The energy required for the auroral substorm has accumulated before the onset in the near Earth magnetotail regions. The formation of localized parallel electric fields in the double layer and the subsequent acceleration of charged particles forming substorm auroral arcs trigger a further release of the tail energy in the M-I coupling system.

The special plasma regions characterized by the localized parallel electrostatic field and the associated potential drop have been studied extensively as double layers (see for example, the review by Block, 1978; Carlqvist, 1982; Lysak et al., 2020). Several mechanisms for supporting the parallel electric fields have been proposed (e.g., Fälthammar, 1983). Based on the theoretical considerations described above, we suggest that the parallel electrostatic field which forms the double layer structure and most efficiently accelerates charged particles results from the displacement current that becomes significant under locally enhanced magnetic shears which are not balanced by FACs. The acceleration of the charged particles may also be carried out in multiple double layers in the system. The double layer formation and the associated quasi-electrostatic acceleration mechanism is of a dynamical nature. In rapidly varying plasmas, the existence time of an individual DL may be short compared to the observed explosive energy release phenomena such as the substorm auroral intensification, but the acceleration of charge particles could last much longer time due to a large number of DL potential steps that the particles experience and the continuous formation of DLs in the area. For example, Mozer et al. (2013) reported that huge numbers of double layers carrying electric fields parallel to the local magnetic field line have been observed in connection with relativistic electron acceleration in the Earth's outer radiation belt. It is very likely that similar multiple DL acceleration also occurs in the auroral region.

While the acceleration of charged particles by double layers plays an important role in the explosive energy release that starts the expansion onset of auroral substorm, the formation of double layers may have occurred for some times before the onset, accelerating electrons down to the ionosphere and causing preexisting auroral arcs. The intensification of preexisting arcs has been observed as a wave-like action with large longitudinal wave number between 100 and 300 (e.g., Donovan et al., 2006; Liang et al., 2008). The expansion onset often manifests itself as a sudden brightening of a preexisting arc. It has been suggested that the intensification of the preexisting arc that leads to the expansion onset is attributed to some plasma instabilities, such as, among others, the kinetic ballooning instability (Cheng, 2004) and cross-field current instability (Lui, 2004) in near Earth region. More recent observations by THEMIS all sky images and the FAST spacecraft (Jiang et al., 2012) have shown that the precipitating electrons are associated with preexisting arcs which are linked to an equatorial region near 8 to 10 R_E , and are accelerated by field aligned potential drops. It indicates that there have been double layers existing for many minutes before substorm onset. The wave-like intensification of the preexisting auroral arcs mentioned above may also be caused partly by variations of the conditions for the double layer formation (for example, the strength of the magnetic field and thus the azimuthal component of the field, and the plasma density in the flux tube.) due to the variations of parameters and propagation of the instabilities. This possibility needs further study.

6 Conditions sustaining the presence of parallel electric fields in the double layer

6.1 Energy supply by poynting flux

The acceleration of charged particles in double layers converts electromagnetic energy into the kinetic energy of charged particles. An important characteristic of the double layer is that the charged particles that are subject to a strong electric force run away (Fälthammar, 1978). This helps to support a voltage drop, while on the other hand, it carries away the power released in the structure. To replenish the energy and sustain the electric field, a continuous energy supply to the structure supported by the large-scale dynamics is needed. This energy is most easily provided by the Poynting flux associated with the passage of Alfvén waves. In the auroral acceleration region, the propagating Poynting flux of Alfvén waves carries both the free magnetic energy and kinetic energy of plasma bulk flow from the generator to the acceleration region and the ionosphere (Song and Lysak, 2001). Alfvén waves have been observed on auroral field lines carrying sufficient Poynting flux to provide the energy to power the aurora (e.g., Wygant et al., 2000; Lysak and Song, 2003).

6.2 Input from varying magnetic shear

The Poynting flux of electromagnetic energy flowing into the acceleration region is balanced not only by the energy dissipation which includes plasma heating and the kinetic energy

of plasma flows, but also by the increase of the energy of the azimuthal magnetic field, since the magnetic shear may reduce after acceleration of charged particles (energy release) and reduce the displacement current. The input from varying magnetic shear is required for the sustained presence of parallel electric fields, as we can see by taking the time derivative of Ampere's law together with the electron equation of motion assuming $E_{\parallel} \approx (4\pi/\omega_{pe}^2)(\partial J_{\parallel}/\partial t)$ as in Section 3:

$$\left[1 + \left(1/\omega_{pe}^2\right) \frac{\partial^2}{\partial t^2}\right] E_{\parallel} = \left(c/\omega_{pe}^2\right) \frac{\partial}{\partial t} (\nabla \times \mathbf{B})_{\parallel} \quad (4)$$

During the process, we note that there is a distinction between the formation phase and the final state of the DL formation. In the formation phase the linear term in the electron equation of motion is most important, but once the double layer is formed, the nonlinear inertial term becomes important. This term just says that the electron kinetic energy is increased as a function of distance by the double layer electric field. In the electron equation of motion which can be written as

$$nm\partial v_{\parallel}/\partial t + \partial(nmv_{\parallel}^2/2 + p)/\partial z = -eE_{\parallel}/m, \quad (5)$$

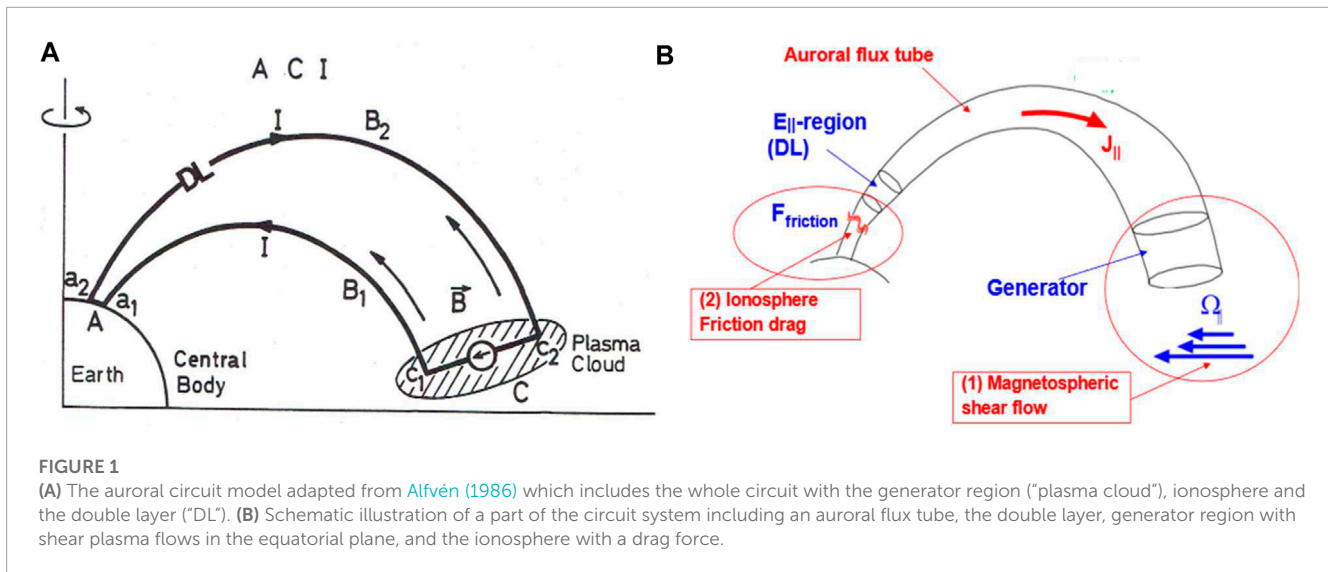
the first term is important in the establishment of the E_{\parallel} , but once the double layer is established, the other terms which include the field-aligned pressure gradients just give the acceleration of the electrons as they pass through the double layer.

6.3 Keeping the current density above a critical value

Sustaining a certain level of the magnetic shear and the resulting current density is important in maintaining the double layer, as the Bohm criterion requires the current density in the double layer to exceed a critical value $J_c = ne(2kT_e/m_e)^{1/2}$ (Block, 1978), i.e., the electron drift velocity must exceed the thermal velocity. This critical current is reduced in regions of low plasma density.

Since the FAC generation is related to the changes of the electric and magnetic fields, and the perturbations of the fields must propagate, the FAC generation is in general not a local process, and an analysis of the whole flux tube and the circuit system is necessary. For example, in the auroral acceleration region the shear Alfvén waves reflected between the generator, the auroral acceleration region and the load will change the generator condition and modify the resultant FAC filaments (e.g., Lysak and Song, 1998). A simplified view of the magnetosphere-ionosphere coupling in auroral regions as an equivalent circuit which includes a magnetospheric generator and the closure of the current in the upper ionospheric layers has been proposed previously (e.g., Marklund et al., 2007; and also see a review by Lysak et al., 2020). Lui (2018) described that the substorm current system consists of two current circuits, with a dynamo for the meridian current circuit and a load for the azimuthal current circuit in the magnetosphere. It is considered that the meridian current circuit is more important in the substorm development (see also, Akasofu, 2013).

Alfvén (1986) has suggested that double layers are elements of a circuit that might include a dynamo region, dissipation regions and



double layers. The energy is transferred first into electromagnetic energy and then to the double layers, where the power is released and charged particles are accelerated. This process may cause explosive phenomena in space plasmas. In studying what happens locally in the double layer, we need to take account of all the regions through which the current flows. Figure 1A shows the auroral circuit model by Alfvén (1986) which includes the whole circuit with the generator region (“plasma cloud”), ionosphere and the double layer (“DL”); while Figure 1B schematically illustrates a part of the circuit system including an auroral flux tube, the double layer, the generator region with shear plasma flows in the equatorial plane, in the M-I coupling system. Akasofu (e.g., Akasofu, 2013; 2015; 2017) has adopted the electric current approach to study substorm processes. In explaining how the expansion phase (the energy release) occurs, he suggested that the magnetosphere tended to become unstable when the accumulated energy reached a certain level. In restoring its stability, the current in the current sheet is reduced by developing plasma instabilities that release the accumulated magnetic energy. In this paper we have provided a possible mechanism for the explosive energy release during auroral substorms by the formation of double layers with quasi-static electric fields in aurora regions that accelerate the auroral particles releasing the accumulated magnetic energy.

Parallel electric fields or potential drops and double layers have been frequently observed in the auroral region (e.g., review by Carlqvist, 1982; Lysak et al., 2020; Haerendel et al., 1976; Mozer et al., 1977; 2013; Shawhan et al., 1978; Borovsky, 1992; Ergun et al., 1998; 2004; McFadden et al., 1999; Marklund et al., 2012, and many others). They are observed closely related to the acceleration of auroral particles. There has been numerous direct observational evidence for multi-kV electrostatic potential structures in the auroral zone. For example, Mozer et al. (1977) measured DL potential at $\sim 1\text{--}10$ keV; Gurnett (1972) estimated the field aligned potential drops observed above the ionospheric auroral region as ~ 6 kV. References in Borovsky (1992) has listed a number of reports that provide evidence for multi-kV DL potential drops. It is very likely that the double layer plays an important role in the auroral particle acceleration during substorms and

causing the explosive energy release at the expansion onset of substorms.

7 Summary and discussion

Auroral physics is important in understanding auroral substorm processes. Energy accumulation during the growth phase is a necessary condition to develop substorms. A physical mechanism is needed to release the energy explosively. The formation of double layers is considered to be a likely mechanism for the explosive energy release and the acceleration of charged particles that causes substorm aurorae and triggers the onset of substorm expansion phase.

The localized parallel electric field that forms the double layers results from the displacement current under certain conditions. We have presented theoretically the required conditions and the physical processes that lead to the formation of double layers. Shear flows caused by earthward flows accumulating magnetic energy in the near-Earth region before the expansion onset produce vorticity which enhance the azimuthal magnetic flux (magnetic shear). The enhanced azimuthal magnetic flux and the drag force at the ionosphere cause a twist of the auroral magnetic flux tube and generate a localized magnetic shear ($\nabla \times \mathbf{B}$). The displacement current would become significant when the magnetic shear cannot be balanced by the charged particle flow and enhances the parallel electric field, forming the double layer where certain conditions are met, for example, in the regions with large B/n , favoring low densities. The displacement current in the structure may change the distribution of the charge density, which leads to the formation of charge separation, causing the quasi-neutrality of the plasma in the structure to be locally violated. Since the plasma tends to be quasi-neutral, the quasi-neutrality condition should hold over most of the field line and quasi-neutrality can be violated only in localized regions. This implies that the existence of double layers (DLs) is localized in areas which satisfy certain plasma conditions. A positive feedback process may generate the parallel electric field rapidly

causing explosive energy release. The Poynting flux carried by Alfvén waves continuously supplies energy to the acceleration region, supporting the DLs, leading to strong long-lasting electrostatic fields dynamically.

It has been shown that the existence of the parallel electric field in double layers in auroral regions is often associated with observations of oppositely directed transverse electric field pairs at higher altitude and larger region (e.g., Mozer et al., 1977; Gurnett, 1972; Ergun et al., 2004, and others). We interpret these fields as being the fields of a kinetic Alfvén wave that carries the Poynting flux necessary to support the particle acceleration due to the parallel electric field (e.g., Mozer et al., 1980). This interpretation is supported by the Polar observations of Wygant et al. (2000) that show that these waves carry sufficient Poynting flux to power the auroral acceleration.

The magnetospheric substorm is a complicated process occurring in a complicated system which has been studied for decades and still not fully understood and causing a lot of controversy. It is difficult to cover every aspect of the problem. In this paper we have focused only on the mechanism of the expansion onset which occurs after the growth phase and the start of the explosive energy release. We have put forward theoretically a mechanism which may initiate the substorm onset by emphasizing the generation of parallel electric fields in the form of double layers. Although it is a qualitative description, the mechanism may also be applicable to explain other similar phenomena with explosive energy release and charged particle acceleration in cosmic plasmas.

Data availability statement

The original contributions presented in the study are included in the article/supplementary materials, further inquiries can be directed to the corresponding author.

References

- Akasofu, S. I. (1964). The development of the auroral substorm. *Planet. Space Sci.* 12 (4), 273–282. doi:10.1016/0032-0633(64)90151-5
- Akasofu, S. I. (2013). The relationship between the magnetosphere and magnetospheric/auroral substorms. *Ann. Geophys.* 31 (3), 387–394. doi:10.5194/angeo-31-387-2013
- Akasofu, S. I. (2015). Auroral substorms as an electrical discharge phenomenon. *Prog. Earth Planet. Sci.* 2 (1), 20–21. doi:10.1186/s40645-015-0050-9
- Akasofu, S. I. (2017). Auroral substorms: search for processes causing the expansion phase in terms of the electric current approach. *Space Sci. Rev.* 212 (1), 341–381. doi:10.1007/s11214-017-0363-7
- Alfvén, H. (1977). Electric currents in cosmic plasmas. *Rev. Geophys. Space Phys.* 15, 271–284. doi:10.1029/rg015i003p00271
- Alfvén, H. (1979). Plasma in laboratory and space. *Le J. de Physique Colloques* 40 (C7), C7-C1-C7-19. doi:10.1051/jphyscol:19797425
- Alfvén, H. (1986). Double layers and circuits in astrophysics. *IEEE Trans. Plasma Sci.* 14 (6), 779–793. doi:10.1109/tps.1986.4316626
- Baker, D. N., Pulkkinen, T. I., Angelopoulos, V., Baumjohann, W., and McPherron, R. L. (1996). Neutral line model of substorms: past results and present view. *J. Geophys. Res.* 101 (A6), 12975–13010. doi:10.1029/95JA03753
- Block, L. P. (1978). A double layer review. *Astrophysics Space Sci.* 55 (1), 59–83. doi:10.1007/bf00642580
- Borovsky, J. E. (1992). Double layers do accelerate particles in the auroral zone. *Phys. Rev. Lett.* 69 (7), 1054–1056. doi:10.1103/physrevlett.69.1054
- Borovsky, J. E., and Joyce, G. (1983). Numerically simulated two-dimensional auroral double layers. *J. Geophys. Res. Space Phys.* 88 (A4), 3116–3126. doi:10.1029/ja088ia04p03116
- Carlqvist, P. (1982). *Double layers in space (No. TRITA-EPP-82-07)*. Stockholm, Sweden: Royal Inst. of Tech.
- Chaston, C. C., Genot, V., Bonnell, J. W., Carlson, C. W., McFadden, J. P., Ergun, R. E., et al. (2006). Ionospheric erosion by Alfvén waves. *J. Geophys. Res.* 111, A03206. doi:10.1029/2005JA011367
- Cheng, C. Z. (2004). Physics of substorm growth phase, onset, and dipolarization. *Space Sci. Rev.* 113, 207–270. doi:10.1023/b:spac.0000042943.59976.0e
- Donovan, E., Mende, S., Jackel, B., Syrjäsoo, M., Meurant, M., Voronkov, I., et al. (2006). The azimuthal evolution of the substorm expansive phase onset aurora. *Proc. ICS* 8, 55–60.
- Ergun, R. E., Andersson, L., Main, D., Su, Y.-J., Newman, D. L., Goldman, M. V., et al. (2004). Auroral particle acceleration by strong double layers: the upward current region. *J. Geophys. Res.* 109, A12220. doi:10.1029/2004JA010545
- Ergun, R. E., Carlson, C. W., McFadden, J. P., Mozer, F. S., Delory, G. T., Peria, W., et al. (1998). FAST satellite observations of electric field structures in the auroral zone. *Geophys. Res. Lett.* 25 (12), 2025–2028. doi:10.1029/98gl00635

Author contributions

YS: Investigation, Writing—original draft, Writing—review and editing. RL: Funding acquisition, Investigation, Writing—review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by NASA Grant 80NSSC20K1269.

Acknowledgments

The authors also thank Naiguo Lin for helpful discussion on the topics and editing of the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Fälthammar, C. G. (1978). Generation mechanisms for magnetic-field-aligned electric fields in the magnetosphere. *J. Geomagnetism Geoelectr.* 30 (4), 419–434. doi:10.5636/jgg.30.419
- Feynman, R. P., Leighton, R. B., and Sands, M. (1963). *The Feynman lectures on physics*, v. 2. Massachusetts, United States: Addison-Wesley Reading.
- Fälthammar, C. G. (1983). *Magnetic-field aligned electric fields*. Stockholm, Sweden: KTH Royal Institute of Technology.
- Génot, V., Louarn, P., and Mottez, F. (2001). Fast evolving spatial structure of auroral parallel electric fields. *J. Geophys. Res. Space Phys.* 106 (A12), 29633–29643. doi:10.1029/2001ja000076
- Gurnett, D. A. (1972). “Electric field and plasma observations in the magnetosphere,” in Proceedings of the COSPAR Symp. on Critical Probl. of Magnetosphere Phys, Washington, DC, USA, May, 1972.
- Haerendel, G. (1992). “Disruption, ballooning or auroral avalanche: on the cause of substorms,” in Proceedings of the First International Conference on Substorms (ICS-1), Kiruna, Sweden, March, 1992, 417–420.
- Haerendel, G. (2015). Substorm onset: current sheet avalanche and stop layer. *J. Geophys. Res. Space Phys.* 120 (3), 1697–1714. doi:10.1002/2014ja020571
- Haerendel, G., Rieger, E., Valenzuela, A., Föppl, H., Stenbaek-Nielsen, H. C., and Wescott, E. M. (1976). First observation of electrostatic acceleration of barium ions into the magnetosphere. *Sounding-Rocket Balloon Res. Auror. Zone*, 203–211.
- Hones, E. W., Jr. (1984). “Plasma sheet behaviour during substorms,” in *Magnetic reconnection in space and laboratory plasmas*, geophys. Monogr. Ser., Editor E. W. Hones (Washington, D. C, USA: AGU), 178–184.
- Hsu, T. S., and McPherron, R. L. (2012). A statistical analysis of substorm associated tail activity. *Adv. Space Res.* 50 (10), 1317–1343. doi:10.1016/j.asr.2012.06.034
- Jiang, F., Strangeway, R. J., Kivelson, M. G., Weygand, J. M., Walker, R. J., Khurana, K. K., et al. (2012). *In situ* observations of the “preexisting auroral arc” by THEMIS all sky imagers and the FAST spacecraft. *J. Geophys. Res. Space Phys.* 117 (A5). doi:10.1029/2011ja017128
- Kan, J. R., and Sun, W. (1996). Substorm expansion phase caused by an intense localized convection imposed on the ionosphere. *J. Geophys. Res.* 101 (A12), 27271–27281. doi:10.1029/96JA02426
- Liang, J., Donovan, E. F., Liu, W. W., Jackel, B., Syrjäso, M., Mende, S. B., et al. (2008). Intensification of preexisting auroral arc at substorm expansion phase onset: wave-like disruption during the first tens of seconds. *Geophys. Res. Lett.* 35 (17). doi:10.1029/2008gl033666
- Lin, N., Frey, H. U., Mende, S. B., Mozer, F. S., Lysak, R. L., Song, Y., et al. (2009). Statistical study of substorm timing sequence. *J. Geophys. Res. Space Phys.* 114 (A12). doi:10.1029/2009ja014381
- Lui, A. T. Y. (1996). Current disruption in the Earth’s magnetosphere: observations and models. *J. Geophys. Res.* 101 (A6), 13067–13088. doi:10.1029/96JA00079
- Lui, A. T. Y. (2004). Potential plasma instabilities for substorm expansion onsets. *Space Sci. Rev.* 113 (1–2), 127–206. doi:10.1023/b:spac.0000042942.00362.4e
- Lui, A. T. Y. (2018). Review on the characteristics of the current sheet in the Earth’s magnetotail. *Electr. Curr. Geosp. beyond*, 155–175. doi:10.1002/9781119324522.ch10
- Lyons, L. R. (1995). A new theory for magnetospheric substorms. *J. Geophys. Res.* 100 (A10), 19069–19081. doi:10.1029/95JA01344
- Lysak, R., Echim, M., Karlsson, T., Marghitu, O., Rankin, R., Song, Y., et al. (2020). Quiet, discrete auroral arcs: acceleration mechanisms. *Space Sci. Rev.* 216 (5), 92–31. doi:10.1007/s11214-020-00715-5
- Lysak, R. L., and Song, Y. (1998). “Dynamics of auroral arc formation during substorms,” in *Substorms-4*. Editors S. Kokubun, and Y. Kamide (Tokyo: Terra Scientific), 35.
- Lysak, R. L., and Song, Y. (2003). Kinetic theory of the Alfvén wave acceleration of auroral electrons. *J. Geophys. Res.* 108 (A4), 8005. doi:10.1029/2002JA009406
- Marklund, G. T., Johansson, S., Lileo, T., and Karlsson, T. (2007). Cluster observations of an auroral potential and associated field-aligned current reconfiguration during thinning of the plasma sheet boundary layer. *J. Geophys. Res.* 112, A01208. doi:10.1029/2006JA011804
- Marklund, G. T., Sadeghi, S., Li, B., Amm, O., Cumnock, J. A., Zhang, Y., et al. (2012). Cluster multipoint study of the acceleration potential pattern and electrodynamic of an auroral surge and its associated horn arc. *J. Geophys. Res. Space Phys.* 117 (A10). doi:10.1029/2012ja018046
- McFadden, J. P., Carlson, C. W., and Ergun, R. E. (1999). Microstructure of the auroral acceleration region as observed by FAST. *J. Geophys. Res. Space Phys.* 104 (A7), 14453–14480. doi:10.1029/1998ja900167
- McPherron, R. L. (1972). Substorm related changes in the geomagnetic tail: the growth phase. *Planet. Space Sci.* 20 (9), 1521–1539. doi:10.1016/0032-0633(72)90054-2
- Mozer, F. S., Bale, S. D., Bonnell, J. W., Chaston, C. C., Roth, I., and Wygant, J. (2013). Megavolt parallel potentials arising from double-layer streams in the Earth’s outer radiation belt. *Phys. Rev. Lett.* 111 (23), 235002. doi:10.1103/physrevlett.111.235002
- Mozer, F. S., Carlson, C. W., Hudson, M. K., Torbert, R. B., Parady, B., Yatteau, J., et al. (1977). Observations of paired electrostatic shocks in the polar magnetosphere. *Phys. Rev. Lett.* 38 (6), 292–295. doi:10.1103/physrevlett.38.292
- Mozer, F. S., Cattell, C. A., Hudson, M. K., Lysak, R. L., Temerin, M., and Torbert, R. B. (1980). Satellite measurements and theories of low altitude auroral particle acceleration. *Space Sci. Rev.* 27, 155–213. doi:10.1007/bf00212238
- Newman, D. L., Goldman, M. V., Ergun, R. E., and Mangeney, A. (2001). Formation of double layers and electron holes in a current-driven space plasma. *Phys. Rev. Lett.* 87 (25), 255001. doi:10.1103/physrevlett.87.255001
- Nishimura, Y., Lyons, L., Zou, S., Angelopoulos, V., and Mende, S. (2010). Substorm triggering by new plasma intrusion: THEMIS all-sky imager observations. *J. Geophys. Res.* 115, A07222. doi:10.1029/2009JA015166
- Ohtani, S., Motoba, T., Gjerloev, J. W., Frey, H. U., Mann, I. R., Chi, P. J., et al. (2022). New insights into the substorm initiation sequence from the spatio-temporal development of auroral electrojets. *J. Geophys. Res. Space Phys.* 127, e2021JA030114. doi:10.1029/2021JA030114
- Rostoker, G., Akasofu, S. I., Foster, J. C., Greenwald, R. A., Kamide, Y., Kawasaki, K., et al. (1980). Magnetospheric substorms—definition and signatures. *J. Geophys. Res. Space Phys.* 85 (A4), 1663–1668. doi:10.1029/ja085ia04p01663
- Roux, A., Perraut, S., Robert, P., Morane, A., Pedersen, A., Korth, A., et al. (1991). Plasma sheet instability related to the westward traveling surge. *J. Geophys. Res.* 96 (A10), 17697–17714. doi:10.1029/91JA01106
- Sergeev, V., Angelopoulos, V., Kubyskhina, M., Donovan, E., Zhou, X. Z., Runov, A., et al. (2011). Substorm growth and expansion onset as observed with ideal ground-spacecraft THEMIS coverage. *J. Geophys. Res. Space Phys.* 116 (A5). doi:10.1029/2010ja015689
- Sergeev, V. A., Angelopoulos, V., and Nakamura, R. (2012). Recent advances in understanding substorm dynamics. *Geophys. Res. Lett.* 39, L05101. doi:10.1029/2012GL050859
- Shawhan, S. D., Fälthammar, C. G., and Block, L. P. (1978). On the nature of large auroral zone electric fields at 1-R_E altitude. *J. Geophys. Res. Space Phys.* 83 (A3), 1049–1054. doi:10.1029/ja083ia03p01049
- Shiokawa, K., Baumjohann, W., Haerendel, G., Paschmann, G., Fennell, J. F., Friis-Christensen, E., et al. (1998). High-speed ion flow, substorm current wedge, and multiple Pi 2 pulsations. *J. Geophys. Res. Space Phys.* 103 (A3), 4491–4507. doi:10.1029/97ja01680
- Song, Y., and Lysak, R. L. (2001). The physics in the auroral dynamo regions and auroral particle acceleration. *Phys. Chem. Earth* 26, 33–42. doi:10.1016/s1464-1917(00)00087-8
- Song, Y., and Lysak, R. L. (2006). Displacement current and the generation of parallel electric fields. *Phys. Rev. Lett.* 96, 145002. doi:10.1103/physrevlett.96.145002
- Wygant, J. R., Keiling, A., Cattell, C. A., Johnson, L. M., Lysak, R. L., Temerin, M., Spann, J., et al. (2000). Polar spacecraft based comparisons of intense electric fields and Poynting flux near and within the plasma sheet-tail lobe boundary to UVI images: An energy source for the aurora. *J. Geophys. Res.* 105 (A8), 18675–18692.