Increasingly Parallel Pressure Anisotropic Ballooning Instability: A New Mechanism for Magnetospheric Substorm Onset

by

Luke Oberhagemann

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Physics
University of Alberta

Abstract

Geomagnetic substorms are a process through which energy stored in Earth's magnetosphere is suddenly and violently released. On the ground, this process is largely observed as a vibrant and dancing display in the aurorae and is well documented, while the sequence of events that occurs in the magnetosphere is less understood and is the subject of ongoing debate. The work presented here offers a novel perspective on this sequence of events, focusing on ballooning instabilities, a type of plasma instability in which a weak portion of a magnetic flux tube bulges outward. Ballooning instabilities have been linked to ground-based auroral observations at the time the main phase of a substorm begins, known as onset, through observations of auroral beads, a periodic brightness feature that exists on the most equatorward auroral arc immediately prior to onset for the vast majority of substorms.

First, a numerical model of ballooning instabilities in magnetospheric conditions is applied to determine under what circumstances these instabilities may form. The main parameters that impact the formation of ballooning instabilities in this model are the plasma β, which is a ratio of the thermal pressure to the magnetic pressure, the pressure gradient, and the pressure anisotropy. The pressure anisotropy is a measure of how different the pressure in the direction parallel to a magnetic field line is from the pressure perpendicular to that field line, with the term parallel anisotropy denoting larger parallel than perpendicular pressure. It is found that, for constant values of the other parameters, an increase in plasma β and an increase in the pressure gradient create conditions that are more ballooning unstable. Significantly, it is also found that a change in the pressure anisotropy *towards* a parallel anisotropic configuration,

without the parallel pressure necessarily exceeding the perpendicular pressure in absolute terms, is destabilising.

Based on these findings, a second numerical model is used to examine the evolution of particle distributions as the geomagnetic tail stretches prior to substorm onset. In particular, the evolution of particles that originate in less stretched, earlier tail configurations and periodically drift around Earth, completing a full orbit to return to the tail in a more stretched configuration immediately prior to substorm onset is examined. It is found that particles follow drift orbits that depend on their pitch angle, the angle between their velocity vector and the local magnetic field, in a well-understood process called drift shell splitting. In the context examined here, drift shell splitting causes a mixing of particles from regions with varying pressures in a manner that leads to increasingly parallel pressure anisotropy, i.e., a more ballooning unstable distribution. In addition, the transport of particles from a less stretched early tail configuration to more stretched, immediate pre-onset field lines causes a decrease in particle energy that is stronger in the direction perpendicular to the magnetic field, further increasing parallel pressure anisotropy.

Therefore, the results presented here combine to form a picture of the time sequence of events leading to substorm onset in which particles drift from less stretched to more stretched tail configurations, resulting in an increasingly parallel pressure anisotropy and, consequently, the triggering of a ballooning instability. The high β and high pressure gradient requirements for triggering ballooning instability are met in the same radially localized equatorial region in Earth's magnetotail that is found to be most conducive to causing increasingly parallel pressure anisotropy for drifting particles. The two models used, therefore, self-consistently predict a region for ballooning instability that agrees with existing predictions. In addition, the prediction that increasingly parallel pressure anisotropies should precede substorm onset and be concurrent

with periods of tail stretching show very good agreement with existing in-situ satellite observations immediately prior to onset. Therefore, the work presented here suggests a self-consistent and observationally testable novel mechanism for substorm onset.

Preface

The research presented here was completed by the author, with chapters 4 and 6 containing manuscripts completed in collaboration with Ian Mann and prepared for submission to Geophysical Research Letters and Journal of Geophysical Research respectively.

Chapter 4 describes the numerical examination of ballooning instability thresholds in the magnetosphere. I was responsible for writing the code for the calculation of these thresholds based on the Chan et al. (1994) model, conducting primary analysis of the results, and manuscript composition.

Chapter 6 describes the numerical tracing of particles through a model magnetic field. I was responsible for writing code based on the Kabin et al. (2011) model, conducting primary analysis of the results, and manuscript composition.

Acknowledgements

I would like to thank my supervisor Ian Mann for his support and guidance throughout the completion of this work. I would also like to thank my friends and family for their continued support.

This work was supported by the NSERC CREATE grant for the International Space Masters (ISM) program.

Table of Contents

List of Figures

Chapter 1: Introduction

The sun exerts a significant influence on Earth. This extends beyond its gravitational field and electromagnetic radiation to the solar wind, a stream of plasma that continuously leaves the sun to fill interplanetary space. The influence of the solar wind on Earth can be highly variable, giving rise to the term "space weather." A notable feature of this is the aurora borealis and australis, the northern and southern lights. These occur as a result of ionised particles which are accelerated in near-Earth space colliding with Earth's ionosphere, generally toward the north and south poles. Another notable aspect is that the solar wind carries with it the interplanetary magnetic field (IMF), and this can be rapidly changing near Earth, resulting in coupling to the Earth's magnetosphere, the magnetic field surrounding Earth. Further, coupling between the solar wind and magnetosphere results in the generation of currents in the ionosphere, which can result in induced currents in the Earth or in other conducting pathways. This can have damaging effects on terrestrial infrastructure, particularly to long conductors such as power lines and pipelines ("Solar Storm Risk to the North American Electric Grid", 2013).

As with terrestrial weather, there are extreme events that can occur as part of space weather. One often cited example is the Carrington event of 1859 in which a solar flare caused aurorae that extended as far south as Honolulu, Hawaii (Cliver and Svalgaard 2004). There have been comparable but smaller events since, extreme space weather posing a significant potential threat to terrestrial technological infrastructure.

Because of this, the solar-terrestrial connection is an area of interest from a scientific as well as technological standpoint. A feature of this connection is a geomagnetic storm in which an increased interaction between the solar wind and Earth's magnetosphere causes a buildup of energy in the magnetosphere. This energy is often suddenly and violently released in a process known as a substorm, which has very vivid and dynamic manifestations in the aurorae wherein dancing lights are seen across a wide region in the sky. Substorms also cause an increase in induced currents on Earth (e.g. Pulkkinen et al., 1999).

Despite their clear visibility in the form of dancing auroral displays, and being the subject of numerous studies over many decades, substorms are not yet fully understood. Particularly, the sequence of events that occurs in the magnetosphere at the moment a substorm begins is

controversial. In light of this, the work presented in this thesis consists of a set of numerical simulations along with theoretical work and literature research that proposes a new mechanism for substorm onset, taking into account the body of evidence compiled by prior researchers who often have varying views on the physics of substorm formation.

In particular, we examine the role of ballooning instabilities, a type of magnetospheric plasma instability where parts of the magnetic field bulge out suddenly, in triggering the energy release during substorms. This follows up on recent observational and numerical work that has highlighted the potential relevance of ballooning instabilities in the substorm process. To do this, a two-step approach is taken. First, in chapters 3 and 4, a numerical model of ballooning instabilities is used to determine under what conditions they may form. Then, in chapters 5 and 6, a numerical model of Earth's magnetic field is applied to determine whether these conditions may result naturally from the development of the magnetosphere prior to substorm onset. We find that particles orbiting Earth as the magnetic field configuration changes during the time prior to a substorm, known as the growth phase, result in a distribution of the pressure in key regions of the field that is highly susceptible to the formation of ballooning instabilities. We propose the basis of a new substorm onset mechanism which is based around the redistribution of plasma pressure in a manner that is increasingly parallel to the background magnetic field. Our model predictions are also compared to observations around substorm onset, and we find good agreement with many known observational constraints but which we link causally in a new way through the new onset model.

The subsequent chapter, chapter 2, describes the background of the solar-terrestrial connection in more detail. Following this, chapters are organized based on two papers that were written as part of this work. Each paper, in chapters 4 and 6, is preceded by a background chapter, giving more detail on the models used and the manner in which they are implemented.

Chapter 2: Background: The Solar-Terrestrial Environment

In combination, Earth's underlying magnetic field and the incident solar wind plasma and magnetic field shape and control the dynamics of the outer regions of the Earth's magnetosphere as shown schematically in Figure 1. Due to the structure of currents in its dynamo core, the near-Earth magnetic field is composed of various multi-pole components, but which are dominated by the dipole term, a dominance that increases with increasing altitude above the Earth's surface (e.g., Tsyganenko, 1995). This dipole magnetic field is then compressed on the dayside and stretched on the night side as it is buffeted by the solar wind (e.g., Shue, 2013).

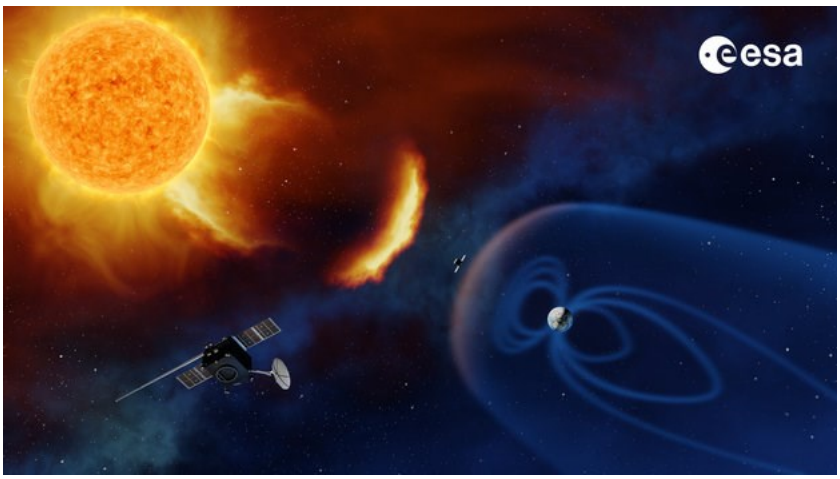

Figure 1: Schematic representation of the sun-Earth system (not to scale). Image credit: ESA https://www.esa.int/Our_Activities/Operations/Space_Situational_Awareness/Monitoring_space_weather

2.1 The Solar Wind

Because of the large pressure difference between the solar corona and the surrounding space, a large amount of plasma is continuously ejected from the sun and travels outward into interplanetary space to form the solar wind (Parker, 1958). Because the sun is rotating, although the solar wind is largely flowing in the radial direction, the solar wind magnetic field does not travel outward in a radial fashion but in an ever-tightening spiral, similar to the locus of water leaving a rotating sprinkler (Parker, 1958). About 95% of this ejected plasma consists of ionized hydrogen, with the remaining components being ionized helium and small amounts of heavier elements (Banks, 1971).

The solar wind is enhanced at times due to activity on the sun. This includes coronal mass ejections (CME's). During these CME events, magnetic reconnection causes large amounts of plasma to be ejected at once from the solar corona, resulting in a significant enhancement of the solar wind density, speed, and magnetic field strength (e.g., Priest 1995). Consequently, the effect of the solar wind on magnetised bodies in the solar system, such as Earth, that it encounters is increased.

At the location of Earth's orbit, the typical solar wind reaches a supersonic speed of around 450 km/s (e.g., Hundhausen 1995), with a typical number density of electrons and ions of around 7 cm^{-3}. Due to this low density, the plasma that makes up the solar wind can be approximated to be free of collisions and, therefore, viscous effects (the collisional mean free path in the solar wind is approximately one astronomical unit) (e.g., Burgess 1995). Because of this, and the resulting high conductivity of the plasma, Ohm's law for ideal magnetohydrodynamics (MHD) which assumes an infinite plasma conductivity and uses a fluid approximation to describe the plasma reduces to (e.g., Freidberg, 1982):

$$\boldsymbol{E} + \boldsymbol{u} \times \boldsymbol{B} = \boldsymbol{0} \tag{2.1}$$

Where $\boldsymbol{E}$ is the electric field, $\boldsymbol{B}$ is the magnetic field and $\boldsymbol{u}$ is the fluid velocity. A consequence of this is that the magnetic flux is frozen in to the fluid, meaning that the magnetic field is carried by the fluid and changes in the magnetic field are based purely on the fluid's motion (e.g., Kivelson, 1995).

The solar wind which expands into the heliosphere therefore carries with it a magnetic field, known as the interplanetary magnetic field (IMF). It is divided into sectors where the magnetic field points outward and others where it is directed inward, with a current sheet separating these areas (e.g., Wilcox & Hundhausen, 1983). This is shown schematically in Figure 2, where the current sheet is shown in purple. The complexity of this field, as compared to the relatively simple, dipolar field of the earth, is due to the more intricate structure of currents on the sun. These produce a magnetic field that is largely dipolar with a weaker quadropole moment (Bruno et al., 1982), but where the orientation of the dipole and solar rotation can produce alternating sectors of inwards and outwards IMF. The orientation of the magnetic field is significant because it determines the manner in which it interacts with the magnetic fields of

individual planets. Two forces that the MHD plasma in the solar wind exerts, which are key to the subsequent discussion in this work, are the magnetic tension and magnetic pressure.

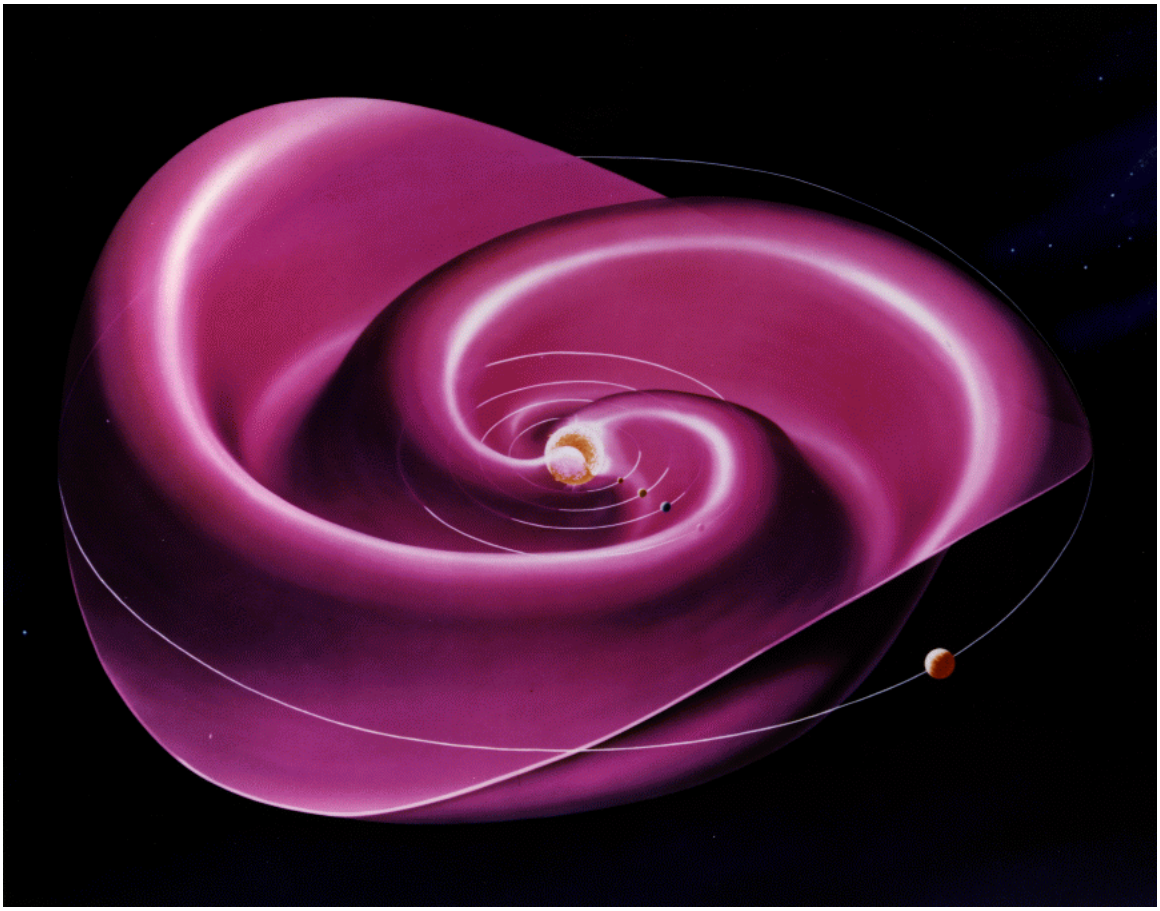

Figure 2: The sector structure of the interplanetary magnetic field, with inward and outward components in the ecliptic plane divided by the current sheet (shown in purple). Image credit: Wilcox Solar Observatory https://web.archive.org/web/20060901124602/http://quake.stanford.edu/~wso/gifs/HCS.html

2.2 Magnetic Tension and Pressure

Two forces that shape the solar wind, and its interaction with magnetised planets, are magnetic tension and magnetic pressure. These result from a combination of the magnetic component of the Lorentz force and Ampere's law (e.g., Parks, 1991):

$$\boldsymbol{J} \times \boldsymbol{B} = \frac{1}{\mu_0}(\nabla \times \boldsymbol{B}) \times \boldsymbol{B}. \tag{2.2}$$

Here the $\boldsymbol{J} \times \boldsymbol{B}$ force represents the force per unit volume on an MHD fluid element, and where $\boldsymbol{J}$ is the volume electric current density. In the isotropic ideal MHD case, this $\boldsymbol{J} \times \boldsymbol{B}$ force can be

decomposed into two terms with an equilibrium against the thermal pressure gradient force (per unit volume) $-\nabla P$ shown in Equation 2.3, and where B is the magnetic field strength, $\mu_0 = 4\pi \times 10^{-7}$ N/A^2 is the permeability of free space, and P is the thermal pressure

$$\frac{(\boldsymbol{B} \cdot \nabla)\boldsymbol{B}}{\mu_0} - \nabla \frac{B^2}{2\mu_0} - \nabla P = 0. \tag{2.3}$$

The first term in Equation 2.3 is the magnetic tension force, per unit volume. This force is analogous to the tension in a string and acts against increases in curvature in the magnetic field lines. The second term is the magnetic pressure force, which acts to push plasma from regions of higher magnetic field strength to regions of lower magnetic field strength. This magnetic pressure also gives rise to a frequently used concept in plasma physics: that of the plasma β, which is the ratio of thermal pressure to magnetic pressure:

$$\beta = \frac{P}{B^2/(2\mu_0)}. \tag{2.4}$$

The third term in Equation 2.3 is the force due to a thermal pressure gradient, per unit volume. Equation 2.3 can be divided into components parallel and perpendicular to the magnetic field. In the parallel direction, the components of the magnetic tension and magnetic pressure along the field lines cancel, resulting in:

$$\nabla_{\parallel} P = 0. \tag{2.5}$$

The consequence of this is the well-known isotropic MHD result that (isotropic) thermal pressure is constant along a magnetic field line. In the perpendicular direction, the remaining components balance as follows:

$$\frac{B^2}{\mu_0 R_c} - \nabla_{\perp} \frac{B^2}{2\mu_0} - \nabla_{\perp} P = 0. \tag{2.6}$$

Here R_c is the radius of curvature of the field line. The simple force balance described here is used as the background for the forces in a magnetic field and is applied in a chapter 6 to determine the pressure in a curved magnetic field geometry. This force balance is significantly complicated by the presence of an anisotropic thermal pressure, a scenario that is described in chapter 4 when addressing instabilities in an anisotropic equilibrium.

2.3 Earth's Magnetosphere

When the IMF reaches Earth, it interacts with Earth's intrinsic magnetic field to contribute to shaping the magnetic field surrounding Earth, known as the magnetosphere. Moving towards Earth along the line between the sun and Earth upstream of the Earth, there is a collisionless shock followed by a stagnation point at which the solar wind reaches zero velocity (Stern, 1979). This point marks the location of the magnetopause along the sun-Earth line, the magnetopause marking the outer boundary of the magnetosphere.

The structure of some regions and important structures within the Earth's magnetosphere are shown schematically in Figure 3. This structure of the magnetosphere arises due to a complex combination of magnetic fields, electric fields, and currents that depends significantly on the conditions in the solar wind and on the level of solar-terrestrial coupling, as well as coupling to the ionosphere and atmosphere below. The ionsphere is a highly conductive layer of ionised gas in the upper atmosphere (e.g., Chapman, 1956). There are several key features that generally exist in the magnetosphere. On the dayside (toward the sun) Earth's magnetic field is compressed by the dynamic pressure of the solar wind and is therefore somewhat flattened from a dipolar shape. The magnetic field lines in the solar wind are draped over Earth's magnetosphere by the flowing solar wind. Oppositely directed IMF and magnetospheric magnetic fields can reconnect on the dayside of the Earth during periods of southward directed IMF, connecting solar wind magnetic field lines to the Earth. These field lines are stretched tailward and can reconnect at a distant location on the nightside, known as the distant neutral line (DNL), which exists at a location around 100 Earth radii down-tail (e.g., Baker et al., 1984). This magnetic reconnection causes particles from the solar wind to be injected into the tail of Earth's magnetic field.

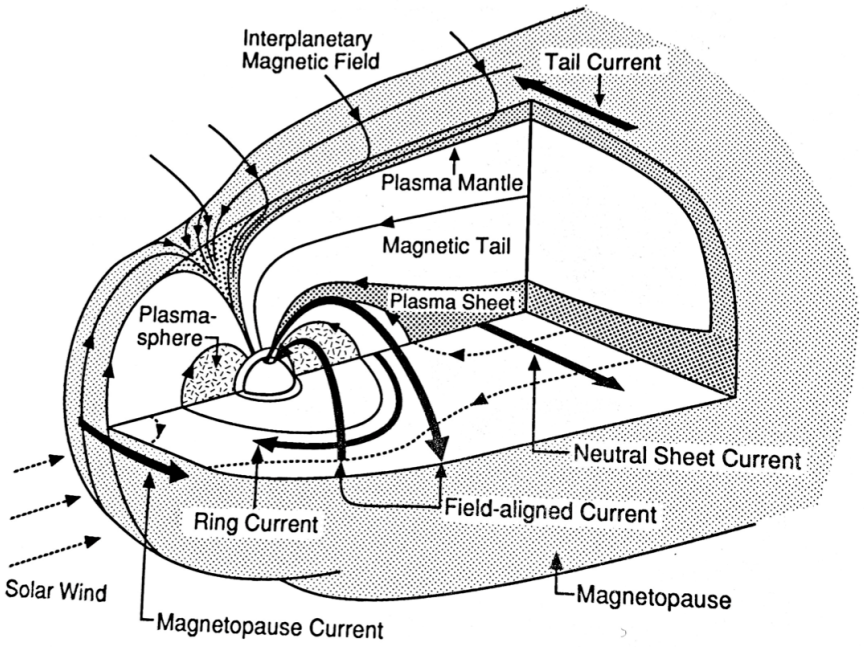

Figure 3: Schematic diagram of the overall structure of the magnetosphere. Image credit: Kivelson and Russell (1995)

In this stretched tail, there exists an electric field that points from dawn to dusk. This electric field causes charged particles that originated from the sun to move earthward in the tail through $\boldsymbol{E} \times \boldsymbol{B}$ drift, which is described below (e.g., Stern, 1979). These particles move Earthward until Earth's magnetic field becomes sufficiently strong to stop further motion. This results in an equatorial region in the tail where the pressure is higher due to particles having accumulated at this point (e.g., Lui, 1992). Towards Earth from this pressure peak, the pressure drops rapidly, but it decreases more slowly as one moves further into the tail. This is illustrated in more detail in the magnetospheric models described in chapter 6. One additional, noteworthy aspect of this high-pressure region in the tail is that, in some models, there is also minimum in the magnetic field strength in this region (e.g., Zhu, 2017). This would result in locally high β values, a feature that could lead to the formation of some types of plasma instability in this region of the tail, including ballooning modes which are a particular focus of this thesis. The particles that are trapped in Earth's magnetosphere also undergo several types of individual

motion, depending on field topology, their energy, etc. The characteristics of these single particle motions are described further below.

2.4 Single Particle Motion

The force on a single charged particle due to electric and magnetic fields is described by the Lorentz force:

$$\boldsymbol{F} = q(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}). \tag{2.7}$$

Here $\boldsymbol{E}$ is the electric field, $\boldsymbol{v}$ is the velocity of the particle, $\boldsymbol{B}$ is the magnetic field, and q is the charge of the particle. The magnetic component of the Lorentz force does not act in the direction of the particle's motion and, therefore, magnetic fields do no work.

There are three primary types of periodic motion in a magnetic field: gyro-, drift, and bounce motion. These are illustrated in Figure 4. Gyro-motion describes the manner in which a particle orbits around magnetic field lines. The Lorentz force causes a centripetal acceleration that results in particles circling a magnetic field line at a distance from the field line defined as the gyroradius. The gyro-motion of a particle also gives rise to the first adiabatic invariant, μ, defined as:

$$\mu = \frac{mv_{\perp}^2}{2B} = \frac{mv \sin^2 \alpha}{2B}. \tag{2.8}$$

Here $v_{\perp}$ is the component of the velocity that is perpendicular to the magnetic field, and α is the pitch angle of the particle, defined as the angle between the magnetic field line and the velocity vector. The first adiabatic invariant is generally conserved if the timescale of the field changes affecting the particle motion are much slower than the period of the gyro-motion. As the gyro-frequency is generally high in the context investigated here, the first adiabatic invariant is assumed to be conserved.

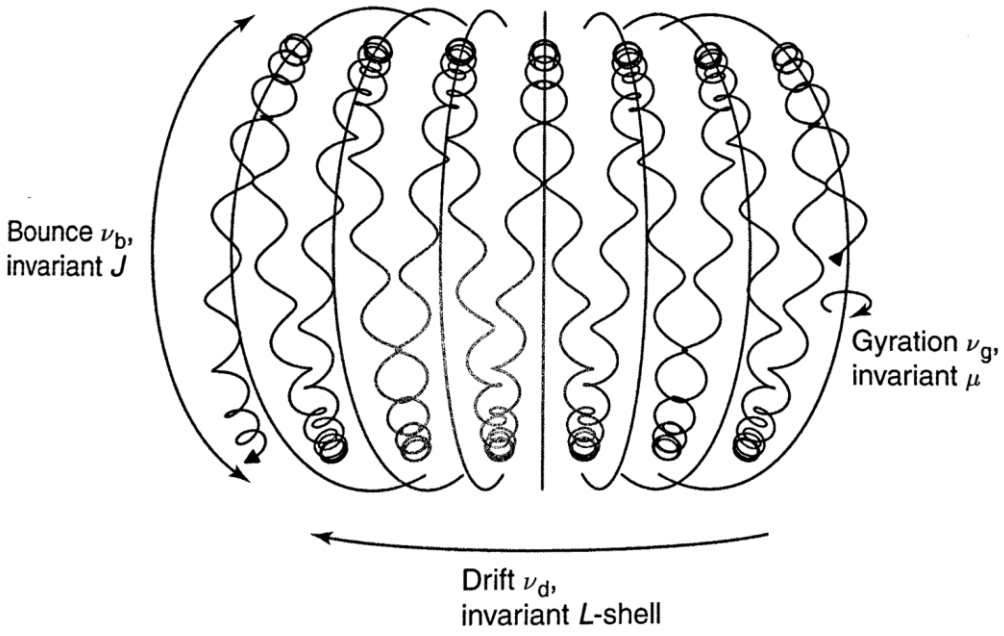

Figure 4: Schematic illustration of three types of periodic charged particle motion in a magnetic field. Image credit: Russell et al. (2016)

If the magnetic field lines are curved, and the particle has a velocity with a component in the direction of the magnetic field, it can undergo changes in its parallel velocity. If the magnetic field increases along the trajectory then there can be a reflection of the particle from the region of increased magnetic field, where the perpendicular velocity increases at the expense of the parallel velocity. The virtual mirror force then reverses the trajectory, and in regions where the field is bounded by two maxima then there can be periodic bounce motion in addition to gyro-motion, as shown schematically in Figure 4. As particles enter regions of higher magnetic field strength, i.e. where field lines are closer together, the Lorentz force acts to push them back toward the region of lower magnetic field strength in which they originated.

This can also be illustrated in terms of the first adiabatic invariant. If the kinetic energy of the particle, i.e., $\frac{1}{2}mv^2$, is constant, conservation of the first adiabatic invariant requires that the pitch angle increases as the magnetic field strength increases. This continues until the pitch angle becomes 90°, where the particle is reflected, also known as the mirror point. If the magnetic field strength does not become sufficiently high to cause a 90° pitch angle, the particle can continue to

move along the field line and does not bounce in the magnetic field. This is relevant in the context of magnetospheric plasmas, as the maximum magnetic field strength along a field line will be at Earth, and those particles that do not bounce at this field strength will collide with the ionosphere and atmosphere. The maximum equatorial pitch angle of trapped particles can be determined through conservation of the first adiabatic invariant as:

$$\alpha_{loss} = \sin\left(\frac{B_0}{B_E}\right)^{1/2}. \tag{2.9}$$

Here B_0 is the equatorial magnetic field strength and B_E is the magnetic field strength at the point at which the field line intersects with the ionosphere. Any particle with a pitch angle less than α_{loss} is not trapped in the magnetosphere, will reach the ionosphere and be lost.

The periodic bounce motion of a particle gives rise to the second adiabatic invariant, J, given by:

$$J = \int_a^b p_{\parallel} ds. \tag{2.10}$$

Here a and b are the mirror points along the field line at which the particle bounces, $p_{\parallel}$ is the momentum of the charged particle in the direction along the magnetic field and s is the distance measured along the field line.

The final type of periodic motion in a magnetic field is drift motion, illustrated in Figure 4. This type of motion can occur when an effective force acts on a charged particle in a direction perpendicular to the magnetic field. This effective force modifies the gyro-motion by decreasing the radius on the side where the particle is pushed toward the magnetic field line and increasing it on the side where it is pushed away. This variable radius prevents the particle from returning to its original location, instead resulting in a motion perpendicular to both the effective force and the magnetic field. In an axisymmetric field, such as a dipole, this can result in periodic azimuthal motion.

This effective force can be caused by a variety of factors, but the ones that are particularly relevant in this context are electric fields, gradients in the magnetic field strength, and the inertia of a particle in a curved magnetic field. If there is an electric field acting

perpendicular to the magnetic field, this results in a type of drift motion known as $\boldsymbol{E} \times \boldsymbol{B}$ drift, the velocity of which is given by:

$$\boldsymbol{v}_E = \frac{\boldsymbol{E} \times \boldsymbol{B}}{B^2}. \tag{2.11}$$

It is noteworthy that this velocity does not depend on the energy of the particle in question, only the strength and direction of electric and magnetic fields at a point. A key feature of the earthward $\boldsymbol{E} \times \boldsymbol{B}$ drift of particles from the distant tail, described above, is the manner in which it affects the pressure anisotropy. Because the first adiabatic invariant is conserved as particles drift earthward and the magnetic field strength increases, their energy in the perpendicular direction must increase, as shown in Equation 2.8. In addition, the length of the field lines decreases, while the second adiabatic invariant is conserved. As Equation 2.10 shows, this requires that the momentum, and consequently the energy, increases in the direction parallel to the magnetic field lines. Therefore, particles are energized in both the parallel and perpendicular directions. However, the energization in the perpendicular direction in an Earth-like field, or in a dipole, is stronger in the perpendicular than parallel directions (e.g., Zaharia 2006). Therefore, the particle distribution in the high-pressure region in the magnetosphere has a temperature and pressure that are perpendicular temperature and pressure anisotropic, i.e. the temperature and pressure in the direction perpendicular to the magnetic field lines are greater than those parallel to the field lines.

The drift velocity due to the gradient of the magnetic field is:

$$\boldsymbol{v}_g = \frac{W_\perp \boldsymbol{B} \times \nabla B}{qB^3}. \tag{2.12}$$

Here $W_\perp = \frac{1}{2} m v_\perp^2$ is the kinetic energy based on the velocity perpendicular to the magnetic field line. The drift velocity due to the inertia of a particle bouncing on a curved field line is:

$$\boldsymbol{v}_c = \frac{2W_\parallel \boldsymbol{B} \times (\boldsymbol{B} \cdot \nabla)\boldsymbol{B}}{qB^4}. \tag{2.13}$$

Here $W_\parallel = \frac{1}{2} m v_\parallel^2$ is the kinetic energy based on the velocity parallel to the magnetic field line. The latter two types of drift are referred to as gradient and curvature drift respectively. These are

often referred to in combination as gradient-curvature drift. Gradient-curvature drift depends both on the kinetic energy of the particle and its pitch angle.

The periodic drift motion of a particle leads to the definition of a third adiabatic invariant, given by (Roederer, 2013):

$$\Phi = \oint A_0 \cdot dl. \tag{2.14}$$

Here A_0 is the vector potential and l is a path along the drift shell of the particle. Φ can be understood as the magnetic flux enclosed by a drift shell. It is only conserved when changes occur at a much lower frequency than the drift frequency. Because this is generally only true when the magnetic field is undergoing very slow time variations, or when the particles of interest have very high energies, the third adiabatic invariant is often not conserved.

2.5 Magnetohydrodynamic waves

In addition to the particle motion described above, another key feature of magnetospheric plasmas for the work presented here is the presence of several types of magnetohydrodynamic (MHD) waves which can be understood using a fluid approximation for the plasma. These waves are important for the discussion of plasma instabilities in the subsequent chapter. A brief introduction for the simple case of MHD waves in a uniform plasma is included here and is expanded upon for the anisotropic case in the subsequent chapter. The wave types described here have low frequencies compared to a period of gyration and scales larger than a gyro-radius.

In a uniform MHD plasma, there are three main types of waves which arise from the three restoring forces due to the magnetic tension, the magnetic pressure, and the thermal pressure (e.g., Abraham-Schrauner, 1966). The first type is the Alfvén wave, which is similar to a transverse wave on a string in that it relies on the restoring force created by the magnetic tension. The dispersion relation for this type of wave is given by:

$$\omega^2 = k^2 v_A^2 \cos^2\theta. \tag{2.15}$$

Here ω is the angular frequency, k is the wave number, and θ is the angle between the wave vector and the magnetic field line. Note that Alfvén waves do not propagate perpendicular to the

magnetic field. $v_A = \frac{B}{\sqrt{\mu_0 \rho}}$ is the Alfvén speed, with ρ being the plasma mass density. The other two types of MHD waves are the slow and fast magnetosonic waves, they correspond to the two solutions of the following dispersion relation:

$$\frac{\omega^2}{k^2} = \frac{c_s^2 + v_A^2 \pm \sqrt{(c_s^2 + v_A^2)^2 - 4c_s^2 v_A^2 \cos^2\theta}}{2}. \quad (2.16)$$

Here c_s is the speed of sound. The larger solution of Equation 2.16 corresponds to the fast magnetosonic wave and the smaller solution corresponds to the slow magnetosonic wave. Perturbations of the magnetic field can take on the form of these waves, and, as is shown subsequently, in an unstable configuration these waves may grow, resulting in the release of free energy.

2.6 Geomagnetic Storms

When the IMF points southward, the field lines in the solar wind reconnect with the northward oriented magnetic field lines of the Earth on the dayside. These field lines then drape over the Earth, reconnect again at the distant neutral and return to the dayside in the Dungey cycle (Dungey, 1961). A consequence of the dayside reconnection is that a large flux of particles into the magnetosphere is possible as the solar wind plasma, into which the magnetic field is frozen, is no longer repelled by Earth's magnetosphere and instead particles can enter the magnetosphere from the solar wind along the reconnected field lines. The related injection of energy into the magnetosphere causes a period of significant activity in the magnetosphere known as a geomagnetic storm. In addition to the southward orientation of the IMF, large enhancements of the solar wind due to, for example, a coronal mass ejection, can significantly increase the severity of these events. A strong enough geomagnetic storm can cause the induction of currents in powerlines and pipelines on Earth that can cause damage on the order of trillions of dollars ("Solar Storm Risk to the North American Electric Grid", 2013). One prominent feature that occurs especially during geomagnetic storms, when the solar wind-magnetosphere interaction can cause significant amounts of field stretching in the tail, is a substorm.

2.7 Substorms

Substorms are characterised by extremely vivid and dynamic displays in the aurorae. These substorm-related auroral displays can be divided into several phases. The first is the growth phase in which discrete auroral arcs in the auroral zones are seen to move equatorward (McPherron, 1970). This corresponds to stretching in the tail, as the field lines corresponding to an arc are pulled further tailward and the point at which they intersect Earth, therefore, moves closer to the equator. This period of tail stretching is followed by substorm onset. This is marked by a brightening of the most equatorward auroral arc (Akasofu, 1964). Subsequently, this arc moves poleward and breaks up in the expansion phase. This is followed by the recovery phase, in which auroral arcs move equatorward from the northern locations they attained during the expansion phase and the aurorae become dimmer.

One currently controversial aspect of substorms is how they begin, i.e., what causes the growth phase to shift into the expansion phase with the consequent rapid release of stored magnetic energy. There are two main opposing schools of thought on the subject. The first maintains that substorm onset occurs as a result of magnetic reconnection at an additional near-Earth neutral line in the tail. Essentially, this view states that, as the tail stretches, opposing magnetic field lines from the southern and northern halves of the magnetotail are forced together to generate sufficient magnetic shear to cause reconnection in the near-Earth tail (e.g., Sergeev et al., 2012). This occurs at a near earth neutral line (NENL) that is located at a distance of ~20-30 Earth radii, closer to Earth than the distant neutral line described previously (e.g., Rae et al., 2009). This view is generally referred to as the out-to-in view because the substorm begins at a more distant location than that suggested by the opposing point of view.

The in-to-out view of substorm onset holds that substorms begin with an instability occurring at $\leq$ 12 Earth radii. Such an instability may cause a disruption of the cross-tail current, triggering the events that lead to a substorm (see e.g., the discussion in Sergeev et al., 2012). Proponents of this view point to the fact that in the aurora substorms begin with the activation and brightening of the most equatorward arc, which maps to a region closer to Earth in the magnetic field, which then moves poleward, corresponding to a region further in the tail (e.g. Rae et al., 2009). There have been a number of instabilities proposed as a candidate for this (Lui,

2004). However, the instability type that is of most interest due to recent findings is the ballooning instability.

Ballooning instabilities are described in more detail in a subsequent chapter, with a brief summary provided here. Ballooning instabilities are a type of instability where a magnetic flux tube bulges out at a weak point (e.g., Ferrière, 2001). They occur in a field geometry where the pressure gradient has a component that points in the same direction as the magnetic field line curvature vector (Chan 1994).

A noteworthy feature of the most equatorward auroral arc during substorm onset is that it occasionally contains periodic structures that are visible without instruments, and to the naked eye, as shown in Figure 5. These structures, known as auroral beads, were optically analysed by Kalmoni et al. (2015). Further analysis by Kalmoni et al. (2017) found that they occur during more than 90% of substorms, with the background arc being too bright in many cases to clearly resolve the intensity modulation associated with the beading. Kalmoni et al. (2018) then calculated a dispersion relation for these beads and found that they align with a dispersion relation for a kinetic Alfvén wave. Viñas and Madden (1986) showed that kinetic Alfvén waves may be triggered in turn by ballooning instabilities in the equatorial plane, the kinetic Alfvén waves coupling the ballooning modes into the ionosphere.

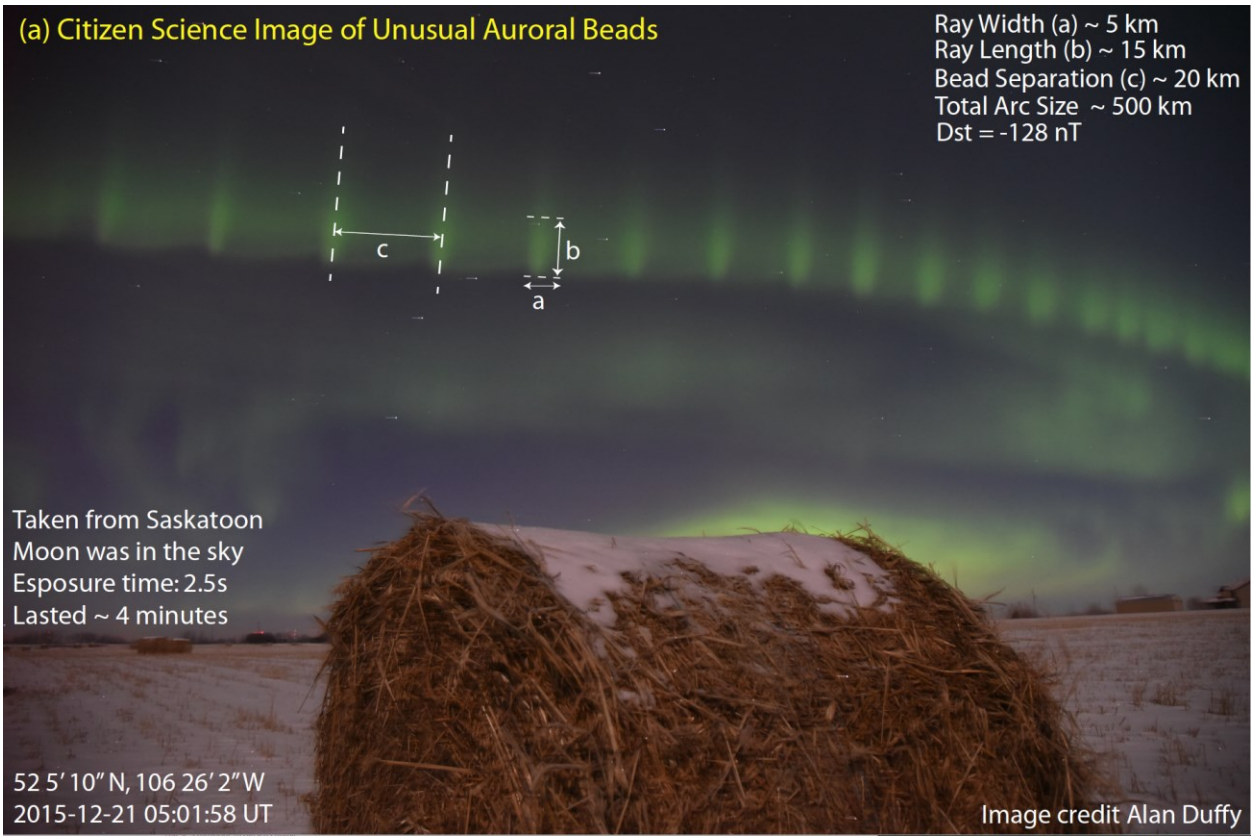

Figure 5: A highly visible occurrence of auroral beads. Source: Alan Duffy, referenced by N. Kalmoni for Aurorasaurus http://blog.aurorasaurus.org/?p=398

Further, there has been some relevant work conducted by Zhu et al. (2017). These authors numerically studied the development of ballooning instabilities in the magnetotail and found that ballooning instabilities, once they form and develop, may trigger magnetic reconnection in the more distant tail. In particular, Zhu et al. (2017) state that reconnection at a two-dimensional near-Earth neutral line may be supressed by the northward component of the magnetic field and is structurally unstable, with the tail being more capable of supporting three-dimensional reconnection. This three-dimensional reconnection takes on the form of a series of points at which magnetic reconnection occurs, i.e. an azimuthal structuring. Zhu et al. (2017) show that this can be triggered by the growth of ballooning instabilities, which are azimuthally structured. These conclusions by Zhu et al. (2017) also agree qualitatively with the observed initial azimuthal structuring of auroral beads. Therefore, this suggests a self-consistent view of substorm onset in which ballooning instabilities play a key role.

The work presented in this thesis examines the conditions under which ballooning instabilities form and whether these conditions exist in the magnetosphere. To this end, chapters 3 and 4 adopt a numerical model created by Chan et al. (1994) to analyse the thresholds for the formation of ballooning instabilities. These findings are then placed into the context of a new mechanism for substorm onset that relies on particle drifts around Earth to create ballooning unstable conditions as a result of the development of increasingly parallel temperature anisotropy which arises from the particle drifts into a stretching magnetotail.

In chapters 5 and 6, this mechanism is examined by tracing particles through a model magnetic field created by Kabin et al. (2011). Changes during the particles' drift orbits are found to change the pressure configuration in a manner which can lead to ballooning unstable modes forming as a result of the development of increasingly parallel pressure anistropy. The predictions made by the mechanism proposed here are compared to several ground-based and in-situ observations of the known time sequence of events in substorms as a way to assess the potential applicability of the model to known substorm features. Based on our results, we not only argue that ballooning instabilities are key to substorm formation, but since the proposed mechanism generates predictions which agree well with these observations the model also offers a novel basis for interpreting and explaining some auroral phenomena and further advances a framework through which substorm onset might also be predicted.

Chapter 3: Review of Plasma Instabilities

This chapter provides background on the ballooning instabilities that are investigated in this thesis. It includes a description of the equilibrium in which they may form, their mathematical description, and the numerical approach taken to determine the thresholds for ballooning instability.

3.1 General Overview of Instabilities

In general, an unstable equilibrium is one in which a small change from the equilibrium results in a reduction of the potential energy. This reduction in potential energy allows the instability to grow and the system to move further from equilibrium. Such unstable equilibria exist throughout nature, from rocks on top of a hill and bifurcations in rivers to the magnetospheric plasmas studied here. The concept is illustrated in general in Figure 6. Both configurations in Figure 6a represent an equilibrium because there are no horizontal forces and the force due to gravity is balanced by the normal force from the surface underneath the pencil. However, as shown in Figure 6b, the equilibrium in which the pencil is standing on its base is more stable than the equilibrium in which the pencil is standing on its tip. This is because a small perturbation from the initial equilibrium results in a moment that pushes the pencil back toward equilibrium when it is standing on its base. This perturbation raises the center of mass slightly, corresponding to an increase in potential energy. A similar size of perturbation leads to movement further from the equilibrium when the pencil is standing on its tip because the perturbation results in a lowering of the center of mass and, consequently, a decrease in the potential energy. Of course, a sufficiently large perturbation can also destabilize the pencil standing on its base, but the case of the pencil standing on its tip is unstable to very small perturbations.

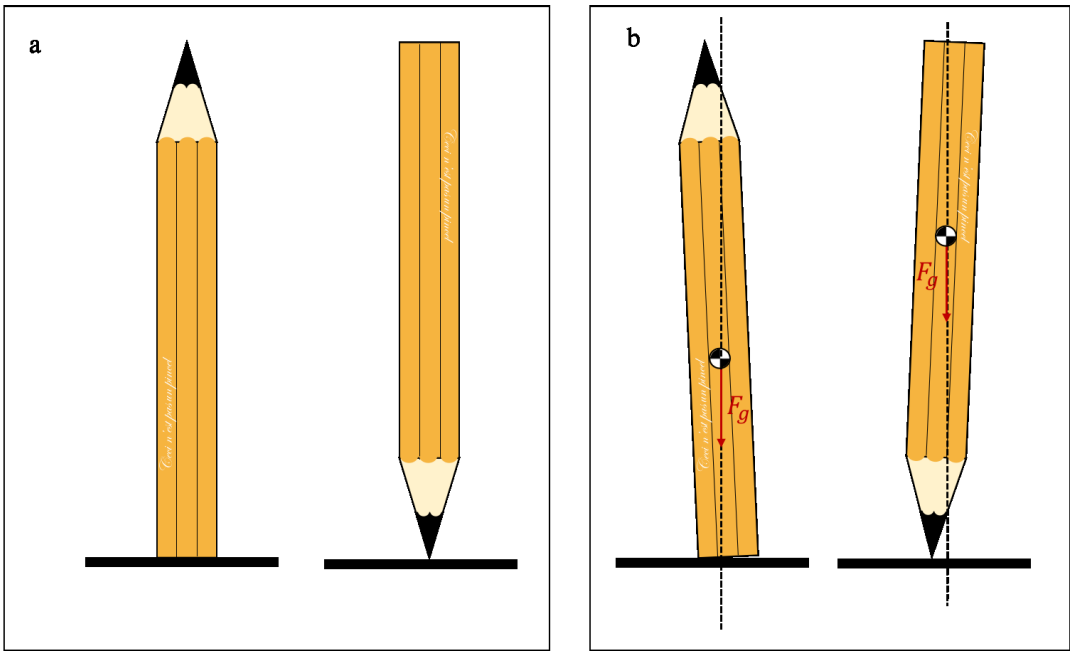

Figure 6: Illustration of the general theory of instabilities showing (a) two equilibria, and (b) the effects of a small perturbation on these equilibria resulting in the destabilization of the pencil standing on its tip but not that standing on its base.

3.1.1 Anisotropic Plasma Equilibria

The MHD forces described in the previous chapter apply to an equilibrium with an isotropic thermal pressure. For the purpose of this work, however, an anisotropic equilibrium is required to examine the effects that occur when the pressure in the direction along magnetic field lines differs from the pressure perpendicular to the magnetic field. Chan et al. (1994) define an anisotropic equilibrium for a collisionless magnetospheric plasma in an axisymmetric magnetic field given by:

$$\mathbf{B} = \nabla\psi \times \nabla\phi. \tag{3.1}$$

Here ϕ is the azimuthal angle and ψ is the flux function. As shown in Equation 3.1, ψ is constant along a magnetic field line and is given in a vacuum dipole magnetic field, in cylindrical coordinates, R, Z, ϕ, where the central axis passes through the central axis of the dipole by:

$$\psi_V = B_E R_E^3 \frac{R^2}{(R^2 + Z^2)^{3/2}}. \tag{3.2}$$

Here $B_E = 31000$ nT is the magnetic field strength at the surface of earth and $R_E = 6371$ km is the radius of Earth. Because ψ is constant along a field line, any location in the magnetic field can be uniquely described by the set of coordinates made up by ψ, ϕ, and the distance measured along a field line, l.

To quantify the pressure anisotropy, Chan et al. (1994) define an anisotropy parameter:

$$\delta = 1 - \frac{P_{\parallel}}{P_{\perp}}. \tag{3.3}$$

This δ parameter takes on a value of zero when the pressure is isotropic, such that parallel and perpendicular pressures are equal, i.e., $P_{\parallel} = P_{\perp}$. While this formulation of an anisotropy parameter is useful for the equations defining equilibrium, it should be noted that it is asymmetric. For strong positive, i.e. perpendicular, anisotropy e.g., where $P_{\perp} = 2P_{\parallel}$, δ is equal to 0.5. For the same amount of parallel anisotropy, $P_{\parallel} = 2P_{\perp}$, δ is equal to -1. This may lead to a misrepresentation of the results in terms of symmetric plots as a function of δ since the perpendicular side in such a plot becomes compressed. Another frequently used anisotropy parameter is $P_{\perp}/P_{\parallel}$ (e.g., Birn et al., 1997, Wang et al., 2013). This presents similar issues. Therefore, the results presented here, apart from those included for comparison with Chan et al. (1994), are plotted using a symmetric instability parameter:

$$\delta_{\text{sym}} = \begin{cases} 1 - \dfrac{\beta_{\parallel}}{\beta_{\perp}} \text{ for } \beta_{\parallel} > \beta_{\perp} \\ -\left(1 - \dfrac{\beta_{\perp}}{\beta_{\parallel}}\right) \text{ for } \beta_{\perp} > \beta_{\parallel} \end{cases} \tag{3.4}$$

A δ_{sym} of 0 is isotropic; when $P_{\perp} = 2P_{\parallel}$, $\delta_{sym} = 1$; and when $P_{\parallel} = 2P_{\perp}$, $\delta_{\text{sym}} = -1$.

Chan et al. (1994) base their magnetospheric equilibrium on the following parallel and perpendicular force balances (e.g., Hasegawa and Sato, 1989, Cheng, 1992). Significantly, as opposed to the isotropic pressure case, the parallel force balance in an anisotropic pressure equilibrium no longer results in a constant pressure along the field line. It is given by:

$$B\left(\frac{\partial P_{\parallel}}{\partial B}\right)_{\psi} = P_{\parallel} - P_{\perp}. \tag{3.5}$$

Note that in the case where $P_{\parallel} = P_{\perp}$ this reduces to the isotropic equation. An additional noteworthy consequence of the presence of a parallel pressure gradient is that the slow magnetosonic and Alfvén waves become coupled. Chan et al. (1994) insert a gyrotropic pressure tensor with a bi-Maxwellian distribution function into Equation 3.5. The bi-Maxwellian distribution function is the product of two Maxwellian distribution functions, one for temperature in the parallel direction and the other for temperature in the perpendicular direction. This results in field aligned pressure distributions:

$$P_{\parallel} = P_{\parallel_0}\left(\frac{1-\delta_0}{1-\delta_0 B_0/B}\right), \tag{3.6}$$

and

$$P_{\perp} = P_{\perp_0}\left(\frac{1-\delta_0}{1-\delta_0 B_0/B}\right)^2. \tag{3.7}$$

Here the subscript 0 denotes values at the equator. In both Equation 3.6 and Equation 3.7 the isotropic case once again reduces to the result that pressure is constant along a field line. Combining Equation 3.6 and Equation 3.7 describes the manner in which the pressure anisotropy changes along a field line:

$$\delta = \delta_0 \frac{B_0}{B}. \tag{3.8}$$

Because the magnetic field strength increases as one moves along a magnetic field line from its equatorial crossing point toward Earth, the anisotropy decreases. This is the case whether the anisotropy is parallel or perpendicular at the equator. Because the pressure is constant along a field line in the isotropic case, the anisotropy effectively controls how concentrated the plasma is near the equator by allowing for the presence of parallel pressure gradients.

The perpendicular force balance is also complicated by the presence of an anisotropic pressure, it is defined by Chan et al. (1994) as:

$$\sigma\Delta^*\psi + \nabla\sigma \cdot \nabla\psi + \mu_0 R^2 \left(\frac{\partial P_\parallel}{\partial \psi}\right)_B = 0. \tag{3.9}$$

Here:

$$\sigma = 1 + \frac{1}{2}(\beta_\perp - \beta_\parallel). \tag{3.10}$$

The parameter σ describes the onset of an additional type of plasma instability, the firehose instability, with $\sigma < 1$ corresponding to a firehose unstable equilibrium. However, the configurations studied here do not reach this instability condition. Therefore, the firehose instability is not included in the subsequent discussion. The operator Δ^* in Equation 3.9 (see Chan et al., 1994) is given by:

$$\Delta^* = \frac{\partial^2}{\partial R^2} - \frac{1}{R}\frac{\partial}{\partial R} + \frac{\partial^2}{\partial Z^2}. \tag{3.11}$$

The first and third terms in Equation 3.9 are similar to the perpendicular force balance in the isotropic case, but the second term is purely the result of the pressure anisotropy. To apply this force balance to the calculation of an equilibrium, Chan et al. (1994) use a multiscale perturbation method approach valid near the field line being examined, which is outlined in general terms here.

The approach begins with the use of a vacuum dipole field as the zero-order term as, in the absence of pressure, Earth's magnetic field is largely dipolar (e.g. Tsyganenko, 1995). An ordering is then undertaken in which $\beta_{\perp V} = 2\mu_0 P_\perp / B_V^2$ is used as a small parameter. The subscript V denotes a value calculated for a vacuum dipole magnetic field, therefore this β value is lower than the β value measured by spacecraft, which is calculated using the magnetic field that is depressed by the presence of a cross-tail current. $\beta_{\perp V}$ and $\beta_\perp$ are related through:

$$\beta_\perp = \frac{\beta_{\perp V}}{(1 - \beta_V/2)^2}. \tag{3.12}$$

This ordering is used to create a perturbation series for ψ in the mild anisotropy and strong anisotropy cases. In the mild anisotropy case, $\beta_{\perp V}$ and $\beta_{\parallel V}$ are of the same order and in the high anisotropy case $\beta_{\parallel V}$ is on the order of $(\beta_{\perp V})^2$. The neglected terms for the mild anisotropy

case are on the order of $(\beta_{\perp V})^2$ and for the strong anisotropy case they are on the order of $(\beta_{\perp V})^{3/2}$. The lower order of the neglected terms would result in lower accuracy for the high anisotropy case for the same value of $\beta_{\perp V}$. However, as will be shown later, the high anisotropy cases require lower $\beta_{\perp V}$ for instability and, therefore, lower $\beta_{\perp V}$ values are used in the calculation, thus increasing the accuracy in the high anisotropy cases. For both mild and strong anisotropy, the magnetic field is described by:

$$\boldsymbol{B} = \boldsymbol{B}_V(1 - \beta_{\perp V}/2) + \text{H.O.T.} \tag{3.13}$$

With the flux function given by:

$$\psi = \psi_V \left(1 + \frac{\beta_{\perp V}}{2(\alpha - 1)}\right). \tag{3.14}$$

Here α is a parameter that describes the pressure gradient by relating the pressure to the dipole L-shell parameter through $P_0 \propto L^{-\alpha}$ (L is the geocentric equatorial distance in Earth radii of the dipole field line, and hence can be used to label field lines). To compare their anisotropic results to the isotropic case, Chan et al. (1994) define an average equatorial plasma β, $\bar{\beta}_0$ given by:

$$\bar{\beta}_0 = 1/3(2\beta_\perp + \beta_\parallel). \tag{3.15}$$

The factor of 2 multiplying $\beta_\perp$is due to there being two directions perpendicular to the magnetic field and one direction parallel to the field. Using a constant $\bar{\beta}_0$ of 0.5, the surfaces of constant flux for this equilibrium are shown in Figure 7. For all types of anisotropy, the surfaces of constant flux are moved radially outward from the vacuum dipole values. In addition, Figure 7 illustrates that more perpendicular anisotropy leads to a greater amount of field line curvature in the equatorial plane than parallel anisotropy. This is significant for the stability analysis that follows.

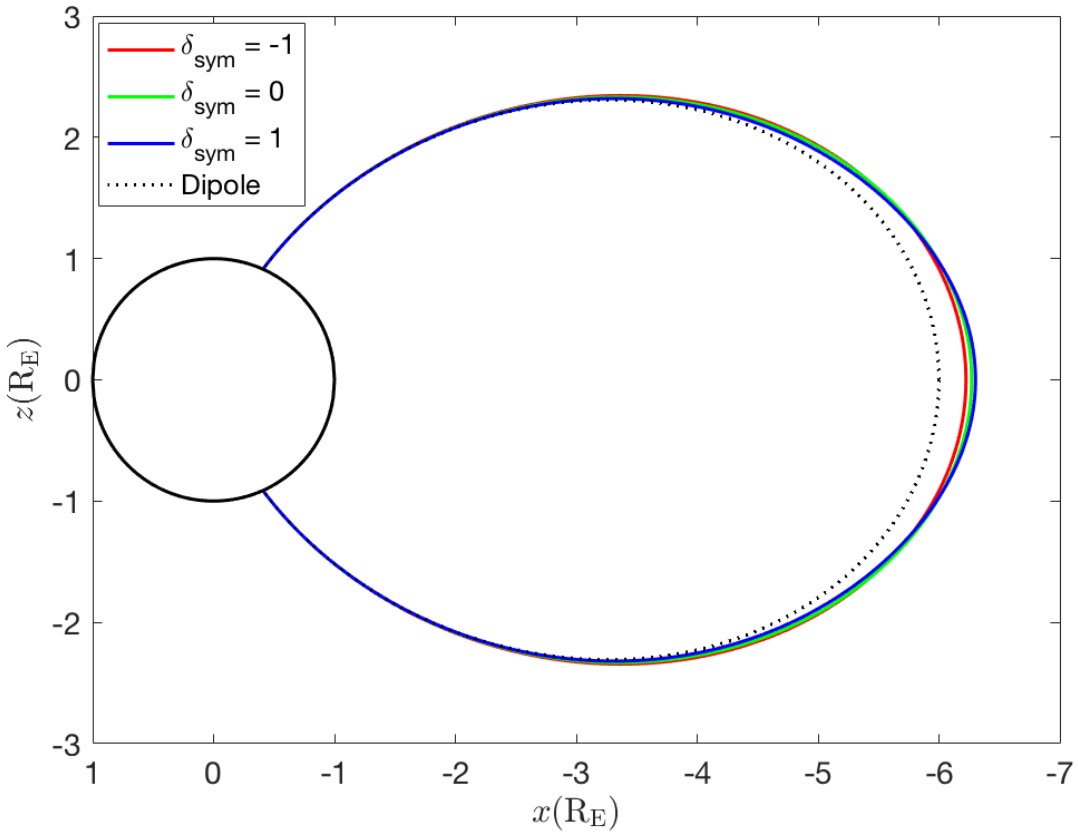

Figure 7: Equilibrium surfaces of constant flux for $\bar{\beta}_0 = 0.5$, modified from the vacuum field line at 6 R_E and using an α value of 6.

3.1.2 MHD Instabilities

In addition to the ballooning instability, the mirror instability is also relevant to the discussion here. A mirror instability can occur in a magnetic field with straight field lines or a magnetic field, such as that in the magnetosphere, in which the field lines are curved. Mirror instabilities can only occur in a perpendicular anisotropic equilibrium. The criterion for their formation is given by the mirror parameter $\tau < 0$, which is defined by Chan et al. (1994) as:

$$\tau = 1 - \frac{\beta_\perp - \beta_\parallel}{1 - \delta}. \tag{3.16}$$

If the condition $\tau < 0$ is met, regions of increased magnetic field strength form next to regions of lower field strength, causing particles to experience the mirror force described earlier. As shown in Figure 8, this results in particles becoming trapped in the regions of lower magnetic

field strength, causing these regions to expand outward while the regions of higher flux contract further.

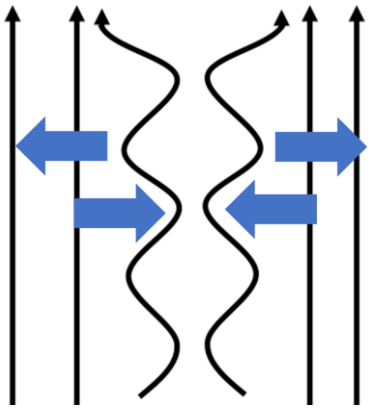

Figure 8: Schematic representation of the development of the mirror instability.

While the conditions for ballooning instabilities to form are slightly different, both the ballooning instability and mirror instability can form in the same magnetic field geometry. For a ballooning instability to form, field line curvature is required (e.g., Ferriére, 2001). Ballooning may occur when a component of the pressure gradient vector aligns with the magnetic field line curvature vector, i.e. when the force due to the thermal pressure, F_p, is pushing against a curved field line as shown in Figure 9.

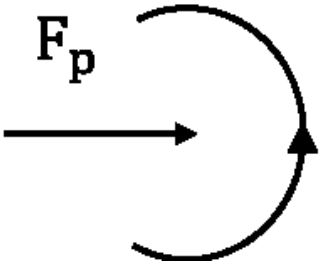

Figure 9: Schematic illustration of the conditions under which a ballooning instability may form.

When a ballooning unstable configuration is reached, the pressure force causes the magnetic flux tube around this field line to bulge out at a weak point (e.g., Ferriére, 2001). Ballooning instabilities are often conflated with interchange instabilities, in which entire adjacent flux tubes are exchanged. However, while both are types of quasi-interchange instabilities, and both occur when a plasma is stratified due to field line curvature, these two types of instabilities connect to different types of MHD waves, with one being an unstable manifestation of a slow

magnetosonic wave and the other being an unstable manifestation of an Alfvén wave (see e.g., Ferriére, 2001 for a more detailed discussion). Which wave mode each type of instability connects to is dependent on the configuration of the field. The stratification that occurs due to curved field lines is analogous to the stratification that occurs in the density of a fluid due to gravity (e.g. Viñas and Madden, 1986, Ferriére, 2001). Generally, in magnetospheric conditions, the stratification due to gravity is negligible and the plasma is stratified purely due to the effects of magnetic field line curvature.

Chan et al. (1994) examine the stability of their anisotropic equilibrium with respect to ballooning and mirror modes. This is done through application of the perpendicular Ampère's law and vorticity equation described in Chen and Hasegawa (1991). Only perturbations of the lowest frequency antisymmetric field-guided mode in ψ are considered as these were shown by Chen and Hasegawa (1991) to be the most unstable. As shown by these authors, and adapted by Chan et al. (1994) for these conditions, the perpendicular Ampère's law is:

$$\tau \delta B_{\parallel} = \frac{2\mu_0 c}{\omega B^2}\left(\mathbf{e}_k \cdot \widetilde{\nabla} P_{\perp}\right)\delta\psi. \tag{3.17}$$

Here $c = 3 \times 10^8$ m/s is the speed of light, ω is the wave frequency, $\mathbf{e}_k \cdot \widetilde{\nabla} P_{\perp}$ is the gradient of the perpendicular pressure in the direction of the wave normal, $\delta\psi$ is the perturbation in the flux function, and $\delta B_{\parallel}$ is the perturbation of the magnetic field in the direction of the magnetic field line vectors. Similarly, the vorticity equation is:

$$\begin{aligned} B\frac{\partial}{\partial l}\left(\frac{\sigma k_{\perp}^2}{B}\frac{\partial \delta\psi}{\partial l}\right) + \frac{\omega^2 k_{\perp}^2}{V_A^2}\delta\psi \\ = -\frac{2\mu_0}{B^2}\left(\mathbf{e}_k \cdot \widetilde{\nabla} P_{\perp}\right)\left(\frac{\omega}{c}\delta B_{\parallel} + \Omega_B \delta\psi\right) - \frac{2\mu_0}{B^2}\left(\mathbf{e}_k \cdot \widetilde{\nabla} P_{\perp}\right)\Omega_k \delta\psi. \end{aligned} \tag{3.18}$$

Here $k_{\perp}$ is the component of the wave vector perpendicular to the field line, V_A is the Alfvén speed, Ω_B is the component of the gradient $\nabla \ln B$ in the direction along the wave vector, and Ω_k is the component of the field line curvature vector in the direction of the wave vector. Chan et al. (1994) assume a large azimuthal mode number so $k_{\perp}^2 = k_{\phi}^2 + k_{\psi}^2$ reduces to $k_{\perp} \approx k_{\phi}$. In this

case, as shown by Chan et al. (1994), Equation 3.17 and Equation 3.18 can be combined to form the following Sturm-Liouville eigenvalue problem:

$$B\frac{\partial}{\partial l}\left(\frac{\sigma}{R^2B}\frac{\partial \delta\psi}{\partial l}\right)+\frac{\beta_E\kappa_\psi}{R^2L_P}\,\delta\psi+\frac{\omega^2}{R^2V_A^2}\delta\psi=0. \tag{3.19}$$

As shown by Chan et al. (1994), this equation describes the stability of the Chan et al. (1994) equilibrium to ballooning modes. Because the perturbations are of the form $e^{-i\omega t}$, a positive imaginary component of the frequency, ω, corresponds to an unstable equilibrium. The variable β_E in Equation 3.19 is an effective β value, given by:

$$\beta_E=\frac{\beta_\perp}{2}\left(1-\delta+\frac{\sigma}{\tau}\right). \tag{3.20}$$

The presence of τ in this equation ensures that equilibria unstable to the mirror instability result in an imaginary eigenvalue. L_P is the pressure gradient length scale given by:

$$L_P=\frac{\psi_V}{\alpha R B_V}. \tag{3.21}$$

κ_ψ is the curvature of the surface of constant flux. Because of the ordering of the terms in the equilibrium described earlier, as shown by Chan et al. (1994), it is sufficiently accurate to use the value of κ_ψ based on the vacuum dipole field.

As a first step to solving Equation 3.19 numerically, Chan et al. (1994) introduce a change of variable from length along a field line to latitude, x. In addition, all lengths are normalized by LR_E, which represents the equatorial crossing point of the given field line in a vacuum dipole field. All magnetic field strengths are normalized using B_E/L^3, which similarly represents the equatorial magnetic field strength for the vacuum dipole case. This converts the equation into the form:

$$\frac{d}{dx}\left(\frac{\sigma}{\tilde{L}_x\tilde{B}\tilde{R}^2}\frac{d\widetilde{\delta\psi}}{dx}\right)+\frac{\tilde{L}_x}{\tilde{B}\tilde{R}^2}\left(\frac{\beta_E\tilde{\kappa}_\psi}{\tilde{L}_P}+\frac{\widetilde{\omega}^2}{\tilde{V}_A^2}\right)\widetilde{\delta\psi}=0 \tag{3.22}$$

where the conversion factor:

$$\tilde{L}_x = \frac{\partial \tilde{l}}{\partial x} = \cos x \sqrt{1 + 3\sin^2 x}. \tag{3.23}$$

The remaining variables are calculated from the latitude using the following equations:

$$\tilde{R} = \cos^3 x \tag{3.24}$$

$$\tilde{B}_V = \frac{\sqrt{1 + 3\sin^2 x}}{\cos^6 x} \tag{3.25}$$

$$\tilde{\kappa}_\psi = \frac{3(1 + \sin^2 x)}{\cos x (1 + 3\sin^2 x)^{3/2}} \tag{3.26}$$

$$\tilde{L}_P = \frac{1}{\alpha} \frac{\cos^3 x}{\sqrt{1 + 3\sin^2 x}} \tag{3.27}$$

Chan et al. (1994) also assume a constant mass density along a field line. Therefore,

$$\tilde{V}_A^2 = \frac{\tilde{B}^2}{\rho/\rho_0} = \tilde{B}^2. \tag{3.28}$$

Through these equations, each variable in Equation 3.22 is defined based on the latitude, which takes on values between 0, at the equator, and $x = \cos^{-1}\sqrt{1/L}$ at the point where the field line intersects Earth's ionosphere which is assumed for simplicity to lie at Earth's surface. The L value used here, as in Chan et al. (1994), is 6.6, corresponding to geostationary orbit at the magnetic equator. The boundary conditions used are $\widetilde{\delta\psi} = 0$ on both ends. This is because of the antisymmetric nature of the field-aligned eigenmode resulting in an antinode at the equator and because of the assumption of a perfectly conducting ionosphere.

The independent variables in Equation 3.22, in addition to the latitude, are therefore: the equatorial perpendicular beta based on the vacuum magnetic field strength, $\beta_{\perp 0V}$, the equatorial value of the anisotropy, δ_0, and the pressure gradient parameter α. Chan et al. (1994) state that α can take on values between 3 and 7, however they show results based only on a value of 6. The lower bound on α is determined by quiet-time magnetic field observations, and the upper bound is chosen to ensure stability with respect to interchange modes. The main focus of the work conducted here consists of calculating instability thresholds for additional values of α in the allowable range and determining the significance of the behavior of the instability thresholds. In addition, our work focuses primarily on the evolution of ballooning instabilities under conditions when the equilibrium develops towards more parallel pressure anisotropy. This evolution also forms the basis of a new model for substorm onset, which is also advanced in this thesis.

3.2 Numerical Approach

In general, Sturm-Liouville eigenvalue problems can be solved analytically if they have the form:

$$\frac{d^2Y}{dX^2} + WY = 0 \tag{3.29}$$

Where W is the eigenvalue. The Chan et al. (1994) equation only has this form for a specific set of parameters and must, therefore, be solved using numerical methods. A shooting method algorithm is used with an integration scheme. For simplicity, the eigenvalue problem is rewritten as:

$$\frac{d}{dx}\left(p(x)\frac{d\widetilde{\delta\psi}}{dx}\right) + q(x)\widetilde{\delta\psi} + w(x)\tilde{\omega}^2\widetilde{\delta\psi} = 0 \tag{3.30}$$

Here we follow the approach in Press et al. (1992), with the adoption of a shooting method which must be applied to solve this problem in the following manner:

1. The 2nd order ODE is divided into two first order ODE's as follows:

$$\frac{d\widetilde{\delta\psi}}{dx} = \frac{u}{p(x)} \tag{3.31}$$

$$\frac{du}{dx} = -(q(x) + w(x)\tilde{\omega}^2)\widetilde{\delta\psi} \tag{3.32}$$

2. The domain is discretized using a mesh with a constant grid size. For finding the instability thresholds, a mesh with 1000 points is used.
3. The boundary conditions at the ionosphere are set using the $\widetilde{\delta\psi} = 0$ condition, this corresponds to $\frac{du}{dx} = 0$ for Equation 3.32. For Equation 3.31, the boundary condition can be set to an arbitrary constant because the differential equation is homogeneous, and so any constant multiple of the eigenfunction is a solution. For this purpose, the value $u = 1$ was chosen, corresponding to $\frac{d\widetilde{\delta\psi}}{dx} = \frac{1}{p(x_I)}$ where x_I is the latitude at the point where the field line intersects the ionosphere assumed to lie at Earth's surface.
4. A guess is made for the eigenvalue $\tilde{\omega}^2$
5. Based on the boundary conditions and the guessed value of $\tilde{\omega}^2$, the differential equation is integrated from the ionosphere to the equator.
6. The value of $\widetilde{\delta\psi}$ at the equator represents the distance by which the guessed eigenvalue has caused the eigenfunction to miss its boundary condition, since the boundary condition there is $\widetilde{\delta\psi} = 0$.
7. Steps 4 to 6 are repeated with different guesses for $\tilde{\omega}^2$, these guesses can be entered into a root finding algorithm that methodically adds new guesses until a value of $\widetilde{\delta\psi}$ at the equator is reached that is sufficiently close to 0 within a defined tolerance. Several root finders were examined and used for this application in the course of the research presented in this thesis:

3.2.1 Selection of a Root Finding Algorithm

As a first step toward exploring the solution space of the eigenvalue problem, a root finder was created that can robustly search for eigenvalues on the complex plane to determine whether any eigenvalue has a positive, imaginary component. This root finder is described in detail in Appendix A1. In general, it uses Cauchy's argument principle to detect the presence of one or multiple roots in a region in the complex plane. If roots are detected in this region, it subdivides the domain into four sub-domains, each of which are tested for the presence of roots.

This is based in general on the work of Wilf (1978). Any sub-domain in which a root is not detected is discarded and any domain in which a root is detected is further subdivided in an iterative process. This is illustrated schematically in Figure 10.

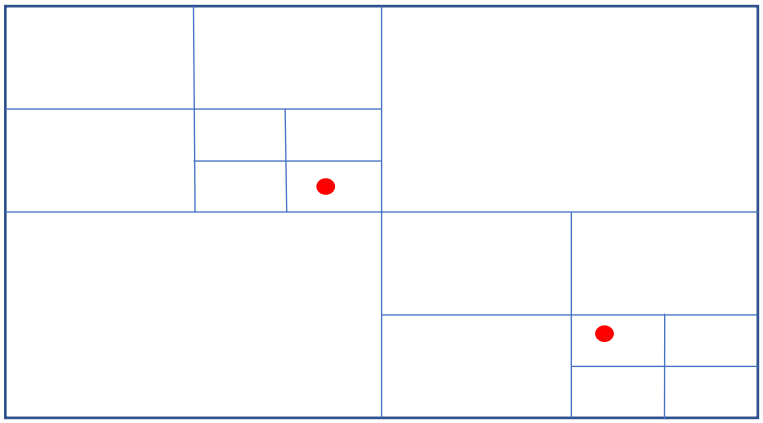

Figure 10: Schematic showing the process through which the complex subdivision scheme locates roots (represented by red dots)

This root finder is useful for mapping the solution space in an efficient way without requiring prior knowledge of the nature of the roots. Therefore, it was used when first exploring the solution space. The results showed that values of $\tilde{\omega}$ are either purely real, for a stable configuration, or purely imaginary for an unstable configuration. Therefore, the eigenvalue $\tilde{\omega}^2$ is always real and moves from positive values toward zero as the field line approaches instability. Because of this straightforward behavior of the eigenvalues, as determined through the complex subdivision scheme, a simple secant method algorithm can be applied for finding the eigenvalues as follows, based on two initial guesses $\tilde{\omega}_1^2$ and $\tilde{\omega}_2^2$:

$$\tilde{\omega}_3^2 = \tilde{\omega}_2^2 - \widetilde{\delta\psi}_0(\tilde{\omega}_2^2)\frac{\tilde{\omega}_2^2 - \tilde{\omega}_1^2}{\widetilde{\delta\psi}_0(\tilde{\omega}_2^2) - \widetilde{\delta\psi}_0(\tilde{\omega}_1^2)} \tag{3.33}$$

Where $\widetilde{\delta\psi}_0(G)$ is the value of $\widetilde{\delta\psi}$ at the equator for the guessed eigenvalue, G. Equation 3.33 is applied until the root is determined to within a defined error threshold.

3.2.2 Code Used for Solving for the Eigenvalue

The structure of the code used to solve for an eigenvalue based on specified values of $\beta_{\perp 0V}$ and δ_0 is shown in Figure 11. The individual functions are described below

Secant Algorithm

This function implements the secant algorithm as described above based on two initial guesses for the eigenvalue $\tilde{\omega}^2$. Generally, 0 and 0.1were used. The threshold for completion of the root finder is $\tilde{\omega}_2^2 - \tilde{\omega}_1^2 < 0.001$.

Variable Definition

This function creates a mesh based on a set number of points, usually 1000, and uses this to calculate all of the variables used in Equation 3.22. These variables are then combined into the parameters $p(x)$, $q(x)$, and $w(x)$ as described in Equation 3.30.

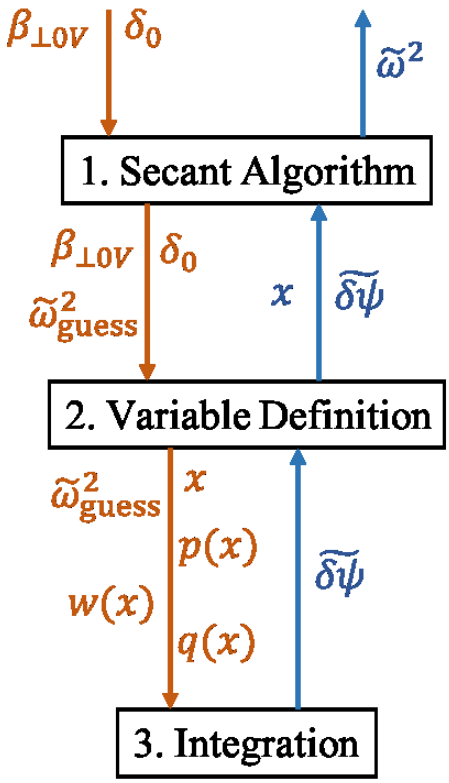

Figure 11: Schematic diagram of the structure of the code used to solve the eigenvalue problem, brown variables are passed downward, and blue variables are passed upward

Integration

The integration function takes in the values of $p(x)$, $q(x)$, and $w(x)$ as well as the guess for $\tilde{\omega}^2$ that comes from the secant algorithm. It uses these variables to implement a fourth order Runge-Kutta scheme to calculate the eigenfunction. The eigenfunction is then passed upward for use in iterating the secant algorithm until it converges within the specified threshold.

3.2.3 Calculations for Stable Eigenmodes

The first application of these functions is that they can be used to determine the eigenmodes for stable configurations of the field line. When the configurations are stable, i.e. $\tilde{\omega}^2 > 0$, Equation 3.22 describes an eigenmode that corresponds to coupled slow magnetosonic and Alfvèn waves in anisotropic cases and simply an Alfvèn wave in an isotropic case. This feature can be used to validate the code because the equation has an analytical solution in the case when there is zero pressure and the density has the following form (Ozeke et al., 2004):

$$\rho/\rho_0 = \frac{(1 + 3\sin^2 x)^2}{\cos^{12} x}. \tag{3.34}$$

The validation through comparison to this analytical solution is shown in Appendix A2. The numerical solution obtained shows good agreement with the analytical solution. Verification through progressive coarsening of the grid was also conducted using the point at which the eigenvalue moves from stable to unstable, described in more detail later, and as shown in Appendix A3.

The eigenfunction $\widetilde{\delta\psi}$ can be related to the components of the magnetic and electric fields through (Chan et al., 1994):

$$\delta B_{\parallel} = \frac{k_\phi c}{\omega} \frac{\beta_\perp}{2\tau L_P} \delta\psi \tag{3.35}$$

$$\delta B_{\psi} = \frac{k_\phi c}{\omega} \frac{\partial \delta\psi}{\partial l} \tag{3.36}$$

$$\delta B_\phi = \frac{B}{k_\phi}\frac{\partial}{\partial l}\frac{\delta B_\parallel}{B} \tag{3.37}$$

$$\delta E_\psi = -\frac{\omega}{k_\phi c}\delta B_\parallel \tag{3.38}$$

$$\delta E_\phi = -ik_\phi \delta\psi \tag{3.39}$$

The component $\delta E_\parallel$ is zero because there is no resistance, i.e., there is a perfect field-aligned conductivity, in a collisionless plasma. The components δE_ψ and δB_ϕ are both negligible because of the large azimuthal mode number assumption that is built into the equations. For comparison with Chan et al. (1994), the eigenmodes are plotted for a configuration with $\beta_{\perp 0V} = 0.5$, $\alpha = 6$, and using the asymmetric δ, in Figure 12. These plots are identical in appearance to the ones in Chan et al. (1994), with the only difference being due to an arbitrary normalization.

As mentioned earlier, the parameter δ is not symmetric and $\delta = 0.5$ corresponds to $P_\perp = 2P_\parallel$ whereas $\delta = -0.5$ corresponds to $P_\parallel = 1.5P_\perp$. Therefore, the eigenmodes show differing degrees of anisotropy in Figure 12. The plot therefore appears to suggest that a parallel anisotropy changes the eigenmode from its isotropic shape to a much lesser extent than an equivalent perpendicular anisotropy, which is incorrect. In addition, the use of a constant $\beta_{\perp 0V}$ for plotting the eigenmodes corresponds to quite different values of $\bar{\beta}_0$. For values of $\delta = -0.5$, $\delta = 0$, and $\delta = 0.5$, then $\beta_{\perp 0V} = 0.5$ corresponds to $\bar{\beta}_0 = 2.1$, $\bar{\beta}_0 = 0.89$, and $\bar{\beta}_0 = 0.75$, respectively. To avoid this issue, eigenmodes are plotted using δ_{sym} and a constant $\bar{\beta}_0 = 0.5$ in the paper developed for publication and which is presented in the following chapter.

Figure 12 does however illustrates the manner in which the parallel eigenmode in the magnetic field becomes significantly localized near the equator when the anisotropy is strongly perpendicular. This is possibly due this perpendicular anisotropic configuration approaching the mirror instability threshold. Therefore, the perturbation in the magnetic field may take on a shape

similar to the one shown in Figure 8. The assembled code can now be used to calculate the point at which the eigenvalue $\tilde{\omega}^2$ becomes less than zero, i.e., the ballooning instability threshold.

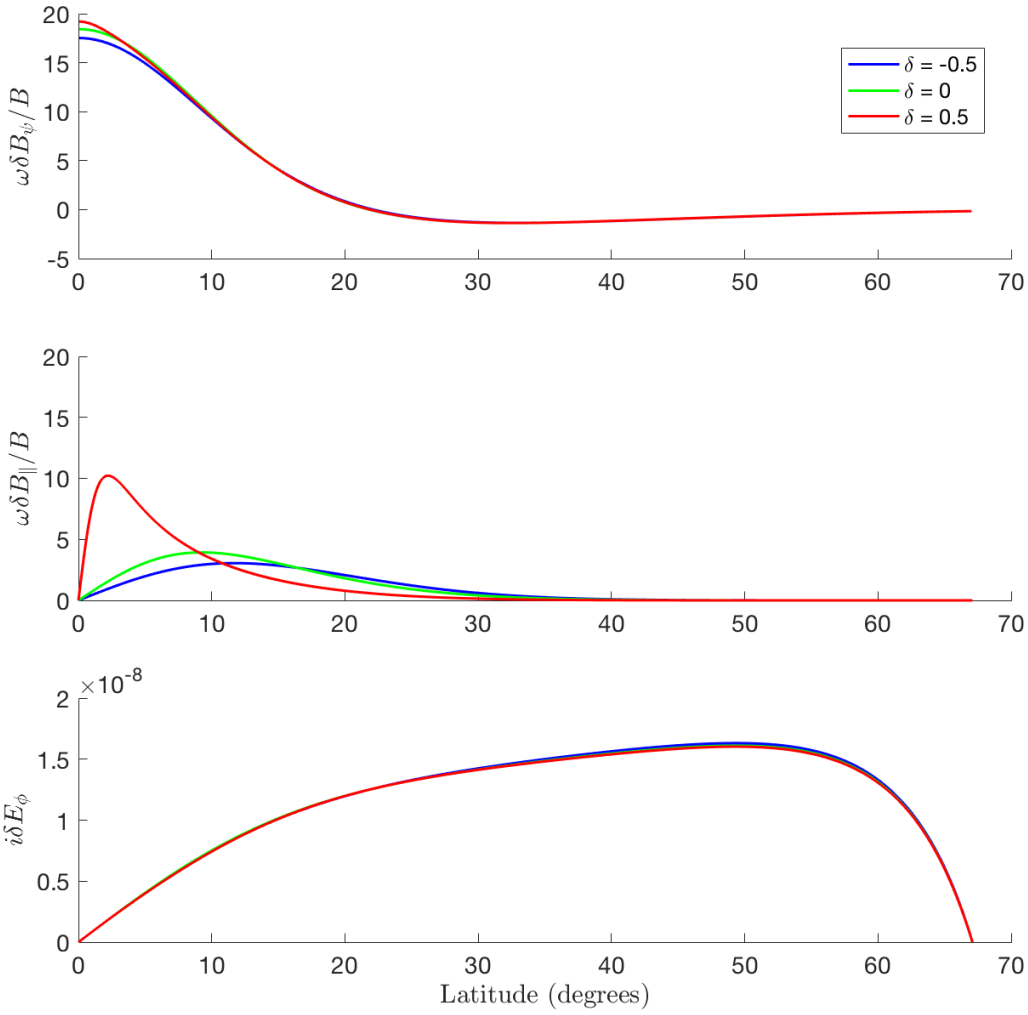

Figure 12: Eigenmodes for $\beta_{\perp 0V} = 0.5$ and $\alpha = 6$ plotted as in Chan et al. (1994)

3.2.4 Calculations of the Ballooning instability thresholds

The method used to calculate the instability threshold simply involves adding another layer to the code structure in Figure 11 as shown in Figure 13. This additional layer includes another secant algorithm which finds the value of $\bar{\beta}_0$ for which the eigenvalue becomes zero, for

a given anisotropy. The tolerance used for finding $\bar{\beta}_0$ is 0.001 as with the other secant algorithm. Since the instability thresholds for a range of δ_{sym} are required, the function contains a for loop that begins at $\delta_{\text{sym}} = -1$ and continues until $\delta_{\text{sym}} = 1$. For each subsequent step in the loop, the threshold $\bar{\beta}_0$ from the previous step is used as an initial guess.

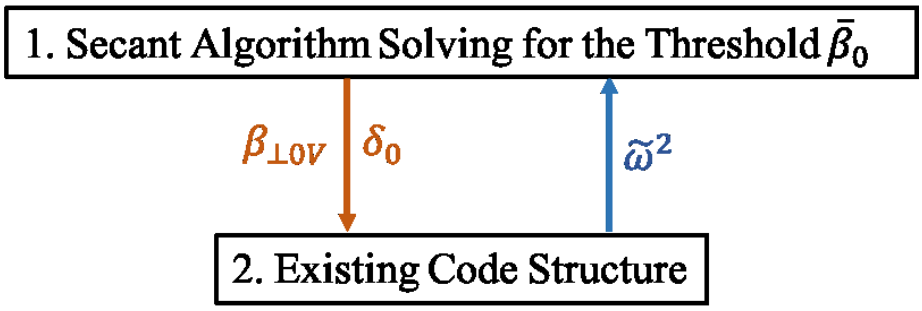

Figure 13: Code structure for finding the instability threshold

The results of these calculations are shown in the paper presented in the subsequent chapter.

Chapter 4: Paper Examining Ballooning Instability Thresholds Prepared for Submission

This chapter consists of a manuscript for a paper to be submitted to the journal Geophysical Research Letters and is formatted accordingly.

A new substorm onset mechanism: Increasingly parallel pressure anisotropic ballooning

L. R. Oberhagemann[1] and I. R. Mann[1]

[1]Department of Physics, University of Alberta, Edmonton, Canada

Corresponding author: Luke Oberhagemann (oberhage@ualberta.ca)

Key Points:

- We propose a new paradigm for substorm onset, where a ballooning instability is triggered by pressure anisotropy becoming more parallel
- This increasingly parallel pressure anisotropy occurs as a direct consequence of tail stretching during the substorm growth phase
- Our paradigm predicts increasingly parallel temperature anisotropies pre-onset, and onset locations in the transition region, as observed

Abstract

Recent observations demonstrating that auroral beads are observed prior to the majority of substorm onsets (Kalmoni et al., 2017) have reinforced the potential importance of ballooning instabilities for near-Earth magnetospheric substorm onset. Here we examine pressure anisotropic ballooning instabilities in stretched magnetotail geometries. Our results show that transition from an initially perpendicular anisotropy toward parallel anisotropy reduces the plasma β threshold for triggering a ballooning instability. Such increasingly parallel anisotropies can form as a direct consequence of tail stretching that occurs during the late substorm growth phase, through the well-known effects of drift shell splitting and the competition between betatron and Fermi processes in the tail. We propose such increasingly parallel ballooning triggers auroral substorm onset on field lines in the transition region between dipolar and tail-like fields, consistent with observational constraints on the location of the onset arc with respect to the ion isotropy boundary in the magnetotail.

Plain Language Summary

The aurorae, the northern and southern lights, occur when charged particles originating from the sun and interacting with Earth's magnetic field impact the atmosphere. At times, extreme events on the sun or in Earth's magnetic field can have remarkable impacts on the aurorae. Here we provide a new potential explanation for a process which can trigger the rapid release of stored energy in the nightside of near-Earth space. Such rapid releases of energy lead to the development of very spectacular bright and dancing displays of the aurorae in a process known as a substorm. While the processes responsible for powering this is well understood, why the lights begin to dance exactly when then do, a moment known as onset of the substorm, is not. It is generally agreed that this onset occurs after a period of stretching of the magnetic field on the nightside of Earth (away from the sun). Our new model suggests that as this stretching occurs, there is more pressure parallel to the magnetic field than across it in a confined region of the field. In our new paradigm, this leads to an instability, called a ballooning instability, which explosively releases energy in the magnetic field, leading to a substorm. The confined region of the field that our work predicts agrees well with known locations where these bright substorm aurorae begin and where possible signatures of ballooning instabilities have been observed.

1. Introduction

The nature of the processes active at the onset of the expansion phase of magnetospheric substorms continues to be a point of significant controversy in the magnetospheric community (see e.g., Angelopoulos et al., 2008; Lui 2004, 2018; Segeev et al., 2012). The sequence of events that is observed in the aurorae during a substorm is generally accepted to consist of an initial growth phase, associated with stretching of the tail and stable growth phase arcs, and subsequent brightening of the most equatorward arc followed by its rapid poleward motion and breakup (e.g., McPherron 1970, Akasofu 1964). Consensus about the processes in the magnetotail that are responsible for these auroral signatures, however, remains elusive. In particular, two opposing views remain to be reconciled: the out-to-in mechanism, whereby the substorm sequence begins with magnetic reconnection at a near-Earth neutral line, and the in-to-out mechanism, whereby the substorm sequence begins closer to the Earth, perhaps due to the triggering of plasma instabilities causing current disruption (see e.g., the review by Sergeev et al., 2012). Specifically, in relation to the latter, a number of fluid and kinetic scale ballooning instabilities have been proposed as a potential trigger for substorm expansion phase onset.

In general, ballooning instabilities are a type of quasi-interchange instability that may form when the local pressure gradient aligns with the magnetic field line curvature vector (e.g., Ferrière, 2001). Their presence in the magnetotail has been documented through in-situ observations (e.g., Panov et al., 2012). Recent results have also added significant support to the hypothesis that ballooning instabilities may play a key role in substorm onset. For example, Kalmoni et al. (2017) used a statistical analysis to demonstrate that auroral beads are an extremely common feature of the most equatorward auroral arc that brightens at the beginning of a substorm. Kalmoni et al. (2018) then linked this phenomenon to kinetic Alfvén waves, which can couple the equatorial magnetotail and the ionosphere, through the construction of an observational dispersion relation. This is relevant to ballooning instabilities since, as demonstrated by Viñas and Madden (1986), the formation of ballooning instabilities in the magnetotail may lead to the excitation of kinetic Alfvén waves.

Numerical simulations by Zhu et al. (2017) further established a mechanism through which ballooning instabilities can lead to magnetic reconnection in the mid-tail. This is significant since it seems likely that magnetic reconnection plays a key role in the development of the expansion phase. As seen, for example, in the poleward expansion of the onset arc and in the development of auroral breakup, regardless of the precise nature of the physical process which triggers expansion phase onset (e.g., Sergeev et al., 2012). In light of these results, in this paper we examine the conditions required for the triggering of pressure anisotropic ballooning instabilities in a model magnetotail with a stretched magnetic field. Our results form the basis of a new substorm onset paradigm advanced here. Specifically, we propose that a shift *toward* parallel temperature anisotropy, which does not necessarily have to reach a state with an overall parallel pressure anisotropy, can trigger ballooning in the magnetotail leading to substorm expansion phase onset.

In what follows, we outline the fundamental elements of the new onset model, demonstrating how stretching of the nightside field during the growth phase can naturally lead to onset without requiring the action of magnetic reconnection at a near-Earth neutral line, although such reconnection likely follows later in the substorm sequence for events where auroral breakup and the release of the westward travelling surge occurs. We further compare our model predictions to well-known observational constraints in relation to the location of the onset arc with respect to various plasma boundaries in the growth phase magnetotail. We also comment on potential implications of our model for understanding the distinction in tail response between pseudo-breakups and tail activations, the latter of which result in the development of full substorms.

2. Methods

We adopt the modelling approach formalised by Chan et al. (1994) to analyse the characteristics of pressure anisotropic ballooning modes with application to the late growth phase magnetotail and to a potential triggering of substorm onset. The Chan et al. (1994) model assumes rotational symmetry near the noon-midnight plane and neglects kinetic effects. A

significant benefit of this model is that it accounts for the effect of pressure anisotropy on instability thresholds, including ballooning as well as firehose and mirror thresholds. Briefly, the Chan et al. (1994) model consists of two parts. The first is an anisotropic pressure equilibrium calculated using a multiscale perturbation method where the zero-order term is given by a vacuum dipole field. This equilibrium is illustrated by the surfaces of constant flux in Figure 1. In our work we adopt a symmetric pressure anisotropy parameter δ_{sym} to parameterise our results, which has the advantage that positive and negative values represent symmetric levels of perpendicular and parallel pressure anisotropy. With this approach, δ_{sym} represents the level of perpendicular ($\delta_{sym} > 0$) or parallel ($\delta_{sym} < 0$) anisotropy and is defined as:

$$\delta_{\mathrm{sym}} = \begin{cases} 1 - \dfrac{\beta_{\parallel}}{\beta_{\perp}} \text{ for } \beta_{\parallel} > \beta_{\perp} \\ -\left(1 - \dfrac{\beta_{\perp}}{\beta_{\parallel}}\right) \text{ for } \beta_{\perp} > \beta_{\parallel} \end{cases} \quad (1)$$

$P_{\parallel}$ and $P_{\perp}$ are the parallel and perpendicular pressures, and $\beta_{\parallel}$ and $\beta_{\perp}$ are the parallel and perpendicular plasma beta. Here, therefore, $\delta_{\mathrm{sym}} = 0$ describes an isotropic pressure, while $P_{\perp} = 2P_{\parallel}$ results in $\delta_{sym} = 1$, and $\delta_{\mathrm{sym}} = -1$ corresponds to $P_{\parallel} = 2P_{\perp}$. This parameter is slightly modified from the asymmetric anisotropy parameter in Chan et al. (1994).

An averaged equatorial plasma beta, $\bar{\beta}_0$, is derived from the parallel and perpendicular components of β through $\bar{\beta}_0 = 1/3(\beta_{\parallel_0} + 2\beta_{\perp_0})$ where the subscript 0 denotes values at the equator. The background pressure gradient parameter α is related to the pressure through $P_0 \propto L^{-\alpha}$ where L represents the dipole L-shell parameter. For the purpose of this work, we examine values of α between 3 and 7. This parameter space is similar to that identified by Chan et al. (1994) and represents a lower bound based on quiet time observations and an upper bound chosen to ensure stability with respect to interchange modes. Figure 1 shows that finite pressure pushes surfaces of constant flux outward and, significantly, the curvature of these surfaces in the equatorial plane increases monotonically with δ_{sym}.

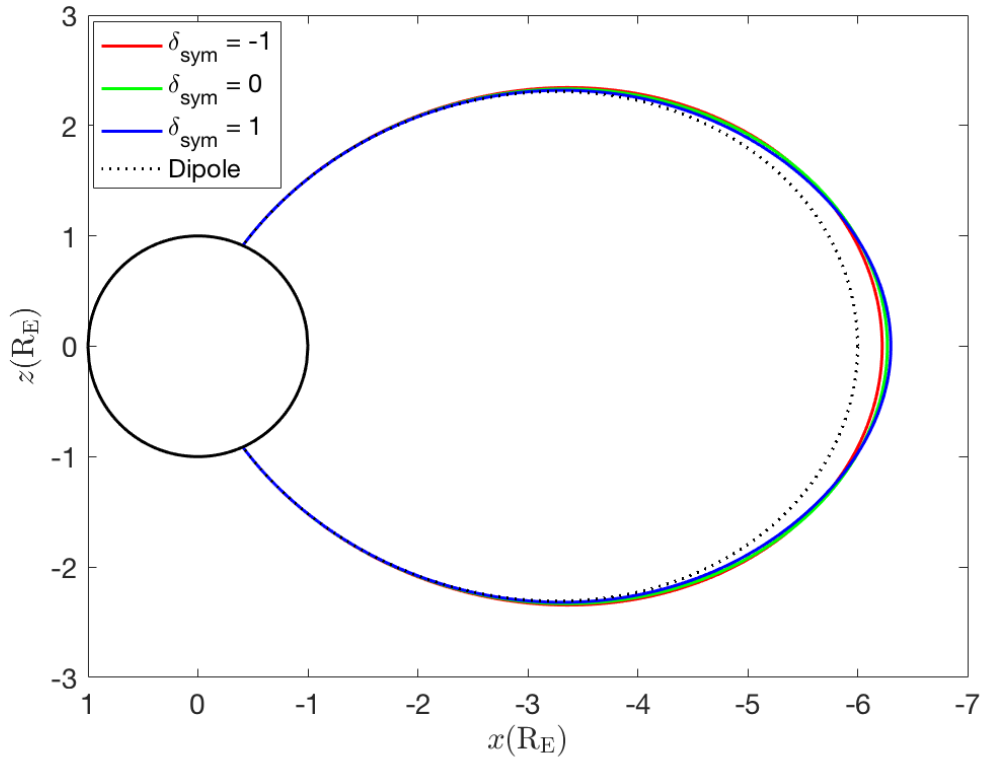

Figure 1: Equilibrium background magnetic field lines showing the effects of a finite, anisotropic pressure on an equilibrium magnetic field configuration with an average equatorial β, $\bar{\beta}_0 =$ 0.75, and a pressure gradient parameter, α=6.

The second part of the Chan et al. (1994) model is a stability analysis of the modes supported by the equilibrium, which is calculated for the lowest antisymmetric mode as this mode is the most unstable (Chen and Hasegawa, 1991). Chan et al. (1994) use a magnetic field defined by $\boldsymbol{B} = \nabla\phi \times \nabla\psi$, where ϕ is the azimuthal angle and ψ is the poloidal flux coordinate. The two parts of the model result in the following Sturm-Liouville eigenvalue problem:

$$B\frac{\partial}{\partial l}\left(\frac{\sigma}{R^2 B}\frac{\partial \delta\psi}{\partial l}\right) + \frac{\beta_E \kappa_\psi}{R^2 L_P}\,\delta\psi + \frac{\omega^2}{R^2 V_A^2}\delta\psi = 0 \qquad (2)$$

Here B is the magnetic field strength, l is the distance along the field line, R is the cylindrical radial coordinate, κ_ψ is the field line curvature in the ψ direction, and V_A is the Alfvén speed. The parameter σ is given by $\sigma = 1 + \frac{1}{2}\beta_\perp \delta$ and represents the firehose instability threshold, where δ is the (asymmetric) anisotropy parameter $\delta = 1 - \frac{\beta_\parallel}{\beta_\perp}$. Negative values of σ represent a

crossing of the firehose instability threshold. This threshold is higher than the ballooning threshold for the set of parameters that we examine here and is therefore not reached. $\beta_E = \frac{\beta_\perp}{2}(1 - \delta + \frac{\sigma}{\tau})$ is defined as an effective β value, while $\tau = 1 - \frac{\beta_\perp \delta}{1-\delta}$ represents the mirror instability threshold. As with the firehose instability, negative values of τ are mirror unstable. The pressure gradient length scale, L_P, is given by:

$$L_P = \frac{\psi_V}{\alpha R B_V} \tag{3}$$

Where the subscript V denotes values for a vacuum dipole magnetic field

Unlike the firehose threshold, the mirror instability threshold is relevant for our analysis since the ballooning instability threshold merges with the mirror threshold for sufficiently large perpendicular pressure anisotropy. The boundary conditions for Equation 2 are $\delta\psi = 0$ at the equator and the ionosphere, due to the antisymmetric nature of the mode that is analyzed and the assumption of a perfectly conducting ionosphere. A positive, imaginary component of the eigenvalue, ω, represents conditions that are ballooning unstable. See Chan et al. (1994) for more details.

In the stable, isotropic regime, the eigenmodes of Equation 2 describe Alfvén waves standing on the background equilibrium magnetic field lines. Figure 2 shows example eigenmodes for the same set of parameters as the surfaces of constant flux shown in Figure 1. A significant feature of these eigenmodes is that the deflection of the magnetic field in the direction parallel to the magnetic field line, shown in Figure 2b, is significantly more peaked near the equator for more perpendicular pressure anisotropies. This behavior suggests a possible observational diagnostic for monitoring change in the equatorial pressure anisotropy, which will be significant in the context of the new paradigm for substorm onset we advance here.

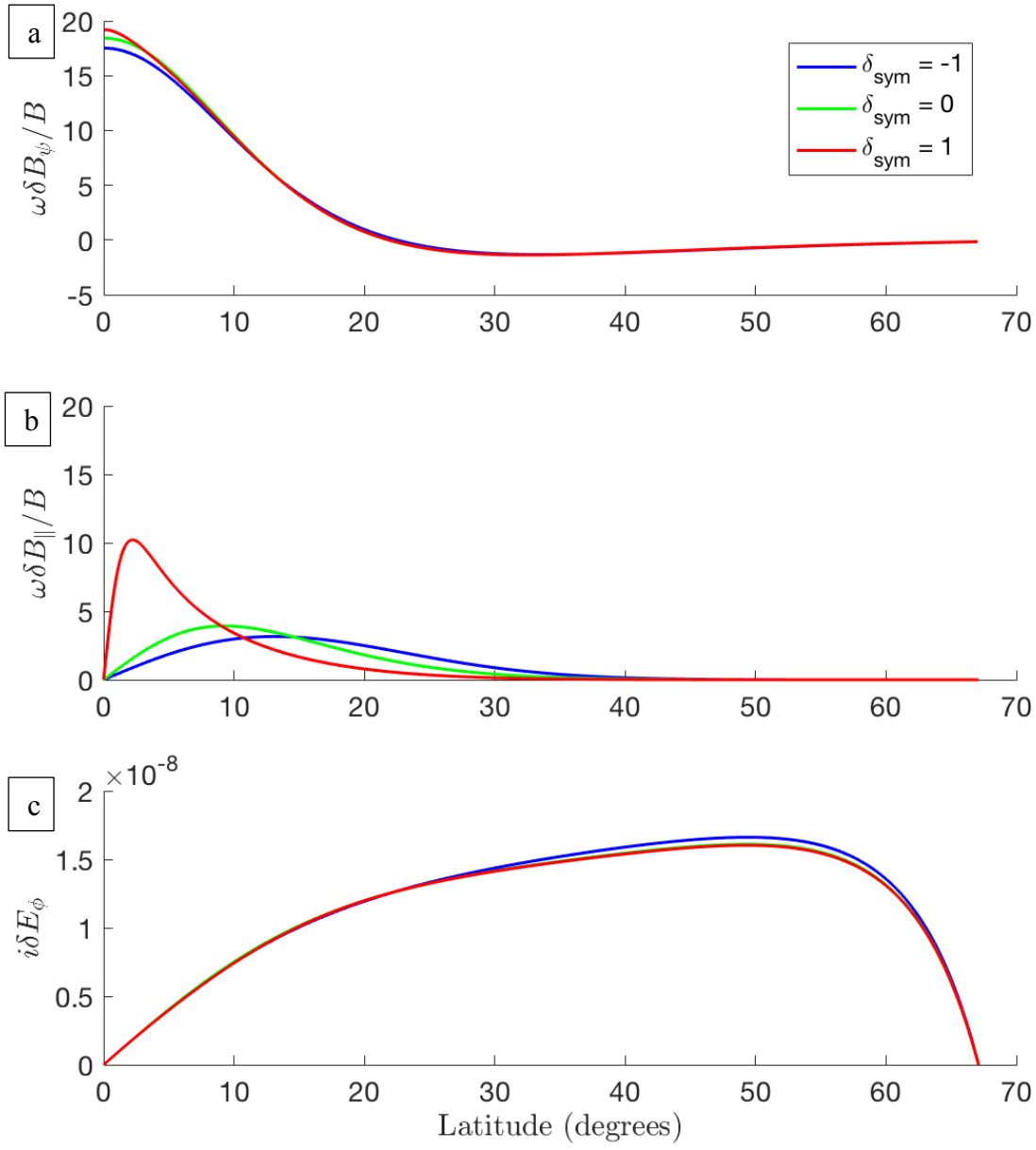

Figure 2: Eigenmodes of equation 2 for $\bar{\beta}_0 = 0.75$ and $\alpha = 6$. *The amplitudes of the eigenmodes are arbitrary, having been normalized by an arbitrary constant value.*

3. Ballooning Instability Thresholds

In their previous work, Chan et al. (1994), while addressing a range of anisotropies, focused their conclusions primarily on the characteristics of Alfven-ballooning modes in the

near-Earth magnetosphere under conditions of strong perpendicular temperature anisotropy. Here we use the same formalism, relying on the ordering used by Chan et al. (1994) for the weak anisotropy case, where $\beta_{\parallel}$ and $\beta_{\perp}$ are of the same order. As described subsequently, this is the regime relevant for our findings. The ordering in Chan et al. (1994) relies on the use of $\beta_{\perp V}$, the value of $\beta_{\perp}$ calculated using the vacuum dipole field, as a small parameter ($\beta_{\perp V}$ is smaller than $\beta_{\perp}$, with $\beta_{\perp V} = 1$ corresponding to $\beta_{\perp} = 4$). This assumption breaks down somewhat for the higher β instability thresholds. However, it is expected that the qualitative shapes of the thresholds are correct, particularly since they are constrained by the analytically defined mirror instability threshold, and by lower β thresholds nearby. For reasons described below, we focus on examining the implications of the transition towards Alfven-Ballooning instability under conditions of decreasing δ_{sym} associated in particular with the stretching of the magnetic field in the tail during the growth phase. We discuss the implications of such instability transitions in the context of a new paradigm for substorm onset in further detail in section 4 below.

The major factors determining whether instability may occur are the pressure gradient parameter, α, the equatorial pressure anisotropy, $\delta_{\mathrm{sym},0}$, and the average equatorial plasma beta, $\bar{\beta}_0$. The ballooning instability thresholds, where areas under the curves represent stable configurations, are shown in Figure 3. In our subsequent discussion, we focus on changes in the system that begin from a mildly perpendicular anisotropic state, since in general the magnetotail is believed to typically be in this state at the start of the substorm growth phase as a result of the fact that betatron acceleration is typically stronger than Fermi acceleration for Earthward drifting particles on quasi-dipolar field lines (e.g., Zaharia, 2006).

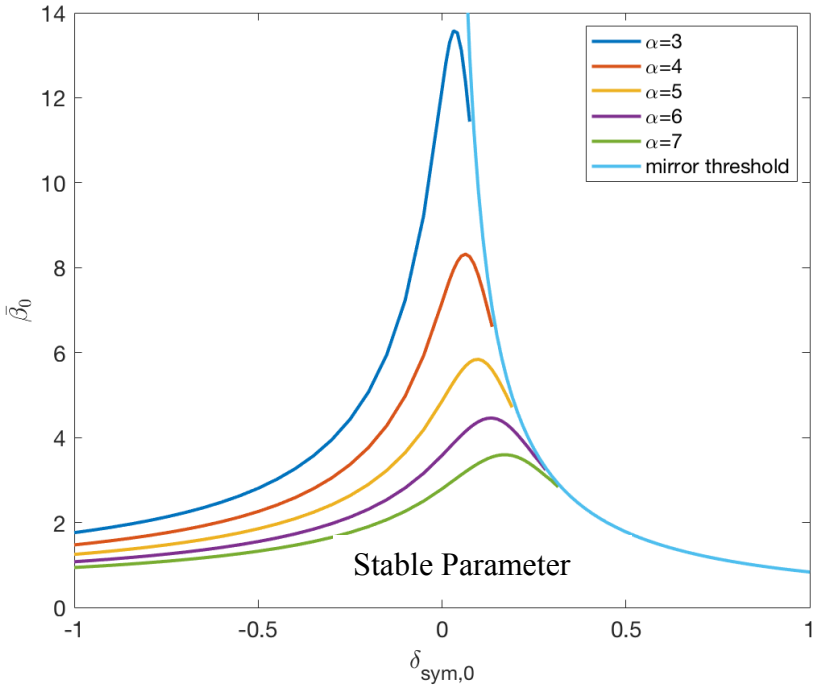

Figure 3: Instability thresholds as a function of δ_{sym} and $\bar{\beta}_0$ for a range of pressure gradients as specified by α.

The $\alpha = 6$ threshold in Figure 3 corresponds to the results presented in Chan et al. (1994). A notable and key feature of the instability thresholds in the context of our onset hypothesis is the presence of a peak at slightly perpendicular pressure anisotropies, i.e., where δ_{sym} is small and greater than zero. As the pressure gradient increases (increasing α), for example as a result of the addition of new magnetic flux into the polar cap during the substorm growth phase, the peak moves further in the direction of perpendicular anisotropy (i.e., where δ_{sym} gets larger and more positive). Significantly, this means that, starting from modest perpendicular temperature anisotropy, the $\bar{\beta}_0$ required to reach the instability threshold for fixed δ_{sym} decreases for increasing α.

Most significantly for our proposed onset mechanism, a change of the anisotropy parameter from an initially mild perpendicular anisotropy in the direction of more parallel anisotropy, without necessarily having to reach a parallel pressure anisotropic state overall (i.e., δ_{sym} can remain

positive), allows for a reduction in the $\bar{\beta}_0$ threshold for instability. As illustrated in the equilibrium shown in Figure 1, a movement toward parallel anisotropy corresponds to a reduction in field line curvature. Because field line curvature has a stabilizing effect, reducing this curvature is destabilizing. In our opinion, conditions where the equilibrium of the magnetotail transitions from an initial state with a modest perpendicular anisotropy towards a more parallel anisotropic state will occur naturally as a result of tail stretching during the growth phase. In our view, the transition to a ballooning unstable configuration as a result of a more parallel anisotropic configuration developing during a period of tail stretching should be identified as defining the end of the growth phase and the triggering of substorm onset.

4. Increasingly Parallel Anisotropic Ballooning as a Substorm Onset Mechanism

Here we examine the hypothesis that movement away from an initially mildly perpendicular pressure anisotropy towards a more parallel anisotropic state until onset is triggered by a transition to ballooning instability causes substorm onset. Specifically, we contextualize this in terms of previous measurements which lend significance to several sets of in-situ observations at geosynchronous orbit. For example, Baker et al. (1978) observed electron pitch angle distributions that were associated with a more parallel temperature anisotropy at geosynchronous orbit immediately prior to substorm onset. Electrons have much smaller gyroradii than ions, such that they are likely to remain adiabatic over a region which extends further downtail than the ions under any given stretched magnetic field configuration due to a lack of adiabatic scattering. The Baker et al. observations hence confirm that tail stretching should indeed increase the level of parallel pressure anisotropy, decreasing the value of δ_{sym} such that for electrons observed at geosynchronous orbit δ_{sym} can reach a minimum immediately prior to onset.

Birn et al. (1997), again using data from geosynchronous orbit, observed a similar significant movement away from an initially mildly perpendicular anisotropy state during the growth phase to a much more pressure anisotropic state close to substorm onset. In the observations reported by Birn et al. (1997), this was especially clear in the electrons but was also seen, but to a somewhat lesser extent, in the ions. Of course, for sufficiently stretched

magnetotail magnetic fields there will be a transition from adiabatic to isotropic distributions in the tail at distances downtail beyond which the field line curvature becomes of the same order as the ion gyroradius in a given energy channel (the so-called ion isotropy boundary; see e.g., Wang et al., 2013).

The location of the Baker et al., (1986) and Birn et al. (1997) observations at geosynchronous orbit is also significant as it falls into the region in the tail where field lines can often transition from being dipolar to tail-like, and where the radial pressure gradients and plasma β values are relatively high. Further, these parameters necessarily increase as the tail stretches during the growth phase. Therefore, these observations are consistent with the concept that during the growth phase, as the pressure gradient α, and plasma beta, $\bar{\beta}_0$, increase, tail stretching also decreases δ_{sym}. If the tail starts from a mildy perpendicular pressure anistropic state, the change in all of the parameters will be in the direction of destabilization until the tail reaches a state which exceeds the ballooning instability threshold and substorm onset occurs. We include a brief discussion of this mechanism here, and reserve a more detailed description, including detailed numerical modelling, for a subsequent paper.

The expectation that the growth phase will be accompanied by a transition toward more parallel pressure anisotropy in the tail is also entirely consistent with the well-known drift shell splitting phenomena. For example, in the context of magnetopause shadowing, higher pitch angle particles drifting from the nightside reach larger distances on the dayside than lower pitch angle particles on similar drift shells. As a result, these higher pitch angle particles are preferentially lost since their drift shells can encounter the magnetopause while those of lower pitch angle particles do not (e.g. Baker et al., 1978, Wang et al., 2013). In the context of our substorm onset model, the relatively rapid stretching of the tail that occurs during the substorm growth phase can occur on timescales comparable to the drift times of the pressure carrying ions. Therefore, particles on such drift paths will have pitch angle dependent paths. This may lead to increasingly parallel pressure anisotropies in the tail as a result of their adiabatic drift-bounce motion, which is described in more detail below. The proposed mechanism whereby increasingly parallel pressure develops in the tail during the growth phase is illustrated in Figure 4.

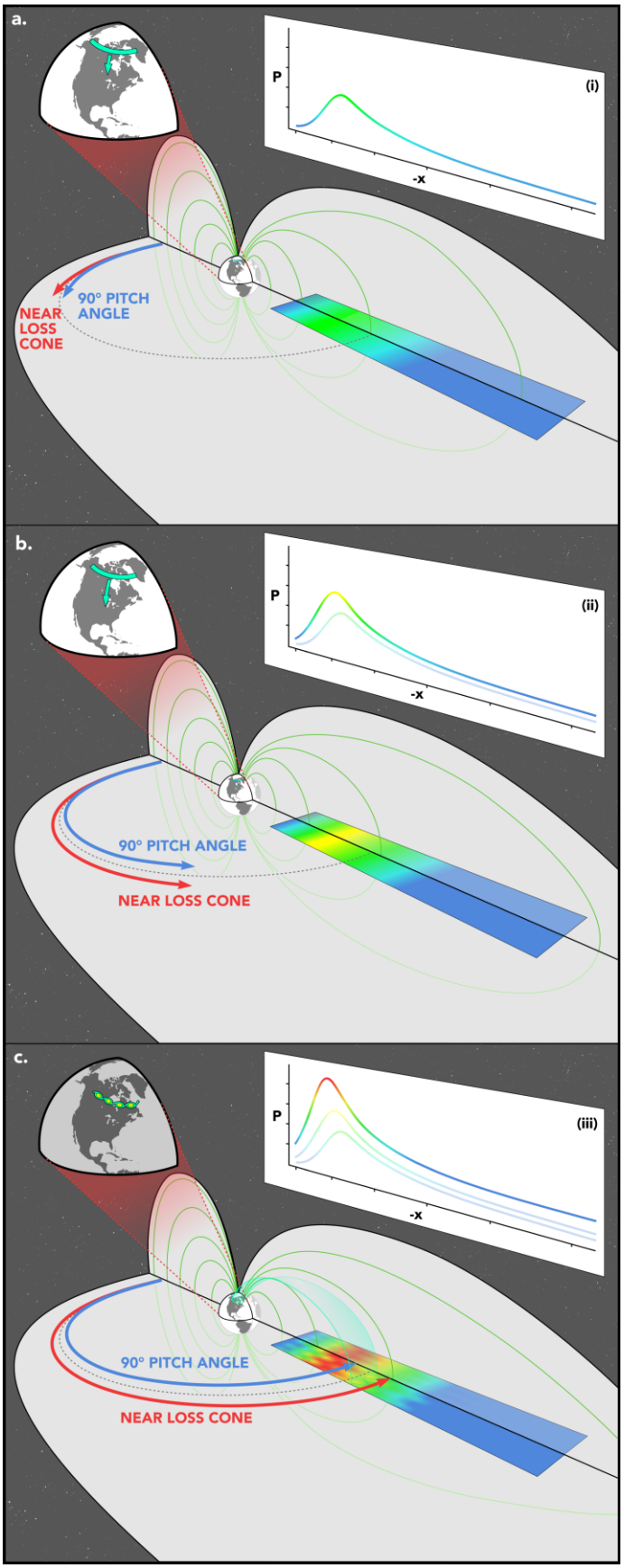
a.
P
(i)
-x
90° PITCH ANGLE
NEAR LOSS CONE
b.
P
(ii)
-x
90° PITCH ANGLE
NEAR LOSS CONE
c.
P
(iii)
-x
90° PITCH ANGLE
NEAR LOSS CONE

Figure 4: Ballooning instability triggered substorm onset through rapid stretching during the substorm growth phase causing increasingly parallel pressure anisotropy. The progression of the aurorae during this time in addition to the change of the nightside equatorial pressure distribution are illustrated in the upper portions of each panel. Panels a. and b. show the pitch angle (p.a.) dependent motion of particles around Earth as the tail stretches; red arrows show an equatorial drift trajectory for pitch angles close to the loss cone, and blue arrows a trajectory at 90 degree p.a. Panel c. shows the consequences of this drift shell splitting as particles reach a tail that is more stretched than the one in which they originated. As more low p.a. particles mix with fewer high p.a. particles in the high pressure gradient region, and the increased length of the field lines causes p.a.'s to decrease in general, the overall p.a. distribution becomes more parallel to the field lines and a ballooning instability may form

Early in the growth phase, particles drifting around the Earth and initially residing in the tail will experience a less stretched tail at the start of their drift orbit. As the growth phase develops, the tail becomes more stretched and its magnetic field strength decreases. This leads to drift shell splitting once the particles re-enter the stretched tail later in the growth phase. The 90-degree pitch angle particles move Earthward as they follow a trajectory of constant magnetic field strength, while particles with lower pitch angles move away from Earth to conserve the second adiabatic invariant as their bounce paths now occur on longer field lines. Because the region of interest here is one with high earthward pressure gradients, this results in higher pitch angle particles originating from a lower pressure region than low pitch angle particles. Therefore, particles mix in a manner that leads to their distributions developing more parallel anisotropies at a fixed location in the high pressure gradient region of the more stretched tail at a later stage in the growth phase. Due to the radially confined nature of the region in which a large Earthward pressure gradient exists, the impact of this for destabilizing ballooning can only occur in a radially localized part of the transition region from dipole-like to tail-like fields. This is also important since observations of the auroral beads associated with substorm onset occur in a latitudinally narrow region, but which can be more extended in the east-west direction in the ionosphere. In our opinion, it would be natural for the latitudinally localized region in which substorm onset occurs to be identified with the narrow region in the tail where the steep Earthward pressure gradient, large plasma beta, and decreasing δ_{sym} combine to trigger a ballooning instability which marks the end of the growth phase, and the onset of the expansion phase of the substorm.

In addition to the effect of drift shell splitting, and more generally, stretching of the tail during the substorm growth phase may cause increasingly parallel anisotropies through a reversal of the usual adiabatic effects on particles moving earthward from the tail. As described earlier, mildly perpendicular anisotropies are typically created as particles travel earthward from the distant tail and encounter increasing magnetic field strengths and less stretched field lines, since the betatron acceleration is more efficient than the Fermi acceleration in such fields. However, in a field which is stretching it may be possible for particles drifting from a less stretched to a more stretched tail to encounter decreased magnetic field strengths and increasingly stretched field lines. If the resulting betatron deceleration is faster than the Fermi deceleration, the effect would be reversed, causing a movement toward more parallel pressure anisotropy.

5. Conclusions

Here we have presented a new paradigm for substorm onset. We advance the hypothesis that movement away from an initially mildly perpendicular pressure anisotropy towards a more parallel anisotropic state continues during the growth phase until expansion phase onset is triggered by the transition to ballooning instability caused by an increasingly parallel pressure anisotropic distribution. Calculation of the instability thresholds in a model field demonstrate the feasibility of this hypothesis. These indicate significantly that there is a peak in the equatorial plasma beta, $\bar{\beta}_0$, threshold, above which ballooning instability is reached, as a function of a symmetric anisotropy parameter, δ_{sym}, at perpendicular values. This highlights a pathway for ballooning-related onset arising from a transition towards a more parallel anisotropic state. Note also that instability does not require the transition to an overall parallel anisotropic state; instead, the point at which the tail transitions to a ballooning unstable state will be dependent upon the local pressure gradients and plasma beta which develop during the course of the growth phase.

Nonetheless, regardless of the specific parameters for any given substorm, our new paradigm for onset successfully predicts that onset should occur Earthward of the ion isotropy boundary for the energy of ions which are carrying the bulk of the thermal pressure, and in the high plasma beta region of the tail characterized by the transition from dipole-like to tail-like fields. Moreover, our model predicts that the unstable region which develops will be latitudinally

narrow consistent with the characteristics of auroral beads which have been identified as possibly the ionospheric signature equatorial ballooning communicated to the ionosphere by kinetic Alfven waves. In our view a natural extension of our paradigm is to associate the difference between events with the characteristics of pseudo-breakups and those of full substorm onsets with whether or not the dynamics associated with a ballooning unstable tail subsequently result in additional nightside magnetic reconnection. In cases with appropriate subsequent magnetic reconnection in the near-tail, the closure of nightside flux could lead to the poleward expansion of the aurora and the release of the westward travelling surge in a full substorm. Ballooning activations where this does not occur might be more closely associated with pseudo-breakups.

We placed our findings in the context of a proposed substorm onset mechanism that occurs as a result of the established stretching of the magnetotail during the substorm growth phase. The mechanism which we propose to generate the required increasingly parallel pressure anisotropy relies on the well-understood processes of drift shell splitting. The pitch angle dependent behavior naturally produces the required transition towards increasingly parallel anisotropic distributions in the onset region, as particles drift between field lines whose tail curvature and stretching increases through the course of the growth phase, because magnetic flux is added to the polar cap by dayside reconnection. This process is not highly dependent on the specific conditions of the magnetic field topology in the tail, requiring only appropriate stretching of the field, and may therefore apply across a broad range of events and tail magnetic field configurations. Our proposed onset mechanism addresses the long unanswered question of why a specific substorm onset occurs when it does. In our view, it also provides an important new framework centered around the level of (mostly ion) pressure anisotropy and through which substorm onsets and related magnetotail and auroral dynamics can be better interpreted. Further detailed modelling in realistic magnetotail geometries, as well as renewed studies of the temporal dynamics of energetic ion anisotropies during the growth and early expansion phases of substorms, should be completed to further test our theory.

Acknowledgements

This was work was supported by an NSERC Discovery Grant to IRM and the NSERC CREATE grant for the International Space Masters (ISM) program. The authors thank Andy Kale for the preparation of Figure 4.

References

Akasofu, S.-I. (1964). The development of the auroral substorm. *Planetary and Space Science*, *12*(4), 273–282. https://doi.org/10.1016/0032-0633(64)90151-5

Angelopoulos, V., McFadden, J.P., Larson, D., Carlson, C.W., Mende, S.B., Frey, H., Phan, T., Sibeck, D.G., Glassmeier, K.H., Auster, U. and Donovan, E., 2008. Tail reconnection triggering substorm onset. *Science*, *321*(5891), pp.931-935.

Baker, D. N., Higbie, P. R., Hones, E. W., & Belian, R. D. (1978). High-resolution energetic particle measurements at 6.6 RE 3. Low-energy electron anisotropies and short-term substorm predictions. *Journal of Geophysical Research: Space Physics*, *83*(A10), 4863–4868. https://doi.org/10.1029/JA083iA10p04863

Birn, J., Thomsen, M. F., Borovsky, J. E., Reeves, G. D., McComas, D. J., & Belian, R. D. (1997). Characteristic plasma properties during dispersionless substorm injections at geosynchronous orbit. *Journal of Geophysical Research: Space Physics*, *102*(A2), 2309–2324. https://doi.org/10.1029/96JA02870

Chan, A. A., Xia, M., & Chen, L. (1994). Anisotropic Alfvén-ballooning modes in Earth's magnetosphere. *Journal of Geophysical Research: Space Physics*, *99*(A9), 17351–17366. https://doi.org/10.1029/93JA03353

Chen, L., & Hasegawa, A. (1991). Kinetic theory of geomagnetic pulsations: 1. Internal excitations by energetic particles. *Journal of Geophysical Research: Space Physics*, *96*(A2), 1503–1512. https://doi.org/10.1029/90JA02346

Ferrière, K. (2001). Interchange, quasi-interchange, ballooning modes: What is their exact definition? *Eos, Transactions American Geophysical Union*, *82*(38), 420–420. https://doi.org/10.1029/01EO00260

Kalmoni, N. M. E., Rae, I. J., Murphy, K. R., Forsyth, C., Watt, C. E. J., & Owen, C. J. (2017). Statistical azimuthal structuring of the substorm onset arc: Implications for the onset mechanism. *Geophysical Research Letters*, *44*(5), 2078–2087. https://doi.org/10.1002/2016GL071826

Kalmoni, N. M. E., Rae, I. J., Watt, C. E. J., Murphy, K. R., Samara, M., Michell, R. G., … Forsyth, C. (2018). A diagnosis of the plasma waves responsible for the explosive energy release of substorm onset. *Nature Communications*, *9*(1), 4806. https://doi.org/10.1038/s41467-018-07086-0

Lui, A.T.Y., 2004. Potential plasma instabilities for substorm expansion onsets. *Space Science Reviews*, *113*(1-2), pp.127-206.

Lui, A.T.Y., 2018, March. Review on the Characteristics of the Current Sheet in the Earth's Magnetotail. In *Electric Currents in Geospace and Beyond* (Vol. 235, pp. 155-175). Hoboken, NJ, USA: John Wiley & Sons, Inc..

Mcpherron, R. L. (1970). Growth Phase of Magnetospheric Substorms. *J. Geophys. Res*.

Panov, E. V., Sergeev, V. A., Pritchett, P. L., Coroniti, F. V., Nakamura, R., Baumjohann, W., … McFadden, J. P. (2012). Observations of kinetic ballooning/interchange instability

signatures in the magnetotail. *Geophysical Research Letters*, *39*(8). https://doi.org/10.1029/2012GL051668

Sergeev, V. A., Angelopoulos, V., & Nakamura, R. (2012). Recent advances in understanding substorm dynamics. *Geophysical Research Letters*, *39*(5), n/a-n/a. https://doi.org/10.1029/2012GL050859

Viñas, A. F., & Madden, T. R. (1986). Shear flow-ballooning instability as a possible mechanism for hydromagnetic fluctuations. *Journal of Geophysical Research: Space Physics*, *91*(A2), 1519–1528. https://doi.org/10.1029/JA091iA02p01519

Wang, C.-P., Zaharia, S. G., Lyons, L. R., & Angelopoulos, V. (2013). Spatial distributions of ion pitch angle anisotropy in the near-Earth magnetosphere and tail plasma sheet. *Journal of Geophysical Research: Space Physics*, *118*(1), 244–255. https://doi.org/10.1029/2012JA018275

Zaharia, S., Jordanova, V. K., Thomsen, M. F., & Reeves, G. D. (2006). Self- consistent modeling of magnetic fields and plasmas in the inner magnetosphere: Application to a geomagnetic storm. *Journal of Geophysical Research: Space Physics*, *111*(A11). https://doi.org/10.1029/2006JA011619

Zhu, P., Bhattacharjee, A., Sangari, A., Wang, Z., & Bonofiglo, P. (2017). Three-dimensional geometry of magnetic reconnection induced by ballooning instability in a generalized Harris sheet. *Physics of Plasmas*, *24*(2), 024503. https://doi.org/10.1063/1.4976994

Chapter 5: Review of Particle Tracing

This chapter provides background for the particle tracing work conducted in the paper in the subsequent chapter. In addition to theoretical background, it describes the numerical approach taken to conduct this particle tracing.

5.1 Tracing Particles Through a Model Magnetic Field

The equations governing particles drifting around Earth, described previously, can be used to determine the location of a particle at any point on its drift orbit. To follow up on, and further test, the conclusions of the previous chapter, we trace particles in a model magnetotail magnetic field. We trace particles backward from the highly stretched state of the magnetotail immediately prior to substorm onset to previous, less stretched tail states to determine where relevant particles at substorm onset originate at earlier times during tail development. This allows us to compare the history of particles with different pitch angles and, thus, to infer whether the particle distribution becomes more parallel anisotropic and, therefore, closer to ballooning instability. The motion of a particle depends on whether the magnetic field in which it is located is stationary or varies with time. Both scenarios are described below, and the quasi-stationary approach that we adopt here is described subsequently.

5.1.1 Particles Moving in a Stationary Magnetic Field

In a stationary magnetic field, particles follow drift orbits that are entirely defined by conservation of the first and second adiabatic invariants. Further, in the stationary case, this conservation is satisfied through conservation of the magnetic field strength at a particle's mirror point, B_m, and the following field geometric integral (Roederer, 1967):

$$I = \int_{s_m}^{s_m} \sqrt{1 - \frac{B}{B_m}}\, ds. \tag{5.1}$$

Here s_m is the location of the mirror point along the field line. These two properties of a particle can be calculated simply by numerically integrating along a field line to find I until the point is reached at which the magnetic field strength matches B_m, which in turn is calculated based on the initial pitch angle of the particle as described in Chapter 2. B_m and I also allow one to determine the radial location of a particle at any azimuthal location, i.e. longitude, in a

magnetic field model without requiring knowledge of its location at every previous azimuthal location. Because of this, it can be applied to two-dimensional magnetic field models, where the location and pitch angle of a particle at, e.g., noon are sufficient to determine its location at midnight because they are sufficient to calculate B_m and I. As is evident through the lack of energy-dependent parameters in Equation 5.1, the path that a particle takes in a stationary magnetic field is not energy dependent.

5.1.2 Particles Moving in a Changing Magnetic Field

If the magnetic field is changing, the first and second adiabatic invariants are necessary, but no longer sufficient to compute the location of a particle, assuming that changes are occurring at a frequency that is lower than both the gyro- and bounce frequencies. There are two extreme cases in which this problem may be simplified (Roederer, 1970): If the field is changing much faster than the drift period of a particle, this particle will remain on a single field line and move with it; if the field is changing much more slowly than a drift period, conservation of the third adiabatic invariant may be added to conservation of the other two to determine a particle's location. Calculating the third invariant requires the integration of magnetic flux along a full drift orbit and, therefore, a three-dimensional magnetic field model.

5.1.3 The Quasi-Stationary Approach

The particle tracing done for this work is conducted in a manner that traces particles that move from a less stretched to a more stretched tail to determine how a particle distribution may change over the course of the tail stretching that occurs during the substorm growth phase. The approach taken is to treat the two tail states as stationary fields and to determine the location of particles in each. This quasi-stationary method applies to particles that originate in a quiet-time tail and are in other areas of the field as it stretches, returning to a more developed growth phase tail. This assumes that there is little change that happens in the tail as the particle is located within it, relying on the azimuthally narrow extent of the stretching tail resulting in most of the particle's drift orbit being in other field sections. Because the $\boldsymbol{E} \times \boldsymbol{B}$ drift caused by an induced electric field affects particles of all energies equally, this assumption should not affect the motion of the particles relative to one another. This relative motion of the particles is what determines the development of the particle distribution.

5.2 Numerical Particle Tracing

We are interested in the history of particles that are located in the transition region between dipolar and tail-like field lines at the point of substorm onset. This region was identified in the previous chapter as having the necessary combination of pressure gradients and plasma β to cause a movement toward ballooning instability. The remaining factor that must be examined is the pressure anisotropy. This can be inferred by tracing particles backward from the highly stretched tail state prior to substorm onset to the less stretched tail states that exist previously. An equatorial particle distribution is defined for this purpose on a field line in the transition region at substorm onset and backward tracing is used to determine where individual particles in this distribution originate in the quiet-time tail. The change in pitch angle of these particles and the pressure at their original location determines the change in the anisotropy of the overall distribution as described subsequently.

The quasi-stationary approach allows for the use of a two-dimensional magnetic field model as only B_m and I need to be calculated. The model chosen here is the one defined in Kabin et.al. (2011), which allows for a significant level of control over the stretching of the tail and the location of the transition region between dipolar and tail-like field lines, i.e. the features that are of interest for this analysis. The model consists of two parts: a set of field lines that transition from a dipolar to stretched state based on a Harris sheet model of the tail, and a divergence-free magnetic field that is mapped onto these field lines. It is described in general terms here, with a more detailed description contained in the subsequent chapter.

This divergence-free magnetic field is created by assuming that the magnetic field is axisymmetric in the vicinity of the tail and making use of the $\boldsymbol{B} = \nabla\psi \times \nabla\phi$ formulation of the magnetic field as in the previous chapter. To determine the magnetic field strength, the field line constant ψ is computed by tracing a field line down to Earth and using the dipole value of ψ at Earth. This makes use of the fact that the magnetic field at Earth is sufficiently strong, particularly in the polar region where the magnetic field lines intersect, that any deflection due to currents in the tail is negligible. Any higher order terms in Earth's magnetic field are neglected. The gradient of ψ can be found by stepping perpendicular to the field line and determining the value of ψ at the additional points that this requires by, again, tracing a field line from these location to Earth. As a result of this, each time the magnetic field strength must be computed,

several field lines need to be traced to Earth. An example of these field lines in the Kabin et al. (2011) model on the tail side is shown in Figure 14. The GSM coordinate system is used in Figure 14, where x points from Earth to the sun, z is a projection of Earth's North-South axis perpendicular to the x direction in the noon-midnight plane, and y is orthogonal to the other two, completing a right-hand coordinate system.

In addition to the field lines in Figure 14, a dipole is used on the dayside for particle tracing, acting as an illustration of where the particles are located when they are at noon. This does not have any impact on the calculated locations of the particles in the tail, as they are based purely on the B_m and I values that are independent of the particle's previous locations and therefore an additional, intermediate location has no effect. Further, it should be noted that the extremely stretched field lines beyond the transition region may be non-physical in their shape. However, the benefit of this model is in its ability to resolve the transition region, which the analysis here is restricted to.

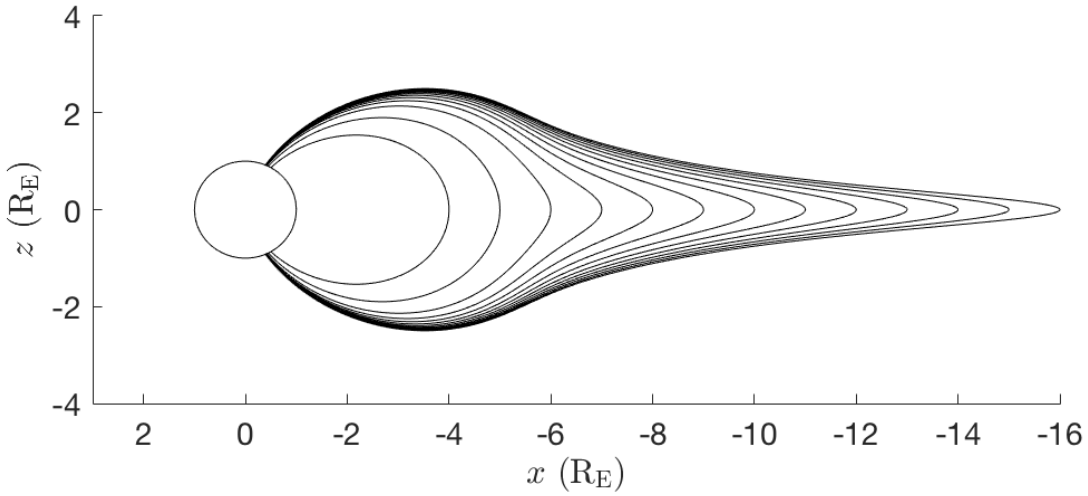

Figure 14: Field lines drawn according to the Kabin et al. (2011) *model in a highly stretched configuration.*

In the remaining sections of this chapter, the procedure for using this magnetic field model to conduct the particle tracing work is described in general terms, with an emphasis on the computational approach used. To determine the history of the particles located on a field line in the most stretched tail case, the tracing is applied to a distribution of particles located on one

field line that covers the whole range of possible equatorial pitch angles, from the loss cone angle to 90°. On Compute Canada servers, backward tracing particles with 0.1 radian increments on three final field lines is completed in 3.5 hours.

5.2.1 Code Structure

The structure of the code used to trace the particles is shown in Figure 15. This set of functions is used on the Compute Canada network, with the shell script included in Appendix B1 used to input individual particles into the code structure in parallel. This approach is taken rather than using the parallel functionality in MATLAB due to a lack of stability on the Compute Canada servers.

The variables in Figure 15 not defined previously are: x_i and α_i, the initial x position and pitch angle in the equatorial plane; x and z, vectors defining the coordinates of the current field line; ii, an index describing at which location along the field line the field strength is to be evaluated; t, the step size used to draw the field line; g, the integrand in Equation 5.1; x_{final}, z_{final}, B_{final}, vectors describing the coordinates and field strength along the field line that a particle was traced to; and α_f, the equatorial pitch angle of the particle on this field line. The individual functions are as follows:

Results Output

This function takes as its input the initial position and pitch angle of a particle and feeds it into the particle tracing function. It writes the results to a text file and accesses text files with previous results when tracing a particle's return from the dayside to night, as these are required to determine B_m and I. For running locally, the particle distribution can be contained in a for loop in this function.

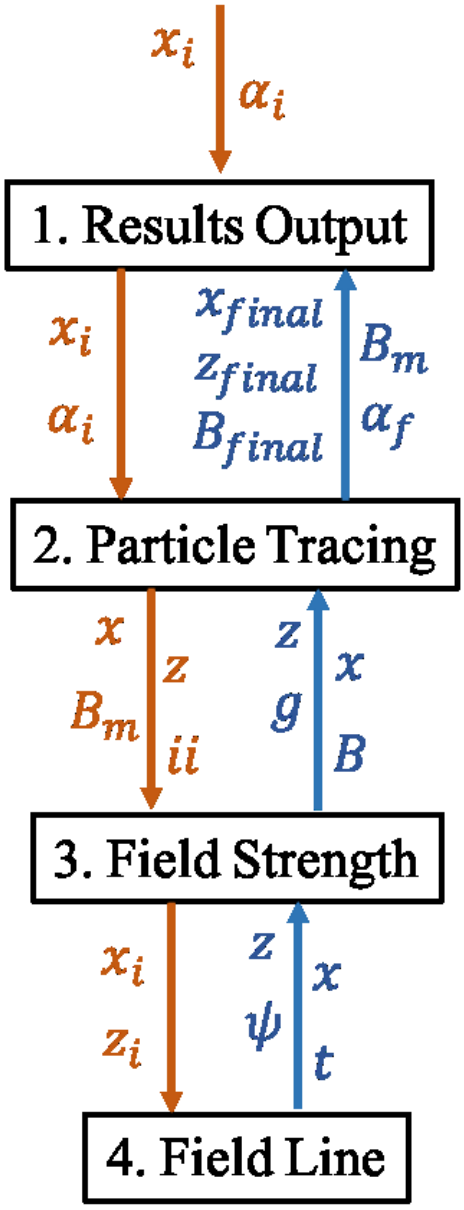

Figure 15: Schematic diagram of the code structure used to trace particles through the Kabin et al. (2011) *magnetic field model.*

Particle Tracing

The mirror magnetic field strength is calculated using Equation 2.9 from chapter two, solving for the magnetic field strength at the mirror point based on a defined pitch angle rather than using the field strength at Earth. The integral I is computed for the particle in its original location by moving along the field line and integrating using a while loop that continues until the mirror magnetic field strength is reached. The integration scheme used here is a fifth order

Runge-Kutta scheme with adaptive mesh refinement. This scheme is highly accurate while also allowing for an increase in step size when finer step sizes are not required. Because each point used during the integration requires the tracing of several field lines to Earth, this significantly reduces the amount of computational time required to complete a calculation. The formula for the fifth order method is (Press et al., 1992):

$$y_{n+1} = y_n + \frac{37}{378}k_1 + 0k_2 + \frac{250}{621}k_3 + \frac{125}{594}k_4 + 0k_5 + \frac{512}{1771}k_6 \tag{5.2}$$

Where the values of k_n are given by:

$$k_1 = hf(x_n, y_n) \tag{5.3}$$

$$k_2 = hf(x_n + \frac{1}{5}h, y_n + \frac{1}{5}k_1) \tag{5.4}$$

$$k_3 = hf(x_n + \frac{3}{10}h, y_n + \frac{3}{40}k_1 + \frac{9}{40}k_2) \tag{5.5}$$

$$k_4 = hf(x_n + \frac{3}{5}h, y_n + \frac{3}{10}k_1 - \frac{9}{10}k_2 + \frac{6}{5}k_3) \tag{5.6}$$

$$k_5 = hf(x_n + h, y_n - \frac{11}{54}k_1 + \frac{5}{2}k_2 - \frac{70}{27}k_3 + \frac{35}{27}k_4) \tag{5.7}$$

$$k_6 = hf(x_n + \frac{7}{8}h, y_n + \frac{1631}{55296}k_1 + \frac{175}{512}k_2 + \frac{575}{13824}k_3 + \frac{44275}{110592}k_4 + \frac{253}{4096}k_5) \tag{5.8}$$

In the equations above, h is a chosen step size. Using these same k factors that were used for a fifth order solution, a fourth order solution for the integration may be found using:

$$y_{n+1}^{*} = y_n + \frac{2825}{27648}k_1 + 0k_2 + \frac{18575}{48384}k_3 + \frac{13525}{55296}k_4 + \frac{277}{14336}k_5 + \frac{1}{4}k_6 \tag{5.9}$$

The results for the fourth order and fifth order schemes may then be compared to obtain an estimate for the error:

$$\Delta \equiv y_{n+1} - y_{n+1}^{*} \tag{5.10}$$

Based on a step size, h_1, and its associated error Δ_1, the step size for a desired error, Δ_0, can be found using:

$$h_0 = h_1 \left|\frac{\Delta_0}{\Delta_1}\right|^{0.25} \tag{5.11}$$

This may then be used to adjust the step size and repeat the integration for the same starting point y_n. The criterion used to determine Δ_0 here is designed to result in constant fractional errors except near zero crossings, and is given by:

$$\Delta_0 = \epsilon \left[y(i) + h\frac{dy}{dx}(i)\right] \tag{5.12}$$

Where $y(i)$ and $\frac{dy}{dx}(i)$ are the values of the integral, in this case I, and the integrand at the current location. The factor ϵ is set to 1×10^{-5}. The second term in Equation 5.12 ensures that a tighter restriction on the error is imposed when the step size is smaller. This is necessary because smaller step sizes require more steps, and therefore more intervals that each have their own errors that add to produce a large error for the integral overall. Because the value of I at the final point of the integration is required, minimizing the accumulation of error over the whole interval is crucial. The integration over an interval is repeated and the step size is adjusted until the error criterion has been met.

One modification is made to this method due to the discretized nature of the field lines, i.e., the domain over which the integration is performed. The steps taken here must be on the correct field line. To ensure this, the field line is drawn at the beginning and the integration is

performed at points along it. This means that the step size h must be a multiple of the step size used to draw the field line that allows for all of the fractional steps given in Equations 5.3-5.8 to be taken. For this purpose, the value of h derived from the equations above is rounded down to the nearest 40 increments along the field line, the lowest common multiple of the factors dividing h. This condition also imposes a minimum step size of 40 increments for the adaptive mesh refinement. A maximum step size is also imposed that ensures that no points go beyond the length of the field line.

To find the field line of the particle as traced into a specified tail state, a bisection algorithm is used to locate the field line where $I - I_{start} = 0$. This algorithm is somewhat slower than other alternatives, however its stability ensures that code that is run unsupervised on Compute Canada servers converges without issues. The convergence criterion used here is that the size of the window in the equatorial location of the particle is less than 0.005 Earth radii.

Field Strength

To calculate the field strength, $\nabla\psi$ is computed by stepping perpendicular to the field line at the point where the strength is required. A fourth order central scheme is used here. This requires the tracing of four additional field lines, but allows for a fairly high accuracy at a large step size, which is required to allow for significantly different values of ψ at the points used for the scheme. As illustrated in Figure 14, field lines that are far apart intersect Earth very near each other. Since ψ is calculated at Earth, this may cause a difference too small for the accurate calculation of $\nabla\psi$, with additional issues caused by the discretized nature of the field lines. A larger step size avoids this.

Field Line

The field line is drawn according to the Kabin et al. (2011) model. An analytically defined dipole is used on the dayside.

5.2.2 Step Sizes

Three different step sizes are required for drawing the field lines, calculating the field strength, and performing the integration. The step size used to draw the field lines is $t = 1 \times 10^{-5}$ Earth radii. This fine step size is required to ensure that the equatorial crossing point of the field lines is determined accurately, as any error prior to this point accumulates. It was found that

a relatively constant error in the magnetic field strength while moving along a field line can be produced by using a step size derived through the formula:

$$h_{strength} = 1 \times 10^4 t \left(\frac{r}{x_0}\right)^3 \tag{5.13}$$

Where r is the radial distance to the point in spherical coordinates, x_0 is the radial location of the equatorial crossing of the field line, and t is the step sized used to draw the field line. With $t = 1 \times 10^{-5}$ Earth radii, this results in a step size of 0.1 Earth radii at the equator and which decreases as one follows the field line toward Earth. The cubic power is chosen to match the cubic power of r in the definition of the dipole magnetic field strength. The step size for the integration scheme is defined by the adaptive mesh refinement algorithm, but an initial step size of 1000 increments along the field line is used.

Validation and verification of this code are included in Appendices B2 and B3. A further examination of Kabin et al. (2011) is included in Appendix B4, and a method for calculating the pressure in this model using variables defined relative to a dipole field is outlined in Appendix B5.

Chapter 6: Paper Examining the Effects of Drift Orbits on Particle Distributions

This chapter consists of a manuscript for a paper to be submitted to the journal Journal of Geophysical Research and is formatted accordingly.

Increasingly parallel pressure anisotropic ballooning: Substorm growth phase drift orbits resulting in a locally ballooning unstable geomagnetic tail

L. R. Oberhagemann[1] and I. R. Mann[1]

[1]Department of Physics, University of Alberta, Edmonton, Canada

Corresponding author: Luke Oberhagemann (oberhage@ualberta.ca)

Key Points:

- We examine the substorm onset paradigm of increasingly parallel pressure anisotropic ballooning by tracing particle distributions in a model magnetosphere
- During tail stretching, particle distributions evolve towards increasingly parallel pressure anisotropy, which can result in a ballooning unstable tail
- The results can explain why onset occurs when it does, and suggest a potential for forecasting, and a test for assessing the role of ballooning at onset

Abstract

The controversy surrounding magnetospheric substorms has historically been divided into two schools of thought: those who are proponents of the out-to-in view state that substorms begin with mid-tail reconnection, while those who support the in-to-out view claim that an instability near Earth causes substorm initiation. The work presented here examines the recent findings by Oberhagemann and Mann (2019) showing that ballooning instabilities may be triggered by a pressure distribution that moves from an initially mild perpendicular anisotropy toward more parallel anisotropy. We trace particles through a two-dimensional model magnetic field to examine the evolution of particle distributions as they periodically drift through a tail which changes from a less stretched, quiet time to a more stretched late growth phase configuration. Our findings show that these particle pitch angle distributions become more parallel anisotropic through a combination of drift shell splitting and preferential de-energization in the perpendicular direction. Drift shell splitting causes low pitch angle particles from a high pressure gradient region to mix with high pitch angle particles from a low pressure region in a radially localized area in the tail transition region from quasi-dipolar to more tail-like fields where high earthward pressure gradients exist. Preferential de-energization occurs due to a transfer of particles from less stretched to more stretched field lines that reverses the usual balance between Fermi and betatron acceleration and causes particle pitch angle distributions to become more field-aligned. We then examine these findings in the context of ground based and in-situ observations of the observational time sequence of events through the growth phase to onset, finding significant parallels between these observations and the predictions made by our mechanism. Notably, the most unstable region to increasingly parallel anisotropic ballooning lies in the radially narrow region associated with auroral beads, which accompany the onset of the expansion phase of many substorms.

1. Introduction

The controversial topic of what triggers the onset of magnetospheric substorms has been under discussion since the time sequence of events in the aurorae was first described by Akasofu (1964). It is generally agreed upon that the auroral signatures of substorms after expansion phase onset consist of brightening of the most equatorward auroral arc followed by its poleward motion

and breakup. Before onset, this arc moves toward the equator, corresponding to stretching of the magnetic tail during the growth phase (e.g., McPherron 1970).

A remaining point of controversy regarding substorms is what triggers expansion phase onset and whether this originates due to mid-tail reconnection or due to an instability that forms closer to Earth (e.g., Sergeev et al. 2012). These two viewpoints are generally referred to as the out-to-in and in-to-out paradigms respectively. Recent work on auroral beads, periodic brightness features which develop along the onset auroral arc (e.g., Kalmoni et al., 2017), has shown that they are a nearly ubiquitous feature of the most equatorward auroral arc immediately pre-onset. They are perhaps connected by the propagation of kinetic Alfvén waves, to ballooning instabilities in the near-Earth magnetotail (Kalmoni et al. 2017, 2018; Viñas and Madden 1986). In addition, Zhu et al. 2017 recently conducted numerical work using MHD modelling which indicates that ballooning instabilities may also subsequently cause magnetic reconnection in the mid-tail, suggesting a possible connection between the in-to-out and out-to-in schools of thought.

In this context, Oberhagemann and Mann (2019) examined the thresholds for the destabilisation of anisotropic pressure ballooning modes in the magnetotail. They showed that a change in the particle distribution toward a more field-aligned pressure lowers the plasma $\boldsymbol{\beta}$ at which ballooning instabilities are triggered. Crucially, a change from the usual, mildly perpendicular anisotropic quiet time pressure *toward* more parallel anisotropy without having to reach fully parallel anisotropy is destabilizing. This finding is then placed in the context of a substorm onset mechanism in which particles that gradient-curvature drift around Earth experience a change toward more parallel pressure anisotropies, and which according to Oberhagemann and Mann (2019) would naturally occur in the late growth phase as a result of tail stretching.

Here, we examine this hypothesis by tracing particles through a two-dimensional model magnetotail magnetic field. We show that a combination of drift shell splitting and preferential de-energization in the perpendicular direction in the stretched late growth phase magnetic field further drives particle distributions towards more parallel pressure anisotropy. Consistent with the theoretical proposition by Oberhagemann and Mann (2019), such distributions can be ballooning unstable and could be responsible for triggering substorm onset. Further, we examine

our results in the context of several ground-based and in-situ observations of the substorm sequence and illustrate multiple applications for interpreting these observations through our proposed model.

2. Methods

We examine the evolution particle distributions undergo as they drift around Earth during an interval of tail magnetic field stretching in the substorm growth phase. As identified by Oberhagemann and Mann (2019), the primary region of interest in relation to ballooning instability is the transition region between dipolar and tail-like magnetic field lines in the highly stretched tail in the late growth phase prior to substorm onset. We use a quasi-stationary approach in which we backward trace particles from the highly stretched tail that exists during the late growth phase to earlier less stretched states. Each state of the tail is treated as a stationary magnetic field and therefore effects due to induced electric fields are neglected. This approach is valid for particles that drift rapidly enough that little change in magnetic topology occurs in the tail while they are located within it, but which experience an evolution in parallel and perpendicular energy as a result of the conservation of the first and second adiabatic invariants in the slowly evolving growth phase during repeated entries into the stretching magnetotail.

2.1 Magnetotail Magnetic Field Model

For the purpose of this work, we use the self-consistent empirical magnetotail equilibrium magnetic field model from Kabin et al. (2011). In their work, this model was used to examine magnetic field changes that occur during the dipolarisation that follows a substorm as the tail changes from a more stretched to a more dipolar configuration. Our application uses the same model to examine growth phase evolution of the tail that starts from a less stretched state and becomes more stretched with time. A significant benefit of this model is its ability to capture the narrow transition region from dipolar to tail-like fields in the stretched tail. The model

consists of a vector field onto which a divergence-free magnetic field is mapped. The vector field is based on a Harris current sheet and is defined in Cartesian GSM coordinates by:

$$\boldsymbol{b} = \boldsymbol{B}_{dip} + \boldsymbol{e}_x f(x) B_0 \tanh\left(\frac{z}{L_z}\right). \quad (1)$$

Here $\boldsymbol{B}_{dip}$ is the dipole magnetic field, L_z is the thickness of the magnetotail, and B_0 is a measure of how strongly the current sheet deflects the magnetic field. The function $f(x)$ controls the radial distance and width of the transition region between dipolar and tail-like fields, given by:

$$f(x) = \frac{1}{2}\left(1 - \tanh\left(\frac{x + R_t}{\delta}\right)\right). \quad (2)$$

Here R_t is the geocentric equatorial radial distance to the transition region and δ is its width. An example magnetic field described by this model is shown in Figure 16. For illustration purposes, we also show a dipole magnetic field on the dayside on the left-hand side of the plot at positive x. Note that this model assumption does not capture the effects of dayside compression. However, as we conserve the first and second adiabatic invariants the dayside configuration through which the particles travel does not impact the evolution of the particle distributions between successive entries into the stretching tail magnetic field.

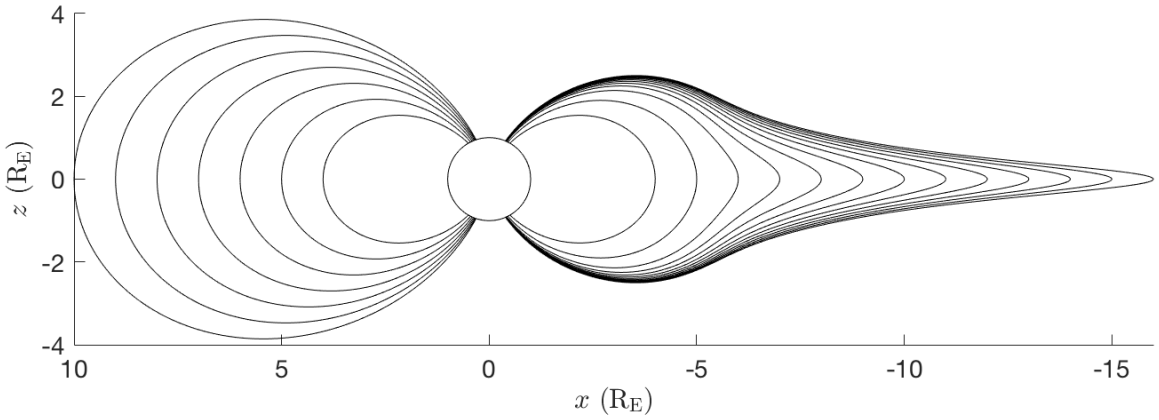

Figure 16: Example magnetic field lines defined by the Kabin et al. (2011) *model, with the parameters* $B_0 = -240$ nT, $L_z = 0.6\ R_E$, $R_t = 6\ R_E$, and $\delta = 1\ R_E$.

To map a divergence-free magnetic field onto this vector field, the field lines are used as the basis of a magnetic field, $\boldsymbol{B}$, of the form $\boldsymbol{B} = \nabla\psi \times \nabla\phi$ where ϕ is the azimuthal angle and ψ is the flux function. Assuming rotational symmetry for a small range of azimuths near the noon-midnight plane, the magnetic field strength is found by tracing individual field lines to Earth and determining the value of ψ at Earth using the dipole value. As the value of x is constant along a field line, it allows $\nabla\psi$ to be calculated at any location in the tail.

As shown in Equation 1 and Equation 2, the shape of the tail is controlled by the parameters B_0, L_z, R_t, and δ. We trace particles backward from a strongly stretched tail configuration, characterised with appropriate values of these parameters in the Kabin et al. (2011) model assumed to represent the late growth phase. In a strongly stretched field configuration in Kabin et al. (2011), the parameters are -240 nT, 0.25 R_E, 6 R_E, and 1 R_E respectively. We adopt the same set of parameters with a modification of the L_z value, which we change from 0.25 R_E to 0.6 R_E due to the former resulting in a region of non-physical negative plasma pressure. This change is justified observationally because L_z can have values between $0.1R_E$ and 2 R_E (Wanliss et al., 2000).

To examine the effects of tail stretching, we trace particles backwards from the $R_t = 6\ R_E$ tail to three less stretched configurations representative of earlier times in the growth phase, with R_t values of 6.25, 6.5, and 6.75. In doing so, we obey the $B_0 \propto R_t^{-3}$ scaling described in Kabin et al. (2011), such that B_0 takes on values of -210 nT, -190 nT and -170 nT, respectively, for these three configurations. L_z and δ are assumed to remain constant. These tail states are chosen arbitrarily to examine the effects of different degrees of tail stretching during a single drift orbit. We expect that a different choice of parameters, which still represent earlier stretched tail configurations, would result in similar results regarding the effects of tail stretching, including, for example, changing L_z and δ. The equatorial magnetic field strengths

for the three chosen configurations are shown in Figure 17. Note the decrease in magnetic field strength in the transition region (between approximately $x = -5$ and $x = -8$ in this model).

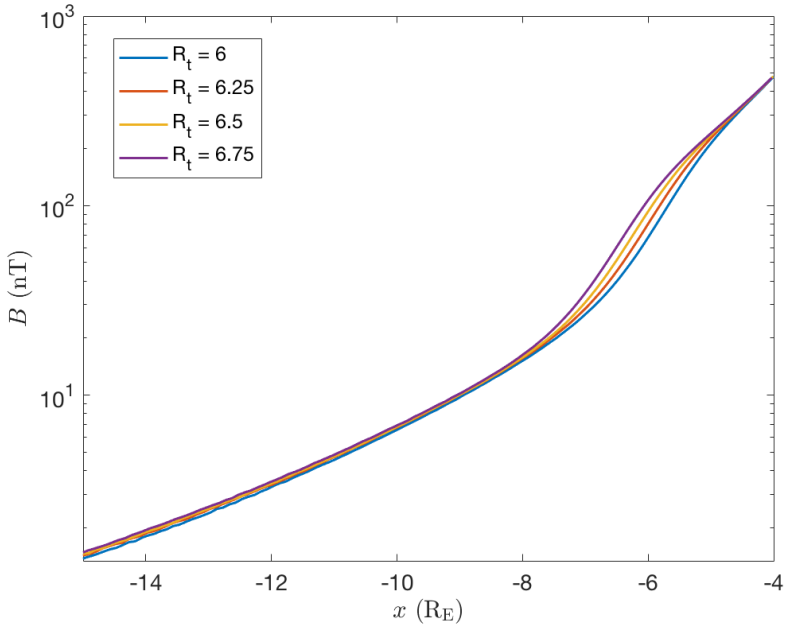

Figure 17: Equatorial magnetic field strengths in the Kabin et al. (2011) *model for the chosen configurations. In the* $R_t = 6\ R_E, 6.25 R_E, 6.5 R_E,$ *and* $6.75 R_E$ *configurations* $B_0 = -240, -210, -190,$ *and* -170 nT *respectively*. In all configurations $L_z = 0.6\ R_E$ and $\delta = 1\ R_E$.

We also calculate the pressure profiles implied by this model for these three configurations using an MHD force balance, the results of which are shown in Figure 18. Conforming to expectations, the pressure increases as the tail stretches and, consequently, so do the pressure gradients. The magnitudes of the calculated pressures and the shapes of the pressure

profiles show reasonable agreement with measured values in the tail (e.g., Lui and Hamilton, 1992).

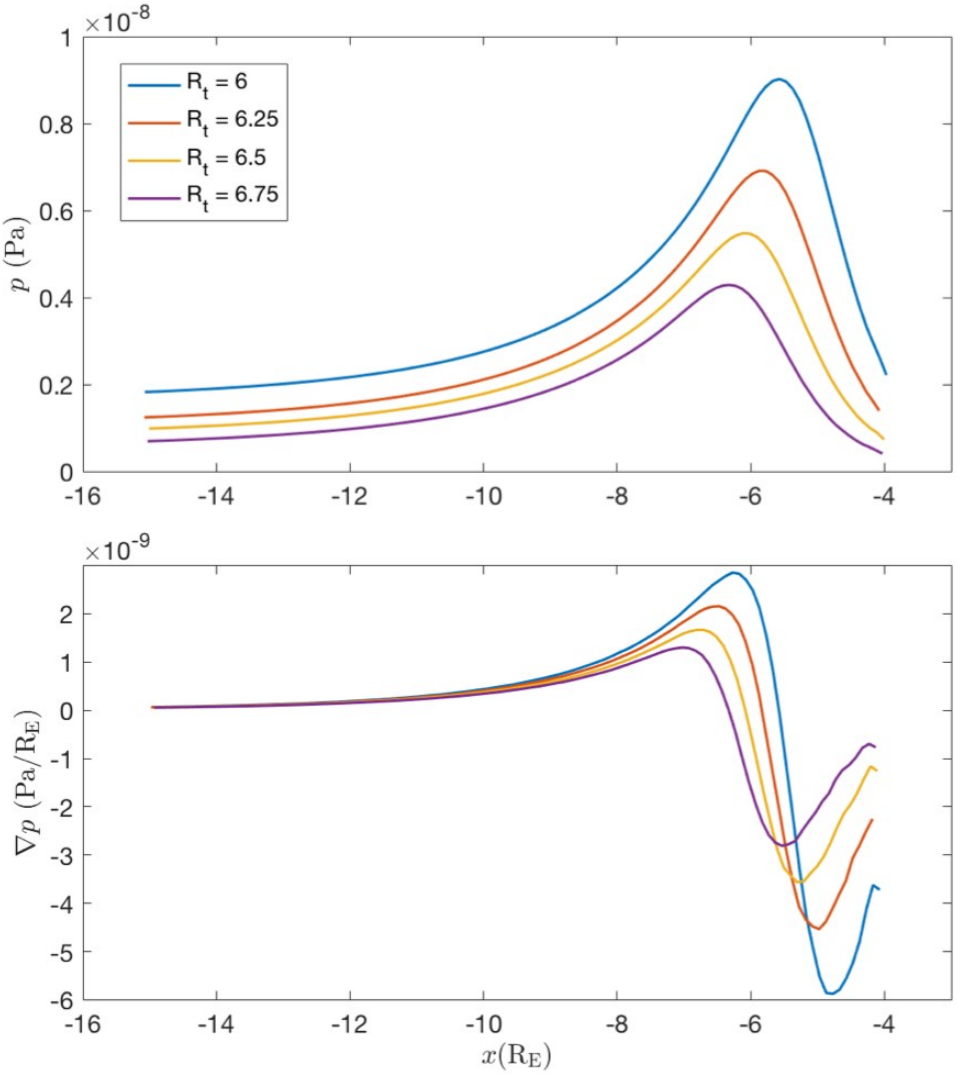

Figure 18: Pressure profiles as a function of x *(top), and pressure gradients* ∇p *(bottom) calculated using the Kabin et al.* (2011) *model. Note smaller volume of* R_E *represents a more stretched tail configuration. As previously, in the* $R_t = 6\ R_E, 6.25 R_E,\ 6.5 R_E,$ *and* $6.75 R_E$ *configurations* $B_0 = -240, -210, -190,$ *and* -170 nT *respectively, and for all cases* $L_z = 0.6\ R_E$ and $\delta = 1\ R_E$.

The pressure gradient profile for the most stretched case in the bottom panel of Figure 18 shows a local peak at $x \approx -6.2\ \mathrm{R_E}$. Therefore, this region may be the most conducive to the triggering of ballooning instabilities, since these modes are destabilised by high pressure

gradients (see e.g., Oberhagemann and Mann, 2019). To examine the evolution of particle parallel and perpendicular velocity for drift orbits that pass near to this peak in the tail, we trace particles backwards from three tail field lines with equatorial crossing points at $x = -6$, -6.2, and -6.4 R_E.

Figure 19 shows the equatorial plasma β values calculated using the Kabin et al. (2011) model for the same parameters as in Figure 18. This illustrates the manner in which β increases significantly as the tail stretches due to the combination of increased pressure and reduced magnetic field strength. This lends further significance to the effect of tail stretching in creating conditions susceptible to enabling ballooning instabilities to form, since higher β values are a key component in reaching instability (e.g., Chan et al. 1994).

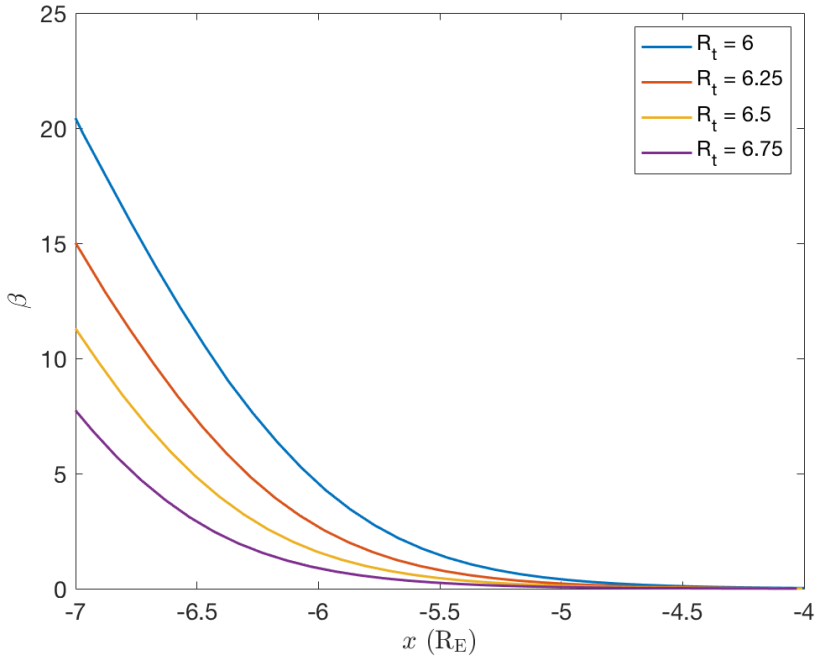

Figure 19: Equatorial plasma β values in the near-Earth region in the Kabin et al. (2011) *magnetic field model. As previously, in the $R_t = 6\ R_E$, $6.25 R_E$, $6.5 R_E$, and $6.75 R_E$ configurations $B_0 = -240, -210, -190, and -170$ nT respectively, and for all cases $L_z = 0.6\ R_E$ and $\delta = 1\ R_E$.*

2.2 Particle Tracing

We trace the evolution of particle parallel and perpendicular velocities with periodic entries into a stretching tail assuming conservation of the first and second adiabatic invariants. In a stationary magnetic field, these quantities are conserved through conservation of the magnetic field strength at a particle's mirror point, $\boldsymbol{B_m}$, and the field geometric integral $\boldsymbol{I}$ (Roederer et al., 1967):

$$I = \int \sqrt{1 - \frac{B}{B_m}}\, ds. \tag{3}$$

In our approach, $\boldsymbol{I}$ is calculated by integrating along the field using a fifth order Runge-Kutta scheme with adaptive mesh refinement to account for the possibility of rapid changes in the shape of the integrand as one moves along a field line from the equator to the mirror point, and the possibility for a large variation of the size of the domain when tracing particles with different pitch angles. In the Kabin et al. (2011) model, the magnetic field strength at the surface of Earth remains the dipole value, as currents in the tail are assumed to be too distant to affect the magnetic field at Earth's surface. Therefore, the conservation of $\boldsymbol{B_m}$ in this model results in mirror points near the top of the ionosphere remaining at this location for all states of the tail, preventing particles from entering the loss cone. Because $\boldsymbol{B_m}$ is dependent on pitch angle, we trace a set of 16 particles with pitch angles evenly spaced between $\boldsymbol{\pi/2}$ and the loss cone angle. This allows us to capture the effect on particles with mirror points that range between the equator and the ionosphere.

3. Results of Particle Tracing

Figure 20 shows the results of tracing particles through a quasi-static evolution of the Kabin model magnetic field. The leftmost column indicates the starting and ending configurations of the field, with each subsequent row describing a greater degree of stretching during a particle's drift orbit from one tail entry to the next. The set of figures in column A (left) show the change in pitch angle of the individual particles, $\boldsymbol{\Delta_\alpha = \alpha_f - \alpha_i}$, as a function their pitch angle in the final, more stretched field configuration, $\boldsymbol{\alpha_f}$. This is shown for each of the three final field lines with equatorial crossing points, $\boldsymbol{x_f}$, of $\boldsymbol{-6}$, $\boldsymbol{-6.2}$, and $\boldsymbol{-6.4}$. The set of

figures in column B shows the equatorial location of particles in the initial, less stretched configuration, $\boldsymbol{x_i}$, as a function of their final pitch angle. Therefore, it illustrates the origin of particles over the entire pitch angle distribution on the final field line.

Figure 20 column A (left) shows that pitch angles over the vast majority of the distribution decrease as they move from a less stretched to a more stretched tail. This effect becomes stronger for a larger difference in the amount of stretching between the initial and final tail configurations and most strongly affects particles with final pitch angles between $\boldsymbol{\pi/8}$ and $\boldsymbol{\pi/4}$. Particles with $\boldsymbol{\alpha_f = \pi/2}$ conserve their pitch angle as a consequence of conserving $\boldsymbol{I}$ and therefore show no change in pitch angles as expected. Low pitch angle particles also show little change in pitch angle. This may be due to the effects of tail stretching being strongest at the equator and therefore affecting a smaller proportion of these particles' bounce paths. The relatively larger error in the particles with final pitch angles near $\boldsymbol{\pi/2}$ is due to the arcsine function used to calculate the initial pitch angles amplifying small errors in the magnetic field strength for these pitch angles to a greater extent than an equivalent error for lower pitch angles. This is because the function becomes very steep when $\boldsymbol{B}$ is near $\boldsymbol{B_m}$, resulting in a small error in $\boldsymbol{B}$ corresponding to a larger difference in the calculated pitch angle.

The decrease in pitch angle in Figure 20 would result in pressure anisotropies that are more parallel anisotropic in the more stretched configuration than they were initially, this effect being most significant at moderate pitch angles. As suggested by Oberhagemann and Mann (2019), this growth phase change towards a more parallel anisotropic distribution may be attributed to the well-known effect that causes slightly perpendicular pressure anisotropies in a quiet-time magnetotail (e.g. Wang et al., 2013). Namely, in the quiet-time magnetotail, as particles $\boldsymbol{E \times B}$ drift Earthward from the distant tail they are energized due to conservation of the first and second adiabatic invariants, with energization in the perpendicular direction as a result of betatron acceleration being stronger than Fermi acceleration as particles move Earthward in typical quiet time tail magnetic field topologies (e.g., Zaharia et al., 2006). In our case, however, rather than $\boldsymbol{E \times B}$ drifting into regions with shorter field lines and higher magnetic field strengths, during growth phase tail stretching particles gradient-curvature drift into a more stretched tail with longer field lines and lower magnetic field strengths. Thus, the same process may act in reverse, causing a slight de-energization that acts more strongly in the perpendicular

direction and therefore moves the particle distribution toward a pressure that is more strongly field aligned than previously. While this de-energization results in a reduction of plasma pressure, there is a concurrent reduction in the magnetic field strength in the magnetotail during growth phase stretching. Therefore, the reduction in energy may not have a strong effect on the plasma $\boldsymbol{\beta}$ which, as opposed to the pressure, is a parameter that controls the ballooning stability of the tail (Oberhagemann and Mann, 2019).

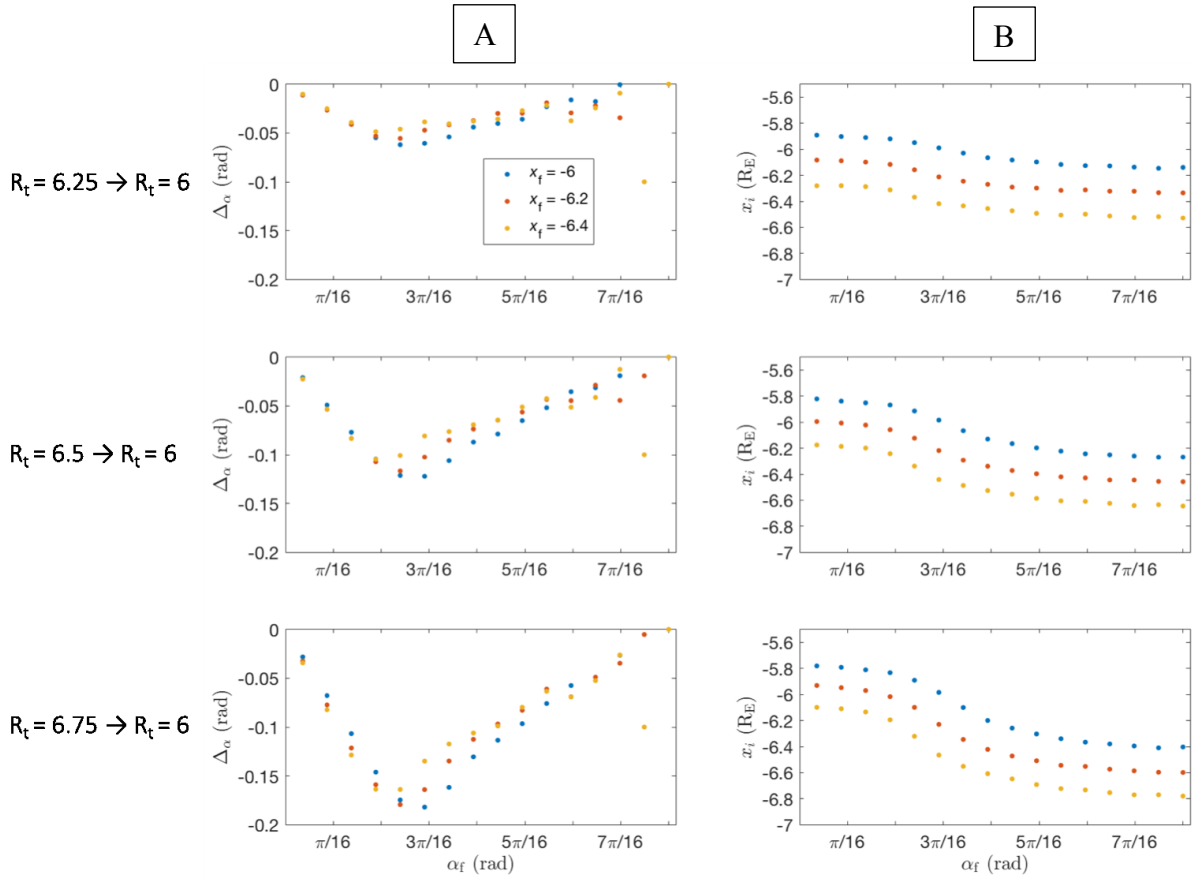

Figure 20: Results of tracing particles backward from the more stretched $R_t = 6$ tail configuration to the three less stretched configurations. As previously, in the $R_t = 6\ R_E$, $6.25 R_E$, $6.5 R_E$, and $6.75 R_E$ configurations $B_0 = -240, -210, -190,$ and -170 nT respectively, and for all cases $L_z = 0.6\ R_E$ and $\delta = 1\ R_E$.

Figure 20 column B (right) shows a pitch angle dependence for the drift orbits that is characteristic of drift shell splitting (e.g., Sibeck et al., 1987) during tail stretching. Due to a decrease in the magnetic field strength, particularly near the equator, particles with $\boldsymbol{\pi/2}$ pitch

angles move radially inward to maintain a constant $\boldsymbol{B_m}$. Meanwhile, particles with low pitch angles move radially outward to conserve $\boldsymbol{I}$. For each final field line and degree of stretching, there is an intermediate pitch angle, with an $\boldsymbol{\alpha_f}$ near $\boldsymbol{3\pi/16}$, where $\boldsymbol{x_i = x_f}$. As is expected, a greater difference in initial and final tail stretching configurations results in greater drift shell splitting and, thus, a more spread out region from which the particles originate. This effect may cause particle mixing where final pitch angle distributions are sourced from particles from different radial distances and which results in more parallel pressure anisotropies as shown below in Figure 21.

Figure 21 shows the case in which $x_f = -6.4$ particles are traced backward from the more stretched final $R_t = 6$ configuration to the earlier less stretched $R_t = 6.75$ configuration. The black dots indicate these particles' mirror points and the equatorial pressure gradients are shown with white contour lines plotted at equidistant constant pressures in the equatorial (colour) profile of tail plasma pressure. The starting location of these particles, in Figure 21a (see especially the inset), illustrates the manner in which drift shell splitting causes mixing of different P.A. particles which originate on different field lines but which meet on the same final field line in Figure 21c. Because we have chosen particles which are in the high pressure gradient region of the magnetic field associated with potentially locally unstable ballooning, the lower pitch angle particles that originate nearer to Earth start in a higher pressure region than the higher pitch angle particles that move inward during the course of their drift orbit. This may lead to an increasingly parallel pressure anisotropy by mixing a greater number particles with more field aligned pitch angles with fewer particles with less field aligned pitch angles. As opposed to the pitch angle reduction, which occurs due to a general stretching of the tail, this effect is more radially localized. If the final field line is located too close to Earth, the low pitch angle particles may originate on the opposite side of the pressure peak from the high pitch angle particles, thus resulting in similar pressures for both. Conversely, if the final field line is located at too large a radial distance, low pressure gradients may result in lower differences in pressure. In between is the radially localised potentially optimally unstable region for ballooning triggered by increasingly parallel pressure anisotropy.

The dayside particle locations, shown in Figure 21b, illustrate an additional effect that may result in increasingly parallel pressure anisotropies and butterfly pitch angle distributions in the tail. Particles with high pitch angles move radially outward on the dayside, which may subject them to preferential magnetopause shadowing (e.g. Fritz et al., 2003). As these particles collide with the magnetopause and are removed from the distribution, this would result in particles that contribute mainly to the pressure in the perpendicular direction to be removed from the distribution, thereby adding to the generation of parallel anisotropy.

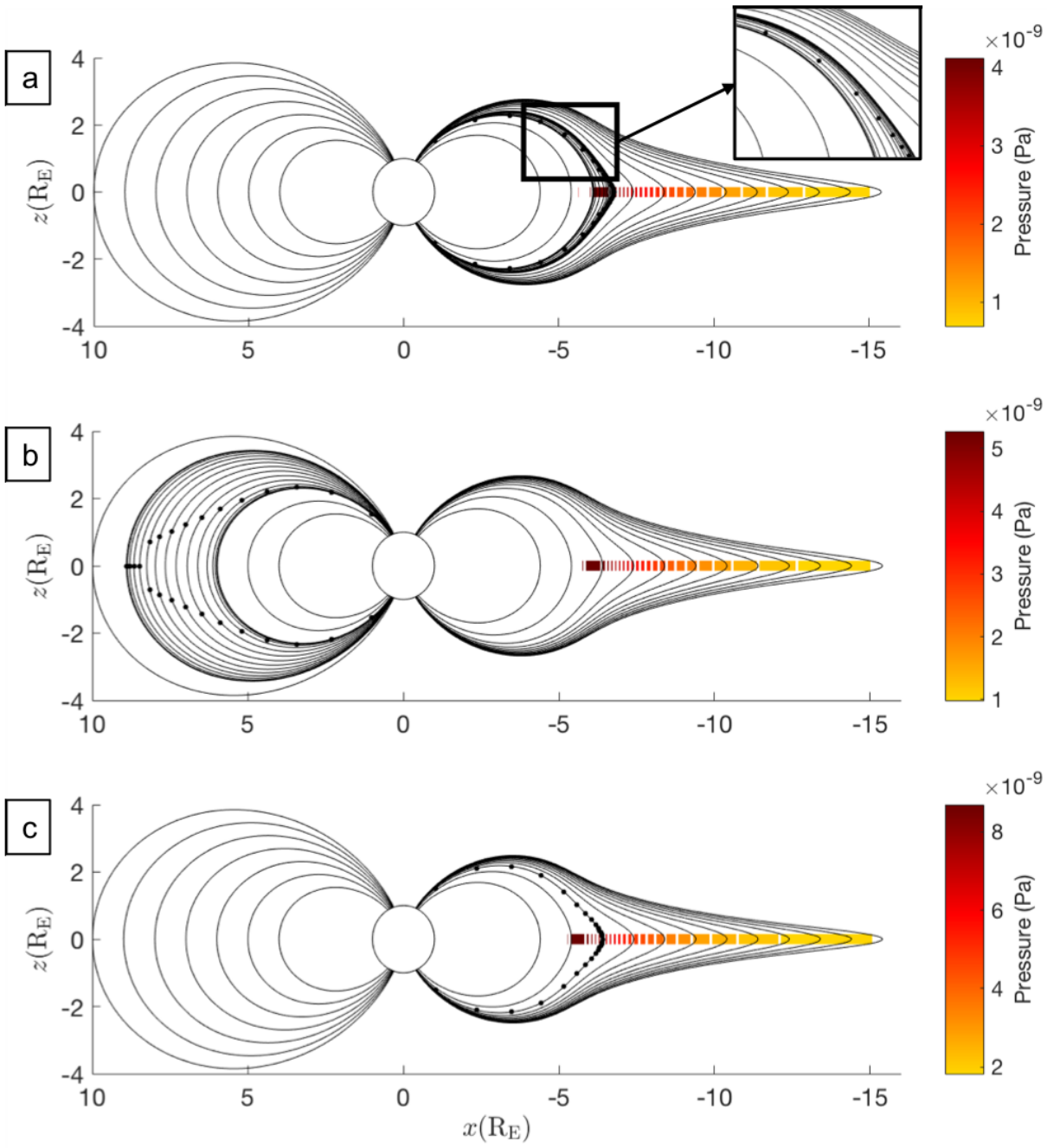

Figure 21: Particle motion during gradient-curvature drift from (a) the initial less stretched $R_t = 6.75$ configuration to (b) the dayside, shown here using an assumed dipole field on the dayside and the $R_t = 6.5$ tail configuration for illustration purposes, and finally (c) the field line that intersects the equator at $x_f = -6.4\ R_E$ in the final and more stretched $R_t = 6$ magnetotail magnetic field configuration, with the additional parameters for these states as described previously.

4. Discussion and Connections to Substorm Observations

The two effects shown here, reduction in pitch angle and radial particle mixing due to drift shell splitting, point to a significant role for the growth phase in creating the conditions identified by Oberhagemann and Mann (2019) for the formation of ballooning instabilities as a result of the generation of increasingly parallel pressure anisotropy in hot ion species in the tail. Because the effects are stronger when the difference between the initial and final tail configurations is larger, tail stretching that occurs on the timescale of the growth phase which is comparable to a particle's drift period is crucial. In addition, identifying the radially localized region in which the radial mixing effect may occur with the region where ballooning instabilities may preferentially form is also natural, since both require the presence of a high earthward pressure gradient. Therefore, our results are consistent with the hypothesis of Oberhagemann and Mann (2019) that localised increases in parallel pressure anisotropy can trigger the destabilisation of ballooning. This can hence also offer a natural explanation for the manner in which the magnetotail may transition from its quiet state to a substorm onset triggered by ballooning. All that is required is a stretching growth phase tail field geometry with a localised transition region from a dipole-like to a tail-like field.

Therefore, the effects we identify and the association with ballooning, although calculated quantitatively for a specific magnetic field model are likely to be generally applicable to tail-like fields. This is because drift shell splitting and preferential energization in the

perpendicular direction are well-known and well-established concepts that we show here to have effects in the specific circumstances of growth phase tail stretching preceding substorm onset.

The substorm onset mechanism advanced here and supported by our results, therefore, consists of the following elements:

1. An initially stable, quiet time, weakly stretched pre-growth phase tail contains ion distributions characterised by perpendicular anisotropic pressures;

2. Particles gradient-curvature drifting into and out of the stretching growth phase tail configuration evolve increasingly parallel pressure anisotropy, particularly in a localised part of the tail in the transition region;

3. A high pitch angle subset of these particles may be lost to the magnetopause on the dayside due to magnetopause shadowing, perhaps further increasing parallel anisotropy;

4. The particles return, at the end of their drift orbit, to a tail that is significantly more stretched than when they originated. Through a combination of pitch angle decrease and drift shell splitting, the returning particles cause the pressure anisotropy in the high earthward pressure gradient region of the tail to become increasingly parallel;

5. This increasingly parallel anisotropy, in combination with increased plasma $\boldsymbol{\beta}$ values and higher pressure gradients in the more stretched tail, causes ballooning instabilities to form in the manner suggested by Oberhagemann and Mann (2019);

6. These ballooning instabilities may also trigger kinetic Alfvén waves as described by Viñas and Madden (1986), the signatures of which might be visible in ground-based optical observations of auroral beads. Auroral beads are associated with the localized region of ballooning in the equatorial magnetotail and are now known to be observed during the vast majority of substorm onsets (Kalmoni et al., 2017, Kalmoni et al., 2018);

7. As discussed for example by Zhu et al. (2017), such near-Earth ballooning could also trigger near-Earth magnetic reconnection further tailward, leading to the expansion

phase dipolarisation in the time of the poleward and westward motion of the expansion phase aurora, and the release of the westward travelling surge.

4.1 In-situ Observations of Increasingly Parallel Anisotropies

The increasingly parallel pressure anisotropies suggested by our model have been observed rather extensively in prior studies using in-situ satellite particle measurements. As noted by Oberhagemann and Mann (2019), Baker et al. (1978) observed more field-aligned pitch angle distributions in electrons near the geomagnetic equator at geosynchronous orbit immediately prior to substorm onset. This agrees with the work presented here because, while electrons do not carry as much pressure as the ions, and therefore will have a lesser overall effect on the plasma pressure distribution, they also drift around Earth during the growth phase and therefore are expected to experience similar effects at radial distances where their first and second adiabatic invariants are conserved, and the effects of pitch angle scattering and isotropisation from field line curvature are weak.

Birn et al. (1997) also observed the evolution of particle distributions toward parallel temperature, and consequently pressure, anisotropy around substorm onset in electron distributions in geostationary orbit and, to a slightly lesser extent, in ion distributions. The lesser movement toward parallel anisotropies in the ions may be partly due to measurements being limited to ion energies less than 40 keV, therefore representing a distribution that disproportionately consists of lower energy ions. Such lower energy ions may have drift periods that are too large to experience significant evolution of their distributions on the timescale of field rapid stretching during the substorm growth phase. Therefore, there may be higher energy ions that experience the effects we describe here but these were not studied by Birn et al. (1997).

The anisotropy observed by Birn et al. (1997) moves towards a more parallel distribution with only minor fluctuations, for approximately one hour prior to substorm onset for both the electrons and ions. This agrees very well with the concept advanced here that the evolution towards more parallel distributions is explained by evolution of particle distributions during growth phase tail stretching. The point at which the most parallel anisotropy was observed by Birn et al. (1997) at geosynchronous orbit was occasionally prior to substorm onset but, on average, it is generally observed a few minutes after the point of substorm onset. This may be

due to the measurements at geosynchronous orbit not being coincident with the location in the tail at which the ballooning instabilities occurred. This may especially be true if, as indicated by our results, the instability is only triggered in a radially narrow region.

Another connection between our results and the measurements reported by Birn et al. (1997) is that, as our results predict, the movement toward parallel anisotropies shows remarkable alignment with increases in the magnetic field elevation angle, which corresponds to a greater amount of tail stretching. Birn et al. (1997) show magnetic field elevation angles for one set of events on February 23, 1990 in which three substorms occur. The point of onset in each case corresponds to a very strong movement toward parallel anisotropy in the electrons that results in the parallel electron temperature becoming almost twice as large as the perpendicular electron temperature in one case. Each of these onsets is preceded by a period of tail stretching that occurs at the same time as the gradual movement toward parallel anisotropy. At the point of the first onset, the tail stretching abruptly ends and the magnetic elevation angle decreases sharply following onset, possibly corresponding to dipolarization. As the magnetic field elevation angle reaches its minimum, the temperature anisotropy also reaches its most perpendicular range. This is then followed by another period of tail stretching that corresponds to another period of the development of increasingly parallel anisotropy with the most stretched tail configuration and the most parallel anisotropy occurring at the point of the second onset. A similar sequence of events happens prior to the third onset. This close link between tail stretching and increasingly parallel temperature anisotropies strongly supports our findings and the sequence of events that we propose occurs prior to the triggering of ballooning associated with substorm onset.

Further, as a point of comparison, the drift period of 40 keV ions, the largest energy for which the anisotropy is measured by Birn et al. (1997), in a dipole field is on the order of 2-4 hours (Parks, 1991). This agrees well with the timescale for the substorm growth phase, as well as the fact that when substorms occur in succession, the most frequent time between substorms is 2.75 hours (Borovsky et al., 1993). The ions have to have sufficiently short drift periods that they can return to the tail and then cause the onset of increasingly parallel pressure anisotropic ballooning. Further, this establishes a requirement for the rate of tail stretching prior to substorm onset. Stretching that occurs too slowly does not cause significant changes in the tail

configuration during the timescale of a particle's drift orbit, and therefore results in only mild changes in anisotropy. This appears to represent a testable hypothesis for our proposed onset mechanism and may also offer an explanation for why in some events tail stretching is not followed by substorm onset.

Indeed, there continues to be interest in understanding the difference in tail conditions which result in periods of steady magnetospheric convection (SMC) rather than a sequence of substorms under strong driving. Our substorm onset model suggests that the difference is explained by the extent of temporal changes in tail magnetic field topology. If during SMC events the rate at which new magnetic flux is added on the dayside is balanced by flux closure on the nightside, then the extent of tail stretching would remain approximately constant. In our onset model, such quasi-static tail field topologies would not become unstable to ballooning, thereby preventing ballooning from triggering substorm onset.

4.2 Proton Aurorae: A Case Study for the Application of the Increasingly Parallel Anisotropic Ballooning Mechanism

In addition to showing good agreement with in-situ measurements of increasingly parallel anisotropy, the mechanism of substorm onset through increasingly parallel anisotropic ballooning advanced by Oberhagemann and Mann (2019 and discussed further here offers a new lens through which to view established phenomena that occur prior to substorm onset. One such phenomenon is the presence of proton aurorae around the onset arc. This occurrence has been described by Samson et al. (1992) and consists of proton aurorae that surround the onset arc prior to the beginning of an auroral substorm.

Samson et al. (1992) observed a monotonically increasing energy of proton precipitation with decreasing latitude in the proton auroral observations. This suggests that higher energy precipitating ions originate on field lines with equatorial crossing points nearer to Earth. In combination with the fact that the event studied by Samson et al. (1992) occurred when the tail was quite stretched prior to substorm onset, this has led to the conclusion that these proton aurorae occur due to stochastic ion scattering due to field line curvature (e.g. Gilson et al., 2012). Stochastic scattering of a particle's pitch angle occurs when its gyroradius becomes comparable to the field line radius of curvature. Strong scattering of particles into the loss cone is observed

when the ratio of the field line radius of curvature to the gyroradius, $R_c/\rho \leq 8$ (e.g., Gilson et. al, 2012). Because field lines closer to Earth are generally less curved, as shown in e.g., our Figure 16, and higher energy particles have larger gyroradii, this explains the cause of the monotonically increasing energy in the proton auroral observations reported by Samson et al. (1992).

An issue that has been identified with this explanation for the proton aurora at onset is that a significant amount of stretching does not occur during the time in which an increase in the proton aurora are observed (Liang et al., 2018). If one considers only the state of particles that remain in the tail during this time, this low degree of stretching is not sufficient to produce a large enough increase in field line curvature to cause a significant increase in scattering into the loss cone. However, the mechanism we advance here whereby substorm onset is triggered by ballooning destabilized by increasingly parallel pressure anisotropy offers a solution to this issue through the drift transport of particles from less stretched to increasingly more stretched tail configurations by gradient-curvature drift during the growth phase. Due to this transport, a high-energy portion of the drifting particles, which was able to exist on the original, less curved field lines without experiencing stochastic scattering, may subsequently encounter more stretched field lines where the strong scattering criterion is met upon completion of its drift orbit. This is because the relevant difference in field line curvature is not the change that occurs during the increase in proton aurorae, but the much larger change that occurs in the time that these particles complete a drift orbit. This is especially relevant because the high energy particles that are susceptible to stochastic scattering cannot be treated as though they remain in the tail, due to their relatively fast drift orbits. The onset of the proton aurora in this case would occur at the moment these high energy drifting particles return to the tail.

Note that stochastic scattering only affects a small, high energy, subset of the total particle distribution and therefore does not significantly affect our conclusions regarding the movement of the distribution as a whole toward increasingly parallel anisotropy. As Wang, et al. (2013) show, using several magnetic field models, the strong scattering condition generally applies at distances around geosynchronous orbit at energies higher than approximately 100 keV. Therefore, the majority of ion energies do not have gyroradii for which the strong scattering condition applies in the transition region. In addition, Birn, et al. (1997) also reported on the

phenomenon of proton aurorae using their in-situ measurements. They also conclude that the stochastic scattering effect occurs only for high energy protons. In addition, they note that this is due to the addition of a new population to the existing distribution, an observation that agrees very well with our hypothesis that the scattered particles represent the addition of a distribution that drifts into the more stretched, late growth phase tail.

5. Conclusions

We have shown that the destabilizing effect of increasingly parallel pressure anisotropies, identified by Oberhagemann and Mann (2019) as a mechanism for triggering substorm onset with a ballooning instability, can occur as a natural consequence of the evolution of tail magnetic field topology during the substorm growth phase. As particles drift from a less stretched, early growth phase tail, changes occur on the nightside during the time the particles' drift orbits are outside the tail. The field stretching and change in tail field topology results in the well-known drift shell splitting effect when these particles return, in this case leading to higher pitch angle particles moving radially inward and low pitch angle particles moving radially outward. When this occurs in the high pressure gradient region that is also conducive to the formation of ballooning instabilities in the region of locally earthward pressure gradients, it results in a larger number of low pitch angle particles mixing with a lower number of high pitch angle particles, the radial gradients causing a shift toward more parallel pressure anisotropies. In addition, the growth phase transport from less stretched to more stretched field lines causes a de-energization that acts preferentially in the perpendicular direction, resulting in more parallel anisotropy over a wider region. In addition to self-consistently predicting the region in the tail where auroral substorms are seen to originate with the brightening of the equatorward arc on closed field lines, corresponding in the equatorial plane to the region where high pressure gradients exist in the

transition between dipolar and tail-like fields, the onset mechanism proposed here also shows remarkable agreement with observations. Namely:

1. The ballooning instabilities that form as a consequence of this mechanism have been linked to the observations of auroral beads described by Kalmoni et al. (2018)

2. Observations of dipolarization can be tied to reconnection, which may be directly caused by, and subsequent to, the onset of ballooning instabilities (Zhu et al., 2017)

3. Observations of increasingly parallel anisotropy in electrons and ions prior to substorm onset align with the onset mechanism proposed here (Baker et al., 1978, Birn et al., 1997)

4. Tail stretching and a movement toward parallel anisotropy occur concurrently (Birn et al., 1997)

5. The well-established phenomenon of growth phase tail stretching that occurs before substorm onset, as first described by McPherron, (1970), causes the increasingly parallel anisotropies that allow ballooning instabilities to form. This explains why substorm onset follows such a period of tail stretching and distinguishes periods of more rapid stretching from those with slower tail stretching that do not cause substorm onset.

The explanation that this mechanism offers for the occurrence of proton aurorae around the substorm onset arc illustrates one other testable prediction that may be derived from it. Namely, the return of particles to a more stretched tail that is key to triggering increasingly parallel anisotropic ballooning leads also and concurrently to the precipitation of a higher-energy part of the particle distribution at locations around the substorm onset arc due to stochastic scattering through field line curvature effects.

Therefore, this work establishes a new substorm onset mechanism that offers significant benefits for understanding the events around substorms through testable predictions. It suggests a new framework to be further examined using in-situ measurements and more detailed numerical work. Moreover, not only does it offer a physics-based framework for substorm onset prediction,

it also presents a potential physical answer to the long-standing question of why a substorm onset begins when it does.

Acknowledgements

This was work was supported by an NSERC Discovery Grant to IRM and the NSERC CREATE grant for the International Space Masters (ISM) program.

References

Akasofu, S.-I. (1964). The development of the auroral substorm. *Planetary and Space Science*, *12*(4), 273–282. https://doi.org/10.1016/0032-0633(64)90151-5

Baker, D. N., Higbie, P. R., Hones, E. W., & Belian, R. D. (1978). High-resolution energetic particle measurements at 6.6 RE 3. Low-energy electron anisotropies and short-term substorm predictions. *Journal of Geophysical Research: Space Physics*, *83*(A10), 4863–4868. https://doi.org/10.1029/JA083iA10p04863

Birn, J., Thomsen, M. F., Borovsky, J. E., Reeves, G. D., McComas, D. J., & Belian, R. D. (1997). Characteristic plasma properties during dispersionless substorm injections at geosynchronous orbit. *Journal of Geophysical Research: Space Physics*, *102*(A2), 2309–2324. https://doi.org/10.1029/96JA02870

Borovsky, J. E., Nemzek, R. J., & Belian, R. D. (1993a). The occurrence rate of magnetospheric-substorm onsets: Random and periodic substorms. *Journal of Geophysical Research: Space Physics*, *98*(A3), 3807–3813. https://doi.org/10.1029/92JA02556

Borovsky, J. E., Nemzek, R. J., & Belian, R. D. (1993b). The occurrence rate of magnetospheric-substorm onsets: Random and periodic substorms. *Journal of Geophysical Research: Space Physics*, *98*(A3), 3807–3813. https://doi.org/10.1029/92JA02556

Chan, A. A., Xia, M., & Chen, L. (1994). Anisotropic Alfvén-ballooning modes in Earth's magnetosphere. *Journal of Geophysical Research: Space Physics*, *99*(A9), 17351–17366. https://doi.org/10.1029/93JA03353

E. Borovsky, J., J. Nemzek, R., & D. Belian, R. (1993). The occurrence rate of magnetospheric-substorm onsets - Random and periodic substorms. *Journal of Geophysical Research*, *98*, 3807–3813. https://doi.org/10.1029/92JA02556

Gilson, M. L., Raeder, J., Donovan, E., Ge, Y. S., & Kepko, L. (2012). Global simulation of proton precipitation due to field line curvature during substorms. *Journal of Geophysical Research: Space Physics*, *117*(A5). https://doi.org/10.1029/2012JA017562

Kabin, K., Spanswick, E., Rankin, R., Donovan, E., & Samson, J. C. (2011). Modeling the relationship between substorm dipolarization and dispersionless injection. *Journal of Geophysical Research: Space Physics*, *116*(A4). https://doi.org/10.1029/2010JA015736

Kalmoni, N. M. E., Rae, I. J., Murphy, K. R., Forsyth, C., Watt, C. E. J., & Owen, C. J. (2017). Statistical azimuthal structuring of the substorm onset arc: Implications for the onset mechanism. *Geophysical Research Letters*, *44*(5), 2078–2087. https://doi.org/10.1002/2016GL071826

Kalmoni, N. M. E., Rae, I. J., Watt, C. E. J., Murphy, K. R., Samara, M., Michell, R. G., … Forsyth, C. (2018). A diagnosis of the plasma waves responsible for the explosive energy release of substorm onset. *Nature Communications*, *9*(1), 4806. https://doi.org/10.1038/s41467-018-07086-0

Liang, J., Donovan, E., Gillies, D., Spanswick, E., & Connors, M. (2018). Proton auroras during the transitional stage of substorm onset. *Earth, Planets and Space*, *70*(1). https://doi.org/10.1186/s40623-018-0899-0

Lui, A. T. Y., & Hamilton, D. C. (1992a). Radial profiles of quiet time magnetospheric parameters. *Journal of Geophysical Research: Space Physics*, *97*(A12), 19325–19332. https://doi.org/10.1029/92JA01539

McPherron, R. L. (1970). Growth Phase of Magnetospheric Substorms. *J. Geophys. Res*.

Oberhagemann, L.R, & Mann, I.R. (2019). A new substorm onset mechanism: Increasingly parallel pressure anisotropic ballooning. Manuscript in preparation

Parks, G. K. (1991). *Physics of Space Plasmas*. Addison-Wesley Publishing Company.

Roederer, J. G. (1967). On the adiabatic motion of energetic particles in a model magnetosphere. *Journal of Geophysical Research (1896-1977)*, *72*(3), 981–992. https://doi.org/10.1029/JZ072i003p00981

Samson, J. C., Lyons, L. R., Newell, P. T., Creutzberg, F., & Xu, B. (1992). Proton aurora and substorm intensifications. *Geophysical Research Letters*, *19*(21), 2167–2170. https://doi.org/10.1029/92GL02184

Sergeev, V. A., Angelopoulos, V., & Nakamura, R. (2012). Recent advances in understanding substorm dynamics: FRONTIER. *Geophysical Research Letters*, *39*(5), n/a-n/a. https://doi.org/10.1029/2012GL050859

Sibeck, D. G., McEntire, R. W., Lui, A. T. Y., Lopez, R. E., & Krimigis, S. M. (1987). Magnetic field drift shell splitting: Cause of unusual dayside particle pitch angle distributions during

storms and substorms. *Journal of Geophysical Research: Space Physics*, *92*(A12), 13485–13497. https://doi.org/10.1029/JA092iA12p13485

Viñas, A. F., & Madden, T. R. (1986). Shear flow-ballooning instability as a possible mechanism for hydromagnetic fluctuations. *Journal of Geophysical Research: Space Physics*, *91*(A2), 1519–1528. https://doi.org/10.1029/JA091iA02p01519

Wang, C.-P., Zaharia, S. G., Lyons, L. R., & Angelopoulos, V. (2013). Spatial distributions of ion pitch angle anisotropy in the near-Earth magnetosphere and tail plasma sheet. *Journal of Geophysical Research: Space Physics*, *118*(1), 244–255. https://doi.org/10.1029/2012JA018275

Wanliss, J., C. Samson, J., & Friedrich, E. (2000). On the use of photometer data to map dynamics of the magnetotail current sheet during substorm growth phase. *Journal of Geophysical Research*, *105*, 27673–27684. https://doi.org/10.1029/2000JA000178

Zaharia, S., Jordanova, V. K., Thomsen, M. F., & Reeves, G. D. (2006). Self- consistent modeling of magnetic fields and plasmas in the inner magnetosphere: Application to a geomagnetic storm. *Journal of Geophysical Research: Space Physics*, *111*(A11). https://doi.org/10.1029/2006JA011619

Zhu, P., Bhattacharjee, A., Sangari, A., Wang, Z., & Bonofiglo, P. (2017). Three-dimensional geometry of magnetic reconnection induced by ballooning instability in a generalized Harris sheet. *Physics of Plasmas*, *24*(2), 024503. https://doi.org/10.1063/1.4976994

Chapter 7: Conclusions and Future Work

This section contains a general summary of the work presented in this thesis. Additional and more detailed conclusions are also presented in the papers in chapters 4 and 6.

A new substorm onset mechanism was proposed in which particles gradient-curvature drift from a less stretched to a more stretched magnetotail as the growth phase progresses. The increased length of the field lines on which these orbits terminate, relative to the starting field lines, results in a preferential de-energization of the particles in the direction perpendicular to the field lines. In addition, particles with pitch angles that are more perpendicular to the field lines originate in a lower pressure region than those with pitch angles that are more parallel. The combined effect of this is to cause the pressure distribution to change from mildly perpendicular anisotropic toward more parallel anisotropy. In conjunction with the increased plasma β and pressure gradient that forms in the more stretched tail prior to onset, this corresponds to a movement towards ballooning instability which is triggered by a transition towards increased parallel pressure anisotropy. The triggering of a ballooning instability may then cause substorm onset, appearing as the signature of auroral beads on the onset arc. The ballooning may perhaps also result in the triggering of magnetic reconnection further down the tail, in addition to creating the auroral beads. Such a hypothesis would then identify the triggering of ballooning with substorm expansion phase onset, with the reconnection also subsequently being responsible for the well-known dipolarisation and the poleward and westward motion of the aurora associated with auroral breakup.

The region in the tail which is identified as most likely to be associated with the destabilistion of ballooning indicated by the results presented here, namely the transition region between dipolar and tail-like magnetic field lines where relatively large pressures and pressure gradients exist, agrees with existing proposals for the location of near-Earth substorm onset (e.g. Samson et al., 1992). The time required for 40 keV ions to complete a drift orbit of Earth in this region is on the order of 2-4 hours (Parks et al., 1991), which agrees well with the most common time between substorms of 2.75 hours (Borovsky et al., 1993). Most significantly, the observed connection between magnetic field stretching and an increasingly parallel anisotropy in electrons represents a behaviour that is strongly reflective of that predicted by the mechanism of increasingly parallel anisotropic ballooning advanced here (e.g., Birn et al., 1997). Therefore, the

proposed onset mechanism which we have developed and advanced in this thesis appears to show good agreement with existing observations.

Physical arguments and the details of the onset mechanism advanced here were developed through the use of the model presented by Chan et al. (1994) to determine the thresholds for ballooning instability, and by using the two-dimensional magnetic field model created by Kabin et al. (2011) to trace the evolution of the distributions of particles drifting around Earth. While these models are useful in capturing the key predictions arising from the model, and identifying the key physical effects, they provide the basis for further future data and modelling interpretations rather than accurately directly simulating the development of ballooning instabilities in a realistic evolving growth phase magnetotail. Therefore, a three-dimensional and more accurate magnetic field model in which particles can be traced as they orbit Earth, and in which ballooning instabilities can develop directly, should be a key component of any future work. In addition, as called for by Kalmoni et al. (2018), further observational study of the transition region is required through in-situ measurements to examine its dynamics in the time surrounding substorm onset. This would require significant satellite coverage because of the apparently radially narrow region in which these events occur in order to allow for the successful in-situ observation of ballooning instabilities and kinetic Alfvén waves, should they exist in this region. In addition, it would allow for the possibility of observing characteristics of the out-to-in mechanism such as the arrival of flows of plasma from the more distant tail, for example as might occur from subsequent magnetic reconnection.

Overall, the results presented in this thesis represent the genesis of a new model for substorm onset, which appears to be consistent with many well-known observational constraints. It can explain not only the time sequence of the events at onset, but also establish a basis for understanding their causal relationships. The predictions of the model are testable, and if correct can provide an explanation for not only how substorms are triggered, but also address the long-standing question of why a specific magnetotail configuration triggers a substorm at the time and location that it does.

References

Abraham-Shrauner, B. (1967). Propagation of hydromagnetic waves through an anisotropic plasma. *Journal of Plasma Physics*, *1*(3), 361–378. https://doi.org/10.1017/S0022377800003354

Akasofu, S.-I. (1964). The development of the auroral substorm. *Planetary and Space Science*, *12*(4), 273–282. https://doi.org/10.1016/0032-0633(64)90151-5

Baker, D. N., Bame, S. J., Birn, J., Feldman, W. C., Gosling, J. T., Hones, E. W., … Sibeck, D. G. (1984). Direct observations of passages of the distant neutral line (80-140 RE) following substorm pnsets: ISEE-3. *Geophysical Research Letters*, *11*(10), 1042–1045. https://doi.org/10.1029/GL011i010p01042

Baker, D. N., Higbie, P. R., Hones, E. W., & Belian, R. D. (1978). High-resolution energetic particle measurements at 6.6 RE 3. Low-energy electron anisotropies and short-term substorm predictions. *Journal of Geophysical Research: Space Physics*, *83*(A10), 4863–4868. https://doi.org/10.1029/JA083iA10p04863

Banks, P. M. (1971). Interplanetary hydrogen and helium from cosmic dust and the solar wind. *Journal of Geophysical Research (1896-1977)*, *76*(19), 4341–4348. https://doi.org/10.1029/JA076i019p04341

Birn, J., Thomsen, M. F., Borovsky, J. E., Reeves, G. D., McComas, D. J., & Belian, R. D. (1997). Characteristic plasma properties during dispersionless substorm injections at geosynchronous orbit. *Journal of Geophysical Research: Space Physics*, *102*(A2), 2309–2324. https://doi.org/10.1029/96JA02870

Borovsky, J. E., Nemzek, R. J., & Belian, R. D. (1993). The occurrence rate of magnetospheric-substorm onsets: Random and periodic substorms. *Journal of Geophysical Research: Space Physics*, *98*(A3), 3807–3813. https://doi.org/10.1029/92JA02556

Bruno, R., Burlaga, L. F., & Hundhausen, A. J. (1982). Quadrupole distortions of the heliospheric current sheet in 1976 and 1977. *Journal of Geophysical Research: Space Physics*, *87*(A12), 10339–10346. https://doi.org/10.1029/JA087iA12p10339

Burgess, D. (1995). Collisionless Shocks. In M. G. Kivelson & Russell, C.T. (Eds.), *Introduction to Space Physics* (pp. 129–163). Cambridge University Press.

Chan, A. A., Xia, M., & Chen, L. (1994). Anisotropic Alfvén-ballooning modes in Earth's magnetosphere. *Journal of Geophysical Research: Space Physics*, *99*(A9), 17351–17366. https://doi.org/10.1029/93JA03353

Chapman, S. (1956). The electrical conductivity of the ionosphere: A review. *Il Nuovo Cimento (1955-1965)*, *4*(4), 1385–1412. https://doi.org/10.1007/BF02746310

Chen, L., & Hasegawa, A. (1991). Kinetic theory of geomagnetic pulsations: 1. Internal excitations by energetic particles. *Journal of Geophysical Research: Space Physics*, *96*(A2), 1503–1512. https://doi.org/10.1029/90JA02346

Cheng, C. Z. (1992). Magnetospheric equilibrium with anisotropic pressure. *Journal of Geophysical Research: Space Physics*, *97*(A2), 1497–1510. https://doi.org/10.1029/91JA02433

Cliver, E. W., & Svalgaard, L. (2004). The 1859 Solar–Terrestrial Disturbance And the Current Limits of Extreme Space Weather Activity. *Solar Physics*, *224*(1), 407–422. https://doi.org/10.1007/s11207-005-4980-z

Dungey, J. W. (1961). Interplanetary Magnetic Field and the Auroral Zones. *Physical Review Letters*, *6*, 47–48. https://doi.org/10.1103/PhysRevLett.6.47

E. Borovsky, J., J. Nemzek, R., & D. Belian, R. (1993). The occurrence rate of magnetospheric-substorm onsets - Random and periodic substorms. *Journal of Geophysical Research*, *98*, 3807–3813. https://doi.org/10.1029/92JA02556

E.R. Priest. (1995). The Sun and its Magnetohydrodynamics. In M. G. Kivelson & Russell, C.T. (Eds.), *Introduction to Space Physics* (pp. 58–89). Cambridge University Press.

Ferrière, K. (2001). Interchange, quasi-interchange, ballooning modes: What is their exact definition? *Eos, Transactions American Geophysical Union*, *82*(38), 420–420. https://doi.org/10.1029/01EO00260

Freidberg, J. P. (1982). Ideal magnetohydrodynamic theory of magnetic fusion systems. *Reviews of Modern Physics*, *54*(3), 801–902. https://doi.org/10.1103/RevModPhys.54.801

Gilson, M. L., Raeder, J., Donovan, E., Ge, Y. S., & Kepko, L. (2012). Global simulation of proton precipitation due to field line curvature during substorms. *Journal of Geophysical Research: Space Physics*, *117*(A5). https://doi.org/10.1029/2012JA017562

Hasegawa, A., & Sato, T. (1989). *Space plasma physics: I - Stationary processes*. Retrieved from http://adsabs.harvard.edu/abs/1989spp..book.....H

Hundhausen, A. J. (1995). The Solar Wind. In M. G. Kivelson & Russell, C.T. (Eds.), *Introduction to Space Physics* (pp. 91–128). Cambridge University Press.

Kabin, K., Kalugin, G., Donovan, E., & Spanswick, E. (2017). Particle energization by a substorm dipolarization. *Journal of Geophysical Research: Space Physics*, *122*(1), 349–367. https://doi.org/10.1002/2016JA023459

Kabin, K., Spanswick, E., Rankin, R., Donovan, E., & Samson, J. C. (2011). Modeling the relationship between substorm dipolarization and dispersionless injection. *Journal of Geophysical Research: Space Physics*, *116*(A4). https://doi.org/10.1029/2010JA015736

Kalmoni, N. M. E., Rae, I. J., Murphy, K. R., Forsyth, C., Watt, C. E. J., & Owen, C. J. (2017). Statistical azimuthal structuring of the substorm onset arc: Implications for the onset mechanism. *Geophysical Research Letters*, *44*(5), 2078–2087. https://doi.org/10.1002/2016GL071826

Kalmoni, N. M. E., Rae, I. J., Watt, C. E. J., Murphy, K. R., Forsyth, C., & Owen, C. J. (2015). Statistical characterization of the growth and spatial scales of the substorm onset arc. *Journal of Geophysical Research: Space Physics*, *120*(10), 8503–8516. https://doi.org/10.1002/2015JA021470

Kalmoni, N. M. E., Rae, I. J., Watt, C. E. J., Murphy, K. R., Samara, M., Michell, R. G., … Forsyth, C. (2018). A diagnosis of the plasma waves responsible for the explosive energy release of substorm onset. *Nature Communications*, *9*(1), 4806. https://doi.org/10.1038/s41467-018-07086-0

Kivelson, M. G. (1995). Physics of Space Plasmas. In M. G. Kivelson & Russell, C.T. (Eds.), *Introduction to Space Physics* (pp. 27–55). Cambridge University Press.

Liang, J., Donovan, E., Gillies, D., Spanswick, E., & Connors, M. (2018). Proton auroras during the transitional stage of substorm onset. *Earth, Planets and Space*, *70*(1). https://doi.org/10.1186/s40623-018-0899-0

Lloyds of London. (2013). *Solar Storm Risk to the North American Electric Grid*. Retrieved from https://www.lloyds.com/news-and-risk-insight/risk-reports/library/natural-environment/solar-storm

Lui, A. T. Y. (2004). Potential Plasma Instabilities For Substorm Expansion Onsets. *Space Science Reviews*, *113*(1/2), 127–206. https://doi.org/10.1023/B:SPAC.0000042942.00362.4e

Lui, A. T. Y., & Hamilton, D. C. (1992). Radial profiles of quiet time magnetospheric parameters. *Journal of Geophysical Research: Space Physics*, *97*(A12), 19325–19332. https://doi.org/10.1029/92JA01539

McPherron, R. L. (1970). Growth Phase of Magnetospheric Substorms. *J. Geophys. Res*.

Ozeke, L. G., & Mann, I. R. (2004). Modeling the properties of guided poloidal Alfvén waves with finite asymmetric ionospheric conductivities in a dipole field. *Journal of Geophysical Research: Space Physics*, *109*(A5). https://doi.org/10.1029/2003JA010151

Panov, E. V., Sergeev, V. A., Pritchett, P. L., Coroniti, F. V., Nakamura, R., Baumjohann, W., … McFadden, J. P. (2012). Observations of kinetic ballooning/interchange instability signatures in the magnetotail. *Geophysical Research Letters*, *39*(8). https://doi.org/10.1029/2012GL051668

Parker, E. N. (1958). Dynamics of the Interplanetary Gas and Magnetic Fields. *The Astrophysical Journal*, *128*, 664. https://doi.org/10.1086/146579

Parks, G. K. (1991). *Physics of Space Plasmas*. Addison-Wesley Publishing Company.

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. (1992). *Numerical Recipes in FORTRAN 77: Volume 1, Volume 1 of Fortran Numerical Recipes: The Art of Scientific Computing* (2 edition). Cambridge England ; New York: Cambridge University Press.

Pulkkinen, A., Lindahl, S., Viljanen, A., & Pirjola, R. (2005). Geomagnetic storm of 29–31 October 2003: Geomagnetically induced currents and their relation to problems in the Swedish high-voltage power transmission system. *Space Weather*, *3*(8). https://doi.org/10.1029/2004SW000123

Rae, I. J., Mann, I. R., Angelopoulos, V., Murphy, K. R., Milling, D. K., Kale, A., … Donovan, E. F. (2009). Near-Earth initiation of a terrestrial substorm. *Journal of Geophysical Research: Space Physics*, *114*(A7). https://doi.org/10.1029/2008JA013771

Roederer, J. G. (1970). *Dynamics of Geomagnetically Trapped Radiation*. Retrieved from https://www.springer.com/gp/book/9783642493027

Roederer, Juan G. (1967). On the adiabatic motion of energetic particles in a model magnetosphere. *Journal of Geophysical Research (1896-1977)*, *72*(3), 981–992. https://doi.org/10.1029/JZ072i003p00981

Roederer, Juan G., & Zhang, H. (2013). *Dynamics of Magnetically Trapped Particles: Foundations of the Physics of Radiation Belts and Space Plasmas* (2nd ed.). Retrieved from https://www.springer.com/gp/book/9783642415296

Russell, C. T., Luhmann, J. G., & Strangeway, R. J. (2016). *Space Physics: An Introduction* (Har/Psc edition). New York, NY: Cambridge University Press.

Samson, J. C., Lyons, L. R., Newell, P. T., Creutzberg, F., & Xu, B. (1992). Proton aurora and substorm intensifications. *Geophysical Research Letters*, *19*(21), 2167–2170. https://doi.org/10.1029/92GL02184

Sergeev, V. A., Angelopoulos, V., & Nakamura, R. (2012). Recent advances in understanding substorm dynamics: FRONTIER. *Geophysical Research Letters*, *39*(5), n/a-n/a. https://doi.org/10.1029/2012GL050859

Shue, J.-H., & Chao, J.-K. (2013). The role of enhanced thermal pressure in the earthward motion of the Earth's magnetopause. *Journal of Geophysical Research: Space Physics*, *118*(6), 3017–3026. https://doi.org/10.1002/jgra.50290

Sibeck, D. G., McEntire, R. W., Lui, A. T. Y., Lopez, R. E., & Krimigis, S. M. (1987). Magnetic field drift shell splitting: Cause of unusual dayside particle pitch angle distributions during storms and substorms. *Journal of Geophysical Research: Space Physics*, *92*(A12), 13485–13497. https://doi.org/10.1029/JA092iA12p13485

Stern, D. P. (1979). The electric field and global electrodynamics of the magnetosphere: Review and Quadrennial Report to the IUGG. *Reviews of Geophysics*, *17*(4), 626–640. https://doi.org/10.1029/RG017i004p00626

Tsyganenko, N. A. (1995). Modeling the Earth's magnetospheric magnetic field confined within a realistic magnetopause. *Journal of Geophysical Research: Space Physics*, *100*(A4), 5599–5612. https://doi.org/10.1029/94JA03193

Viñas, A. F., & Madden, T. R. (1986). Shear flow-ballooning instability as a possible mechanism for hydromagnetic fluctuations. *Journal of Geophysical Research: Space Physics*, *91*(A2), 1519–1528. https://doi.org/10.1029/JA091iA02p01519

Wang, C.-P., Zaharia, S. G., Lyons, L. R., & Angelopoulos, V. (2013). Spatial distributions of ion pitch angle anisotropy in the near-Earth magnetosphere and tail plasma sheet. *Journal of Geophysical Research: Space Physics*, *118*(1), 244–255. https://doi.org/10.1029/2012JA018275

Wanliss, J., C. Samson, J., & Friedrich, E. (2000). On the use of photometer data to map dynamics of the magnetotail current sheet during substorm growth phase. *Journal of Geophysical Research*, *105*, 27673–27684. https://doi.org/10.1029/2000JA000178

Wilcox, J. M., & Hundhausen, A. J. (1983). Comparison of heliospheric current sheet structure obtained from potential magnetic field computations and from observed polarization coronal brightness. *Journal of Geophysical Research*, *88*(A10), 8095. https://doi.org/10.1029/JA088iA10p08095

Wilf, H. S. (1978). A Global Bisection Algorithm for Computing the Zeros of Polynomials in the Complex Plane. *J. ACM*, *25*(3), 415–420. https://doi.org/10.1145/322077.322084

Zaharia, S., Jordanova, V. K., Thomsen, M. F., & Reeves, G. D. (2006). Self- consistent modeling of magnetic fields and plasmas in the inner magnetosphere: Application to a geomagnetic storm. *Journal of Geophysical Research: Space Physics*, *111*(A11). https://doi.org/10.1029/2006JA011619

Zhu, P., Bhattacharjee, A., Sangari, A., Wang, Z., & Bonofiglo, P. (2017). Three-dimensional geometry of magnetic reconnection induced by ballooning instability in a generalized Harris sheet. *Physics of Plasmas*, *24*(2), 024503. https://doi.org/10.1063/1.4976994

Appendix A: Supplementary Material for Chapter 3

Appendix A1: A Robust Method for Finding Complex Eigenvalues

This section consists of coursework completed for the PHYS 580 Computational Physics course in the fall of 2016, but whose results were used in the course of the research work presented in this thesis.

While a secant method is an effective means of root finding in the complex plane, it requires some knowledge of the solution space to be able to make effective initial guesses. Because of this, a robust root finding scheme that can find complex roots without requiring accurate initial guesses is useful. For real roots this is simple because a bisection algorithm can be used, where the only requirement is that the root is bracketed. Therefore, to explore the solution space I used a root finder that works analogously to the bisection algorithm for the complex plane.

The most direct way to map the solution space would be to directly solve the eigenvalue problem for a grid of possible eigenvalues in the complex plane. Because my integration starts at the ionosphere, the difference between the value of the eigenfunction at the equator and its value according to the boundary condition could be evaluated. Plotting the inverse of this difference would result in sharp peaks where the boundary condition is met.

By reading this plot, one may find a rough approximation of the eigenvalues that can be used to run a secant algorithm. The issue with this method is that to resolve all of the eigenvalues over a large area, the integral must be evaluated for many guessed eigenvalues. For a 4x15 grid in the complex plane, this takes about 30 seconds for one case. If the eigenvalues are close together, as occurs at the onset of instability, a higher resolution is required so this will take more time. This is in addition to the time required for the user to run the subsequent set of secant algorithms. Therefore, a faster way is required.

Another option is to use Nyquist integration, using the equation:

$$\oint_{\Gamma} dz \; z\frac{f'}{f} = 2\pi i \sum_{l=1}^{N} z_l^p \tag{A1.1}$$

Where Γ is a closed, counter clockwise contour around a group of roots, f is the function value on that contour, f' is its derivative, and z is the location on the contour.

This would remove the requirement for user input, resulting in a much faster solution time. However, to use this method for the large area over which the eigenvalues may be spread, and to find their values with sufficient accuracy, a large number of points would be required to achieve a precise integration around the perimeter of the contour. This also requires a large amount of computational resources.

To take advantage of the lack of user input in Nyquist integration, a root finder was constructed that uses Cauchy's argument principle, given by:

$$\oint_{\Gamma} dz \, \frac{f'}{f} = 2N\pi i \tag{A1.2}$$

where N is the number of roots inside of the contour. A search for a bisection method in the complex plane yielded a paper by H. Wilf that describes evaluating Cauchy's argument principle around a region in the complex plane divided into four sub regions (Wilf 1978). The region is then shrunk to include only the sub regions that contain roots, and the refinement can be continued until the root is bracketed closely. The method described here uses this general idea.

An additional refinement to this method could be implemented in which Nyquist integration is employed once the perimeter is sufficiently small. However, due to the speed with which this method was found to converge, this is not necessary.

The Root Finding Scheme

The root finding scheme evaluates the integral in Cauchy's argument principle around four quadrants in a specified domain, shrinks the domain, and breaks this smaller domain into four new quadrants until it circles individual roots very closely as illustrated in Figure 10. Figure 23 shows a block diagram that describes the algorithm. The individual sections are described below.

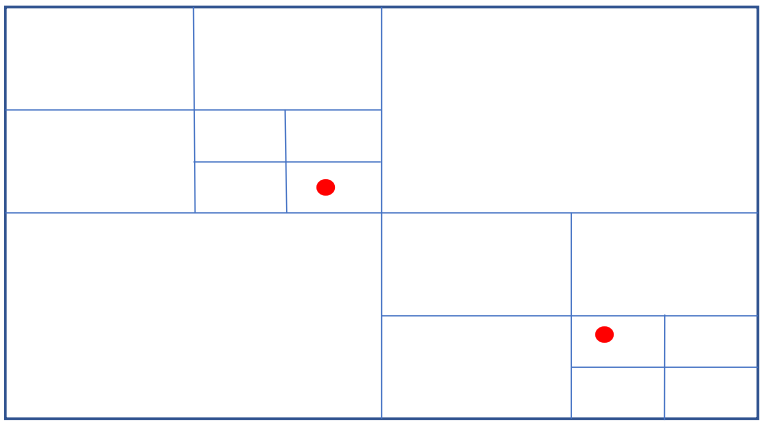

Figure 22: Schematic illustration of the root finding scheme, red points represent the roots

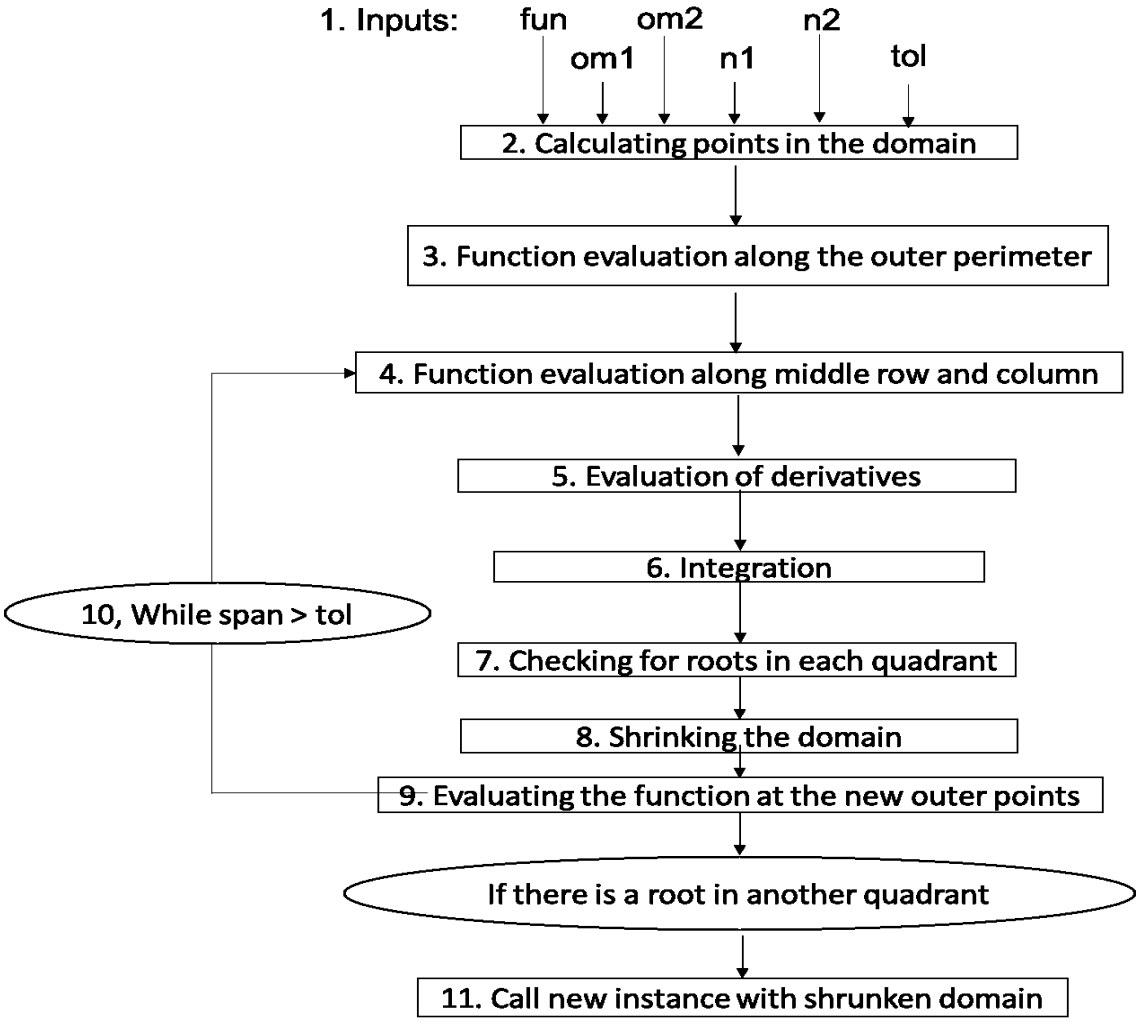

Figure 23: Block diagram describing the root finder

1. Inputs

The inputs to this function are as follows:

a. fun: the function on which the root finding scheme should operate
b. om1: the corner of the domain with the smallest real and imaginary part
c. om2: the corner of the domain with the largest real and imaginary part
d. n1: the number of points placed in the domain along the imaginary direction
e. n2: the number of points placed in the domain along the real direction
f. tol: the specified tolerance, being the width or the height of the domain, whichever is largest, below which the root finder terminates

2. Calculating the points in the domain

The domain is placed between the corners given by om1 and om2 as shown in Figure 24. om1 is chosen so that the domain spans the imaginary axis. In Figure 24, n1 is 7 and n2 is 5. Points are interpolated between these ends and placed in a matrix, with values in the same row having the same imaginary component and values in the same column having the same real component. For example, the second point in row 1 would be given by $om1 + \frac{real(om2)-real(om1)}{n2-1} * 1 = 1 - 2i$

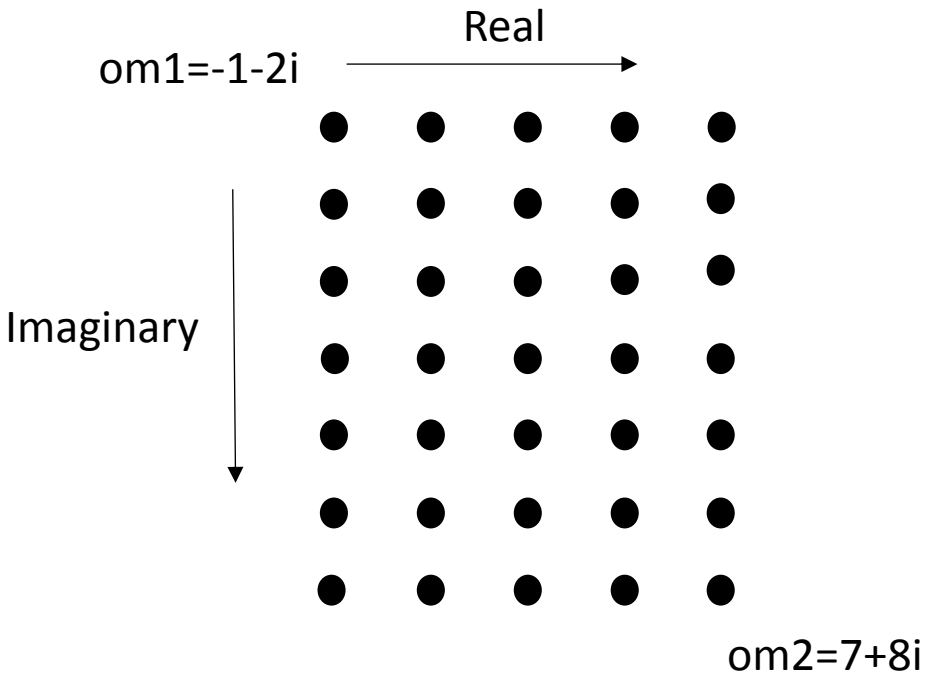

Figure 24: Illustration of the distribution of points in the example interval. See text for details.

3. Function evaluation along the outer perimeter

Each of the points along the outer perimeter is put into the specified function, fun, and its corresponding value is calculated.

4. Function evaluation along the middle row and column

The same process is repeated along the middle row and column, shown in red in Figure 25. Note that for there to be a middle row and column, n1 and n2 must be odd numbers. To avoid issues where there is a root on the boundary of the integration, the domain should be slightly asymmetric about the real line so that the middle row does not fall on it.

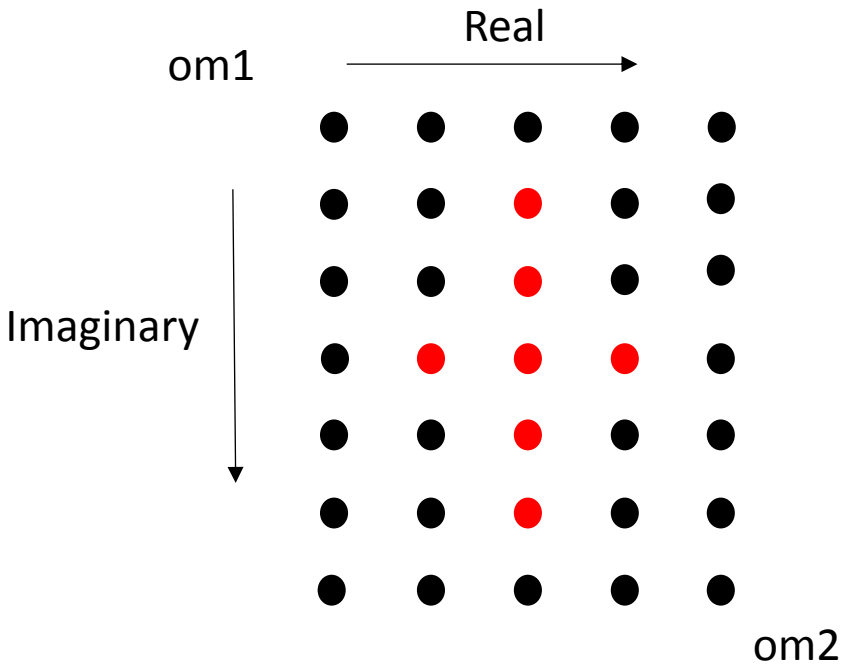

Figure 25: Illustration fo the root funding process involving horizontal and vertical middle rows (red). See text for details.

5. Evaluation of derivatives

To evaluate derivatives at every point, including the corners, a combination of forward and backward differencing schemes was used. This is illustrated for the outer perimeter in Figure 26, where rightward and downward arrows represent forward differencing and leftward and upward arrows represent backward differencing. The value of the derivative is assigned to the

point at the root of each arrow. The derivatives along the middle row and column are evaluated using forward differencing.

The value for the number of roots found using Cauchy's argument principle only needs to be roughly approximate to an integer, since the root finder only needs to be able to tell the difference between an interval containing a root and one that doesn't. This is further described in section 7. Therefore, first order approximations for the derivatives are appropriate as they can be evaluated quickly.

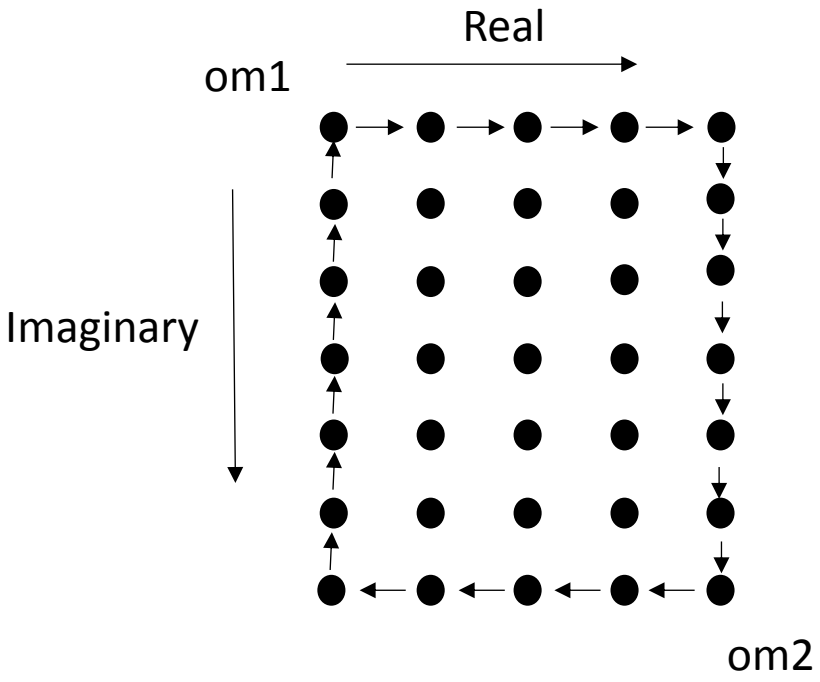

Figure 26: Schematic of differentiation along the outer perimeter in the root finder algorithm. See text for details.

6. Integration

Next, integrals for use in Cauchy's argument principle are evaluated between all the points on the outer perimeter and along the horizontal and vertical middle rows.

Trapezoidal rule integration was used for the integration required for Cauchy's argument principle. This is due to the simplicity with which this integration scheme can be evaluated and the fact that it is conducive to progressive grid refinement, as described in section 9.

7. Checking for roots in each quadrant

The integrals between the points are combined to form integrals around the four closed contours that bound the four quadrants. These are then tested for the presence of roots.

There is no root in a subinterval if the integral around it is less than $2\pi i * 0.1$. This was empirically determined to allow for coarse integration and differentiation around the intervals while still eliminating those intervals that did not contain any roots.

8. Shrinking the domain

The domain is then reduced to only the points in the first quadrant that is found to contain a root. Following this, additional points are added between the old points to keep the number of points constant and progressively refine the grid spacing. Figure 27 shows the points that comprise the new domain in red, with the added points shown in green.

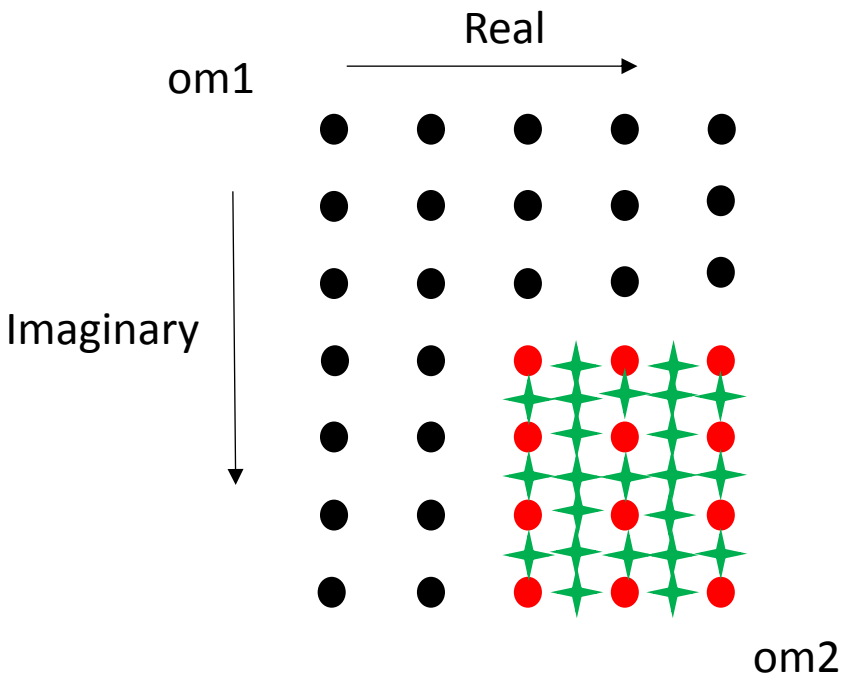

Figure 27: Interpolation (green) of grid points (red) in the shrunken domain. See text for details.

9. Evaluating the function at the new points

The function is then evaluated at the newly added points to allow for the process to be repeated.

This progressive grid refinement has a strong impact on the number of function evaluations required. With grid refinement this is:

$$2(n1-1)+2(n2-1)+(n2-2)+(n1-3) \tag{A1.3}$$

for the first step and

$$(n1-1)+(n2-1)+(n1-2)+(n2-3) \tag{A1.4}$$

for each subsequent one. By contrast, if the grid were not refined, the initial number of points would have to be such that the smallest grid is n1× n2 as these are the smallest dimensions for which this scheme works. This number of points is:

$$\big(2(n1-1)+2(n2-1)+(n2-2)+(n1-3)\big)*2^{I} \tag{A1.5}$$

where I is the number of iterations. Subsequent iterations would also require extra function evaluations along the new middle row and column. For an 11x11 smallest possible grid undergoing 15 iterations, the scheme with grid refinement would require 617 function evaluations whereas without, the first step alone would take 1867776 evaluations.

10. While span > tol

This loop continues as long as either the difference between the largest and smallest real components or the largest and smallest imaginary components is greater than the user specified tolerance (tol). This allows roots to be determined to a greater degree of accuracy by lowering the tol input.

11. Call new instance with shrunken domain

If there is a root in a different subinterval from the one in which the first root was found, the root finder calls another iteration of itself, with the same inputs apart from om1 and om2 which now correspond to the corners of the subinterval in which the additional root was found.

Testing

The root finder was tested on the following function:

$$y = (x-1)(x-5-6i)(x-5-6.01i) \tag{A1.6}$$

This function has a root on the real axis at x=1 and two closely spaced roots at x=5+6i and x=5+6.01i. This emulates the way in which the eigenvalue problem could produce real eigenvalues as well as ones that have the same real components and only slightly different imaginary components.

The root finder was used with the inputs:

a. fun=y
b. om1=0-12.5i
c. om2=15+10i
d. n1=11
e. n2=11
f. tol=0.001

The roots, rounded to the same decimal place as the tolerance, that were found were: 5.000+6.010i, 5.000+6.000i, and 1.000+0.000i, thus matching the expected values. The respective number of iterations for each root were: 5, 13, and 15, where the second two values represent the iterations that were performed in the additional instance that the root finder opened. The time to find all of the roots was ~2s. The reported n1 and n2 are the smallest numbers for which convergence could be achieved. This indicates that this method can significantly reduce the number of function evaluations required, both by requiring few points to begin and reusing old values where possible.

Applying the Complex Subdivision Scheme to the Eigenvalue Problem

Using this scheme for a variety of configurations applied to the ballooning equation, Equation 3.22, both stable and unstable, shows that in the stable cases the eigenvalue, $\tilde{\omega}^2$, is such

that there are three positive frequencies, $\widetilde{\omega}$, that are solutions to the Sturm-Liouville problem. Because the cases are stable, all three frequencies lie on the real line. As the configuration moves toward an instability threshold, the three positive frequencies corresponding to the eigenvalue become smaller and the difference between them shrinks. At the point of the instability threshold, the smallest frequency becomes 0. As the configuration moves further away from the threshold into the unstable region, the formerly lowest frequency takes on increasing values along the imaginary line. The higher eigenvalues follow the same trend until all eigenvalues are on the imaginary line with their magnitudes being reversed.

In agreement with Chan et al. (1994), only purely real or purely imaginary frequencies were found. This is significant for the application of the secant algorithm because, in terms of the eigenvalue, $\widetilde{\omega}^2$, this simply corresponds to a change from a positive to a negative eigenvalue. Therefore, the secant algorithm is applied with initial guesses for the eigenvalue of 0 and 0.1. This causes it to remain along the real line and to converge onto the first eigenvalue at the point when it transitions from positive to negative, corresponding to the onset of instability.

Appendix A2: Validation Through Comparison with an Analytical Solution

To validate the model, I compare my results to an analytical solution. This requires the use of a zero pressure case, because an analytical solution can only be found under these conditions. In the zero pressure case, the Chan et al. (1994) equation reduces to:

$$\frac{d}{dx}\left(\frac{1}{\tilde{L}_x \tilde{B}_V \tilde{R}^2}\frac{d\widetilde{\delta\psi}}{dx}\right) + \frac{\tilde{L}_x}{\tilde{B}_V \tilde{R}^2}\left(\frac{\widetilde{\omega}^2}{\tilde{B}_V^2}\right)\frac{\rho}{\rho_0}\,\widetilde{\delta\psi} = 0 \tag{A2.1}$$

Ozeke et al. (2004) show that this can be reduced to an analytically solvable form through a change in variables. First, $Z = \sin x$ is defined, which reduces the equation to:

$$\frac{d}{dZ}\left(\frac{1}{1+3Z^2}\frac{d\widetilde{\delta\psi}}{dZ}\right) + \frac{\cos^{12} x}{1+3Z^2}\widetilde{\omega}^2\frac{\rho}{\rho_0}\,\widetilde{\delta\psi} = 0 \tag{A2.2}$$

A second change of variables can then be undertaken in which $Y = Z + Z^3$ is defined. This then results in the equation:

$$\frac{d\widetilde{\delta\psi}}{dY} + \frac{\cos^{12} x}{(1+3Z^2)^2}\,\tilde{\omega}^2\,\frac{\rho}{\rho_0}\,\widetilde{\delta\psi} = 0 \tag{A2.3}$$

Finally, to cancel the remaining variables in the second term, the density profile ρ/ρ_0 is set to have the form

$$\frac{\rho}{\rho_0} = \frac{(1+3Z^2)^2}{\cos^{12} x} \tag{A2.4}$$

The analytical solution to the problem then becomes:

$$\widetilde{\delta\psi} = A\cos(\tilde{\omega}Y) + B\sin(\tilde{\omega}Y) \tag{A2.5}$$

Because $Y = 0$ and $\widetilde{\delta\psi} = 0$ at the equator, the constant $A = 0$ and, since any multiple of the eigenfunction is also a solution to the equation, the constant B is arbitrarily set to 1. $\tilde{\omega}$ can then be solved for by using the boundary condition at the ionosphere:

$$\widetilde{\delta\psi}(x_I) = 0 = \sin(\tilde{\omega}\,(\sin(x_I) + \sin^3(x_I)) \tag{A2.6}$$

This can be rearranged to:

$$\tilde{\omega} = \frac{\pi}{\sin(x_I) + \sin^3(x_I)} = 1.845 \tag{A2.7}$$

With this value for comparison, the code is validated by setting the pressure terms $\beta_{\perp 0V} = \alpha = 0$ and the Alfvèn speed $\tilde{V}_A^2 = \frac{\tilde{B}_V^2}{\rho/\rho_0} = \frac{\tilde{B}_V^2 \cos^{12} x}{(1+3\sin^2 x)^2}$. Evaluating this for several grid sizes results in Figure 28:

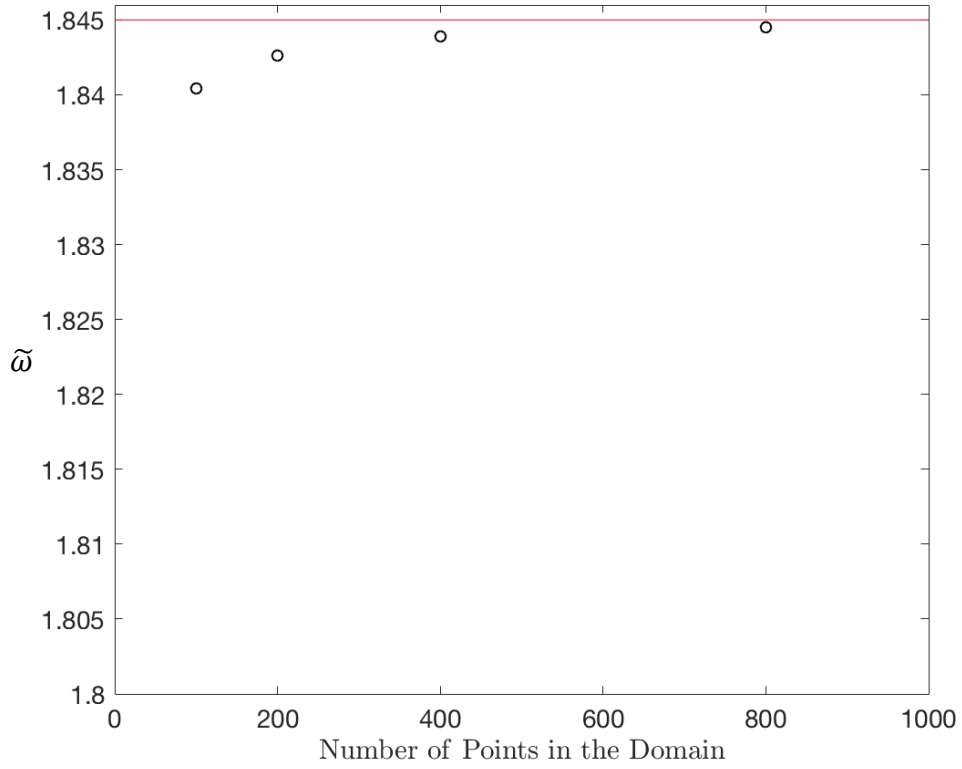

Figure 28: Behaviour of the eigenvalue of Equation 3.22 in the validation case, the red line denotes the analytical solution.

The results in Figure 28 show good agreement for all grid sizes, within 1% of the analytical eigenvalue, and the solution approaches this value asymptotically for a larger number of points, n, in the integration domain. A comparison of the analytical and numerical eigenfunctions is shown for the n=800 case in Figure 29. Both eigenfunctions are normalized to a maximum value of 1. The two overlap very closely, so a close-up view of the comparison is shown in Figure 30. Again, the difference between the two eigenfunctions is significantly less than 1%.

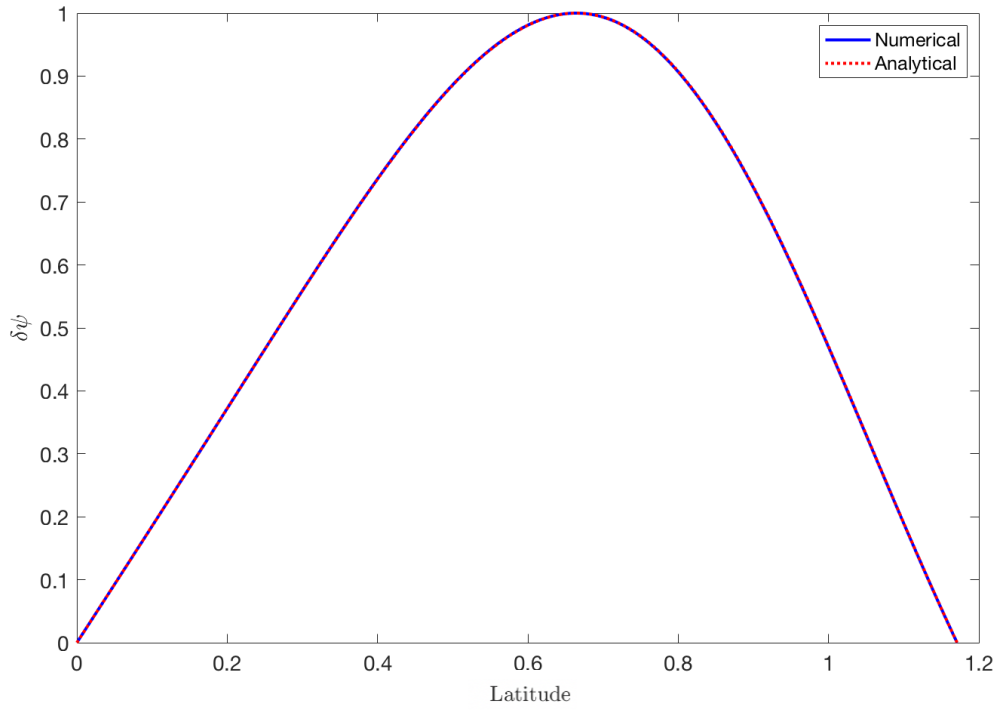

Figure 29: Comparison of the analytically and numerically calculated eigenfunctions, normalized to a maximum value of 1. See text for details.

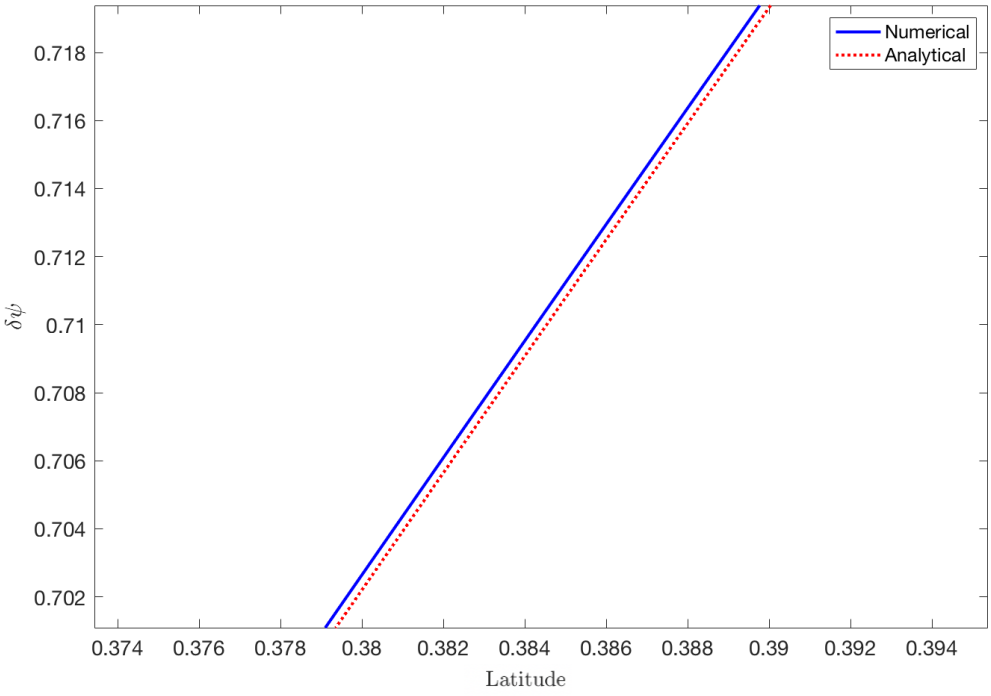

Figure 30: Close-up view of the eigenfunction comparison shown in Figure 8, for a narrow latitudinal range.

Appendix A3: Verification Through Progressive Mesh Coarsening

The numerical model was also verified using progressive mesh coarsening while observing the behaviour of the $\bar{\beta}_0$ value at the stability threshold. This was possible because this value of $\bar{\beta}_0$ corresponds the point at which the eigenvalue becomes zero. The eigenvalue represents the solution of an integration carried out over the whole domain from the equator to the intersection of the field line with Earth and, therefore, captures effects throughout it.

Because n=1000 grid points were used for the calculation of the instability thresholds, this is the finest mesh used in this validation, from which the grid was progressively coarsened using a factor of 1.5 each time. Therefore, the three related grid sizes are n=1000, n=666, and n=444. This progressive coarsening is conducted for the threshold $\bar{\beta}_0$ using three different anisotropies and a constant pressure gradient parameter, $\alpha = 6$. This pressure gradient is the one used in Chan et al. (1994) and therefore allows for comparison between the results presented here and theirs. The isotropic case, with $\delta_0 = 0$, is shown in Figure 31. The second case, shown in Figure 32, uses a mild perpendicular anisotropy. This is because at large perpendicular

anisotropies the ballooning threshold merges with the mirror threshold and can be calculated analytically. The value of δ_0 used in Figure 32 is 0.1, which corresponds approximately to the peak in the threshold diagram. Finally, a strong parallel anisotropy case with $\delta_0 = -1$ is shown in Figure 33. In all three figures, the red line represents the threshold $\bar{\beta}_0$ shown in Chan et al. (1994). Because these values are read from a plot, there is some error in their exact determination. Nonetheless there appears to be reasonable agreement in all cases.

These verification cases further show that for each of the anisotropies chosen, the convergence of the solution appears to be asymptotic, with a difference between the finest mesh and the coarsest mesh of less than 10%. The solutions also fall within 10% of the results in Chan et al. (1994).

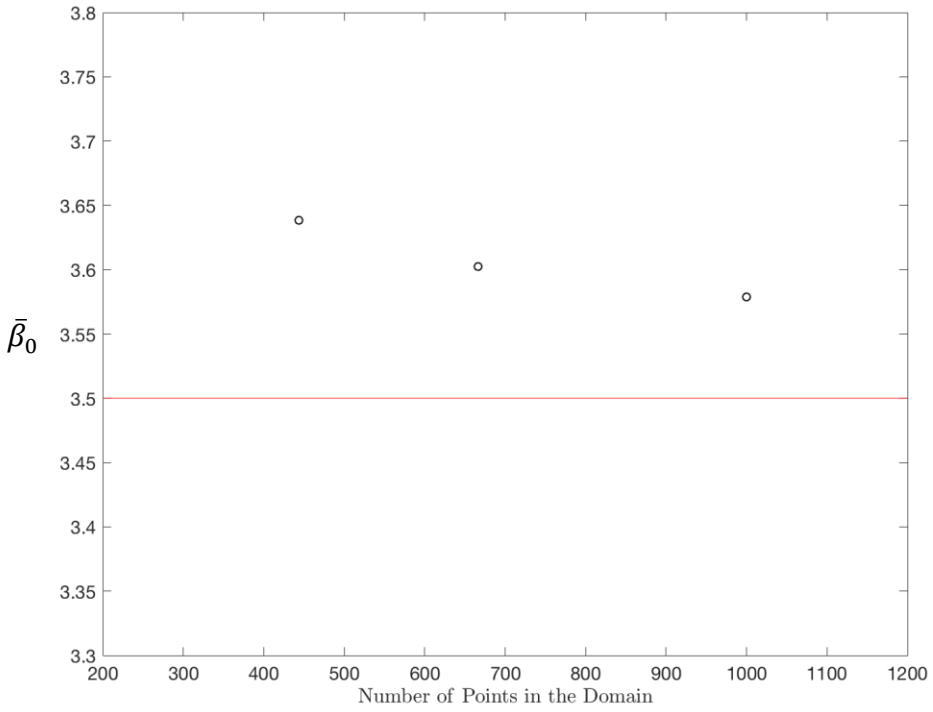

Figure 31: Verification of the $\bar{\beta}_0$ value for the isotropic case of ballooning, $\delta_0 = 0$.

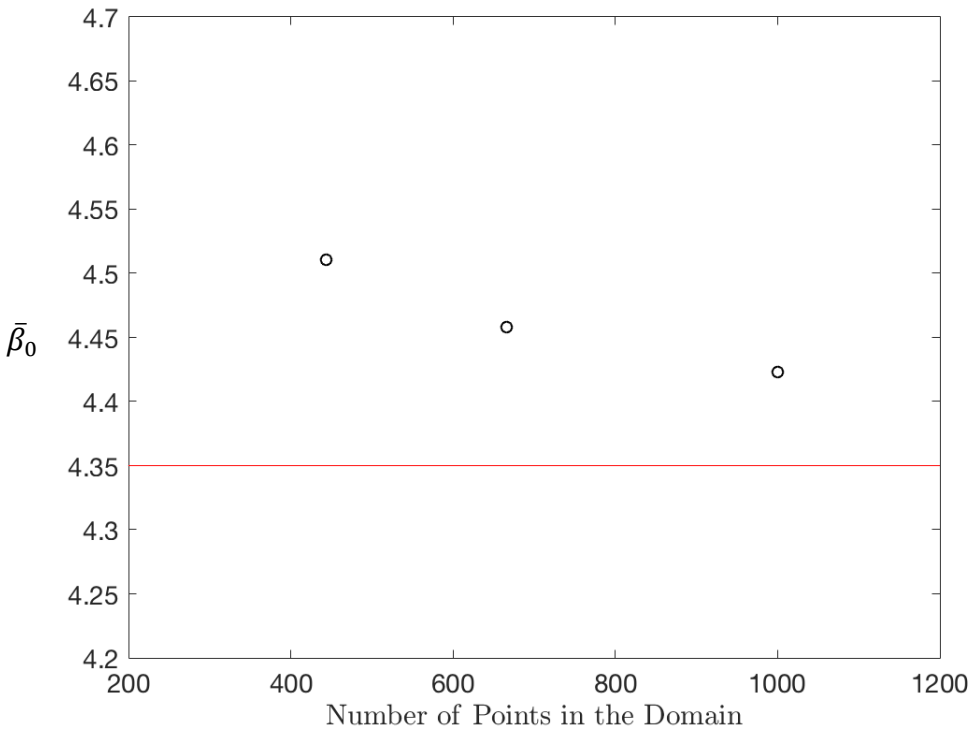

Figure 32: Verification for mild perpendicular anisotropy, $\delta_0 = 0.1$

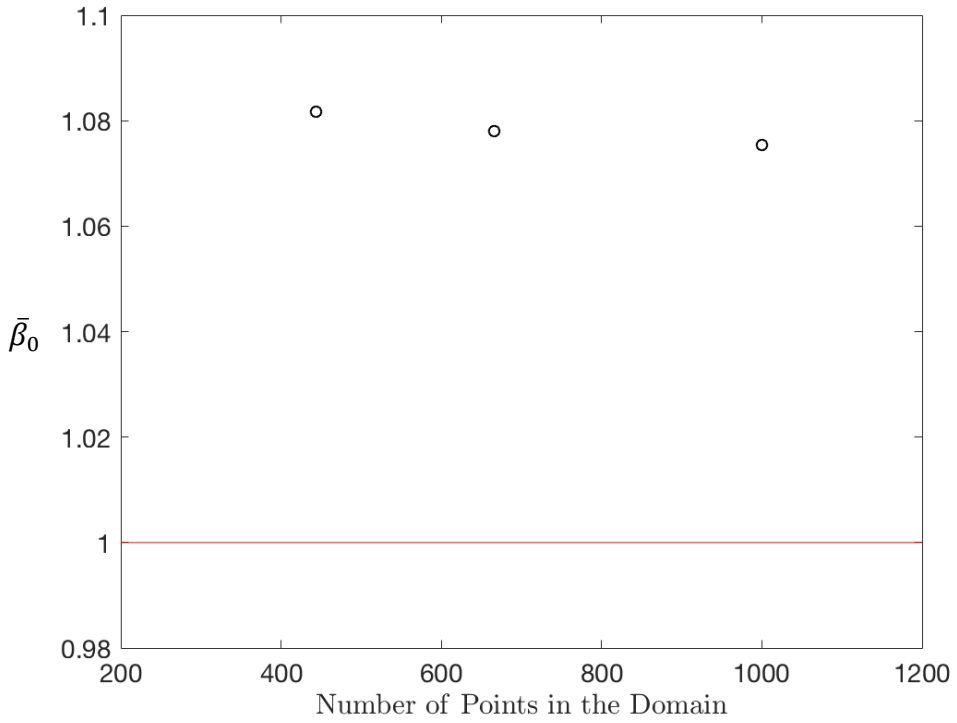

Figure 33: Verification for strong parallel anisotropy, $\delta_0 = -1.0$

Appendix B: Supplementary Material for Chapter 5

Appendix B1: Shell Script Used to Implement Tracing Functions

The following shell script is used to compile and run MATLAB code on Compute Canada servers:

```
#!/bin/bash
#SBATCH --time=0-04:00:00
#SBATCH --ntasks=48
#SBATCH --mem-per-cpu=4000M
#SBATCH --account=def-mann
#SBATCH --mail-user=oberhage@ualberta.ca
#SBATCH --mail-type=BEGIN
#SBATCH --mail-type=FAIL
#SBATCH --mail-type=END
#
# compiling matlab functions
mcc -m -R -nodisplay lossconeanglenight.m
mcc -m -R -nodisplay f_single_particle_nighttodaystep.m
mcc -m -R -nodisplay f_single_particle_daytonightstep.m
# loading matlab module for running compiled functions
module purge
module load mcr/R2018a
# adding compiled code to path
setrpaths.sh --path lossconeanglenight
setrpaths.sh --path f_single_particle_nighttodaystep
setrpaths.sh --path f_single_particle_daytonightstep
#
# for loop that traces particles from 3 field lines
for x_i in -6 -6.2 -6.4
do
# running first compiled function with x_i as an input
run_mcr_binary.sh lossconeanglenight $x_i
alphaloss=$(cat "alpha_loss_xi""$x_i"".txt")
alpha=1.5708
while (( $(echo "$alpha >= $alphaloss" |bc -l) ));
do
( run_mcr_binary.sh f_single_particle_nighttodaystep $x_i $alpha ;
run_mcr_binary.sh f_single_particle_daytonightstep $x_i $alpha ) &
alpha=$(echo "scale=8; $alpha - 0.1" |bc)
# limiting number of jobs to number of cores
[ $( jobs | wc -l) -ge 48 ] && wait[1]
sleep 15s
done
done
wait
```

[1] This line of code is required to ensure that the number of cores made available on Compute Canada servers is not exceeded, it was found from: https://stackoverflow.com/questions/6593531/running-a-limited-number-of-child-processes-in-parallel-in-bash

Appendix B2: Validation

To validate the code used to trace particles through a changing magnetic field, a particle distribution is traced from an analytically defined dipole on the dayside to a night side field in which all tail-like terms are removed, resulting in a numerical dipole. The change in radial distance from Earth of the equatorial crossing point for the different pitch angle particles on the final field line is shown in Figure 34. Because the dayside and night side are both dipolar, the expected change should be zero and any difference corresponds to error. The error shown here is within the 0.005 R_{E} tolerance used for the root finder. Therefore, the code is accurately tracing the particles.

The change in pitch angle relative to the initial pitch angle is shown for every particle in the distribution in Figure 35. The particles with $\pi/2$ pitch angles have no error as their pitch angle is analytically defined as remaining constant through conservation of the second adiabatic invariant. For the remaining particles, the error is less than 1% and increases with higher pitch angle. This may be due to the fact that the final pitch angle is computed using an arcsin function, $\alpha_f = \sin^{-1}\sqrt{\frac{B}{B_m}}$, which amplifies the same error in the magnetic field strength more when the field strength is close to the magnitude of the magnetic field at the mirror point.

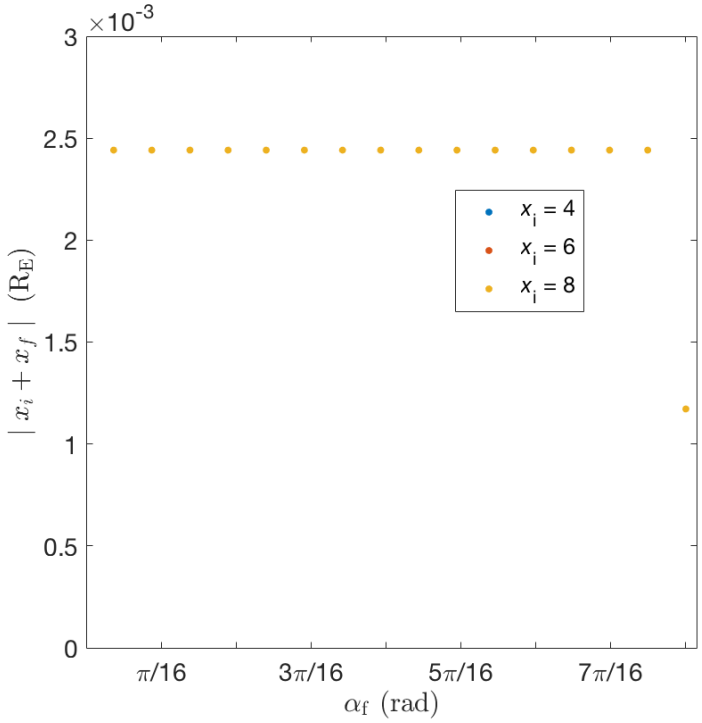

Figure 34: Change in equatorial position as a function of pitch angle (results for the three starting field lines overlap)

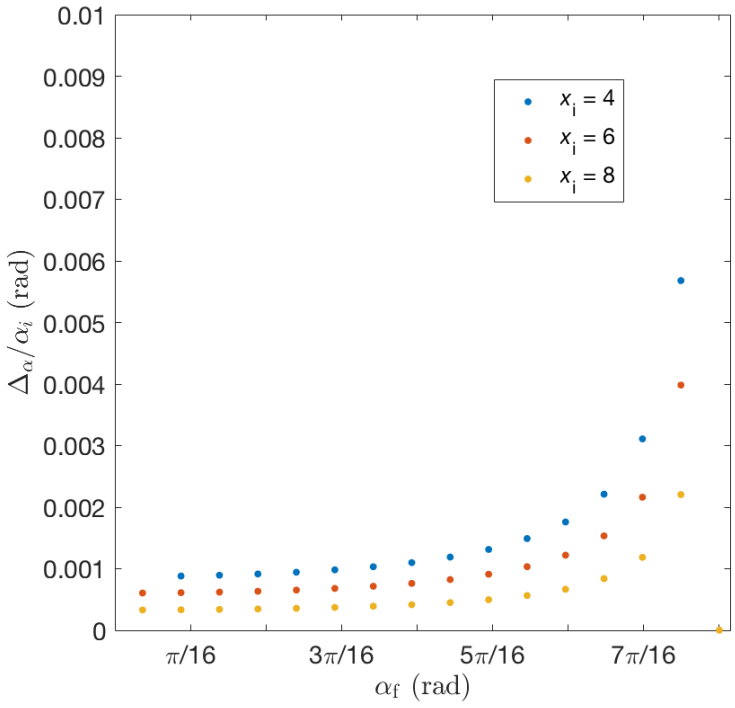

Figure 35: Relative change in pitch angle for each particle in the distribution

Appendix B3: Verification

To test the grid independence of the solution, particles are traced from their final position in the most stretched, $R_t = 6\ R_E$, configuration to the least stretched, $R_t = 6.75\ R_E$, configuration. This captures the widest range of tail configurations possible and it is assumed that the verification therefore holds for intermediate tail states. Particles with final pitch angles of 0.078 radians are used as these have the longest bounce paths and therefore reflect the behaviour of a large portion of the model. The final equatorial location used is $x_f = -6.2\ R_E$.

One factor that complicates the progressive coarsening of the mesh is that there are three different step sizes, one for drawing the field lines, one for calculating the field strength, and one for integrating along a field line, the last of which is adaptively refined. For the purpose of this analysis, the step size for drawing the field lines is progressively coarsened as the other two are derived from it. However, the adaptive mesh refinement scheme results in the simple coarsening factor of 1.5 adopted in Figure 36 not directly applying to the integration itself. Rather, it applies to the values on which the integration is based.

The results of this progressive coarsening are shown in Figure 36. The error increases with increasing step size, however the manner in which it does so is not asymptotic. This can be attributed to the adaptive mesh refinement scheme. The difference in the results between the coarsest and finest refinement is within 1%.

In addition to this, an additional verification is conducted by using a fixed step size for the integration along a field line. The value used is 4000 times the step size for drawing the field lines. Coarsening the field line step size as before, this gives the convergence behaviour shown in Figure 37. The asymptotic behaviour of this convergence shows that the minimum step size used for the adaptive mesh refinement scheme of 40 times the field line step size is sufficiently small.

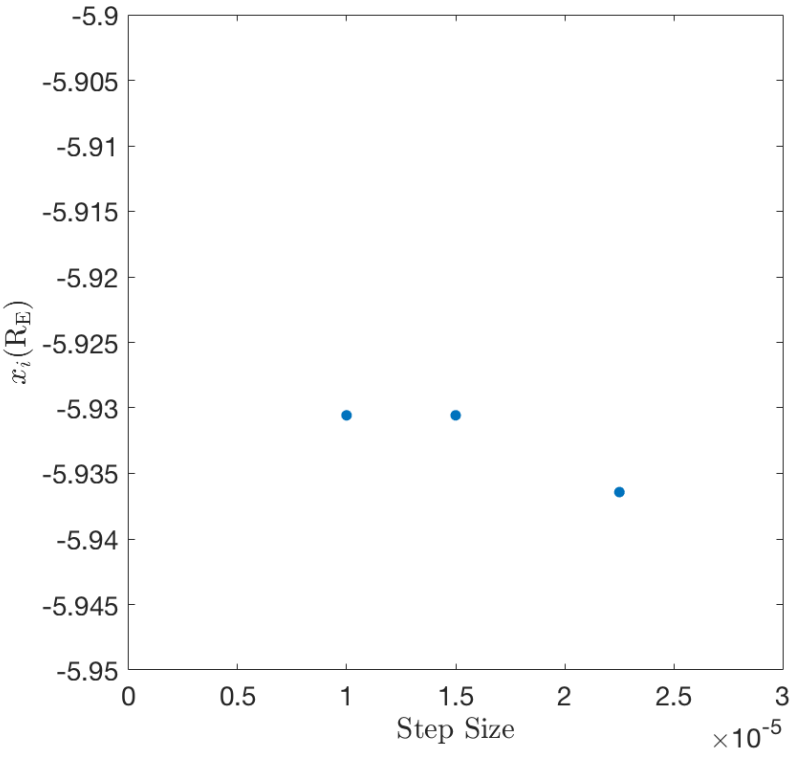

Figure 36: Initial position of near loss cone particles as a function of field line step size when using an adaptive integration scheme

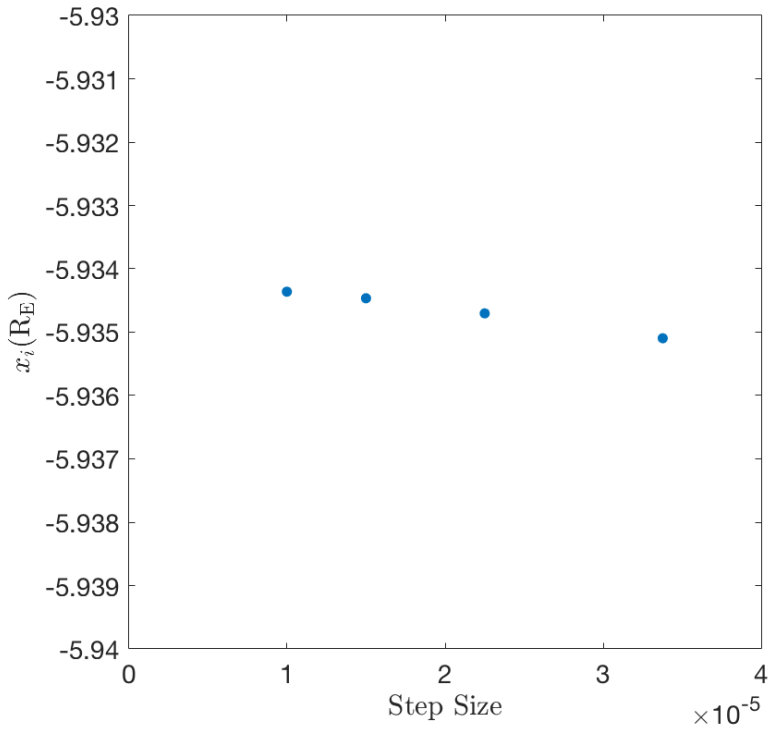

Figure 37: Initial position of near loss cone particles as a function of field line step size when using a fixed step size integration scheme

Appendix B4: Examination of the results in Kabin et al. (2011)

When applying the Kabin et al. (2011) model, a few minor issues were found in their paper that were corrected for the implementation presented in this thesis. Namely, a factor of -1 was required to ensure that the magnetic field vectors point into the north rather than the south pole, and the equatorial magnetic field strength at Earth was missing as a multiple in determining ψ. Both of these were corrected in Kabin et al. (2017). After this, several comparisons were made to the plots in Kabin et al. (2011) to check that the model had been implemented correctly. While good agreement was found in plots of the field lines and the equatorial magnetic field strength, there were significant differences when calculating the current in the magnetotail. As this is related to the pressure in the magnetic field model, an important parameter for our work, several approaches were attempted to ensure that there were no mistakes in the Kabin model's implementation. These were conducted in addition to reviewing the code and adjusting step sizes and are enumerated below:

1. Using the Grad-Shafranov equation to find the pressure gradient

This method is described in the appendix in Kabin et al. (2011). The Grad-Shafranov equation in an axisymmetric case is given by:

$$\nabla \cdot \left(\frac{\nabla \psi}{r^2 \sin^2 \theta} \right) = -\mu_0 \frac{\partial p}{\partial \psi} \tag{B1}$$

The current can be easily found from this as it is given by (Kabin et al., 2011):

$$\mu_0 \mathbf{J} = -\mathbf{e}_\phi r \sin \theta \left(\nabla \cdot \left(\frac{\nabla \psi}{r^2 \sin^2 \theta} \right) \right) \tag{B2}$$

This method resulted in a plot of the current at the equator that did not agree with the one shown in figures in Kabin et al. (2011). Therefore, other methods were attempted, with the pressure gradient in the other methods used to compare the results of these methods to the results of the Grad-Shafranov equation. For this purpose, a change of variables was used to convert $\frac{\partial p}{\partial \psi}$ to $\frac{\partial p}{\partial x}$ as both represent the gradient in the equatorial plane, albeit in different coordinate systems.

2. Determining the pressure gradient through calculation of the magnetic pressure and tension

The magnetic pressure and tension, defined in Chapter 2, are calculated numerically and then an equilibrium force balance is used to find the pressure gradient. The magnetic pressure is simply a function of the magnetic field strength and the magnetic tension requires the calculation of a field geometric second derivative to determine the field line curvature. This was calculated both in cartesian GSM coordinates and based on an analytical first derivative along the field line. The results of this calculation were compared with method 1 and found to be the same.

3. Normalizing the field strengths using a constant factor

A constant factor of 3100, representing the magnetic field strength at the equator on Earth, is removed from the magnetic field strength and the calculation is repeated with the same results as the previous methods.

4. Normalizing the field strengths using a normalization based on dipole values

This method is described in greater detail in the subsequent section of this appendix. The results agree with the previous methods.

5. Calculating the pressure directly using the adaptive mesh refinement scheme

A fifth order Runge-Kutta scheme with adaptive mesh refinement is used to calculate the pressure, from which the pressure gradient is then calculated. The integration is performed by starting near Earth and using zero pressure as the initial condition. The results agree with the previous methods.

6. Calculating the pressure directly using a shooting method

Because the pressure evanescently goes to 0 at Earth, an integration is attempted starting at a distance of 20 R_E. Because the pressure is not known here but is known near Earth, where it is zero, a shooting method algorithm is used to iterate until this final condition is met. The results agree with the previous methods.

7. Performing a check to ensure that the magnetic field has zero divergence

The divergence of the magnetic field is calculated along the equator to ensure that the divergence-free condition is met. It is found that the divergence does not exceed 30 nT/R_E, a maximum which occurs at approximately $x = -3\ R_E$ where the magnetic field strength is over 1000 nT. In addition, the divergence oscillates with a very small mean value of -1 nT/R_E over the interval between $x = -3\ R_E$ and x$= -8\ R_E$, indicating that any nonzero divergence values are due to random numerical error. Therefore, it is concluded that the field is divergence free as required.

8. Calculating the current directly through the curl of the magnetic field

Using the curl of the magnetic field to find the current directly results in a current that does not match that shown by Kabin et al. (2011). When using this current to calculate the pressure gradient, the results are found to be the same as the other methods.

These various methods gave the same results, which were different from those in Kabin et al. (2011). Because the magnetic field lines and magnetic field strength agree with the one shown in Kabin et al. (2011), and they are described by relatively simple functions, the issue is unlikely to lie in a mistake in this calculation. The author of this paper was contacted, who acknowledged that there may be a mistake in the plot of the current.

Based on the multiple tests we are confident in our results.

Appendix B5: Magnetic field normalization using dipole values

As described in Chapter 2, the equilibrium force balance in isotropic MHD is given by:

$$\nabla p = -\frac{1}{2}\nabla\left(\frac{B^2}{\mu_0}\right) + \frac{B^2}{\mu_0 R_c} \tag{B3}$$

In the equatorial plane in the Kabin et al. (2011) model, using Cartesian GSM coordinates this reduces to:

$$\frac{\partial p}{\partial x} = -\frac{1}{2}\frac{\partial}{\partial x}\left(\frac{B^2}{\mu_0}\right) + \frac{B^2}{\mu_0 R_c} \tag{B4}$$

Where:

$$\mathrm{R_c} = \left|\frac{\left(1 + \left(\frac{dx}{dz}\right)^2\right)^{\frac{3}{2}}}{-\frac{d^2x}{dz^2}}\right| = \left|\frac{1}{\frac{d^2x}{dz^2}}\right| \tag{B5}$$

Because $\frac{dx}{dz} = 0$ in the equatorial plane due to symmetry. Therefore, Equation B4 can be written as:

$$\mu_0 \frac{\partial p}{\partial x} = -\frac{1}{2}\frac{\partial}{\partial x}(B^2) + B^2\left|\frac{d^2x}{dz^2}\right| \tag{B6}$$

This equation involves changes due to pressure happening on top of a vacuum dipole magnetic field, the field strength of which varies with $1/x^3$ in the equatorial plane. Therefore, the rate of change for the derivatives being calculated can vary significantly. To avoid this, a normalization may be undertaken using the variables:

$$\delta_B = \frac{B}{B_V} \tag{B7}$$

And

$$\Delta_x = x - x_V \tag{B8}$$

Where the subscript V denotes values calculated for a vacuum dipole field. The difference in shape of the function using normalized variables is shown in Figure 38 and Figure 39.

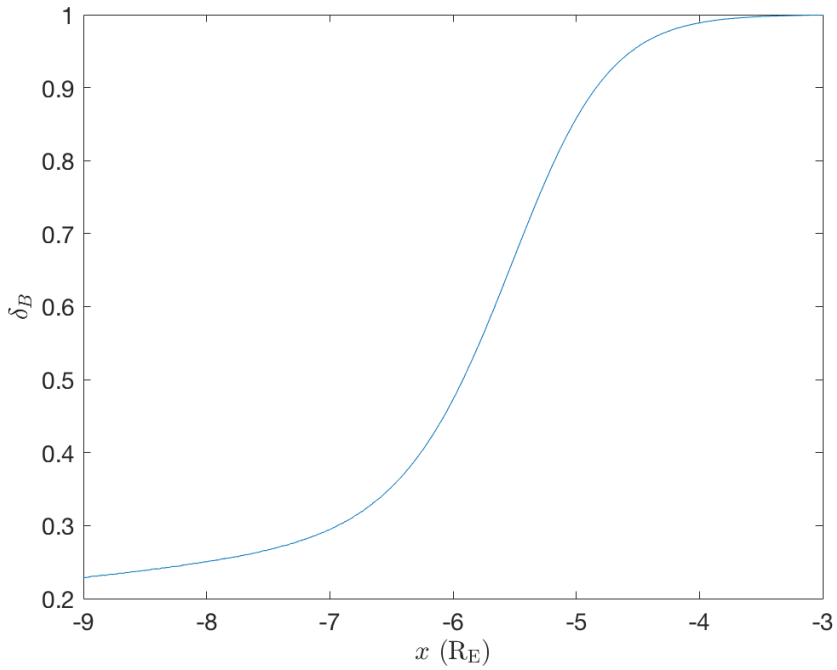

Figure 38: B normalized by the vacuum dipole value

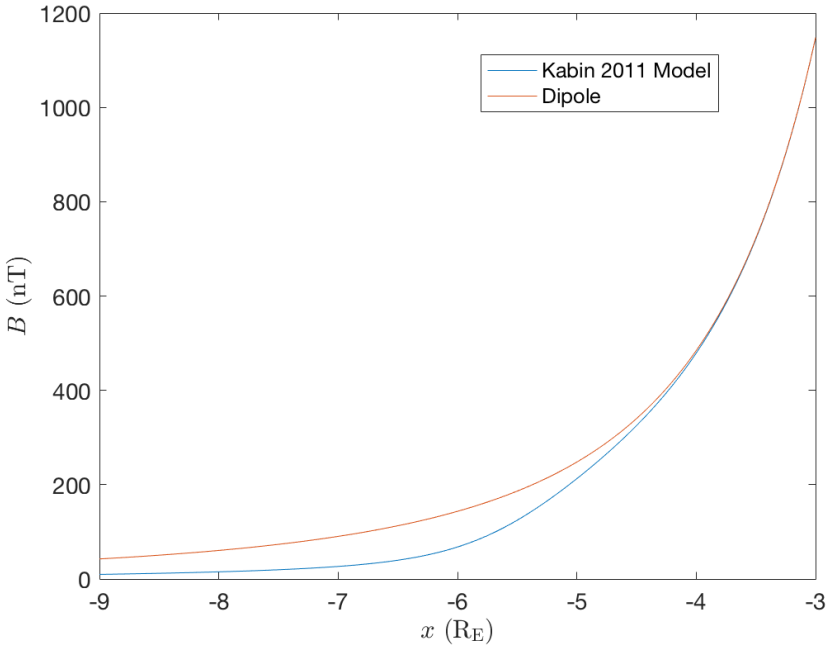

Figure 39: B in the Kabin et al. (2011) *model for the most stretched case shown with the dipole magnetic field strength*

Using these variables, the force balance can be rewritten as:

$$\mu_0 \frac{\partial p}{\partial x} = -\frac{1}{2}\frac{\partial}{\partial x}[(\delta_B B_V)^2] + (\delta_B B_V)^2 \left| \frac{d^2\Delta_x}{dz^2} + \frac{d^2 x_V}{dz^2} \right| \tag{B9}$$

Which can be rearranged to form:

$$\mu_0 \frac{\partial p}{\partial x} = B_V^2 \left(-\frac{1}{2}\frac{\partial \delta_B^2}{\partial x} + \delta_B^2 \frac{d^2\Delta_x}{dz^2} \right) + \delta_B^2 \left(-\frac{1}{2}\frac{\partial B_V^2}{\partial x} + B_V^2 \frac{d^2 x_V}{dz^2} \right) \tag{B10}$$

The second term in this equation is equal to the pressure gradient in a vacuum dipole field, and therefore equal to zero, resulting in:

$$\mu_0 \frac{\partial p}{\partial x} = B_V^2 \left(-\frac{1}{2}\frac{\partial \delta_B^2}{\partial x} + \delta_B^2 \frac{d^2\Delta_x}{dz^2} \right) \tag{B11}$$

To calculate the pressure in the Kabin et al. (2011) model, the pressure gradient may be integrated, or this equation can be integrated by parts, resulting in:

$$\frac{\mu_0}{B_E^2}p = \frac{1}{x^6}\left(-\frac{1}{2}\left(\delta_B(x)\right)^2 + \int_{-3}^{x}\delta_B^2\frac{d^2\Delta_x}{dz^2}dx'\right) + \frac{-1}{3^5}\int_{-3}^{x}\frac{1}{(x')^7}\left(\frac{1}{2}\delta_B^2 - \int_{-3}^{x}\delta_B^2\frac{d^2\Delta_x}{dz^2}dx'\right)dx' \tag{B12}$$

Where integration starts at $x = -3$ because the magnetic field is dipolar at this location, and the pressure is therefore zero.