RICE UNIVERSITY

Hamiltonian Theory and Stochastic Simulation Methods for Radiation Belt Dynamics

by

Xin Tao

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE

Anthony A. Chan, Professor, Chair
Physics and Astronomy

Frank Toffoletto, Associate Professor
Physics and Astronomy

Gerald R. Dickens, Professor
Earth Science

HOUSTON, TEXAS
MAY 2009
UMI Number: 3362419

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI®

UMI Microform 3362419
Copyright 2009 by ProQuest LLC
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346
This thesis describes theoretical studies of adiabatic motion of relativistic charged particles in the radiation belts and numerical modeling of multi-dimensional diffusion due to interactions between electrons and plasma waves.

A general Hamiltonian theory for the adiabatic motion of relativistic charged particles confined by slowly-varying background electromagnetic fields is presented based on a unified Lie-transform perturbation analysis in extended phase space (which includes energy and time as independent coordinates) for all three adiabatic invariants. First, the guiding-center equations of motion for a relativistic particle are derived from the particle Lagrangian. Covariant aspects of the resulting relativistic guiding-center equations of motion are discussed and contrasted with previous works. Next, the second and third invariants for the bounce motion and drift motion, respectively, are obtained by successively removing the bounce phase and the drift phase from the guiding-center Lagrangian. First-order corrections to the second and third adiabatic invariants for a relativistic particle are derived. These results simplify and generalize previous works to all three adiabatic motions of relativistic magnetically-trapped particles.
Interactions with small amplitude plasma waves are described using quasi-linear diffusion theory, and we note that in previous work numerical problems arise when solving the resulting multi-dimensional diffusion equations using standard finite difference methods. In this thesis we introduce two new methods based on stochastic differential equation theory to solve multi-dimensional radiation belt diffusion equations. We use our new codes to assess the importance of cross diffusion, which is often ignored in previous work, and effects of ignoring oblique waves, which are omitted in the parallel-propagation approximation of calculating diffusion coefficients. Using established wave models we show that ignoring cross diffusion or oblique waves may produce large errors at high energies. Results of this work are useful for understanding radiation belt dynamics, which is crucial for predictability of radiation in space.
Acknowledgements

First and foremost, I would like to thank my thesis advisor Professor Anthony Chan for his guidance and inspiration on work leading to this thesis. I am indebted to him for his essential contributions from physical ideas to research methods. I have also benefited a lot from his patience and the freedom he gave me when choosing research subjects that were of interest to me.

I thank Dr. Alain Brizard for his help with the Hamiltonian theory of adiabatic motions. He has impressed me with his excellent mathematical skill. I greatly enjoyed collaborations with Dr. Jay Albert on the modeling of radiation belt dynamics, his great contribution to the project by providing many helpful discussions, comments and suggestions on both numerical techniques and theoretical analysis, and, finally, his frankness and sense of humor. I am grateful to Professor Frank Toffoletto and Professor Gerald Dickens for their serving as my thesis committee members.

My appreciation goes out to the space physics scientists in ISR-1 division in Los Alamos National Laboratory. Specifically, I would like to thank Dr. Josef Koller for offering me the opportunity to work there and for teaching me data assimilation methods. I also
thank Dr. Yue Chen for helpful discussions on data processing and magnetic field models. I enjoyed talks with Dr. Reiner Friedel, who gave lots of helpful advice from running to choosing a career, and for his encouraging me to talk with some great people that I have benefited from.

I feel fortunate to have met lots of great faculty members and fellow students in the space physics program and many friends during my five years at Rice. I cannot name all of them here due to the limits of this section, but still I would like to express my gratitude to Robert Spiro and Stanislav Sazykin for their support to my organization of the presentation practice group. I would like to thank Asher Pembroke for polishing this acknowledgment (and for his organization of great game nights) and Daniel Stark for being my first teacher on scientific writing when he was my English tutor.

Finally, I am grateful to my family for their long term support, encouragement and always being there for me.

Xin Tao

Feb 4, 2009
## Contents

Abstract ii

Acknowledgements iv

List of Figures xi

List of Tables xii

1 Introduction 1

1.1 Background ................................................................. 2

1.1.1 Magnetosphere and radiation belt dynamics ....................... 2

1.1.2 Stochastic differential equations ................................. 7

1.2 Review of previous work .................................................. 9

1.2.1 Theory of adiabatic motion ......................................... 9

1.2.2 Important plasma waves in the magnetosphere .................. 12

1.2.3 Modeling of radiation belt dynamics using quasi linear theory . 14

1.2.4 Using stochastic theory to solve a diffusion equation .......... 18
1.2.5 Other related work ................................................. 19
1.3 Thesis organization ................................................. 19

2 Hamiltonian theory of adiabatic motion of relativistic charged particles 21
2.1 Introduction ......................................................... 21
2.2 Relativistic guiding-center dynamics ............................. 25
  2.2.1 Background-field orderings ................................... 25
  2.2.2 Preliminary coordinate transformation ....................... 26
  2.2.3 Guiding-center Lagrangian for a relativistic particle ....... 27
  2.2.4 Guiding-center equations of motion ......................... 28
2.3 Hamiltonian theory of bounce-center dynamics ................. 32
  2.3.1 Preliminary coordinate transformation ....................... 32
  2.3.2 Coordinate Transformation from \((p_\parallel, s)\) to \((J_b, \psi_b)\) ........... 34
  2.3.3 Lie transform in extended phase-space coordinates .......... 36
  2.3.4 Bounce-center motion in coordinates \((Y, J_b, \psi_b; t, K_b)\) .......... 37
2.4 Hamiltonian theory of drift-center dynamics .................. 41
  2.4.1 Preliminary coordinate transformation ....................... 43
  2.4.2 Lie Transform from \((J_{d0}, \psi_{d0}; t, K_b)\) to \((J_d, \psi_d; t, K_d)\) ........ 45
2.5 Summary and discussion ........................................... 47

3 Stochastic modeling of multi-dimensional diffusion in the radiation belts 50
3.1 Introduction ......................................................... 50
3.2 The SDE method .............................................. 53
  3.2.1 Itô stochastic differential equations ................. 53
  3.2.2 Probabilistic representation of solutions of diffusion equations ... 54
3.3 Application .................................................. 57
  3.3.1 Application to pitch-angle and energy diffusion equations .... 58
  3.3.2 Comparisons with Albert and Young [2005] results ........ 61
  3.3.3 Effects of parallel propagation approximation .............. 65
3.4 Summary and discussion ................................... 71

4 Modeling of multi-dimensional diffusion in radiation belts using layer methods 75
  4.1 Introduction ................................................ 75
  4.2 The Milstein layer methods ................................ 77
    4.2.1 One-step representation of solutions using the SDE method .... 77
    4.2.2 A simple layer method algorithm ..................... 78
  4.3 Application ................................................ 80
    4.3.1 Comparison with Albert and Young [2005] results .......... 82
    4.3.2 Effects of ignoring cross diffusion in Li et al. [2007] chorus wave model ........................................... 83
    4.3.3 Evolution of electron fluxes using a model of fast magnetosonic waves and hiss ......................... 90
  4.4 Summary ................................................... 92
5 Summary and future work

5.1 Summary ................................................................. 96

5.2 Discussion and future work ......................................... 98

A Time-forward SDE method ........................................... 102

B One-step error of the layer method ................................. 105

C Relationship between finite difference methods and layer methods ................................. 107
List of Figures

1.1 Earth's magnetosphere ............................................. 3
1.2 The Van Allen radiation belts ...................................... 4
1.3 Three adiabatic motions ............................................. 5
1.4 Change of electron fluxes during a storm ......................... 8
1.5 Sample trajectories of two stochastic differential equations (SDEs) ............................................. 10
1.6 A MLT distribution of plasma waves in the magnetosphere ............ 15

3.1 Diffusion coefficients (D) repeated from Albert and Young [2005] ........ 62
3.2 Comparisons between the SDE method results and previous results ........ 64
3.3 Effects of cross diffusion using the Horne et al. [2005] chorus wave model . 66
3.4 D of the Horne et al. [2005] wave model with parallel propagation .......... 67
3.5 Comparisons between fluxes obtained from $D^\parallel_d$ and $D_d$ ............... 69
3.6 Comparisons between fluxes obtained from $D^\parallel_d$ and $D$ ................. 70
3.7 Comparisons between fluxes obtained from $D^\parallel$ and $D$ .................. 72

4.1 Convergence of solutions of layer methods with $N_{\alpha_0}$ and $N_y$ .............. 84
4.2 Comparisons between the layer method results and previous results ........ 85
4.3 D of the Li et al. [2007] chorus wave model with oblique waves .......... 87
4.4 Flux evolution using the Li et al. [2007] chorus wave model .............. 88
4.5 Effects of cross diffusion using the Li et al. [2007] chorus wave model ... 89
4.6 D of the combined magnetosonic wave and hiss (MH) wave model ...... 92
4.7 Flux evolution using the MH wave model .................................... 93
4.8 Effects of cross diffusion using the MH wave model ...................... 95
A.1 Local effects of ignoring off-diagonal terms ................................. 104
List of Tables

3.1 Parameters of the Horne et al. [2005] wave model . . . . . . . . . . . . . . . . . . . . . . . . . 63

4.1 Parameters of the Horne et al. [2005] and Li et al. [2007] wave models . . 82
Chapter 1

Introduction

This dissertation describes theoretical research on adiabatic relativistic charged particle motion and numerical modeling of relativistic electron dynamics in the Earth’s radiation belts. The overall purpose is to improve understanding of physical processes responsible for changes of electron fluxes in the radiation belts. Specifically, the purpose of this thesis is to strengthen the foundation of theory and to develop numerical codes to model changes of electron fluxes due to interactions with plasma waves. The increase of energetic electron fluxes are potentially hazardous to satellites and astronauts in space [see Baker et al., 1994], and the precipitation of electrons into the atmosphere can cause changes of chemistry of the atmosphere and ozone destruction [e.g. Thorne, 1977].

The main part of the thesis is divided into three chapters. In Chapter 2, we show a Hamiltonian theory of three relativistic adiabatic motions in slowly varying electromagnetic fields. The theory of adiabatic motion is the foundation of understanding radiation
belt dynamics. In Chapter 3 we show the stochastic differential equation (SDE) method and Chapter 4 the layer method of modeling radiation belt dynamics using multidimensional diffusion theory. While each method has its own advantages and disadvantages, both methods can be used to solve multi-dimensional diffusion equations with cross diffusion terms. The cross diffusion terms complicate the numerical problem and were ignored by most of the previous work, thus in both chapters we explored their importance in radiation belt modeling using different plasma wave models.

1.1 Background

In this section, we will first briefly describe the dynamics of Earth’s magnetosphere and radiation belts. Then we will introduce stochastic differential equations, which are basis of the two numerical codes that we will develop later in the thesis.

1.1.1 Magnetosphere and radiation belt dynamics

The Earth’s magnetosphere is the region of space that is dominated by the Earth’s magnetic field, whose main component can be described as a dipole field [Kivelson and Russell, 1995]. Because of the interaction with the solar wind, which is a magnetized plasma that flows supersonically from the Sun, the magnetosphere is compressed in the dayside and stretched in the tail side, as sketched in Figure 1.1. The magnetosphere can be very dynamic due to changes in solar wind conditions. A characteristic dynamic process is the geomagnetic storm, which is a temporary large scale disturbance of the magnetosphere.
Figure 1.1: Schematic illustration of the Earth's magnetosphere with Van Allen belts explicitly shown. The figure was adapted from http://www.astro.uiuc.edu/-kaler/emag-r.jpg on Jan 05, 2009
A geomagnetic storm is usually indicated by a duration of negative values of the $Dst$ index. This index measures a southward perturbation of Earth's near-equatorial magnetic field caused by the ring current from charged particles drifting around the Earth. During a geomagnetic storm, the particle fluxes in the Van Allen radiation belts can also become very dynamic.

The Van Allen radiation belts (shown in Figure 1.2) were first discovered by James Van
Allen and his colleagues in the late 1950s [Van Allen and Frank, 1959]. The electron radiation belts are two toroidal regions in space containing energetic ($\gtrsim 0.5$ MeV) electrons separated by a slot region. The inner belt results primarily from collisions between upper atmospheric gas atoms and solar protons or galactic cosmic-ray particles and is relatively stable [Schulz and Lanzerotti, 1974]. The outer belt is the area of interest in this thesis, partly because it is very dynamic and partly because of the damaging effects of outer radiation belt electrons on humans and technology in space. The outer radiation belt contains primarily energetic electrons, whose sources are still under intensive research.

Charged particles trapped in the radiation belts display three quasi-periodic motions (shown in Figure 1.3) with largely separated periods. Each quasi-periodic motion corresponds to an adiabatic invariant. This can be readily understood using action-angle vari-
ables in Hamiltonian dynamics [Schulz and Lanzerotti, 1974]. The fastest motion is a gyromotion around a single magnetic field line, and the corresponding adiabatic invariant is defined as $\mu = p_\perp^2 / 2mB$. Here $p_\perp$ is the momentum perpendicular to the local magnetic field $B$, and $m$ is the particle mass. The second adiabatic motion is the north-south bounce motion between two mirror points, where the particle changes its direction of motion along the field line. The second adiabatic invariant is $J = \oint p_\parallel ds$, where $p_\parallel$ is the momentum parallel to $B$ and the integral is along the bounce trajectory. The third adiabatic invariant is due to the drift motion around the Earth, and is represented by the magnetic flux enclosed by a particle’s drift path $\Phi = \oint_S B \cdot dS$. Another useful quantity related to the drift motion is the Roederer $L$-shell: $L \equiv -2\pi k_0 / \Phi R_E$, where $k_0$ is the magnetic dipole moment and $R_E$ the radius of the Earth [Roederer, 1970]. Thus if $\Phi$ is an adiabatic invariant, so is $L$. In a dipole magnetic field, $L = r / R_E$, where $r$ is a particle’s radial distance to the center of the Earth.

Electromagnetic fields that vary on a time scale comparable to one of the three periods can violate the corresponding adiabatic invariant. A possible result of this is stochastic changes of adiabatic invariants and thus the diffusion of electrons in phase space. It is customary to discuss diffusion using pitch angle, which is the angle between the particle’s momentum and magnetic field, its energy, and $L$. Stochastic changes of particle’s pitch angle could cause losses of particles into the atmosphere, thus pitch angle diffusion is generally considered as a loss mechanism [Schulz and Lanzerotti, 1974]. Because phase space density usually has a negative gradient in energy, energy diffusion, on the other
hand, is considered as an important acceleration mechanism. Violating the third adiabatic invariant causes changes of a particle’s $L$. If the first and second invariants are conserved, the particle’s energy increases if it moves inward to the Earth [see Schulz and Lanzerrotti, 1974, Chapter III.1].

The quiet time radiation belt structure, including the slot region between the two belts, has been explained by Lyons and Thorne [1973] as an equilibrium between loss by pitch angle scattering and inward radial diffusion. Changes of electron fluxes during a magnetic storm are shown in the upper panel of Figure 1.4. We see that radiation belt electron fluxes varied by several orders of magnitude during this storm. However, physical mechanisms for acceleration and loss of energetic electrons are not well understood and are under research. Reeves et al. [2003] analyzed 276 magnetic storms from 1989 through 2000, and concluded that the dynamics of radiation belts are a complicated balance between electron loss and acceleration. Possible mechanisms for loss and acceleration of electrons in the radiation belts have been reviewed in Li and Temerin [2001]; Friedel et al. [2002]; Shprits et al. [2008a] and Shprits et al. [2008b].

1.1.2 Stochastic differential equations

A stochastic differential equation (SDE) is used to describe a stochastic process, which is utilized in this thesis to solve diffusion equations. In contrast to ordinary differential equations (ODEs), which are used to describe deterministic processes, SDEs contain stochastic
Figure 1.4: Adopted from *Kim and Chan* [1997]: (a) Hourly averaged electron flux variation at geosynchronous orbit \((r \approx 6.6 \, R_E)\) for November 2-8, 1993, measured by the CPA instrument on the LANL spacecraft 1984-129 (LT = UT + 0.5) for four high energy channels and (b) *Dst* variation for the same period.
terms. A simple one dimensional SDE can be written as

$$dX = b(t, x)dt + \sigma(t, x)dW,$$  \hspace{1cm} (1.1)

where we use capitalized characters ($X$) to denote stochastic process and the corresponding lower case characters ($x$) to denote the value of the stochastic process. Here $b(t, x)$ is the advection coefficient (also called the "drift coefficient") and $\sigma$ the diffusion coefficient of the SDE. The first term on the right hand side is the part we also see in ODEs, and the second term contains stochastic motion. Here $dW$ is an increment of a 1D Brownian motion [Gardiner, 1985]. It equals $\sqrt{dt}\mathcal{N}(0, 1)$, where $\mathcal{N}(0, 1)$ is a standard Gaussian random number with zero mean and unit variance. Sample trajectories of stochastic processes $X$ with $b = 0$ (dotted) and 1 (solid) and $\sigma = 1$ are shown in Figure 1.5.

### 1.2 Review of previous work

In this section, I will review previous work that is related to this thesis, including theory of adiabatic motion, wave particle interactions, important plasma waves in the magnetosphere and numerical methods used in previous work to solve diffusion equations.

#### 1.2.1 Theory of adiabatic motion

The guiding-center equations of motion have been used to model the gyro-averaged motion of a charged particle in electromagnetic fields. Because the time step used to resolve guiding-center motion is much larger than the one used to resolve gyromotion, the numer-
Figure 1.5: Sample trajectories of stochastic processes $X(t)$ described by SDE $dX = b dt + dW$ with $b = 1$ (solid) and $b = 0$ (dotted) using the same sequence of pseudo-random numbers $dW$ for comparison and $X(0) = 0$. 
ical efficiency is increased usually by orders of magnitude. Northrop [1963] developed the guiding-center equations of motion using a non-Hamiltonian method. A small parameter $\epsilon \equiv \rho/l$ was used to order Northrop’s guiding-center equations. Here $\rho$ is the particle’s gyro radius, and $l$ is the scale length of background magnetic fields. While this method is easy to understand and straightforward, the resulting equations of motion do not conserve total energy in static fields because higher order terms in $\epsilon$ are ignored. Littlejohn [1981, 1983] developed a new theory of guiding-center motion using a non-canonical Hamiltonian formulation using Lie transform perturbation analysis [Cary and Littlejohn, 1983] for non-relativistic particles. The resulting guiding-center equations conserve total energy in static fields, and also conserve the extended Hamiltonian and phase-space volume in time-varying fields. These conservation properties are useful for checking numerical accuracy. Another advantage of the method is that the Lie transform analysis is systematic and can (in principle) be carried to arbitrary order. Brizard and Chan [2001] extended the work of Littlejohn [1981] on guiding-center motion to relativistic particles in static fields. Further averaging the guiding-center equation over the bounce-center phase angle will give us bounce-center motion. Littlejohn [1982] derived the Hamiltonian theory of bounce-center motion (or called by Littlejohn [1982] the guiding center bounce motion) for non-relativistic particles. In this thesis, we will develop relativistic guiding-center and bounce-center Hamiltonian theory and extend that to include the drift center motion in time-varying electromagnetic fields.
1.2.2 Important plasma waves in the magnetosphere

Violations of adiabatic invariants can give rise to a variety of dynamical effects on the radiation belts. Understanding the different processes responsible for electron acceleration and loss is the main objective of radiation belt research.

One important process is radial diffusion, especially when enhanced by ULF (Ultra Low Frequency, wave frequency \( f < 3 \) Hz) waves through drift resonance. The drift resonance condition in a symmetric magnetic field is \( \omega - m \omega_d = 0 \), where \( \omega \) is the wave angular frequency, \( \omega_d \) the particle drift frequency and \( m \) a positive integer. Radial diffusion has been proposed as an acceleration mechanism of radiation belt electrons during storm times [e.g., Schulz and Lanzerotti, 1974; Elkington et al., 1999; Hilmer et al., 2000]. However, observations also show that during storm times, phase space densities of electrons can peak around \( L \sim 5 \), and this cannot be explained by radial diffusion alone [Brautigam and Albert, 2000; Green and Kivelson, 2004; Chen et al., 2007].

Resonant interactions with ELF (Extremely Low Frequency, \( 3 < f < 3000 \) Hz) and VLF (Very Low Frequency, \( 3 < f < 30 \) kHz) waves have been invoked as an important mechanism for electron local acceleration and precipitation via cyclotron resonances in the radiation belts. Horne and Thorne [1998] explored possible wave modes for electron acceleration and loss via resonant wave-particle interactions using the resonance condition

\[
\omega - k_\parallel v_\parallel = n\Omega_e.
\]  

(1.2)

Here \( \omega \) is the wave frequency, \( \Omega_e \) is the electron relativistic gyrofrequency, and \( k_\parallel \) and \( v_\parallel \) are wave number and velocity parallel to \( B \), respectively. Harmonic number \( n = \)
0, ±1, ±2, · · · with \( n = 0 \) the Landau resonance and \( n \neq 0 \) cyclotron resonance. By calculating the minimum resonant energy needed for interacting with waves under storm-time conditions, Horne and Thorne [1998] concluded that whistler mode waves and highly oblique magnetosonic waves are possible candidates for accelerating electrons to the MeV energy range, while electromagnetic ion cyclotron (EMIC) waves may only resonate with highly relativistic electrons and contribute to loss via pitch-angle scattering. Other wave modes possible for accelerating electrons include LO, RX and Z mode waves [see also, Xiao et al., 2006, 2007], but both more theoretical work and observations are needed to determine their roles in radiation belt electron dynamics.

Plasmaspheric hiss waves and EMIC waves are under intensive research for their important roles in loss of electrons in the radiation belts. Plasmaspheric hiss waves are whistler mode, highly turbulent waves [Thorne et al., 1973]. These extremely low frequency (ELF) hiss waves can cause electron loss into the atmosphere by pitch-angle scattering, which has been included in several models [Abel and Thorne, 1998a, b; Li et al., 2007; Beutier and Boscher, 1995; Bourdarie et al., 1996; Meredith et al., 2007]. EMIC waves can also resonant with relativistic electrons and cause strong pitch-angle scattering [Albert, 2003; Summers and Thorne, 2003; Li et al., 2007; Albert, 2004; Khazanov and Gamayunov, 2007]. The main difference between hiss and EMIC waves shown by Li et al. [2007] is that EMIC waves tend to cause electrons to diffuse into the loss cone while hiss waves tend to scatter electrons from higher pitch angles to lower pitch angles but not necessarily into the loss cone.
Chorus waves are whistler mode waves. Different from hiss and EMIC waves, they have been associated with both loss and energization of radiation belt electrons through pitch-angle diffusion and energy diffusion. Specifically, chorus waves have been related to microburst precipitation of electrons [Lorentzen et al., 2001; O'Brien et al., 2003, 2004; Thorne et al., 2005], but they are also possible candidates for energizing electrons during storms [Horne and Thorne, 1998; Summers et al., 1998; Meredith et al., 2003a, b; Albert and Young, 2005; Shprits et al., 2006b]. A magnetic local time (MLT) distribution of the above plasma waves in the magnetosphere is shown schematically in Figure 1.6.

Recent observations show that fast magnetosonic waves (also known as equatorial noise) might be related to electron acceleration in the radiation belts [Horne et al., 2007]. These waves propagate almost perpendicularly to background magnetic fields and thus interact with electrons mainly through Landau resonance \( (n = 0) \) [Horne et al., 2007; Albert, 2008].

**1.2.3 Modeling of radiation belt dynamics using quasi linear theory**

Quasi linear diffusion theory is often used to model interactions between electrons and waves. The term “quasi linear” is used because the time rate of change of the lowest order distribution function is calculated nonlinearly from first order quantities, which are calculated using a linear theory. Kennel and Engelmann [1966] derived the quasi linear diffusion equations for non-relativistic particles interacting via cyclotron resonances with small amplitude broad band oblique waves. Lyons [1974a] converted the diffusion coefficients of
Figure 1.6: Adopted from Shprits et al. [2006a]: Schematic picture of the MLT distribution of plasma waves able to resonate with energetic electrons near $L = 4.5$ during the main phase of a storm. Light and dark blue arrows show the convective injection trajectories of ring current particles. Ring current electrons excite whistler mode chorus waves (blue). Whistler mode hiss waves (yellow) are confined to the plasmasphere and plume. Interaction of ring current ions in the high-density cold plasmasphere and regions of plumes excite EMIC waves (red). The wavy line denotes ULF waves, and the red arrow shows a circular trajectory of relativistic electrons in the radiation belt.
Kennel and Engelmann [1966] written in $v_{\parallel}$ (velocity parallel to background magnetic field $B$) and $v_{\perp}$ (velocity perpendicular to $B$) to pitch angle ($\alpha$) and velocity ($v$) diffusion coefficients. This work became the basis of later work of modeling of radiation belt dynamics using quasi linear diffusion theory [Shprits et al., 2006b; Lam et al., 2007; Li et al., 2007; Horne et al., 2007]. Typically, quasilinear diffusion in 3D is described by a Fokker-Planck diffusion equation of the form [Schulz and Lanzerotti, 1974; Brizard and Chan, 2004]

$$\frac{\partial f}{\partial t} = \sum_{i,j=1}^{3} \frac{\partial}{\partial J_i} \left( D_{ij} \frac{\partial f}{\partial J_j} \right)$$

(1.3)

when written in three adiabatic invariants ($J_1, J_2, J_3$) [see also Roederer, 1970, Appendix VIII]. Here $D_{ij}$ are diffusion coefficients, and terms $D_{ij}$ with $i \neq j$ are called cross diffusion, or off-diagonal terms (when viewed as a matrix). Including cross diffusion in calculations, however, makes numerical calculations more complicated. Albert [2004] showed that numerical problems arise when solving diffusion equations using simple finite difference methods with cross diffusion terms included; solutions might be negative and thus unphysical.

Albert and Young [2005] gives the first solution of a 2D radiation belt diffusion equation with cross diffusion included. The first step of their method is to write the usual 2D diffusion equation in a new set of coordinates $(Q_1, Q_2)$, which are chosen in a way such that the cross diffusion vanishes in $(Q_1, Q_2)$; i.e., $D_{Q_1 Q_2} = 0$. Then the diffusion equation written in $(Q_1, Q_2)$ is solved by simple finite difference methods. The solution of $f$ in more physical coordinates, e.g., equatorial pitch angle $\alpha_0$ and energy $E$, is obtained by another coordinate transformation from $(Q_1, Q_2)$ to $(\alpha_0, E)$. From now on, we will call this method
the transformation method for simplicity.

Using the transformation method, Albert and Young [2005] solved a 2D diffusion equation using a chorus wave model from Horne et al. [2005] to show acceleration of electrons by chorus waves. Their results show that chorus waves can accelerate electrons to MeV energy on time scales of one day. To evaluate effects of cross diffusion, Albert and Young [2005] then compared fluxes calculated with and without off-diagonal terms. They concluded that for the Horne et al. [2005] chorus wave model, ignoring cross diffusion does not change flux profiles qualitatively, but quantitatively leads to an overestimate of energy diffusion.

The transformation method is a nice method to solve 2D radiation belt diffusion equations; however, it does not generalize very well to 3D with all nine diffusion coefficients in equation (1.3) included [personal communication, Jay M. Albert, 2007]. Also Albert and Young [2005] only considers interactions between chorus waves and electrons, mainly because only a chorus wave model was available at the time. More wave models [Li et al., 2007; Horne et al., 2007] have been proposed since then, and they should be used to calculate effects of cross diffusion with different wave models. Furthermore, we want to verify the results of Albert and Young [2005] with an independent approach and a method is needed to solve general 3D radiation belt diffusion equations. All the above reasons motivate the second and third parts of this thesis; i.e., to solve multidimensional diffusion equations using stochastic differential equations.
1.2.4 Using stochastic theory to solve a diffusion equation

A diffusion equation describes the evolution in time of a particle density function when the underlying motion of the particle is stochastic. Accordingly, one can derive the Fokker-Planck diffusion equation from stochastic differential equations that describe the particle motion [Gardiner, 1985]. Also, from the Fokker-Planck equation we can obtain the corresponding stochastic process, which can then be used to obtain a solution of the diffusion equation.

There have been at least two methods to solve diffusion equations using SDEs. The first method is to convert the diffusion equation to corresponding SDEs, which show stochastic changes of phase space coordinates of particles. Using test particle simulations with these SDEs, we can solve the corresponding diffusion equation [e.g., Alanko-Huotari et al., 2007; Albright et al., 2003; Yamada et al., 1998]. Because this method describes particles moving forward in time, we will call this method the “time forward” SDE method. Secondly, mathematical theory shows that the solution of a diffusion equation can be written as the expectation value of a stochastic process evaluated under specific conditions [e.g., Freidlin, 1985; Costantini et al., 1998; Bossy et al., 2004; Zhang, 1999]. This method is a generalized method of characteristics, which is used to solve advection equations. Because the method samples trajectories of a stochastic process backward in time in the same sense as the method of characteristics, we call this method the “time backward” SDE method. Since we will use the “time backward” SDE method as the main numerical method in this thesis, we will just use the term “SDE method” to refer to the “time backward” SDE method.
Based on the SDE method, Milstein [2002]; Milstein and Tretyakov [2001] and Milstein and Tretyakov [2002] developed layer methods, which are deterministic and more efficient when solutions on a large number of grid points are needed. This is the third numerical method that we will use in this thesis.

1.2.5 Other related work

The diffusion equation approach mentioned above is not the only method to model radiation belt dynamics. Test particle simulation approaches, which solve the guiding-center equations of motion or the full particle equation of motion, have also been widely used [Elkington et al., 1999; Kress et al., 2007; Albert, 2002]. On the other hand, recent observations of large amplitude chorus waves [Cully et al., 2008; Cattell et al., 2008] have raised questions about whether the quasi linear diffusion approach is suitable to describe interactions between electrons and these large amplitude chorus waves. Nonlinear interactions, like phase trapping, have been explored and applied to wave-particle interactions in Earth's magnetosphere by several authors [e.g., Bell, 1984; Albert, 1993, 2000; Bortnik et al., 2008]. However, in this thesis, we will assume that interactions between electrons and chorus waves are described by the quasi linear diffusion theory.

1.3 Thesis organization

This remainder of this thesis is organized as follows. In Chapter 2, we show development of a Hamiltonian theory of three adiabatic motions using the Lie-transform perturbation
method and the guiding-center equations of motion for a relativistic particle in slowly-varying electromagnetic fields. The stochastic differential equation (SDE) method of solving multi-dimensional diffusion equations is described in Chapter 3, with a 2D stochastic code developed to solve a bounce-averaged pitch-angle and energy diffusion equation. In Chapter 4, we show the layer method, which is based on the SDE method but is deterministic, to solve multi-dimensional diffusion equations. Also, as an application of the layer code, we show effects of including cross diffusion on evolution of electron fluxes using a chorus wave model and a combined magnetosonic and hiss wave model. We summarize our results and discuss possible future work in Chapter 5.
Chapter 2

Hamiltonian theory of adiabatic motion
of relativistic charged particles

This chapter has been published in Physics of Plasmas [Tao et al., 2007].

2.1 Introduction

The concept of the adiabatic motion of a charged particle in magnetic fields is important to research in space plasma physics and fusion physics [Northrop, 1963; Brizard and Hahm, 2007]. Depending on the confining magnetic geometry, a particle may display three quasi-periodic or periodic motions. The fastest of these three motions is the gyromotion about a magnetic field line (with frequency \( \omega_g \)). The second motion exists when a particle bounces along a magnetic field line between two mirror points (with frequency \( \omega_b \)), because of nonuniformity along magnetic field lines. The slowest motion is
the drift motion across magnetic field lines (with frequency $\omega_d$) caused by perpendicular magnetic gradient-curvature drifts. In space physics (and especially in radiation-belt physics [Schulz and Lanzerotti, 1974]), these frequencies are widely separated such that $\omega_g : \omega_b : \omega_d \sim \epsilon^{-1} : 1 : \epsilon$, where $\epsilon \ll 1$ is a small dimensionless ordering parameter to be defined below. Associated with each periodic orbital motion, there exists a corresponding adiabatic invariant. We use $\mu$, $J_b$ and $J_d$ for the three invariants constructed in this work; to be consistent, we may also use $J_g = (mc/q)\mu$, where $m$ is the particle’s rest mass and $q$ its charge.

The theory of the adiabatic motion of charged particles in electromagnetic fields has been well developed by Northrop [1963]. However, the non-Hamiltonian method used by Northrop resulted in dynamical equations that do not possess important conservation properties like energy conservation in static fields, because of the absence of higher-order terms from Northrop’s equations. In later work, Littlejohn [1983] used a noncanonical phase-space transformation method, based on Lie-transform perturbation analysis, to obtain the Hamiltonian formulation of guiding-center dynamics for nonrelativistic particles. By asymptotically removing the dependence on the gyrophase, the first invariant $J_g = (mc/q)\mu$ is obtained from the guiding-center Lagrangian by Noether’s theorem.

The resulting Hamiltonian guiding-center equations of motion conserve total energy for motion in static fields. In the present work, we use the Lie-transform perturbation analysis to develop a systematic Hamiltonian theory for relativistic guiding-center motion in weakly time-dependent electromagnetic fields. Our relativistic guiding-center equations of
motion are expressed in semi-covariant form [Boghosian, 1987], which simplifies the previous work by Grebogi and Littlejohn [1984] (who extended their relativistic guiding-center equations to include ponderomotive effects associated with the presence of high-frequency electromagnetic waves) and generalizes earlier work by Brizard and Chan [1999] (who considered guiding-center motion of a relativistic particle in static magnetic fields).

Phase-space Lagrangians are used in our perturbation analysis. Compared with the usual configuration space Lagrangian of $N$ independent variables, the corresponding phase space Lagrangian contains $2N$ independent variables. Physical equations of motion can be obtained by applying the variational principle to a phase space Lagrangian. The use of phase-space Lagrangians is important to the Lie perturbation analysis due to their linearity in the time derivatives [Littlejohn, 1983].

The derivation of relativistic guiding-center dynamics begins with the removal of the gyrophase dependence from the particle phase-space Lagrangian. Since the condition for these periodic motions to exist is that the time variations of the forces a particle experiences should be slow compared to the particle's motion, we assume first that the electromagnetic fields vary on the drift timescale. Thus we shall construct the first and second adiabatic invariants from the particle's motion. While this ordering is not the most general case, it is the most common one in practice [Littlejohn, 1983]. This procedure gives us the reduced six-dimensional guiding-center Lagrangian and the first invariant $J_g$. Based on the guiding-center Lagrangian, we further remove the bounce phase and obtain the bounce-averaged guiding-center (or bounce-center) motion. The bounce-center Lagrangian for nonrelativis-
tic particles has been derived by Littlejohn [1982], who at the same time constructs the second invariant \( J_b \) and the first-order correction to the second adiabatic invariant. The present work generalizes results of Littlejohn [1982] and Brizard [1990, 2000] for relativistic particles. After we obtain the bounce-center Lagrangian, we change the time-scale ordering of the background fields so that the perturbation analysis can be applied to the drift motion. We assume that the background fields vary on a time scale much slower than the drift time period when we derive the drift-center motion. By drift averaging the bounce-center Lagrangian and removing the drift phase, we obtain the drift invariant \( J_d \) and the first-order correction to the third adiabatic invariant.

The remainder of the chapter is organized as follows. In section 2.2, we derive the guiding-center equations of motion and the guiding-center Lagrangian for relativistic particles moving in slowly-varying electromagnetic fields. This section generalizes Littlejohn [1983] for nonrelativistic particles and earlier work by Brizard and Chan [1999] for relativistic particles moving in static magnetic fields only. In addition, by introducing effective covariant potentials, we also simplify the relativistic guiding-center equations of motion of Grebogi and Littlejohn [1984]. In section 2.3, we extend the work in section 2.2 and use the Lie-transform method to obtain the bounce-center Lagrangian. The first-order correction to the second adiabatic invariant is automatically obtained in this process. In section 2.4, we assume that the electromagnetic fields vary on a time scale much slower than drift period and use a third Lie transform to remove the drift-phase dependence of the system and obtain the first-order correction to the third adiabatic invariant. A summary and comments
on further work are given in section 2.5.

2.2 Relativistic guiding-center dynamics

This section presents the guiding-center equations of motion for a relativistic particle mov­
ing in slowly-varying background electromagnetic fields derived by the Lie-transform method. As the first step of the Lie transform, we show the ordering of the background fields, then we obtain the guiding-center Lagrangian which is later used to derive the guiding-center equations of motion and also to obtain the bounce-center Lagrangian in section 2.3.

2.2.1 Background-field orderings

Following the work of Littlejohn [1983], we use the small parameter $\epsilon \equiv \rho_0/L_0 \ll 1$ to order the background fields, where $\rho_0$ is the typical gyroradius and $L_0$ is the scale length of background fields. In dimensional units, $\epsilon$ scales as $m/q$. We introduce the small parameter $\epsilon$ by denoting the physical electric field by $E_{ph}$, and we assume that the $E_{ph} \times B$ drift is of $O(\epsilon)$ compared to the particle's thermal speed [Littlejohn, 1983; Grebogi and Littlejohn, 1984], and that the background fields $E_{ph}$ and $B$ vary on a time scale comparable to the drift period; i.e., $\partial/\partial t \sim O(\epsilon)$. To indicate the order of a term explicitly by its $\epsilon$ factor, we set $E_{ph} = \epsilon E$, $\Phi_{ph} = \epsilon \Phi$ and $t_1 = \epsilon t$, where $\Phi_{ph}$ is the physical electric potential. Thus $E \times B \sim O(1)$, $\partial/\partial t_1 \sim O(1)$, and physical results are obtained by setting $\epsilon = 1$. 
2.2.2 Preliminary coordinate transformation

With the ordering of background fields given above, the particle phase-space Lagrangian one-form [Cary and Littlejohn, 1983] in slowly-varying background electromagnetic fields is written in terms of extended (position, momentum; time, energy) phase-space coordinates \( z \equiv (x, p; t, W_p) \) as

\[
\Gamma_p = \left[ \frac{1}{e \gamma c} A(x, t_1) + p \right] \cdot dx - W_p dt - H_p d\sigma,
\]

(2.1)

where subscript 'p' denotes particle variables and \( H_p = H_p - W_p \) is the extended particle Hamiltonian, with \( H_p = \gamma mc^2 + q\Phi(x, t_1) \) the Hamiltonian in regular phase space. Here, the physical dynamics takes place on the surface \( H_p = 0 \), the guiding-center relativistic factor is \( \gamma \equiv \sqrt{1 + |p/mc|^2} \), and \( \sigma \) is an orbit parameter.

To show the dependence of \( \Gamma_p \) on the gyrophase explicitly, we decompose the relativistic momentum \( p \) according to

\[
p = p_{\parallel 0} \hat{b} + \sqrt{2m\mu_0 B} \hat{e},
\]

(2.2)

where \( p_{\parallel 0} \equiv p \cdot \hat{b} \) is the component of the relativistic momentum parallel to \( B \), \( \mu_0 \equiv p_{\perp 0}^2/2mB \) will be shown to be the lowest-order term in the asymptotic expansion of an invariant and \( \hat{e} \) is the perpendicular unit vector. The local momentum coordinates \( (p_{\parallel 0}, \mu_0, \theta_0) \) [Brizard and Chan, 1999] are then defined, where \( \theta_0 \) is the instantaneous gyrophase implicitly defined by the following relations [Littlejohn, 1983]

\[
\begin{align*}
\hat{a} &= \cos \theta_0 \hat{e}_1 - \sin \theta_0 \hat{e}_2 \\
\hat{e} &= -\sin \theta_0 \hat{e}_1 - \cos \theta_0 \hat{e}_2
\end{align*}
\]

(2.3)
where \( \hat{e} \) is defined by equation (2.2), \( \hat{a} = \hat{b} \times \hat{e} \), and \((\hat{e}_1, \hat{e}_2, \hat{e}_3)\) forms an arbitrary right-handed unit-vector set with \( \hat{e}_3 \equiv \hat{b} \).

Substituting equation (2.2) into equation (2.1) yields the Lagrangian written in local momentum coordinates

\[
\Gamma_p = \left[ \frac{1}{\epsilon c} A(x, t_1) + p_{||0} \hat{b} + \sqrt{2\mu_0 B} \hat{e} \right] \cdot dx - W_p dt - \mathcal{H}_p d\sigma.
\]

(2.4)

Now \( \gamma \equiv \sqrt{1 + 2\mu_0 B/mc^2 + (p_{||0}/mc)^2} \) and \( \Gamma_p \) is a function of the preliminary phase-space coordinates \((x, p_{||0}, \mu_0, \theta_0; t, W_p)\). Next, we will use a Lie transform to remove the gyrophase dependence from the particle Lagrangian (2.4).

### 2.2.3 Guiding-center Lagrangian for a relativistic particle

A Lie transform from the preliminary coordinates \( z = (x, p_{||0}, \mu_0, \theta_0; t, W_p) \) to the guiding-center coordinates \( Z = (X, p_{||}, \mu, \theta; t, W_g) \) is used to remove the gyrophase dependence of \( \Gamma_p \). Here we use subscript 'g' to refer to guiding-center dynamics. For brevity, the steps of the guiding-center Lie transform are omitted here, but the interested reader may consult Brizard [1995] for details. The resulting guiding-center Lagrangian in extended guiding-center phase-space coordinates \((X, p_{||}, \mu, \theta; t, W_g)\) is

\[
\Gamma_g = \left[ \frac{1}{\epsilon c} A(X, t_1) + p_{||} \hat{b}(X, t_1) + \mathcal{O}(\epsilon) \right] \cdot dX + \frac{mc}{q} \mu d\theta - W_g dt - \mathcal{H}_g d\sigma,
\]

(2.5)
where the extended guiding-center Hamiltonian \( \mathcal{H}_g = H_g - W_g \), with the lowest-order regular Hamiltonian

\[
H_g = \gamma mc^2 + q\Phi(X, t_1).
\]

(2.6)

\[
\equiv mc^2 \sqrt{1 + 2\mu B/mc^2 + (p_\parallel/mc)^2} + q\Phi(X, t_1).
\]

Here, the guiding-center coordinates are related to the preliminary coordinates and are given to lowest order in \( \epsilon \) by \( p_\parallel = p_\parallel^0, \mu = \mu_0, \theta = \theta_0, W_g = W_p \), and

\[
X = x - \epsilon \rho,
\]

(2.7)

where

\[
\rho(\mu_0, \theta_0) \equiv \frac{c}{q} \sqrt{\frac{2m\mu_0}{B}} \alpha
\]

(2.8)

is the gyroradius vector in guiding-center coordinates. Note that because of the slow-time dependence, the differences between the guiding-center Lagrangian (2.5) and that of Brizard and Chan [1999] are the electric potential and the time-changing variables, which give us extra second-order terms in the guiding-center equations of motion.

### 2.2.4 Guiding-center equations of motion

Having found the relativistic guiding-center Lagrangian (2.5), we now solve for guiding-center equations of motion using Euler-Lagrange equations [Goldstein, 1980]. For a Lagrangian \( \mathcal{L}_g \), which is related to \( \Gamma_g \) in equation (2.5) by \( \Gamma_g \equiv \mathcal{L}_g d\sigma \), written in extended guiding-center phase-space coordinates \( Z^\nu \), the Euler-Lagrange equation is

\[
\frac{d}{d\sigma} \left( \frac{\partial \mathcal{L}_g}{\partial Z^\nu} \right) - \frac{\partial \mathcal{L}_g}{\partial Z^\nu} = 0,
\]

(2.9)
where $\dot{Z}^\nu = dZ/d\sigma$. The equations of motion for $t$ and $W_g$ are

$$\dot{t} = \frac{dt}{d\sigma} = -\frac{\partial H_g}{\partial W_g} = +1,$$

(2.10)

which indicates that $t$ and $\sigma$ can be identified, and the time rate of change of energy

$$\dot{W}_g = q \frac{\partial \Phi^*}{\partial t_1} - \frac{q}{c} \dot{X} \cdot \frac{\partial A^*}{\partial t_1},$$

(2.11)

where we replaced $\sigma$ by $t$ because of equation (2.10). Here, the effective potentials $\Phi^*$ and $A^*$ are defined as

$$\begin{pmatrix} \Phi^* \\ \epsilon^{-1} A^* \end{pmatrix} = \begin{pmatrix} \Phi \\ \epsilon^{-1} A \end{pmatrix} + \frac{mc}{q} \gamma c \begin{pmatrix} \gamma c \\ \gamma v|| \hat{b} \end{pmatrix},$$

(2.12)

where the second term on the right side represents the covariant parallel two-flat decomposition of the relativistic guiding-center four-velocity [Boghosian, 1987].

Applying the Euler-Lagrange equation (2.9) to other guiding-center phase-space coordinates $(X, p_\parallel, \mu, \theta)$, we first have $\hat{b} \cdot \ddot{X} = p_\parallel/(\gamma m)$, showing the parallel motion of the guiding center; secondly, $\dot{\theta} = \epsilon^{-1} q B/(\gamma mc)$, showing the fast gyromotion, and $\dot{\mu} = 0$, which proves that $\mu$ is an invariant of the guiding-center motion (here, a dot means a total derivative with respect to $t$). Finally, the relativistic guiding-center equations for $\dot{X}$ and $\dot{p}_\parallel$ are

$$\dot{X} = \frac{p_\parallel}{\gamma m} \frac{B^*}{B^*_\parallel} + \epsilon E^* \times \frac{\epsilon b}{B^*_\parallel},$$

(2.13)

$$\dot{p}_\parallel = q E^* \cdot \frac{B^*}{B^*_\parallel},$$

(2.14)

where the effective fields $(E^*, B^*)$ are defined in terms of the potentials (2.12) as

$$B^* \equiv \nabla \times A^* = B + \frac{\epsilon p_\parallel}{q} \nabla \times \hat{b},$$

(2.15)
and

\[ E^* = -\frac{1}{c} \frac{\partial A^*}{\partial t} - \nabla \Phi^* = E - \frac{e}{q} \left( p_{\parallel} \frac{\partial \hat{b}}{\partial t_1} + mc^2 \nabla \gamma \right), \quad (2.16) \]

where

\[ B_{\parallel}^* \equiv B^* \cdot \hat{b} = B + \varepsilon (cp_{\parallel}/q) \hat{b} \cdot \nabla \times \hat{b}, \quad (2.17) \]

and \( \nabla \gamma = (\mu/\gamma mc^2) \nabla B \). Equation (2.13) shows that the guiding-center velocity consists of the parallel motion along a field line, the \( E \times B \), gradient-\( B \) and curvature drifts. The curvature drift here is hidden in the first term on the right side of equation (2.13) and the gradient drift and the \( E \times B \) drift are contained in the second term. Equation (2.14) represents the parallel force along a field line, which according to equation (2.16) consists of two parts: one from the parallel electric field and the other from the magnetic mirror force. Note that the first-order term in equation (2.16) gives second-order terms in the guiding-center equation of motion, which are important to the conservation properties of the guiding-center motion.

We immediately note the simplicity of the relativistic guiding-center equations of motion (2.13) and (2.14), expressed in terms of the covariant effective potentials (2.12), compared to the relativistic guiding-center equations of motion of Grebogi and Littlejohn [1984] (GL), who used the scalar potential \( \Phi \) instead of the covariant potential \( \Phi^* \). We recover the GL relativistic guiding-center equations of motion by substituting \( qE^* = qE_{\text{GL}}^* - \varepsilon (\mu/\gamma) \nabla B \) in equations (2.13)-(2.14). We also point out that, in contrast to Boghosian's manifestly-covariant formulation for relativistic guiding-center motion [Boghosian, 1987], our "1 + 3" semi-covariant formulation treats time separately from the other phase-space
coordinates and uses an energy-like Hamiltonian (instead of the Lorentz-invariant covariant Hamiltonian).

If the fields are static, then equation (2.11) shows conservation of energy automatically. Also, the relativistic guiding-center equations (2.13)-(2.14) satisfy the Liouville theorem

\[ \frac{\partial B^*}{\partial t} + \nabla \cdot \left( B^* \dot{X} \right) + \frac{\partial}{\partial p_\parallel} \left( B^*_\parallel \dot{p}_\parallel \right) = 0, \quad (2.18) \]

which ensures that guiding-center phase-space volume is conserved by the guiding-center dynamics. We prove equation (2.18) explicitly as follows. First, we easily obtain from equations (2.13) - (2.14)

\[ \frac{\partial B^*_\parallel}{\partial t} = \hat{b} \cdot \frac{\partial B^*}{\partial t} + B^* \cdot \frac{\partial \hat{b}}{\partial t}, \quad (2.19) \]

\[ \nabla \cdot \left( B^* \dot{X} \right) = c \left( \hat{b} \cdot \nabla \times E^* - E^* \cdot \nabla \times \hat{b} \right) + \frac{p_\parallel}{m} B^* \cdot \nabla (\gamma^{-1}), \quad (2.20) \]

\[ \frac{\partial}{\partial p_\parallel} \left( B^*_\parallel \dot{p}_\parallel \right) = \frac{q}{c} \left( \frac{\partial E^*}{\partial p_\parallel} \cdot B^* + E^* \cdot \frac{\partial B^*}{\partial p_\parallel} \right). \quad (2.21) \]

Next, we insert

\[ \frac{\partial E^*}{\partial p_\parallel} = -\frac{p_\parallel}{mq} \nabla (\gamma^{-1}) - \frac{1}{q} \frac{\partial \hat{b}}{\partial t}, \quad (2.22) \]

\[ \frac{\partial B^*}{\partial p_\parallel} = \frac{c}{q} \nabla \times \hat{b}, \quad (2.23) \]

and

\[ \frac{\partial B^*}{\partial t} = -c \nabla \times E^* \quad (2.24) \]

in equations (2.19)-(2.21) to find that equation (2.18) is satisfied exactly, where we have set \( \epsilon = 1 \) in equations (2.18)-(2.24).
2.3 Hamiltonian theory of bounce-center dynamics

To obtain the bounce-center Lagrangian, we perform a Lie transform on the relativistic guiding-center Lagrangian (2.5) to remove the bounce-phase dependence. This Lie transform leads to construction of the second adiabatic invariant and gives the first-order correction to the second adiabatic invariant directly. The nonrelativistic bounce-center Lagrangian has been derived by Littlejohn [1982], and the present work generalizes previous results to the relativistic case.

2.3.1 Preliminary coordinate transformation

We first drop the term $\epsilon (mc/q) \mu d\theta$ in the extended guiding-center Lagrangian (2.5), which means we are now considering a six-dimensional system parametrized by constant-$\mu$ surfaces. Also we separate the extended Hamiltonian $H_g$ from the symplectic part of the extended phase-space Lagrangian (2.5) (i.e., the first and third terms on the right hand side, as in equation (2.26) below). We then perform a coordinate transformation from $X$ to $(\alpha, \beta, s)$, where $(\alpha, \beta)$ are the usual Euler potentials such that $B = \nabla \alpha \times \nabla \beta = B \hat{b}$, and $s$ is the position along a field line labeled by $(\alpha, \beta)$, with $\hat{b} = \partial X / \partial s$. We choose the vector potential $A = \alpha \nabla \beta$, write

$$dX = \frac{\partial X}{\partial \alpha} d\alpha + \frac{\partial X}{\partial \beta} d\beta + \hat{b} ds + \frac{\partial X}{\partial t_1} dt_1,$$

(2.25)

and we write the (symplectic part of the) guiding-center Lagrangian (2.5) order by order as

$$\Gamma_g = \frac{1}{\epsilon} \sum_{n=0}^{\infty} \epsilon^n \Gamma_{gn},$$

(2.26)
where
\[ \Gamma_{g0} = \frac{q}{c} \alpha d\beta - K_g dt, \] (2.27)
and the modified guiding-center energy coordinate
\[ K_g = W_g + \frac{q}{c} \alpha \frac{\partial \beta}{\partial t}. \] (2.28)

Equation (2.28) introduces a change to the extended Hamiltonian \( H_g = H_g - K_g \), where the lowest-order ordinary Hamiltonian \( H_g \) is \( H_{g0} = q\Phi^* + (q/c)\alpha \partial \beta / \partial t \). It is also useful to follow Littlejohn [1982] by using a 2-vector \( y \) with \( y_1 = \alpha, y_2 = \beta \), together with the two-dimensional Levi-Civita symbol \( \eta_{ab} \), where \( a, b \) runs overs 1 and 2. The components of \( \eta_{ab} \) are given by \( \eta_{11} = \eta_{22} = 0 \) and \( \eta_{12} = -\eta_{21} = 1 \).

Before considering the first-order term \( \Gamma_{g1} \) in equation (2.26) written in coordinates \((y, s)\), we make the usual assumption about the lowest-order motion that, with coordinates \((y, t)\) frozen, the bounce motion in \((s, p_\parallel)\) space is periodic [Brizard, 2000]. Thus, using the Hamilton-Jacobi theory [Goldstein, 1980], we construct the action-angle canonical variables \((J_{b0}, \psi_{b0})\) corresponding to the periodic bounce motion. Then
\[ J_{b0}(\alpha, \beta, \mu; t, K_g) = \frac{1}{2\pi} \int p_\parallel ds, \] (2.29)
and \( \omega_{b0} \) is the lowest-order angular bounce frequency, defined by \( \omega_{b0}^{-1} = \partial J_{b0} / \partial K_g \). The bounce-phase angle \( \psi_{b0} \) is canonically conjugate to \( J_{b0} \). Also, the following relation holds for the true motion (the "unfrozen" motion):
\[ \frac{\partial s}{\partial \psi_{b0}} \frac{\partial p_\parallel}{\partial J_{b0}} - \frac{\partial s}{\partial J_{b0}} \frac{\partial p_\parallel}{\partial \psi_{b0}} = 1, \] (2.30)
since the transformation from \((p, s)\) to \((J_{b0}, \psi_{b0})\) is canonical. At lowest order, \(J_{b0}\) is an invariant of motion. When higher-order terms are included and \((y, t_1)\) are allowed to evolve, we will show that \(J'_{b0} = \mathcal{O}(\epsilon)\). The symmetry of the unperturbed motion has been pointed out by Littlejohn [1983], and we will directly use this result to simplify the expression of the first-order correction to the second adiabatic invariant.

Using the coordinates \(Z_0^\mu = (y, J_{b0}, \psi_{b0}, t_1, W_{b0})\), the first order guiding-center Lagrangian in equation (2.26) has the components (here \(\mu\) is a "dummy" variable index)

\[
\Gamma_{g1\mu} = p_\parallel \hat{\mathbf{b}} \cdot \frac{\partial \mathbf{X}}{\partial Z_0^\mu} \equiv p_\parallel b_\mu, \quad (2.31)
\]

and we will omit subscripts of \(J_{b0}\) and \(\psi_{b0}\) when they themselves are subscripts. Note that in covariant form, \(\mathbf{b} = \nabla s + b_a \nabla y_a\). These expressions will be further simplified with the second coordinate transformation from \((p_\parallel, s)\) to \((J_{b0}, \psi_{b0})\).

### 2.3.2 Coordinate Transformation from \((p_\parallel, s)\) to \((J_{b0}, \psi_{b0})\)

To simplify components of \(\Gamma_{g1}\) defined in equation (2.31), we perform a coordinate transformation from \((p_\parallel, s)\) to \((J_{b0}, \psi_{b0})\) by adding a gauge term \(dG_b\) to \(\Gamma_{g1}\) (i.e., \(\Gamma'_{g1} = \Gamma_{g1} + dG_b\)) to eliminate the \(J\)-component \((\Gamma'_{g11} = 0)\). Thus we choose \(G_b\) to be

\[
G_b = - \int_0^{J_{b0}} dJ' \Gamma_{g11}(t_1, \psi_{b0}, J', y). \quad (2.32)
\]

Also, the \(\psi\)-component \(\Gamma'_{g1\psi} = \Gamma_{g1\psi} + \partial G_b/\partial \psi_{b0}\) becomes

\[
\Gamma'_{g1\psi} = \int_0^{J_{b0}} dJ' \left[ \frac{\partial (p_\parallel b_{\psi})}{\partial J'} - \frac{\partial (p_\parallel b_{J})}{\partial \psi_{b0}} \right] = J_{b0}, \quad (2.33)
\]
where we have used equation (2.30) to get the second equality and used \( \Gamma_{g1J} = 0 \) at \( J_{b0} = 0 \) (since \( p_{\parallel} = 0 \)).

Similarly, the other components of \( \Gamma'_{g1} \) are

\[
\Gamma'_{g1a} = \int_{0}^{J_{bo}} dJ' \left[ \frac{\partial \left( p_{\parallel} b_a \right)}{\partial J'} - \frac{\partial \left( p_{\parallel} b_J \right)}{\partial y_a} \right] = F_a, \tag{2.34}
\]

and

\[
\Gamma'_{g1t_1} = \int_{0}^{J_{bo}} dJ' \left[ \frac{\partial \left( p_{\parallel} b_t \right)}{\partial J'} - \frac{\partial \left( p_{\parallel} b_J \right)}{\partial t_1} \right] = F_{t_1}. \tag{2.35}
\]

Combining equations (2.33) - (2.35), the new first-order guiding-center Lagrangian is

\[
\Gamma'_{g1} = F_a \, dy_a + J_{b0} \, d\psi_{b0} + F_{t_1} \, dt_1. \tag{2.36}
\]

The Euler-Lagrange equation [see equation (2.9)] for \( J_{bo} \) obtained from \( \Gamma'_g = \Gamma_0 + \epsilon \Gamma'_{g1} \equiv \mathcal{L}'_g \, d\sigma \) is

\[
\frac{dJ_{bo}}{dt} = \frac{\partial F_a}{\partial \psi_{b0}} \dot{y}_a + \epsilon \frac{\partial F_{t_1}}{\partial \psi_{b0}} + \mathcal{O}(\epsilon^2), \tag{2.37}
\]

where to lowest order

\[
\dot{y}_a = \epsilon \frac{c}{q} \eta_{ab} \frac{\partial H_{g0}}{\partial y_b}. \tag{2.38}
\]

For later use, we now write the first two terms of the guiding-center Lagrangian [equation (2.26)] in coordinates \((\alpha, \beta, J_{b0}, \psi_{b0}; t_1, K_g)\) with the prime dropped,

\[
\Gamma_{g0} = \frac{q}{c} \alpha d\beta - K_g dt_1, \tag{2.39}
\]

\[
\Gamma_{g1} = F_a \, dy_a + J_{b0} \, d\psi_{b0} + F_{t_1} \, dt_1, \tag{2.40}
\]

and the lowest order Hamiltonian is

\[
\mathcal{H}_g = H_{g0} - K_g. \tag{2.41}
\]
With these coordinate transformations and Lagrangian, we do a Lie transform to remove the bounce-phase dependence from \( \Gamma_g \) and obtain the bounce-center Lagrangian \( \Gamma_b \).

### 2.3.3 Lie transform in extended phase-space coordinates

The bounce-center dynamics are obtained using the Lie transform in extended phase-space coordinates. The bounce-center Lagrangian and Hamiltonian are constructed order by order

\[
\Gamma_b = \Gamma_{b0} + \epsilon \Gamma_{b1} + \epsilon^2 \Gamma_{b2} + \cdots, \quad (2.42)
\]

\[
\mathcal{H}_b = \mathcal{H}_{b0} + \epsilon \mathcal{H}_{b1} + \epsilon^2 \mathcal{H}_{b2} + \cdots, \quad (2.43)
\]

where the terms on the right-hand side of equation (2.42) are

\[
\Gamma_{b0} = \Gamma_{g0}, \quad (2.44)
\]

\[
\Gamma_{b1} = \Gamma_{g1} - i_1 \cdot \Omega_{g0} + dS_1, \quad (2.45)
\]

\[
\Gamma_{b2} = -i_2 \cdot \Omega_{g0} - i_1 \cdot \Omega_{g1} + \frac{i_1}{2} \cdot d(i_1 \cdot \Omega_{g0}) + dS_2, \quad (2.46)
\]

and the first two terms in equation (2.43) are

\[
\mathcal{H}_{b0} = \mathcal{H}_{g0}, \quad (2.47)
\]

\[
\mathcal{H}_{b1} = \mathcal{H}_{g1} - g_1 \cdot d\mathcal{H}_{g0}. \quad (2.48)
\]

The term \( i_n \cdot \Omega_g = g_{\mu}^\nu (\Omega_g)_{\mu\nu} dZ^\nu \) in equations (2.45) - (2.46) and the term \( g_n \cdot d\mathcal{H}_g = g_{\mu}^\nu \cdot \partial \mathcal{H}_g / \partial Z^\mu \) in equation (2.48) are expressed in terms of the \( n^{th} \)-order Lie-transform
generating vector $g_n$ and gauge function $S_n$, where

$$(\Omega_g)_{\mu\nu} \equiv [Z^\mu, Z^\nu] = \frac{\partial \Gamma_{g\nu}}{\partial Z^\mu} - \frac{\partial \Gamma_{g\mu}}{\partial Z^\nu}$$  \hspace{1cm} (2.49)$$

is the Lagrange bracket between $Z^\mu$ and $Z^\nu$.

### 2.3.4 Bounce-center motion in coordinates $(Y, J_b, \psi_b; t, K_b)$

Following the Lie-transform procedure described in equations (2.44)-(2.48), we first have the lowest-order Lagrangian and Hamiltonian

$$\Gamma_{b0} = \frac{q}{c} Y_\alpha dY_\beta - K_b dt_1,$$  \hspace{1cm} (2.50)$$

$$H_{b0} = H_{b0} - K_b,$$  \hspace{1cm} (2.51)$$

where $(Y_\alpha, Y_\beta)$ represent the bounce-center coordinates $Y$, and $H_{b0}$ has the same functional dependence on the bounce-center coordinates $(Y, J_b, t)$ as $H_{g0}$ on the guiding-center coordinates $(y, J_{b0}, t)$.

The first-order bounce-center Lagrangian (2.45) then becomes

$$\Gamma_{b1} = (-g_{1b}^k \Omega_{ab0} + F_a) dy_a + J_{b0} d\psi_{b0} + (F_{t1} + g_{1K}^K) dt_1,$$  \hspace{1cm} (2.52)$$

where $\Omega_{ab0} = -(q/c) \eta_{ab}$ and we choose $S_1 = 0$ in equation (2.45). Requiring $\Gamma_{b1a} = 0$ and $\Gamma_{b1t_1} = 0$ gives us

$$g_{1a} = \frac{c}{q} \eta_{ab} F_b,$$  \hspace{1cm} (2.53)$$

$$g_{1K} = -F_{t1}.$$  \hspace{1cm} (2.54)$$
The first-order Hamiltonian then is given by

$$H_{b1} = -g_1^a \frac{\partial H_{g0}}{\partial y_a} - g_1^J \omega_{b0} + g_1^K,$$  
(2.55)

since $H_{g1} = 0$ in equation (2.48). With equations (2.53), (2.54), and the requirement $\tilde{H}_b = 0$, where a tilde in this section denotes the bounce-phase oscillatory part, we have

$$\tilde{g}_1^J = \frac{1}{\omega_{b0}} \left( -\tilde{F}_{t1} + \frac{c}{q} \eta_{ab} \tilde{F}_b \frac{\partial H_{g0}}{\partial y_a} \right),$$  
(2.56)

and

$$H_{b1} = \frac{c}{q} \eta_{ab} \langle F_b \rangle \frac{\partial H_{g0}}{\partial y_a} - \langle g_1^J \rangle \omega_{b0} - \langle F_{t1} \rangle,$$  
(2.57)

where $\langle \cdots \rangle$ denotes a bounce-phase average. It has been shown in Littlejohn [1982] that $F_a$ and $F_{t1}$ are odd in $\psi_{b0}$, and thus we have $\langle F_a \rangle = 0$ and $\langle F_{t1} \rangle = 0$. Equation (2.57) then becomes

$$H_{b1} = -\langle g_1^J \rangle \omega_{b0}.$$  
(2.58)

To obtain $g_1^\psi$ and the bounce-phase averaged part of $g_1^J$ needed in equation (2.58), we need to go to the second-order Lie transform of the Lagrangian. The $\psi_{b0}$ part and the $J_{b0}$ part of the second order Lagrangian $\Gamma_{b2}$ are

$$\Gamma_{b2\psi} = \frac{\partial S_2}{\partial \psi_{b0}} - g_1^J - \frac{1}{2} \frac{c}{q} \eta_{ab} F_b \frac{\partial F_a}{\partial \psi_{b0}},$$  
(2.59)

$$\Gamma_{b2J} = \frac{\partial S_2}{\partial J_{b0}} + g_1^\psi - \frac{1}{2} \frac{c}{q} \eta_{ab} F_b \frac{\partial F_a}{\partial J_{b0}}.$$  
(2.60)

To make $J_b$ the exact invariant, we require that $\Gamma_{b2\psi} = 0$. Taking the bounce-averaged part of equation (2.59) and using $\langle S_2 \rangle = 0$, we have $\langle g_1^J \rangle = 0$, since $\partial F_a / \partial \psi_{b0}$ is even in $\psi_{b0}$, and thus $\langle F_b \partial F_a / \partial \psi_{b0} \rangle = 0$. This result indicates that equation (2.58) becomes $H_{b1} = 0$. 
The bounce-phase dependent part of equation (2.59) is solved to give the gauge function

\[ S_2 = \int d\psi_0 \left( g_1^{j'} + \frac{1}{2} \eta_{ab} F_a \frac{\partial F_a}{\partial \psi_0} \right), \]  

(2.61)

where \( g_1^{j'} = \tilde{g}_1^{j'} \) is given in equation (2.56).

Inserting \( S_2 \) into equation (2.60) and requiring \( \Gamma_{b2j} = 0 \), such that \( J_b \) and \( \psi_b \) are exact conjugate coordinates, gives us

\[ g_1^{\psi} = \frac{1}{2} \eta_{ab} F_a \frac{\partial F_a}{\partial J_{b0}} - \frac{\partial S_2}{\partial J_{b0}}. \]

(2.62)

Thus we obtain the bounce-center Lagrangian and Hamiltonian

\[ \Gamma_b = \frac{1}{\epsilon c} \left( Y_a dY_b + J_b d\psi_b - K_b dt \right), \]  

(2.63)

\[ \mathcal{H}_b = \mathcal{H}_{b0} + O(\epsilon^2), \]  

(2.64)

where the bounce-center coordinates \( (Y, J_b, \psi_b; t, K_b) \) are given by

\[ Y_a = y_a - \frac{\epsilon}{q} \eta_{ab} F_b + O(\epsilon^2), \]  

(2.65)

\[ J_b = J_{b0} + \epsilon g_1^{j'} + O(\epsilon^2), \]  

(2.66)

\[ \psi_b = \psi_{b0} + \epsilon g_1^{\psi} + O(\epsilon^2), \]  

(2.67)

\[ K_b = K_g - \epsilon F_{t1} + O(\epsilon^2), \]  

(2.68)

with time \( t \) an invariant under the transformation. Note from equation (2.65) that \( F_a \) in bounce-center dynamics plays a role similar to the gyroradius vector \( \rho \) in guiding-center dynamics [equation (2.7)]; i.e., \( F_a \) may be interpreted as a "bounce radius" 2-vector. Also from equation (2.68), \( F_{t1} \) is the oscillatory part of the guiding-center energy coordinate \( K_g \).
The bounce-center equations are then

\[
\begin{align*}
\dot{Y}_a &= \frac{c}{q} \eta_{ab} \partial \mathcal{H}_b \frac{\partial}{\partial Y_b}, \\
\dot{\psi}_b &= \frac{\partial \mathcal{H}_b}{\partial J_b}, \\
\dot{J}_b &= 0, \\
\dot{K}_b &= \frac{\partial \mathcal{H}_b}{\partial t}.
\end{align*}
\]

(2.69) (2.70) (2.71) (2.72)

Thus we see that \( J_b \) is the exact invariant for the bounce motion.

In equation (2.56), \( g_1^I = g_1^I \) denotes the first-order correction to the second adiabatic invariant \( J_{b0} \). This first-order correction can also be directly obtained from Northrop [1963], where Northrop shows that

\[
\frac{dJ_{b0}}{dt} = \omega_{b0}^{-1} \left[ \frac{q}{c} \left( \langle \dot{\alpha} \dot{\beta} - \dot{\alpha} \langle \dot{\beta} \rangle \right) + \left( \dot{K}_g - \langle \dot{K}_g \rangle \right) \right], \tag{2.73}
\]

written using our notation. Since

\[
\frac{dJ_b}{dt} = \frac{dJ_{b0}}{dt} + \epsilon \frac{dJ_{b1}}{dt} + \cdots = 0, \tag{2.74}
\]

and to lowest order, we have \( \langle dJ_{b0}/dt \rangle = 0 \), thus

\[
\tilde{J}_{b1} = -\omega_{b0}^{-1} \int \frac{dJ_{b0}}{dt} d\psi_{b0}
= -\omega_{b0}^{-2} \int d\psi_{b0} \left[ \frac{q}{c} \left( \langle \dot{\alpha} \dot{\beta} - \dot{\alpha} \langle \dot{\beta} \rangle \right) \right.
+ \left. \left( \dot{K}_g - \langle \dot{K}_g \rangle \right) \right]. \tag{2.75}
\]

Since we have shown that \( J_{b1} \) (i.e., \( g_1^I \)) is purely oscillatory, we have \( J_{b1} = \tilde{J}_{b1} \). Littlejohn [1982] has shown, for the nonrelativistic case, that the right hand side of equation (2.75)
is equal to the right hand side of equation (2.56). This result also holds for the relativistic case because the equations have the same functional form. We also note, as pointed out by Littlejohn [1982], that the Lie-transform approach is more straightforward than the method used in Northrop et al. [1966], which derives the first-order correction to the second adiabatic invariant for nonrelativistic particles in a static magnetic field.

2.4 Hamiltonian theory of drift-center dynamics

Starting from the bounce-center Lagrangian (2.63), we now derive the drift-averaged bounce-center Lagrangian, or the drift-center Lagrangian. Similar to the analysis given in section 2.3, this procedure leads to the first-order correction to the third adiabatic invariant automatically.

To apply the adiabatic theory to the drift motion, electromagnetic fields must vary on a time scale much slower than the drift period; i.e., $\partial / \partial t \sim \epsilon^2$. We start from the bounce-center Lagrangian (2.63) with term $\bar{J}_d \psi_b$ dropped, which means we are now considering a two-dimensional motion parametrized by the constants $J_g$ and $J_b$. We set $t_2 \equiv \epsilon^2 t$ and the resulting bounce-center Lagrangian is

$$\Gamma_b = \frac{1}{\epsilon^2} \alpha d \beta - \frac{1}{\epsilon^2} \bar{K}_b d t_2 \equiv \frac{1}{\epsilon^2} (\epsilon \bar{\alpha} d \beta - K_b d t_2),$$

(2.76)

where we henceforth use $Y = (\alpha, \beta)$ and replaced $q \alpha / c$ by $\bar{\alpha}$ in the last expression.

We now make the usual assumption for the lowest-order motion that in a static field, or with parameter $t_2$ frozen, the orbit of the particle is closed and hence the drift motion
of the particle is periodic [Northrop, 1963]. Thus the coordinates \((\bar{\alpha}, \beta)\) play a role in drift-center dynamics similar to that of the coordinates \((p, s)\) in bounce-center dynamics. The Hamilton-Jacobi theory again gives us the action-angle variables from canonical coordinates \((\bar{\alpha}, \beta)\) as

\[
J_{d0}(K_b, t) = \frac{1}{2\pi} \oint \bar{\alpha} d\beta, \tag{2.77}
\]

and \(\omega_{d0}^{-1} = \partial J_{d0}/\partial K_b\) is the lowest-order angular frequency of the drift motion. Here we use 'd' to represent drift motion variables. The canonically-conjugate coordinate of \(J_{d0}\) is the drift phase \(\psi_{d0}\). The change from coordinates \((\bar{\alpha}, \beta)\) to \((J_{d0}, \psi_{d0})\) is canonical, thus we have

\[
\frac{\partial \bar{\alpha}}{\partial J_{d0}} \frac{\partial \beta}{\partial \psi_{d0}} - \frac{\partial \bar{\alpha}}{\partial \psi_{d0}} \frac{\partial \beta}{\partial J_{d0}} = 1, \tag{2.78}
\]

which is also valid for the true motion.

For the lowest-order motion, \(J_{d0}\) is a constant, but with time \(t\) unfrozen and higher-order terms included in the true motion, \(J_{d0}\) is no longer an invariant for the drift motion and it will be shown that \(dJ_{d0}/dt = \mathcal{O}(\epsilon^2)\). Thus we first do a coordinate transformation from \((\bar{\alpha}, \beta)\) to \((J_{d0}, \psi_{d0})\) and then use a Lie transform to construct the true invariant \(J_d\) for the drift motion.
2.4.1 Preliminary coordinate transformation

Similar to the construction of the bounce-center dynamics in section 2.3, we first change coordinates from \((\hat{\alpha}, \beta)\) to \((J_{d0}, \psi_{d0})\). Substituting

\[
\frac{d\beta}{dt} = \frac{\partial \beta}{\partial J_{d0}} dJ_{d0} + \frac{\partial \beta}{\partial \psi_{d0}} d\psi_{d0} + \frac{\partial \beta}{\partial t_2} dt_2,
\]

which is similar to equation (2.25), into equation (2.76) gives

\[
\epsilon^2 \Gamma_b = \epsilon \hat{\alpha} \frac{\partial \beta}{\partial J_{d0}} dJ_{d0} + \epsilon \hat{\alpha} \frac{\partial \beta}{\partial \psi_{d0}} d\psi_{d0} - \left[ K_b - \epsilon \hat{\alpha} \frac{\partial \beta}{\partial t_2} \right] dt_2,
\]

which gives the lowest- and first-order bounce-center Lagrangians

\[
\Gamma_{b0} = -K_b dt_2, \quad (2.81)
\]
\[
\Gamma_{b1} = \hat{\alpha} \frac{\partial \beta}{\partial J_{d0}} dJ_{d0} + \hat{\alpha} \frac{\partial \beta}{\partial \psi_{d0}} d\psi_{d0} + \hat{\alpha} \frac{\partial \beta}{\partial t_2} dt_2. \quad (2.82)
\]

Similar to equation (2.32), we perform a gauge transformation on \(\Gamma_{b1} \); i.e., \(\Gamma'_{b1} = \Gamma_{b1} + dG_d\), such that

\[
\Gamma_{b1,J} = 0 \text{ and } \Gamma'_{b1,\psi} = J_{d0}, \quad (2.83)
\]

where we have again omitted the subscripts of \(J_{d0}\) and \(\psi_{d0}\) when they themselves are subscripts. From equation (2.83), we choose \(G_d\) as

\[
G_d = - \int_0^{J_{d0}} \Gamma_{b1,J} dJ' + f(\psi_{d0}, t_2), \quad (2.84)
\]

which is similar to equation (2.32), and \(f(\psi_{d0}, t_2)\) is a function that is determined from the condition

\[
\frac{\partial G_d}{\partial \psi_{d0}} + \Gamma_{b1,\psi} = J_{d0}. \quad (2.85)
\]
Since $\Gamma_{b1\psi}$ can also be written as
\[ \Gamma_{b1\psi} = \int_0^{J_{b0}} \frac{\partial \Gamma_{b1\psi}}{\partial J'} dJ' + \Gamma_{b1\psi} \big|_{J_{b0}=0}, \] (2.86)
the equation for $f(\psi_{d0}, t_2)$ becomes
\[ \frac{\partial f}{\partial \psi_{d0}} = -\bar{\alpha} \frac{\partial \beta}{\partial \psi_{d0}} \bigg|_{J_{b0}=0}. \] (2.87)

Note the difference between equations (2.84) and (2.32), since we generally do not have
\[ \Gamma_{b1\psi} \big|_{J_{b0}=0} = 0. \] Finally the gauge transformation (2.84) yields the new $t_2$-term
\[ \Gamma'_{b1t_2} = -\int_0^{J_{d0}} \frac{\partial \Gamma_{b1J}}{\partial t_2} dJ' + \frac{\partial f}{\partial t_2} + \bar{\alpha} \frac{\partial \beta}{\partial t_2} \]
\[ \equiv F_{t_2}(J_{d0}, \psi_{d0}, t_2). \] (2.88)

Now we have our zeroth- and first-order bounce-center Lagrangian
\[ \Gamma_{b0} = -K_b dt_2, \] (2.89)
\[ \Gamma_{b1} = J_{d0} d\psi_{d0} + F_{t_2} dt_2, \] (2.90)
where we have dropped the prime, with the extended Hamiltonian
\[ \mathcal{H}_{b0} = H_{b0} - K_b. \] (2.91)

From $\Gamma_{b0}$ and $\Gamma_{b1}$, we obtain the Euler-Lagrange equation for $J_{d0}$
\[ \frac{dJ_{d0}}{dt} = \epsilon^2 \frac{\partial F_{t_2}}{\partial \psi_{d0}} + \mathcal{O}(\epsilon^3). \] (2.92)

Comparing equations (2.89)-(2.91) with equations (2.39)-(2.41), we find that the bounce-center and guiding-center equations are very similar, except that in equation (2.89) we do
not have the \( d\beta \) term and in equation (2.90) we do not have the \( dy_a \) term. Thus the Lie transform from the bounce-center coordinates to the drift-center coordinates will be very similar to the Lie transform from the guiding-center coordinates to the bounce-center coordinates.

### 2.4.2 Lie Transform from \((J_{d0}, \psi_{d0}; t, K_b)\) to \((J_d, \psi_d; t, K_d)\)

The lowest-order drift-center Lagrangian and Hamiltonian are given by \( \Gamma_{d0} = \Gamma_{b0} \) and \( \mathcal{H}_{d0} = \mathcal{H}_{b0} \). The first-order Lagrangian is given by

\[
\Gamma_{d1} = \Gamma_{b1} - i_1 \cdot \Omega_{b0} + dS_1. \tag{2.93}
\]

Choosing \( S_1 = 0 \) and substituting \( i_1 \cdot \Omega_{b0} = -g_1^K dt_2 \) and \( \Gamma_{b1} \) from equation (2.90) into equation (2.93), we obtain

\[
\Gamma_{d1} = J_{d0} \psi_{d0} + (F_{t2} + g_1^K)dt_2. \tag{2.94}
\]

The first-order Hamiltonian is then given by

\[
\mathcal{H}_{d1} = \mathcal{H}_{b1} - g_1 \cdot d\mathcal{H}_{b0} = -g_1^I \omega_{d0} + g_1^K, \tag{2.95}
\]

where we have used \( \mathcal{H}_{b1} = 0 \). Note that equations (2.94) - (2.95) look similar to equations (2.52) and (2.55) and we omit the remaining details here.

The first-order coordinate generators from the above Lie transform are

\[
g_1^J = \frac{\partial S_2}{\partial \psi_{d0}} = -\omega_{d0}^{-1} \tilde{F}_{t2}, \tag{2.96}
\]

\[
g_1^\psi = -\frac{\partial S_2}{\partial J_{d0}}, \tag{2.97}
\]

\[
g_1^K = -F_{t2}, \tag{2.98}
\]
where a tilde in this section indicates the drift-phase oscillatory part. Thus the overall coordinate transformation is

\[
J_d = J_{d0} - \epsilon \omega_{d0}^{-1} \tilde{F}_{t_2} + \mathcal{O}(\epsilon^2),
\]

(2.99)

\[
\psi_d = \psi_{d0} + \epsilon \psi_1 + \mathcal{O}(\epsilon^2),
\]

(2.100)

\[
K_d = K_b - \epsilon F_{t_2} + \mathcal{O}(\epsilon^2).
\]

(2.101)

The drift-center Lagrangian written in coordinates \((J_d, \psi_d; t, K_d)\) is

\[
\Gamma_d = \frac{1}{\epsilon} J_d d\psi_d - K_d dt,
\]

(2.102)

and the drift-center Hamiltonian function is

\[
\mathcal{H}_d = \mathcal{H}_{d0} + \epsilon \mathcal{H}_{d1} + \mathcal{O}(\epsilon^2),
\]

(2.103)

where \(\mathcal{H}_{d0} = \mathcal{H}_b\) and \(\mathcal{H}_{d1} = -\langle\langle F_{t_2}\rangle\rangle\), with a drift-phase average denoted as \(\langle\langle \cdots \rangle\rangle\).

The drift-center equations of motion are obtained from the Euler-Lagrange equations

\[
\dot{J}_d = 0,
\]

(2.104)

\[
\dot{\psi}_d = \frac{\partial \mathcal{H}_d}{\partial J_d},
\]

(2.105)

\[
\dot{K}_d = \frac{\partial \mathcal{H}_d}{\partial t}.
\]

(2.106)

Equation (2.99) shows the first-order correction to the third adiabatic invariant. We can also write the oscillatory part of \(F_{t_2}\) in another form by using equation (2.92) and

\[
\frac{dJ_{d0}}{dt} = \frac{\partial J_{d0}}{\partial t} + \frac{\partial J_{d0}}{\partial K_b} \dot{K}_b,
\]

(2.107)
which gives that

\[ \bar{F}_{t_2} = \int d\psi_{d_0}^' \left( \frac{\partial J_{d_0}}{\partial t} + \frac{1}{\omega_{d_0}} \dot{K}_b \right), \]  

(2.108)

where we have set \( \epsilon = 1 \). Northrop [1963] has shown that the first term on the right side of equation (2.107) can also be written as

\[ \frac{\partial J_{d_0}}{\partial t} = -\omega_{d_0}^{-1}\langle \dot{K}_b \rangle. \]  

(2.109)

With \( \partial J_{d_0}/\partial K_b = \omega_{d_0}^{-1} \), we find

\[ \bar{F}_{t_2} = \omega_{d_0}^{-1} \int d\psi_{d_0}^' \left( \dot{K}_b - \langle \dot{K}_b \rangle \right). \]  

(2.110)

Thus, Northrop [1963] implicitly contains the first-order correction term to the third adiabatic invariant (see equation (3.80) on page 64 of Northrop [1963]), but equation (2.99) is an explicit expression.

### 2.5 Summary and discussion

In this work, we have presented the Hamiltonian theory of adiabatic motion of a relativistic charged particle and the derivation of the first-order corrections to the second and third adiabatic invariants. The background electromagnetic fields vary on the drift time scale when we consider the guiding-center motion and the bounce-center motion. The effect of these time-varying background fields on the guiding-center motion is shown by the extra terms in the guiding-center equations (2.13) and (2.14), compared to the guiding-center equations in Brizard and Chan [1999]. The first-order correction to the second adiabatic
invariant of a relativistic particle is then shown in equation (2.66). To apply the adiabatic analysis to the drift motion, we assume that the background fields vary on a time scale much smaller than the drift period. The first-order correction to the third adiabatic invariant is shown in equation (2.99).

This work simplifies previous work on relativistic guiding-center motion, generalizes previous work on bounce-center motion for a relativistic particle in time-varying fields, and extends previous work on drift-center motion using Lie-transform perturbation methods in extended phase space. These results are especially useful in space plasma physics, where adiabatic theory is the foundation for modeling and understanding the dynamics of magnetically-trapped energetic particles.

The hierarchy of the adiabatic motions in this work may be shown as follows

\[ (\mathbf{x}, \mathbf{p}; t, W_p) \xrightarrow{g} (\mathbf{X}, \mathbf{p}^\parallel; t, W_g) \xrightarrow{b} (\alpha, \beta; t, K_b) \xrightarrow{d} (t, K_d) \]

\[ (J_g, \psi_g) \] where \( J_g \) is related to the first invariant \( \mu \) by \( J_g = (mq/c) \mu \) and \( \psi_g = \theta \). The first arrow (g) thus indicates the gyro-phase average process, the second arrow (b) the bounce-phase average and the third arrow (d) the drift-phase average.

In this chapter we have shown how first-order corrections to adiabatic invariants can be obtained using Lie-transform methods. Alternatively, the oscillatory part of the first-order correction to an adiabatic invariant can be obtained as follows. Differenting the exact
invariant

\[ J_k = J_{k0} + \epsilon J_{k1} + \cdots, \quad (2.111) \]

where \( k \) can be \( g, b \) or \( d \), to lowest order gives

\[ \frac{dJ_k}{dt} = \frac{dJ_{k0}}{dt} + \epsilon \omega_{k0} \frac{dJ_{k1}}{d\psi_{k0}} + \cdots = 0. \quad (2.112) \]

Since \( J_{k0} \) satisfies the necessary condition

\[ \left\langle \frac{dJ_{k0}}{dt} \right\rangle_k = 0, \quad (2.113) \]

where \( \langle \cdots \rangle_k \) denotes the fast-angle average canonically conjugate to \( J_k \), and we obtain the oscillatory part of \( J_{k1} \) as

\[ \epsilon \vec{J}_{k1} = -\omega_{k0}^{-1} \int \frac{dJ_{k0}}{dt} d\psi_{k0}. \quad (2.114) \]

The phase-independent part of \( J_{k1} \) can be obtained by using the Lie-transform method.

The use of Hamiltonian theory in describing adiabatic motions results in equations that satisfy energy conservation for time-independent fields and preserve phase-space volume naturally, in contrast to the results of Northrop [1963]. These conservation laws are very useful for checking numerical accuracy in simulations. Based on this work, fluctuations of electromagnetic fields can be added to the background fields and equations of motion in the presence of electromagnetic waves can be derived, as in Brizard and Chan [1999]; Brizard [2000], and Brizard and Chan [2004].
Chapter 3

Stochastic modeling of

multi-dimensional diffusion in the

radiation belts

This chapter has been published in Journal of Geophysical Research [Tao et al., 2008].

3.1 Introduction

The Earth's outer radiation belt is very dynamic, and electron fluxes can vary by several orders of magnitude during storm times, which makes it very hazardous to spacecrafts and astronauts [e.g., Baker et al., 1997]. Quasilinear diffusion theory has been used to evaluate dynamic changes of particle fluxes in the radiation belts [Albert, 2004; Albert and Young, 2005; Horne and Thorne, 2003; Horne et al., 2003]. Using the quasilinear diffusion theory
to model radiation belt dynamics requires at least two kinds of computations: numerical solution of a diffusion equation, which is a one-dimensional or multi-dimensional Fokker-Planck equation, depending on diffusion processes we are interested in, and calculation of diffusion coefficients.

Albert [2004] has shown that numerical problems arise when applying standard finite difference methods to pitch-angle and energy diffusion equations, because of rapidly varying off-diagonal diffusion coefficients. Albert and Young [2005] developed a method for solving the 2D diffusion equation which diagonalizes the diffusion tensor by transforming to a new set of coordinates and solves the transformed equation by simple finite difference methods. In this work we introduce another method, which uses probabilistic representations of solutions of Fokker-Planck equations [Freidlin, 1985; Costantini et al., 1998] via stochastic differential equations (SDEs) and we develop a 2D code for solving pitch-angle and energy diffusion equations. Compared with finite difference methods, the SDE method has three main advantages. First, the SDE method is very efficient when solutions on only a small number of points are desired, particularly when applied to high-dimensional problems, and it is easy to code and parallelize, with parallelization efficiency close to one. Second, with the SDE method we are able to handle complicated boundary geometry, other than constant-coordinate boundaries (see section 3.2.2). Third, generalization of the SDE method to higher dimensions is straightforward and we expect the method to be applicable to general 3D radiation belt diffusion equations. For more applications of similar methods using relations between Fokker-Planck equations and SDEs, see, e.g., Zhang [1999];
Besides solving diffusion equations, correctly calculating quasilinear diffusion coefficients is also important for numerical modeling of the radiation belt dynamics using quasilinear theory. Albert [2005] and Glauert and Horne [2005] have shown full calculations of diffusion coefficients for cyclotron resonant wave-particle interactions, where up to $n = \pm 5$ resonances are included. However, the full calculation of diffusion coefficients is very time consuming. Summers [2005] derived simplified formulae for coefficients with a parallel-propagation approximation (and hence only the $n = -1$ resonance is included), and the computation becomes much faster. Shprits et al. [2006c] calculated bounce-averaged pitch-angle and energy diffusion coefficients $D_{\alpha_0\alpha_0}$ and $D_{pp}$ with the parallel-propagation approximation for $E \leq 1$ MeV particles, and compared them with fully calculated coefficients from the PADIE code of Glauert and Horne [2005]. They concluded that coefficients for field-aligned waves are close to coefficients for waves with mildly oblique wave normal angle distribution from the PADIE code. However, using the wave model from Horne et al. [2005], we compute particle fluxes and we show that for $E = 2$ MeV electrons, $D_{\alpha_0\alpha_0}$ and $D_{pp}$ calculated with the parallel-propagation approximation produce flux differences of about one order of magnitude at some pitch angles, compared to using fully calculated coefficients. Furthermore, we show that by including off-diagonal terms in the calculation, the parallel propagation approximation also produces large errors in fluxes for both $E = 0.5$ MeV and 2 MeV electrons at small pitch angles.
The remainder of this chapter is organized as follows. The SDE method and its numerical implementation are introduced in section 3.2. In section 3.3 we present the application of the SDE method to a bounce-averaged radiation belt pitch-angle and energy diffusion equation. After describing the implementation of the SDE method for the pitch-angle-energy equation (section 3.3.1), we show comparisons between results from the SDE method and the Albert and Young [2005] transformation method (section 3.3.2). Then fluxes calculated from diffusion coefficients with the parallel-propagation approximation [Summers, 2005] are compared with fluxes computed with coefficients from full quasilinear theory [Albert, 2005] (section 3.3.3). We summarize our work and discuss future work in section 3.4.

3.2 The SDE method

Our SDE code is based on mathematical results which show that solutions of diffusion equations can be obtained using an equivalent stochastic process. Thus we first give a description of a stochastic process using Itô stochastic differential equations in section 3.2.1. Then we show how these lead to probabilistic representations of solutions of diffusion equations in section 3.2.2.

3.2.1 Itô stochastic differential equations

Stochastic differential equations (SDEs) are used to describe stochastic processes. They differ from ordinary differential equations by having terms involving random variables
A general \( m \)-dimensional SDE with a \( n \)-dimensional Wiener process is written as

\[
dX(t) = b(X, t) \, dt + \sigma(X, t) \, dW(t),
\]

(3.1)

where the \( m \)-vector \( X \) represents an \( m \)-dimensional stochastic process \( (X_1, X_2, \ldots, X_m) \).

Throughout this work, stochastic processes are indicated by uppercase characters, and their values at a given time are represented by corresponding lower case characters. The \( n \)-vector \( W \) is a \( n \)-dimensional Wiener process \((W_1, W_2, \ldots, W_n)\) and \( dW(t) = W(t+dt) - W(t) \) \cite{Gardiner1985}; an increment of a one-dimensional Wiener process is proportional to a Gaussian random number. The \( m \)-vector \( b \) and the \( m \times n \) matrix \( \sigma \) are coefficients that determine the values of \( X(t) \), they will be directly related to the coefficients of a corresponding diffusion equation in section 3.2.2. Stepping equation (3.1) in time generates a random walk trajectory through \( X \) space.

Note that SDEs may be formulated using two main mathematical methods, the Itô method and the Stratonovich method \cite{Gardiner1985}. In this work we use Itô SDEs, because they are directly related to diffusion equations of interest for the radiation belts, and they are mathematically more convenient \cite{Oksendal1992, Freidlin1985, Costantini1998}.

### 3.2.2 Probabilistic representation of solutions of diffusion equations

To solve a diffusion equation using SDEs, we can first write the diffusion equation in Fokker-Planck form, and then obtain equivalent “time-forward” SDEs from the diffusion
equation. These time-forward SDEs can then be used to simulate particle trajectories using a Monte Carlo technique, and the distribution of particles at any given time can be obtained by binning particles in phase space. This time-forward SDE method is presented in Appendix A to show local effects of off-diagonal terms on the distribution of particles. Alternatively, in this section we present a “time-backward” SDE method, where solutions of diffusion equations are represented by the mean value of a functional of trajectories of a stochastic process [Freidlin, 1985]. This is the method used in our current SDE code. Compared with the time-forward method, the time-backward method is more efficient when solutions on fewer points are of interest, and it is better for handling a variety of boundary conditions.

To introduce the time-backward SDE method, let us first consider a d-dimensional diffusion equation written as

\[
\frac{\partial f}{\partial t} = \sum_{i,j=1}^{d} \frac{1}{2} a_{ij}(t,x) \frac{\partial^2 f}{\partial x_i \partial x_j}(t,x) + \sum_{i=1}^{d} b_i(t,x) \frac{\partial f}{\partial x_i}(t,x) + c(t,x)f(t,x), \tag{3.2}
\]

with initial and Dirichlet boundary conditions

\[
f(0, x) = g_0(x), \quad x \in D, \tag{3.3}
\]
\[
f(t, x) = g_1(t, x), \quad x \in \partial D. \tag{3.4}
\]

Here \(D\) is the domain of the problem with boundary \(\partial D\), and \(g_1(0, x) = g_0(x)\) on \(\partial D\). Note that \(\partial D\) is not restricted to constant-coordinate surfaces in the SDE method [Freidlin, 1985]. The solution \(f(x, t)\) of equation (3.2) is related to the following d-dimensional
stochastic process

\[ dX(s) = b(t - s, X) \, ds + \sigma(t - s, X) \, dW(s), \quad 0 \leq s \leq t, \quad (3.5) \]

where \( X(s = 0) = x \), and \( W(s) \) is a \( d \)-dimensional Wiener process. Here the \( d \times d \) matrix \( \sigma \) is defined by \( \sigma \sigma^T = a \). Note that \( \sigma \) is not uniquely determined by this equation, but according to Levy's theorem [Zhang, 1999; Freidlin, 1985] different choices of \( \sigma \) generate equivalent stochastic processes that yield the same solution of the diffusion equation (3.2). Also note that equation (3.5) is a time-backward SDE: at \( s = 0 \), we evaluate \( b \) and \( \sigma T \) at time \( t \), while at \( s = t \), we evaluate \( b \) and \( \sigma \) at time 0. The solution \( f(x, t) \) is then represented by the stochastic process defined in equation (3.5) as

\[ f(x, t) = \mathbb{E}(F_x), \quad (3.6) \]

where \( \mathbb{E} \) denotes the expectation value, and \( F_x \) is defined by

\[ F_x = \begin{cases} 
  g_0(X|_{s=t}) \exp(Y|_{s=t}), & \tau \geq t, \\
  g_1(t - \tau, X|_{s=t}) \exp(Y|_{s=t}), & \tau < t,
\end{cases} \quad (3.7) \]

where \( \tau \) has the value of \( s \) when the stochastic process \( X(s) \) exits from the boundary \( \partial D \) for the first time, and \( Y(s) \) is defined by

\[ Y(s) = \int_0^s c(t - r, X(r)) \, dr. \quad (3.8) \]

Numerical calculation of \( f \) can be constructed easily from equations (3.6)–(3.8): To obtain \( f(x, t) \), we sample a number of trajectories of the stochastic process defined by equation (3.5) starting from \( x \) and \( s = 0 \), using a Monte-Carlo technique. The simulation of
a trajectory will stop either by reaching the initial condition at \( s = t \) (where time = 0) or by reaching the boundary of the domain \( D \) at \( s = \tau \), whichever comes first, and returns a value defined by equation (3.7). Then we use the average of values returned by all trajectories to approximate \( f(\mathbf{x}, t) \). This process is repeated if we want to calculate \( f \) at other points.

Now let us also consider a particular type of Neumann boundary condition that is commonly encountered in radiation belt diffusion equations:

\[
\nabla f \cdot \mathbf{n} = 0, \quad \mathbf{x} \in \partial_1 D,
\]

where \( \nabla f = (\partial f/\partial x^1, \partial f/\partial x^2, \cdots, \partial f/\partial x^d) \), the boundary \( \partial_1 D \) is the part of \( \partial D \) with the Neumann condition, and \( \mathbf{n} \) is the inward unit normal vector on \( \partial_1 D \). General methods for implementing Neumann boundary conditions in SDE solutions can be found in Freidlin [1985] and Costantini et al. [1998]; here we simply note that condition (3.9) can be enforced in our numerical calculation of \( f(\mathbf{x}, t) \) as follows: Every time a trajectory reaches the Neumann boundary \( \partial_1 D \), we immediately reflect it about the normal vector \( \mathbf{n} \) [Bossy et al., 2004]. This trajectory will later be stopped by either reaching the initial condition or a Dirichlet boundary and at that time the trajectory returns a value defined by equation (3.7).

### 3.3 Application

In this section, we apply the above SDE method to a bounce-averaged pitch-angle and energy diffusion equation [Albert, 2004]. In section 3.3.1 we derive the stochastic process used to solve the diffusion equation. In section 3.3.2 fluxes calculated using the SDE code
are compared with results from Albert and Young [2005] to show that the SDE code is capable of solving the diffusion equation with off-diagonal diffusion coefficients. To show the effect of diffusion coefficients with the parallel-propagation approximation [Summers, 2005] on particle fluxes, we solve the diffusion equation using these diffusion coefficients and in section 3.3.3 we compare with results obtained using fully calculated coefficients.

3.3.1 Application to pitch-angle and energy diffusion equations

We apply the above SDE method to the bounce-averaged pitch-angle and energy diffusion equation written in equatorial pitch angle and momentum \((\alpha_0, p)\)

\[
\frac{\partial f}{\partial t} = \frac{1}{Gp} \frac{\partial}{\partial \alpha_0} G \left( D_{\alpha_0 \alpha_0} \frac{1}{p} \frac{\partial f}{\partial \alpha_0} + D_{\alpha_0 p} \frac{\partial f}{\partial p} \right) + \frac{1}{G} \frac{\partial}{\partial p} G \left( D_{\alpha_0 p} \frac{1}{p} \frac{\partial f}{\partial \alpha_0} + D_{pp} \frac{\partial f}{\partial p} \right),
\]

(3.10)

where \(D_{\alpha_0 \alpha_0}, D_{\alpha_0 p}\) and \(D_{pp}\) are bounce-averaged pitch-angle, mixed and momentum diffusion coefficients [Albert, 2004]. Here \(G\) is a Jacobian factor, \(G = p^2 T(\alpha_0) \sin(\alpha_0) \cos(\alpha_0)\), and \(T(\alpha_0) \approx 1.30 - 0.56 \sin(\alpha_0)\) is the normalized bounce period. Initial and boundary conditions are chosen to be the same as in Albert and Young [2005]. Thus the initial flux is \(j(t = 0) = \exp[-(E - 0.2)/0.1][\sin(\alpha_0) - \sin(\alpha_{0L})]\), where the loss cone angle \(\alpha_{0L} = 5^\circ\).
and flux $j$ is related to phase-space density $f$ by $j = f/p^2$. Boundary conditions are

$$f|_{\alpha_0 = \alpha_{0L}} = 0, \quad \text{(3.11)}$$

$$\frac{\partial f}{\partial \alpha_0} \bigg|_{\alpha_0 = 90^\circ} = 0, \quad \text{(3.12)}$$

$$f|_{E = E_{\text{max}}} = 0, \quad \text{(3.13)}$$

$$f|_{E = E_{\text{min}}} = j(t = 0)|_{E = E_{\text{min}}/p_{\text{min}}^2} \quad \text{(3.14)}$$

where $E_{\text{min}} = 0.2$ MeV and $E_{\text{max}} = 5$ MeV, and $p_{\text{min}}$ is the momentum corresponding to $E_{\text{min}}$ [Albert and Young, 2005].

To solve the equation using the time-backward SDE method, we first write equation (3.10) in the form of (3.2); i.e.,

$$\frac{\partial f}{\partial t} = \frac{D_{\alpha_0 \alpha_0}}{p^2} \frac{\partial^2 f}{\partial \alpha_0^2} + 2 \frac{D_{\alpha_0 p}}{p} \frac{\partial^2 f}{\partial \alpha_0 \partial p} + D_{pp} \frac{\partial^2 f}{\partial p^2}$$

$$+ b_{\alpha_0} \frac{\partial f}{\partial \alpha_0} + b_p \frac{\partial f}{\partial p}, \quad \text{(3.15)}$$

with

$$b_{\alpha_0}(t, \alpha_0, p) = \frac{1}{G p} \frac{\partial}{\partial \alpha_0} \left( \frac{GD_{\alpha_0 \alpha_0}}{p} \right) + \frac{1}{G} \frac{\partial}{\partial p} \left( \frac{GD_{\alpha_0 p}}{p} \right), \quad \text{(3.16)}$$

$$b_p(t, \alpha_0, p) = \frac{1}{G p} \frac{\partial}{\partial \alpha_0} \left( GD_{\alpha_0 p} \right) + \frac{1}{G} \frac{\partial}{\partial p} \left( GD_{pp} \right). \quad \text{(3.17)}$$

Thus the two-dimensional stochastic process defined in equation (3.5) becomes

$$dA_0(s) = b_{\alpha_0}(t - s, A_0, P) \, ds + \sigma_{11} \, dW_1 + \sigma_{12} \, dW_2, \quad \text{(3.18)}$$

$$dP(s) = b_p(t - s, A_0, P) \, ds + \sigma_{21} \, dW_1 + \sigma_{22} \, dW_2, \quad \text{(3.19)}$$

with $A_0(s = 0) = \alpha_0$ and $P(s = 0) = p$. Then, because of the Neumann boundary condition at $\alpha_0 = 90^\circ$, we numerically reflect $A_0$ with respect to $\alpha_0 = 90^\circ$ if it is larger.
than 90°. Here components of the matrix $\sigma$ are defined by

$$
\begin{pmatrix}
\sigma_{11} & \sigma_{12} \\
\sigma_{21} & \sigma_{22}
\end{pmatrix}
\begin{pmatrix}
\sigma_{11} & \sigma_{21} \\
\sigma_{12} & \sigma_{22}
\end{pmatrix} =
\begin{pmatrix}
2D_{a_{0a0}/p^2} & 2D_{a_{0p}/p} \\
2D_{a_{0p}/p} & 2D_{pp}
\end{pmatrix}.
$$

(3.20)

In this work, we choose $\sigma_{12} = 0$ for simplicity and then the other components are

$$
\sigma_{11} = \sqrt{2D_{a_{0a0}/p}},
$$

(3.21)

$$
\sigma_{21} = \sqrt{2D_{a_{0p}}/\sqrt{D_{a_{0a0}}}},
$$

(3.22)

$$
\sigma_{22} = \sqrt{2D_{pp} - \sigma_{21}^2},
$$

(3.23)

where we have used the fact that $D_{a_{0a0}}$ is never zero in equation (3.22).

We have developed a 2D SDE code to solve the diffusion equation (3.10), where SDEs (3.18) and (3.19) are integrated using the Euler-Maruyama method [Kloeden and Platen, 1992]. That is,

$$
A_0(s_{n+1}) = A_0(s_n) + b_{a_0}[t - s_n, A_0(s_n), P(s_n)] \Delta s
+ \sigma_{11}(s_n) \Delta W_1 + \sigma_{12}(s_n) \Delta W_2,
$$

(3.24)

$$
P(s_{n+1}) = P(s_n) + b_p[t - s_n, A_0(s_n), P(s_n)] \Delta s
+ \sigma_{21}(s_n) \Delta W_1 + \sigma_{22}(s_n) \Delta W_2.
$$

(3.25)

Here $\Delta W = \sqrt{s_{n+1} - s_n} N(0, 1)$, where $N(0, 1)$ is a standard Gaussian random number with zero mean and unit variance, generated using the Box-Muller algorithm [Press et al., 2002]. Because the original time-backward SDE method requires fresh samples of trajectories for every different $(\alpha_0, p)$ and traces trajectories back to the initial condition or to a
boundary every time, the current SDE code is less efficient when solutions on many grid points for long times are needed. Improving the efficiency of the SDE code is one of tasks in our future work. In this work, we mainly want to show that the method can be used to solve multi-dimensional diffusion equations. Results from the SDE code are compared with Albert and Young [2005] in the next section.

3.3.2 Comparisons with Albert and Young [2005] results

Albert and Young [2005] solve the diffusion equation (3.10) by first transforming to new coordinates which diagonalize the diffusion tensor, and then applying standard finite difference methods to the transformed diffusion equation. The bounce-averaged diffusion coefficients \(D_{\alpha\alpha\alpha}, D_{\alpha\alpha\beta}, D_{\alpha\beta\beta}\) for storm-time chorus waves were calculated at \(L = 4.5\), with computational methods of Albert [2005]. The wave model used to calculate diffusion coefficients is described in Horne et al. [2005] and Albert and Young [2005]; the wave magnetic field is given by \(B_w^2 = B^2(\omega)g_{\omega}(\tan \theta)\), where the wave power spectral density \(B^2(\omega)\) and the wave normal angle (\(\tan \theta\)) distribution function \(g_{\omega}(\tan \theta)\) are truncated Gaussian functions defined between lower and upper frequency cutoffs \((\omega_{LC} < \omega < \omega_{UC})\) and wave normal angle cutoffs \((\theta_{LC} < \theta < \theta_{UC})\). The latitudinal distribution of the waves and the ratio of electron plasma frequency \(f_{pe}\) to electron cyclotron frequency \(f_{ce}\) are the same as those used by Horne et al. [2005] and Albert and Young [2005] and are shown in Table 3.1. Similar models were used by Li et al. [2007]. Up to \(n = \pm 5\) resonance harmonics were included in the calculation. The calculated diffusion coefficients \(D_{\alpha\alpha\alpha}\) are proportional to
Using the above diffusion coefficients in equation (3.10), we obtain fluxes for $E = 0.5$ MeV and 2.0 MeV electrons with $\alpha_0$ ranging from 6° to 88° with 1° spacing, at $t = 0.1$ and 1 day. We have sampled $N = 9000$ trajectories at each $\alpha_0$ for $E = 0.5$ MeV, and $N = 18000$ trajectories for $E = 2.0$ MeV with $dt = 0.0004$ day. The chosen $dt$
Table 3.1: The latitudinal distribution of the waves and $f_{pe}/f_{ce}$ of the wave model [Horne et al., 2005] used to calculate diffusion coefficients.

gives small relative change in $\alpha_0$ and $E$ per step, compared with scales of the diffusion coefficients and initial phase-space density. Our choices of $N$ and $dt$ might not be optimal, and choosing $N$ adaptively is probably better (Greg Cunningham, personal communication, 2007). Results from the SDE code are compared with those of Albert and Young [2005]. Figure 3.2 shows the comparisons for $E = 0.5$ MeV electrons (upper panel) and $E = 2.0$ MeV electrons (lower panel), with results from the SDE method smoothed using a six-point moving window average in $\alpha_0$ with $\Delta\alpha_0 = 1^\circ$. Within small numerical errors associated with each of the methods, the two sets of results are in excellent agreement, and they demonstrate that our SDE code is able to successfully solve the bounce-averaged pitch-angle and energy diffusion equation.

To show effects of ignoring off-diagonal terms on change of flux, we rerun the SDE code, setting off-diagonal diffusion coefficients to zero. Results are shown in Figure 3.3 for 0.5 MeV (upper panel) and 2 MeV (lower panel) electrons. From Figure 3.3 we see that for 0.5 MeV electrons, while there is a relatively small effect at large pitch angles, ignoring off-diagonal terms overestimates electron fluxes at small pitch angles by a factor of $2 \sim 5$ at $t = 1$ day. For 2 MeV electrons, ignoring off-diagonal terms overestimates fluxes by a
Figure 3.2: Comparisons between results obtained from the SDE method (solid lines) and the Albert and Young [2005] method (dashed lines) for $E = 0.5$ MeV (upper panel) and $E = 2.0$ MeV (lower panel) at $t = 0.1$ day (blue lines) and $t = 1.0$ day (red lines). Here black lines show the initial condition.
factor of $5 \sim 10$ at $t = 1$ day, with larger errors at smaller pitch angles. Thus off-diagonal terms are more important for 2 MeV electrons. We emphasize that these results are for the Horne et al. [2005] wave model, and we note that the peak in flux of 2 MeV electrons near $30^\circ$ may be related to the cutoff in wave power at $35^\circ$ latitude in the Horne et al. [2005] model (see discussion in section 3.4).

### 3.3.3 Effects of parallel propagation approximation

Summers [2005] and Summers et al. [2007a, b] have derived cyclotron-resonance diffusion coefficients for field-aligned waves, where only the $n = -1$ resonance is included (henceforth denoted by $D_{||}$). This assumption of parallel propagation greatly improves the computation efficiency. Bounce-averaged $D_{||}$ are given and compared with diffusion coefficients obtained from the PADIE code [Glauert and Horne, 2005] in Shprits et al. [2006c]. In the present work, we also calculate $D_{||}$ using the methods of Albert [2005], with wave parameters the same as the wave model described in Section 3.3.2, except that $\theta_{LC} = \theta_{UC} = 0$. The resulting diffusion coefficients are the same as those obtained from the PADIE code, and are half of those given by Summers et al. [2007a] (this factor of two difference is discussed in [Albert, 2007]).

Figure 3.4 shows inverse timescales from diffusion coefficients with the parallel wave approximation. Compared with Figure 3.1, we see that the general behavior of $D_{||}$ is quite good, with larger differences for $E > 1$ MeV electrons. The off-diagonal terms of $D_{||}$ are worse approximations than the diagonal terms, with details discussed in Albert [2007].
Figure 3.3: Fluxes for $E = 0.5$ MeV (upper panel) and $E = 2.0$ MeV (lower panel) at $t = 0.1$ day (blue lines) and $t = 1.0$ day (red lines) with and without off-diagonal diffusion terms. Dashed lines are results without off-diagonal diffusion coefficients, and solid lines are results with off-diagonal terms.
Figure 3.4: Same as Figure 3.1, except that diffusion coefficients are calculated with the parallel propagation approximation.
To compare effects of $D^\parallel$ with fully calculated diffusion coefficients $D$, we solve equation (3.10) for 0.5 MeV and 2 MeV electrons using the following four sets of diffusion coefficients: (i) $D^\parallel$; (ii) diagonal terms of $D^\parallel$ (hereafter $D^\parallel_d$); (iii) $D$; (iv) diagonal terms of $D$ (hereafter $D_d$). Results are shown in Figures 3.5 to 3.7.

The upper panel of Figure 3.5 shows the comparison between fluxes calculated using $D^\parallel_d$ and $D_d$ for 0.5 MeV electrons. We see that results from $D^\parallel_d$ agree very well with $D_d$, with slight differences for $\alpha_0$ greater than about 40°. The lower panel of Figure 3.5 shows the same comparison for 2.0 MeV electrons, from which we see that the flux from $D^\parallel_d$ is smaller than that from $D_d$ by up to approximately 5 orders of magnitude at low $\alpha_0 (< 15°)$ at $t = 1.0$ day. This behavior occurs because $D^\parallel_d$ underestimates energy diffusion coefficients for high energy particles at small pitch angles, where $n \neq -1$ resonances also make a significant contribution. Thus $D^\parallel_d$ produces larger differences in fluxes for 2 MeV electrons than 0.5 MeV at small $\alpha_0$, compared with $D_d$.

Figure 3.6 shows comparisons between fluxes calculated using $D^\parallel_d$ and $D$ for 0.5 MeV electrons (upper panel) and 2 MeV electrons (lower panel). The upper panel of Figure 3.6 shows that $D^\parallel_d$ overestimates increase of flux at small pitch angles for 0.5 MeV electrons, which is expected, because $D^\parallel_d$ yields very similar flux increases as $D_d$ for 0.5 MeV electrons. For 2.0 MeV electrons, fluxes from $D^\parallel_d$ are smaller than that from $D$ for $\alpha_0 \lesssim 18°$ and larger for $\alpha_0 \gtrsim 18°$ at $t = 1.0$ day (where the difference can be about 1 to 2 orders of magnitude).

Fluxes calculated from $D^\parallel$ and $D$ (i.e., with off-diagonal terms included) for 0.5 MeV
Figure 3.5: Comparisons between results obtained from diffusion coefficients $D^\parallel_d$ (dashed lines) and $D_d$ (solid lines) for $E = 0.5$ MeV (upper panel) and $E = 2.0$ MeV (lower panel) at $t = 0.1$ day (blue lines) and $t = 1.0$ day (red lines).
Figure 3.6: Comparisons between results obtained from diffusion coefficients $D_d^1$ (dashed lines) and $D$ (solid lines) for $E = 0.5$ MeV (upper panel) and $E = 2.0$ MeV (lower panel) at $t = 0.1$ day (blue lines) and $t = 1.0$ day (red lines).
(upper panel) and 2 MeV (lower panel) electrons are shown in Figure 3.7. Reasonable agreement between $D^\parallel$ and $D$ fluxes is obtained for $\alpha_0 \geq 50^\circ$, but significant differences occur at smaller pitch angles. For 0.5 MeV electrons, $D^\parallel$ underestimates increases of flux at $t = 1.0$ day by approximately an order of magnitude for $\alpha_0 < 20^\circ$. For 2.0 MeV electrons, behavior of $D^\parallel$ is worse at $t = 1.0$ day. We see from the lower panel of Figure 3.7 that $D^\parallel$ underestimates increases of flux by approximately $1 \sim 4$ orders of magnitude for $10^\circ \leq \alpha_0 \leq 35^\circ$. Thus the approximation of parallel propagation produces larger differences at small pitch angles for higher energy particles, especially when off-diagonal terms are included.

3.4 Summary and discussion

In this work a new code, based on the mathematical theory of expressing solutions of diffusion equations in terms of related stochastic processes, has been developed for solving multi-dimensional radiation belt diffusion equations. Two examples are used to show its applications.

First, we apply the SDE code to a bounce-averaged pitch-angle and energy diffusion equation and obtain excellent agreement with a previously developed method [Albert and Young, 2005]. We also confirm that ignoring off-diagonal terms in the diffusion equation overestimates increase of flux, especially at small pitch angles, at $t = 1$ day (by a factor of $2 \sim 5$ for 0.5 MeV, and $5 \sim 10$ for 2 MeV electrons) using the Albert and Young [2005] diffusion coefficients.
Figure 3.7: Comparisons between results obtained from diffusion coefficients $D^\parallel$ (dashed lines) and $D$ (solid lines) for $E = 0.5$ MeV (upper panel) and $E = 2.0$ MeV (lower panel) at $t = 0.1$ day (blue lines) and $t = 1.0$ day (red lines).
Second, by solving the bounce-averaged pitch-angle and energy diffusion equation using fully-calculated diffusion coefficients $D$ [Albert and Young, 2005] and coefficients with the parallel propagation approximation $D^\parallel$ [Summers, 2005; Summers et al., 2007a, b], both calculated using the chorus wave model of Horne et al. [2005], we show that diagonal diffusion coefficients of $D^\parallel$ agree well with that of $D$ only for low energy particles (e.g., $E = 0.5$ MeV). For high energy electrons, the difference between the diagonal terms of $D^\parallel$ and $D$ produces large differences in fluxes at some pitch angles (difference of up to 5 orders of magnitude for 2 MeV electrons at $\alpha_0 \lesssim 15^\circ$, at $t = 1.0$ day). By including off-diagonal diffusion coefficients in our calculation, we show that the off-diagonal terms of $D^\parallel$ can produce difference in fluxes of 4 orders of magnitude for 2 MeV electrons at $t = 1.0$ day. A discussion of the details of different diffusion coefficients and another approximation for a full calculation of diffusion coefficients are presented in Albert [2007].

Note that the above conclusions on the magnitude and location of differences that occur by omitting off-diagonal terms and assuming parallel propagating waves are very likely to be dependent on the wave model used. For example, a different latitudinal distribution of wave power may result in different diffusion coefficients and thus different conclusions. The sensitivity of our results to wave models needs further study. However, before such work is done, it is safer to include both off-diagonal terms and oblique waves in calculations of electron flux.

The SDE method is less efficient when solutions on many grid points are desired. However, when parallel computers are available, computation time can be greatly reduced be-
cause of high parallelization efficiency. Generalization to 3D is straightforward, and our next step is to build a 3D model including pitch-angle, energy and radial diffusion using this method. The SDE method is very promising for providing new insights into the relative roles of local acceleration and radial diffusion as acceleration mechanisms, and the importance of pitch-angle diffusion as a loss process.
Chapter 4

Modeling of multi-dimensional diffusion in radiation belts using layer methods

This chapter has been published in Journal of Geophysical Research [Tao et al., 2009].

4.1 Introduction

In Chapter 3, we developed a stochastic differential equation (SDE) code to solve 2D bounce-averaged pitch angle and energy diffusion equations. The SDE code is very efficient when solutions on a small number of points are needed. However, if solutions are needed on a large computational domain for long times, the SDE code becomes less efficient, for reasons explained in Chapter 3. Milstein [2002], Milstein and Tretyakov [2002] and Milstein and Tretyakov [2001] have used properties of numerical integration of SDEs to develop so-called layer methods, which are deterministic, to solve parabolic equations
successively in time. In this chapter we develop a code using layer methods and show that it is able to solve 2D radiation belt diffusion equations with cross diffusion and it is generalizable to 3D. Although the layer code does not have the high parallelization efficiency compared with the SDE code in Chapter 3, it is more efficient when solving the diffusion equation over a large computational domain for long times. Also our layer code can handle boundary conditions with complicated geometry rather than constant-coordinate boundaries that are typically used in finite difference codes.

Using this layer code, we then explore effects of ignoring off-diagonal terms using two wave models: the chorus wave model from Li et al. [2007], and the combined magnetosonic wave [Horne et al., 2007] and hiss wave [Li et al., 2007] (MH) model. In Chapter 3, we show that ignoring off-diagonal terms causes errors of an order of magnitude for 2 MeV electrons at small pitch angles using the Horne et al. [2005] chorus wave model. Using the Li et al. [2007] chorus wave model and the MH wave model is helpful in understanding the sensitivity of the main conclusions of Chapter 3 to different wave models.

The remainder of this chapter is organized as follows. We introduce the layer methods by using a simple initial-value problem in section 4.2. Details of our 2D layer code to solve a bounce-averaged pitch angle and energy diffusion equation are given in section 4.3. We first show its agreement with Albert and Young [2005] results in section 4.3.1. Then we solve the diffusion equation with diffusion coefficients calculated using the Li et al. [2007] chorus wave model (section 4.3.2) and the MH wave model (section 4.3.3) to show effects of ignoring off-diagonal terms and energization of electrons. Our results are then discussed
and summarized in section 4.4.

### 4.2 The Milstein layer methods

In this section, we will use an initial-value problem to illustrate the layer methods shown in Milstein [2002]. Boundary conditions can be implemented in a similar way as described in Chapter 3, or in Milstein and Tretyakov [2001] and Milstein and Tretyakov [2002].

#### 4.2.1 One-step representation of solutions using the SDE method

Assume that we want to solve the equation

\[
\frac{\partial f}{\partial t} = \sum_{i=1}^{d} b_i(t, x) \frac{\partial f}{\partial x_i}(t, x) + \sum_{i,j=1}^{d} \frac{1}{2} a_{i,j}(t, x) \frac{\partial^2 f}{\partial x_i x_j}(t, x)
\]  

(4.1)

with an initial condition \( f(t_0, x) = g(x) \). First, we discretize time \( t \) equidistantly to \( t_0, \cdots, t_n, t_{n+1}, \cdots \), and assume that we know solutions of all \( f \) at time \( t_n \), which means that now \( f(t_n, x) \) can be considered as an initial condition when solving \( f(t_{n+1}, x) \). Then using the SDE method described in Chapter 3, we have

\[
f(t_{n+1}, x) \approx \mathbb{E}(f(t_n, \bar{x})),
\]  

(4.2)

where \( \mathbb{E} \) is the expectation value and \( \bar{x} \) is given by

\[
\bar{x} = x + b(t_{n+1}, x) \Delta t + \sigma(t_{n+1}, x) \Delta W.
\]  

(4.3)

Here \( \Delta t = t_{n+1} - t_n \), \( \sigma \) is related to \( a \) by \( \sigma \sigma^T = a \) and \( \Delta W \equiv (\Delta W_1, \Delta W_2, \ldots, \Delta W_d) \) is one increment of a Wiener random process [Gardiner, 1985].
Numerically $\Delta W$ can be generated from a vector of standard Gaussian random numbers with zero mean and unit variance, as we did in Chapter 3, or we can choose the probability distribution of the components to be

$$P(\Delta W_i = \pm \sqrt{\Delta t}) = \frac{1}{2}, \quad i = 1, 2, \ldots, d,$$

(4.4)

where $P$ denotes the probability [Milstein, 2002]. Substituting $\Delta W_i$ from equation (4.4) into equation (4.3), we will have $2^d$ possible $\Delta W$'s, thus $2^d$ possible $\bar{x}$'s, each with probability $1/2^d$. The expectation value term in equation (4.2) can then be rewritten as

$$f(t_{n+1}, \bar{x}) \approx \mathbb{E}(f(t_n, \bar{x})) = \frac{1}{2^d} \sum_{j=1}^{2^d} f(t_n, \bar{x}_j).$$

(4.5)

Note that in reality, we usually do not know $f$ at $t_n$ and at an arbitrary point $\bar{x}_j$. This is why we trace trajectories back to the initial condition in Chapter 3. However, as described by Milstein [2002], we can use interpolations to obtain $f(t_n, \bar{x}_j)$ from already known $f$'s at fixed grid points to make a convergent algorithm. In this way, we obtain solutions successively from time layer $t_n$ to $t_{n+1}$, hence the name "layer methods" [Milstein, 2002].

### 4.2.2 A simple layer method algorithm

A simple interpolation method is linear interpolation. Take a 1D diffusion equation for example:

$$\frac{\partial f}{\partial t} = b \frac{\partial f}{\partial x} + \frac{1}{2} \sigma^2 \frac{\partial^2 f}{\partial x^2}.$$  

(4.6)

First, discretizing $x$ equidistantly into $x_0, x_1, x_2, \ldots, x_N$, we have

$$f(t_n, \bar{x}) = \frac{x_{i+1} - \bar{x}}{x_{i+1} - x_i} f(t_n, x_i) + \frac{\bar{x} - x_i}{x_{i+1} - x_i} f(t_n, x_{i+1}) + O(\Delta x^2), \quad x_i \leq \bar{x} \leq x_{i+1},$$

(4.7)
where \( x_i \) and \( x_{i+1} \) are fixed grid points, and \( \Delta x \equiv x_{i+1} - x_i \). Then a simple layer method algorithm for a 1D diffusion problem is

\[
f(t_{n+1}, x_j) = \frac{1}{2} [f(t_n, x_1) + f(t_n, x_2)], \quad j = 1, 2, ..., N - 1, \quad (4.8)
\]

\[
\bar{x}_1 = x_j + b\Delta t + \sigma \sqrt{\Delta t}, \quad (4.9)
\]

\[
\bar{x}_2 = x_j + b\Delta t - \sigma \sqrt{\Delta t}, \quad (4.10)
\]

\[
f(t_0, x_j) = g(t_0, x_j), \quad (4.11)
\]

with \( f(t_n, \bar{x}_{1,2}) \) calculated using equation (4.7). We see from equations (4.7) to (4.11) that negative values of \( f \) cannot arise from this procedure. A proof that the one-step error of the above algorithm is \( O(\Delta t^2) \) is given in Appendix B. Also we show the connection between layer methods with bilinear interpolation and conventional finite difference methods in Appendix C for a simple 2D diffusion equation without cross diffusion.

Dirichlet and Neumann boundary conditions can be implemented in a similar way as described in Chapter 3. For example, if we have a Dirichlet boundary \( f(t, x_0) = g_0(t, x_0) \), we replace \( \bar{x} \) by \( x_0 \) if \( \bar{x} < x_0 \); if we have a Neumann boundary \( \partial f / \partial x(t, x_N) = 0 \), then we replace \( \bar{x} \) by \( 2x_N - \bar{x} \) if \( \bar{x} > x_N \). For more general ways of handling boundary conditions, we refer readers to Milstein and Tretyakov [2001, 2002].
4.3 Application

In this section, we apply the layer method described in section 4.2 to the bounce-averaged equatorial pitch angle \( \alpha_0 \) and momentum \( p \) diffusion equation in the radiation belts

\[
\frac{\partial f}{\partial t} = \frac{1}{Gp} \frac{\partial}{\partial \alpha_0} G \left( D_{\alpha_0 \alpha_0} \frac{\partial f}{\partial \alpha_0} + D_{\alpha_0 p} \frac{\partial f}{\partial p} \right) + \frac{1}{Gp} \frac{\partial}{\partial p} G \left( D_{\alpha_0 p} \frac{\partial f}{\partial \alpha_0} + D_{pp} \frac{\partial f}{\partial p} \right),
\]

(4.12)

where \( D_{\alpha_0 \alpha_0}, D_{\alpha_0 p} \) and \( D_{pp} \) are bounce-averaged pitch-angle, mixed and momentum diffusion coefficients [Albert, 2004]. Here \( G \) is a Jacobian factor, \( G = p^2 T(\alpha_0) \sin(\alpha_0) \cos(\alpha_0) \), and \( T(\alpha_0) \approx 1.30 - 0.56 \sin(\alpha_0) \) is the normalized bounce period. Initial and boundary conditions are chosen to be the same as in Albert and Young [2005] and Chapter 3. Thus the initial flux is

\[
j(t = 0) = \exp[-(E - 0.2)/0.1][\sin(\alpha_0) - \sin(\alpha_0L)],
\]

(4.13)

where the loss cone angle \( \alpha_0L = 5^\circ \) and flux \( j \) is related to phase-space density \( f \) by \( j = f/p^2 \). Boundary conditions are

\[
f|_{\alpha_0 = \alpha_0L} = 0,
\]

(4.14)

\[
\frac{\partial f}{\partial \alpha_0} \bigg|_{\alpha_0 = 90^\circ} = 0,
\]

(4.15)

\[
f|_{E = E_{\max}} = 0,
\]

(4.16)

\[
f|_{E = E_{\min}} = j(t = 0)|_{E = E_{\min}} \Big/ p_{\min}^2
\]

(4.17)

where \( E_{\min} = 0.2 \) MeV and \( E_{\max} = 5 \) MeV, and \( p_{\min} \) is the momentum corresponding to \( E_{\min} \) [Albert and Young, 2005].
We write a code using the layer method to solve the diffusion equation (4.12). Discretize $\alpha_0$ and $y \equiv \log E$ equidistantly into $N_{\alpha_0}$ and $N_y$ grid cells, and thus $\Delta \alpha_0 = (\pi/2 - \alpha_{0L})/N_{\alpha_0}$ and $\Delta y = (y_{\text{max}} - y_{\text{min}})/N_y$. The equation we use to solve $f$ is

$$f(t_{n+1}, \alpha_{0i}, y_j) = \frac{1}{4} \sum_{i,j} f(t_n, \tilde{\alpha}_{0i}(\Delta W_1, \Delta W_2), \tilde{y}_j(\Delta W_1, \Delta W_2)), \tag{4.18}$$

where

$$\tilde{\alpha}_{0i} = \alpha_{0i} + b_{\alpha_0} \Delta t + \sigma_{11} \Delta W_1 + \sigma_{12} \Delta W_2, \tag{4.19}$$

$$\tilde{p}_j = p_j + b_p \Delta t + \sigma_{21} \Delta W_1 + \sigma_{22} \Delta W_2, \tag{4.20}$$

and $\tilde{y}_j$ is then obtained from $\tilde{p}_j$. If $\tilde{\alpha}_{0i} < \alpha_{0L}$ or $\tilde{\alpha}_{0i} > \pi/2$, we replace $\tilde{\alpha}_{0i}$ by $\alpha_{0L}$ or $\pi - \tilde{\alpha}_{0i}$, respectively. If $\tilde{p} < p_{\text{min}}$ or $\tilde{p} > p_{\text{max}}$, we set $\tilde{p} = p_{\text{min}}$ or $\tilde{p} = p_{\text{max}}$, respectively.

In equations (4.18) to (4.20), $\Delta W_1$ and $\Delta W_2$ each take value $\pm \sqrt{\Delta t}$, and the summation in equation (4.18) sums over four different combinations of $(\Delta W_1, \Delta W_2)$. The functions $b$ and $\sigma$ are the same as in Chapter 3:

$$b_{\alpha_0} = \frac{1}{Gp} \frac{\partial}{\partial \alpha_0} \left( \frac{GD_{\alpha_0 \alpha_0}}{p} \right) + \frac{1}{G} \frac{\partial}{\partial p} \left( \frac{GD_{\alpha p}}{p} \right), \tag{4.21}$$

$$b_p = \frac{1}{Gp} \frac{\partial}{\partial \alpha_0} \left( GD_{\alpha p} \right) + \frac{1}{G} \frac{\partial}{\partial p} \left( GD_{pp} \right), \tag{4.22}$$

$$\sigma_{11} = \sqrt{2D_{\alpha_0 \alpha_0}/p}, \tag{4.23}$$

$$\sigma_{12} = 0, \tag{4.24}$$

$$\sigma_{21} = \sqrt{2D_{\alpha p}/\sqrt{D_{\alpha_0 \alpha_0}}}, \tag{4.25}$$

$$\sigma_{22} = \sqrt{2D_{pp} - \sigma_{21}^2}. \tag{4.26}$$

With the choice of bilinear interpolation to obtain $f(t_n, \tilde{\alpha}_{0i}, \tilde{y}_j)$ in equation (4.18) from
### Table 4.1: Local time sector (LTS) distribution, latitudinal distribution (|\( \lambda \)|) of the waves, equatorial ratio of electron plasma to gyro frequencies (\( f_{pe}/f_{ce} \)) and wave magnetic field amplitude of wave models (\( B_w \)) \cite{Horne2005} and \cite{Li2007}

| LTS (MLT)          | |\( \lambda \)| | \( f_{pe}/f_{ce} \) | \( B_w \) |
|---------------------|-----------------|----------------|-------------------|----------|
| \textit{Horne et al. [2005]} | 2300-0600 | 0 to 15° | 3.43 | 50 pT |
|                     | 0600-1200 | 15° to 35° | 4.0 | 100 pT |
|                     | 1200-1500 | 10° to 35° | 6.72 | 50 pT |
| \textit{Li et al. [2007]} | 0000-0600 | 0° to 15° | 3.8 | 50 pT |
|                     | 0600-1200 | 0° to 35° | 4.6 | \( 10^{0.75+0.04\lambda} \) pT |

its neighboring grid points, the above algorithm has a global error of \( O(\Delta t) \) when \( \Delta \alpha = c_\alpha \Delta t, \Delta y = c_y \Delta t \), where \( c_\alpha \) and \( c_y \) are two constants \cite{Milstein2002}.

#### 4.3.1 Comparison with Albert and Young [2005] results

To show that layer methods can be used to solve the diffusion equation (4.12), we compare results calculated using our layer code with results of \cite{Albert2005} using the same diffusion coefficients. The diffusion coefficients are calculated using the \textit{Horne et al. [2005]} storm time chorus wave model at \( L = 4.5 \). Wave parameters are shown in Table 4.1 for a comparison with the \textit{Li et al. [2007]} storm time chorus wave model (see section 4.3.2).

We choose \( \Delta t = 4 \times 10^{-4} \) day to give a relatively small change of \( \Delta \alpha_0 \) and \( \Delta y \), com-
pared with the computational domain. We plot fluxes of 0.5 MeV and 2 MeV at $t = 2 \text{d}$ in Figure 4.1 to show the convergence of solutions with respect to $N_{ao}$ and $N_y$, and this leads to our choice of $N_{ao} = 1400$ and $N_y = 1500$. Comparisons with results from Albert and Young [2005] are shown in Figure 4.2 for $E = 0.5 \text{ MeV}$ (top) and $E = 2 \text{ MeV}$ (bottom) electrons. Considering the small errors associated with each method, we conclude that the two sets of results agree very well with each other and our layer code is capable of solving the bounce-averaged pitch-angle and momentum diffusion equation (4.12) with cross diffusion. In the next section, we apply our layer code to the diffusion equation with diffusion coefficients calculated using the Li et al. [2007] chorus wave model.

4.3.2 Effects of ignoring cross diffusion in Li et al. [2007] chorus wave model

Li et al. [2007] used a new chorus wave model and calculated changes of electron fluxes due to cyclotron resonances with chorus waves by solving a 2D bounce-averaged pitch angle and energy diffusion equation. However, cross diffusion is not included in their calculation [Li et al., 2007]. In this work, we calculate diffusion coefficients including cross diffusion using the Li et al. [2007] main phase storm time chorus wave model. We also use a wave normal angle distribution from Horne et al. [2005], in contrast to Li et al. [2007], who use a parallel propagation approximation. The resulting diffusion coefficients are shown in Figure 4.3. By solving the diffusion equation with off-diagonal terms using our layer code, we show in this section effects of ignoring off-diagonal terms on electron fluxes using the
Figure 4.1: Fluxes for (top) $E = 0.5 \text{MeV}$ and (bottom) $E = 2.0 \text{MeV}$ at $t = 2 \text{d}$ with different choices of $N_{\alpha_0}$ and $N_p$, using Horne et al. [2005] chorus wave model. In the top panel the three lines are very close together, in contrast to the larger separations shown in the bottom panel.
Figure 4.2: Comparisons between results obtained from the layer method (solid lines) and the Albert and Young [2005] method (dashed lines) for $E = 0.5$ MeV (upper panel) and $E = 2.0$ MeV (lower panel) at $t = 0.1$ day (blue lines) and $t = 1.0$ day (red lines). Here black lines show the initial condition.
Li et al. [2007] chorus wave model.

For comparison with conclusions in Chapter 3, the boundary and initial conditions are the same as equations (4.13)-(4.17), thus they are different from those in Li et al. [2007]. Figure 4.4 shows color plots of fluxes calculated using the diffusion coefficients from Li et al. [2007] chorus waves at \( t = 0.1, 1 \) and 2 days. Both results with (left column) and without (right column) cross diffusion are shown for comparison. We see from Figure 4.4 that at high energies ignoring cross diffusion overestimates fluxes at lower pitch angles and creates a peak in fluxes around 20°. This can be seen more clearly from Figure 4.5, which shows line plots of fluxes calculated with and without cross diffusion at \( t = 0.1, 1 \) and 2 days for 0.5 MeV and 2 MeV electrons. At \( t = 0.1 \) day, the error caused by ignoring cross diffusion is small for 0.5 MeV electrons at all pitch angles and 2 MeV electrons at high pitch angles. For 2 MeV electrons at low pitch angles, however, the error is about a factor of \(~ 10\). At \( t = 1 \) and 2 days, at 0.5 MeV, ignoring cross diffusion overestimates fluxes at small pitch angles by only a factor of \( 2 \sim 3 \). However, at 2 MeV, ignoring cross diffusion causes an error of about two orders of magnitude at small pitch angles. Thus, similar to results in Chapter 3, ignoring off-diagonal terms has a relatively small effect on fluxes for lower energy electrons at higher pitch angles, but it introduces larger errors for larger energy electrons at lower pitch angles.

To understand the similarity between conclusions obtained here and in Chapter 3, we now discuss features of the Li et al. [2007] and Horne et al. [2005] wave models, whose parameters are listed in Table 4.1. First, both Li et al. [2007] and Horne et al. [2005] wave
Figure 4.3: Inverse time scales in units of $s^{-1}$ from diffusion coefficients calculated using the \textit{Li et al.} [2007] chorus wave model with the wave normal angle distribution from \textit{Horne et al.} [2005]. The last panel shows the sign of the cross diffusion coefficients.
Figure 4.4: Fluxes calculated by the layer code using the Li et al. [2007] chorus wave model at $t = 0.1, 1$ and 2 days. The left column shows fluxes with cross diffusion and the right column without cross diffusion.
Figure 4.5: Fluxes for (top) $E = 0.5\text{ MeV}$ and (bottom) $E = 2.0\text{ MeV}$ at $t = 0.1\text{ day}$ (blue lines), $t = 1\text{ day}$ (green lines) and $t = 2\text{ day}$ (red lines) with and without off-diagonal diffusion terms, calculated using the Li et al. [2007] chorus wave model. Black lines show the initial condition. Dashed lines are results without off-diagonal diffusion coefficients, and solid lines are results with off-diagonal terms.
models have similar latitudinal cutoffs of chorus wave power. For Li et al. [2007]: \(|\lambda| < 35°\) on the dayside, and \(|\lambda| < 15°\) on the nightside, for Horne et al. [2005]: \(15° < |\lambda| < 35°\) in the prenoon sector, \(10° < |\lambda| < 35°\) for the afternoon sector, and \(|\lambda| < 15°\) on the nightside.

We see larger errors at smaller pitch angles with both models. Second, even though the Li et al. [2007] chorus wave model has a dayside wave power increasing with latitude, which gives a more abrupt cutoff at the maximum latitude than the Horne et al. [2005] wave model, the actual values of the wave amplitude are not very different. At \(\lambda = 0°\), the drift-averaged wave amplitude for the Li et al. [2007] wave model is \(6/24 \times 5.6 \text{pT} + 6/24 \times 50 \text{pT} = 13.9 \text{pT}\), while the Horne et al. [2005] model gives \(7/24 \times 50 \text{pT} = 14.6 \text{pT}\). At \(\lambda = 35°\), the drift-averaged wave amplitude for the Li et al. [2007] model in main phase is \(6/24 \times 141.25 \text{pT} + 6/24 \times 50 \text{pT} = 47.8 \text{pT}\), while the Horne et al. [2005] model gives \(6/24 \times 100 \text{pT} + 3/24 \times 50 \text{pT} = 31.25 \text{pT}\). Thus we see that both models assume zero amplitude above 35° latitude and have comparable wave power levels, so it is not too surprising to see similar conclusions on ignoring cross diffusion from the two wave models.

4.3.3 Evolution of electron fluxes using a model of fast magnetosonic waves and hiss

Interactions with fast magnetosonic waves have been recently suggested by Horne et al. [2007] to be a possible important acceleration mechanism. Because these interactions typically only involve the Landau resonance \((n = 0)\), coupling of diffusion in \(\alpha_0\) and \(p\) is expected to be especially important for them. For the wave model given by Horne et al.
[2007], the quasi-linear diffusion coefficients of the magnetosonic waves are non-zero only over a limited range of pitch angle and energy [see Figure 9, Albert, 2008]. Thus we combine the magnetosonic wave model of Horne et al. [2007] outside the plasmasphere with the main phase hiss wave model in plumes from Li et al. [2007].

The MLT averaged diffusion coefficients from magnetosonic waves (60%) and hiss waves (15%) are shown in Figure 4.6. A similar numerical experiment as Figure 4.1 is used to determine that \( N_{ao} = 1400 \) and \( N_y = 1500 \) is necessary to obtain accurate solutions. The resulting evolution of electron fluxes are plotted in Figure 4.7 at \( t = 0.1, 1 \) and 2 day. We see from the left column (results with \( D_{aoP} \)) that the magnetosonic waves can accelerate electrons to MeV on timescales of a day, and the fluxes show a peak around 55°, producing a butterfly distribution, at high energies. Comparing the left column with the right column (results without \( D_{aoP} \)), we see that ignoring cross diffusion overestimates fluxes at larger energies and larger pitch angles (\( > 55^\circ \)), which is different from the effects using the Horne et al. [2005] and Li et al. [2007] chorus wave models. The effects of ignoring cross diffusion can be seen more clearly in Figure 4.8, which shows fluxes versus equatorial pitch angle at \( t = 0.1, 1 \) and 2 days for 0.5 and 2 MeV. We see that for 0.5 MeV electrons, effects of ignoring cross diffusion are small at all pitch angles. However, for 2 MeV electrons, ignoring cross diffusion overestimates fluxes at large pitch angles (\( > 55^\circ \)) by a factor of \( 5 \sim 10 \) at all three times, and by a factor of \( \sim 5 \) at small pitch angles at \( t = 1 \) and 2 days.
Figure 4.6: Inverse time scales in units of $s^{-1}$ from diffusion coefficients calculated using a combined magnetosonic wave [Horne et al., 2007] and hiss wave [Li et al., 2007] model. The last panel shows the sign of the cross diffusion coefficients.

4.4 Summary

In this work, we introduce the layer method, which is based on the SDE method of Chapter 3, to solve multi-dimensional radiation belt diffusion equations. Compared with the SDE
Figure 4.7: Fluxes calculated by the layer code using the MH wave model at $t = 0.1, 1$ and 2 days. The left column shows fluxes with cross diffusion and the right column without cross diffusion.

Method, the layer method is deterministic and more efficient when solutions on a large computational domain are needed for long times. Compared with finite difference methods, the layer methods are less efficient, but generalize to 3D easily and are able to handle
complicated boundary geometries. We apply the layer method to a bounce-averaged pitch-angle and energy diffusion equation and obtain excellent agreement with a previous method [Albert and Young, 2005] using the Horne et al. [2005] chorus wave model. We show that our layer code is able to solve multi-dimensional diffusion equations with cross terms.

We then use the layer code to evaluate effects of ignoring cross diffusion using the Li et al. [2007] chorus wave model, as a comparison with the Horne et al. [2005] chorus wave model used in Chapter 3. The main conclusion is similar to Chapter 3; i.e., ignoring off-diagonal terms produces larger errors at smaller pitch angles and higher energies. We show in section 4.3.2 that this similarity might be due to the fact that both wave models have a latitudinal cutoff at 35° and comparable wave power levels.

In section 4.3.3, we show evolution of electron fluxes using a combined magnetosonic wave [Horne et al., 2005] and hiss wave model [Li et al., 2007]. We show that, despite pitch angle scattering by hiss waves, electrons are energized to MeV in 2 days of simulation. Ignoring cross diffusion overestimates fluxes at larger pitch angles and higher energies, in contrast to the effects of ignoring cross diffusion using the Horne et al. [2005] and Li et al. [2007] chorus wave models. Overall, we conclude that cross diffusion terms are important and should be included when modeling diffusion of electrons in the outer radiation belt.
Figure 4.8: Fluxes for (top) $E = 0.5\,\text{MeV}$ and (bottom) $E = 2.0\,\text{MeV}$ at $t = 0.1\,\text{day}$ (blue lines), $t = 1\,\text{day}$ (green lines) and $t = 2.0\,\text{day}$ (red lines) with and without off-diagonal diffusion terms, calculated using the MH wave model. Black lines show the initial conditions. Dashed lines are results without off-diagonal diffusion coefficients, and solid lines are results with off-diagonal terms.
Chapter 5

Summary and future work

5.1 Summary

This thesis presents research work on a basic theory and numerical methods of modeling of radiation belt dynamics. With the background information and review of previous work given in Chapter 1, the main results of the thesis can be summarized as follows.

In Chapter 2 we developed a non-canonical Hamiltonian theory of relativistic charged particles moving in slowly-varying electromagnetic fields using Lie transform analysis. The guiding-center Lagrangian was obtained from the full Lagrangian using a Lie transform, and the guiding-center equations of motion were then derived. A second Lie transform was used to remove the bounce-phase dependence from the guiding-center Lagrangian to obtain the bounce-center motion and the first-order correction to the second adiabatic invariant. The drift-center equations of motion and the first-order correction to the third...
invariant was obtained by a third Lie transform from the bounce-center Lagrangian. Because of the use of Hamiltonian theory, the resulting equations of motion have conservation properties which are useful for checking numerical accuracy. This work represents the first fully Hamiltonian and relativistic theory of all three adiabatic invariants and it provides a basis for developing theory of interactions between charged particles and waves.

Stochastic modeling of radiation belt dynamics, which results from violation of one or more of the three adiabatic invariants, was studied in Chapter 3. We showed the SDE method to solve diffusion equations and obtained excellent agreement between our 2D SDE code and Albert and Young [2005] results. We then used the SDE code to explore effects of ignoring cross diffusion and oblique waves using the Horne et al. [2005] chorus wave model. We showed that both simplifications lead to large errors of fluxes for 2 MeV electrons at small pitch angles.

In Chapter 4, we presented the layer method, which is based on the SDE method, to solve multi-dimensional diffusion equations. The layer method is deterministic and more efficient than the SDE method when solutions on large number of grid points are needed for long times. Using the layer method, we further explored effects of ignoring cross diffusion using a Li et al. [2007] chorus wave model and a combined magnetosonic [Horne et al., 2007] and hiss [Li et al., 2007] wave model. We showed that for the chorus wave model, cross diffusion is more important at smaller pitch angles and higher energies, while for the MH wave model, the cross diffusion is more important at higher pitch angles and higher energies. We concluded from both Chapter 3 and 4 that cross diffusion can be important
and should be included to increase model accuracy.

5.2 Discussion and future work

This thesis has focused on time-dependent 2D modeling of quasi-linear diffusion in radiation belts. The first extension would be to go from 2D to 3D, with cross diffusion terms included. Varotsou et al. [2005] and Varotsou et al. [2008] have presented results from 3D modeling where radial diffusion is added to a bounce-averaged pitch angle and momentum (or energy) diffusion equation and the resulting 3D diffusion equation is solved without any cross diffusion. We have shown in this thesis that the off-diagonal term $D_{oop}$ is important in 2D; we expect it to be also important in 3D.

To incorporate off-diagonal terms in 3D modeling, the first question is which numerical method to use. Compared with finite difference methods, common advantages of the SDE and layer methods are that they do not generate negative phase space densities, they can handle different boundary geometry, and they are easy to code and generalize to 3D. The SDE method is very efficient when solutions on a small number of grid points are needed, especially in high dimensions. This property makes the SDE method very efficient to calculate phase space densities along trajectories of spacecraft. Also the SDE method has a parallelization efficiency close to 1. The layer method, on the other hand, is deterministic and much faster than the SDE method when solutions on many phase space points are needed. However, the layer method is still less efficient compared with finite difference methods due to the use of large number of grid points needed to reduce numerical error.
The main disadvantage of standard finite difference methods, shown by Albert [2009], is that they might produce negative phase space densities. However, it might still be helpful to revisit finite difference methods to examine the region of negative phase space densities. Specifically, if the region of negative phase space densities is limited to areas where phase space densities are close to zero, and it is expected that solutions at other areas are accurate because the methods are probably convergent (meaning they are stable and result in numerical solutions that approach the analytic solution in limit $\Delta t \rightarrow 0$), then using finite difference methods might be useful.

Implementation of a numerical method requires a choice of phase space coordinates. Current 3D modeling is done in $(\alpha_0, E, L)$ coordinates by superposing radial diffusion to a bounce averaged pitch angle and energy diffusion equation [Shprits et al., 2008b]. The 3D diffusion equation is solved by an operator splitting method in three steps because the radial diffusion part assumes constant $(\mu, J)$ instead of constant $(\alpha_0, E)$: First, the 2D bounce-averaged pitch-angle and energy diffusion equation is stepped in time in the $(\alpha_0, E)$ coordinates. Secondly, the new $(\alpha_0, E)$ coordinates are transformed to $(\mu, J)$ space. Thirdly, the L coordinate is stepped in time at constant $(\mu, J)$, which are then transformed back to $(\alpha_0, E)$ space. This approach is complicated because of the coordinate transformation and because the interpolations used during the transformation introduce more numerical errors. Also it has been suggested recently that in an asymmetric magnetic field, $\alpha_0$ might be dependent on local time itself and thus is not a suitable coordinate [Jay Albert, personal communication, 2009].
Another approach to 3D modeling is to start from first principles and derive the three-dimensional diffusion in three adiabatic invariants \((\mu, J, L)\). This approach has been taken by Brizard and Chan [2004] using \((\mu, E, L)\) coordinates for symmetric fields. Extension from Brizard and Chan [2004] to asymmetric fields might give us expressions of \(D_{\mu L}\) and \(D_{JL}\). In an asymmetric magnetic field, violation of \(\mu\) or \(J\) adiabatic invariant could cause change of particle’s \(L\) value, because of drift-shell splitting [Schulz and Lanzerotti, 1974]. Thus, including \(D_{\mu L}\) and \(D_{JL}\) in the calculation might be important to reduce radiation belt modeling error. Although there are advantages to use \((\mu, J, L)\) coordinates, there are also some difficulties. There are no codes available to evaluate diffusion coefficients of the sort derived by Brizard and Chan [2004]. Also, boundary conditions are usually specified on straight lines in \((\alpha_0, E)\) space, and a straight line in \((\alpha_0, E)\) space would become a curve in \((\mu, J)\) space. Thus it might be harder to handle boundary conditions if using \((\mu, J, L)\) coordinates. Further research is needed to better understand the advantages and disadvantages of different choices of phase-space coordinates.

To increase our model’s accuracy, better wave models and initial and boundary conditions are needed. For example, our calculations of diffusion coefficients \(D\) depend on the wave models we specified; comparing Figures 4.4 and 4.7, it is easy to see that errors could be large if inaccurate wave models are used. Further work should be done with improved wave models as they become available. Also, effects of different initial conditions and boundary conditions, such as relativistic kappa-type functions [Xiao et al., 2008], should be considered. Using available ring current codes (e.g., the Rice Convection Model
[Toffoletto et al., 2003]) to provide dynamic low energy boundary conditions would make radiation belt diffusion codes more realistic.

The quasi-linear theory was developed to describe interactions between electrons and small-amplitude broadband waves. Specifying the parameter range (e.g., wave amplitude and/or frequency spectrum) for validity of quasi-linear theory is important. Nonlinear interactions between electrons and large amplitude or narrow band waves are under intense research, e.g., Bell [1984]; Albert [1993, 2000, 2002]; Bortnik et al. [2008] and Furuya et al. [2008]. Assessing the importance of nonlinear interactions and possibly including them in radiation belt modeling are important future work.

The relativistic Hamiltonian adiabatic theory developed in this thesis is helpful to understanding the theoretical basis of radiation belt modeling. The SDE and layer methods presented are promising in building 3D global radiation belt diffusion models. With more advanced plasma wave models, improved initial and boundary conditions and a 3D diffusion code, it is possible to build a 3D global radiation belt model whose output can be compared with observations from upcoming spacecraft missions. The next several years are likely to be continue to be an exciting and active period in radiation belt research.
Appendix A

Time-forward SDE method

Gardiner [1985] shows that the Fokker-Planck diffusion equation

\[ \frac{\partial f}{\partial t} = -\sum_{i=1}^{d} \frac{\partial}{\partial x_i} [b_i(t,x)f] + \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i \partial x_j} [a_{ij}(t,x)f] \]  
(A.1)

corresponds to a system of stochastic differential equations

\[ dX = b(t,x)dt + \sigma dW(t), \]  
(A.2)

where \( \sigma \sigma^T = a \). The SDEs describe changes of \( x \) with time, and they can be used to solve the Fokker-Planck diffusion equation (A.1) as described in Chapter 3; we call the method "the time-forward" SDE method. To use the time-forward SDE method to solve the bounce-averaged pitch-angle and energy diffusion equation (3.10), we first set \( F = Gf \) and write the diffusion equation in the following form

\[ \frac{\partial F}{\partial t} = \frac{\partial^2}{\partial \alpha_0^2} \left( \frac{D_{\alpha_0 \alpha_0}}{p^2} F \right) + 2 \frac{\partial^2}{\partial \alpha_0 \partial p} \left( \frac{D_{\alpha p}}{p} F \right) \\
+ \frac{\partial^2}{\partial p^2} (D_{pp}F) - \frac{\partial}{\partial \alpha_0} (b_{\alpha_0} F) - \frac{\partial}{\partial p} (b_p F), \]  
(A.3)
where $b_{ao}$ and $b_p$ are defined in equations (3.16) and (3.17). Thus the time-forward stochastic differential equations corresponding to equation (A.3) are [Alanko-Huotari et al., 2007; Yamada et al., 1998; Qin et al., 2005]

\begin{align}
\frac{dA_0(t)}{dt} &= b_{ao}(t, A_0, P) dt + \sigma_{11} dW_1 + \sigma_{12} dW_2, \\
\frac{dP(t)}{dt} &= b_p(t, A_0, P) dt + \sigma_{21} dW_1 + \sigma_{22} dW_2,
\end{align}

where the components of the matrix $\sigma$ are also defined by equations (3.21) to (3.23).

Equations (A.4) and (A.5) are solved to give changes of particle coordinates $(a_0, p)$ using the Euler-Maruyama method [Kloeden and Platen, 1992]; i.e.,

\begin{align}
A_0(t_{n+1}) &= A_0(t_n) + b_{ao}[t_n, A_0(t_n), P(t_n)] \Delta t \\
&\quad + \sigma_{11}(t_n) \Delta W_1 + \sigma_{12}(t_n) \Delta W_2, \\
P(t_{n+1}) &= P(t_n) + b_p[t_n, A_0(t_n), P(t_n)] \Delta t \\
&\quad + \sigma_{21}(t_n) \Delta W_1 + \sigma_{22}(t_n) \Delta W_2.
\end{align}

After a given time period, the distribution of electrons can be obtained.

In Figure A.1, to illustrate the local effects of off-diagonal diffusion coefficients on distributions of particles, we released 9000 particles from $\alpha_0 = 30^\circ, E = 1$ MeV, where $D_{aoP}$ is positive, and $\alpha_0 = 50^\circ, E = 3$ MeV, where $D_{aoP}$ is negative (see Figure 3.1). We chose time periods short enough to ignore boundary effects. The distribution of particles shown in Figure A.1 was obtained after $t = 0.06$ day for $E = 3$ MeV and $t = 0.01$ day for $E = 1$ MeV. We also turned off-diagonal diffusion coefficients on and off to show local effects of ignoring off-diagonal terms. The left panel has $D_{aoP} \neq 0$, and the right panel has
Figure A.1: Local effects of ignoring off-diagonal terms. Lines are contours of particle numbers. Particles are released from $\alpha_0 = 30^\circ$ for $E = 1.0$ MeV and $\alpha_0 = 50^\circ, E = 3$ MeV in both panels. In the left panel, off-diagonal terms are kept. $D_{\alpha_0P}$ is positive at $\alpha_0 = 30^\circ, E = 1.0$ MeV and negative at $\alpha_0 = 50^\circ, E = 3$ MeV. In the right panel, off-diagonal terms are set to zero.

$D_{\alpha_0P} = 0$. We see from Figure A.1 that without $D_{\alpha_0P}$, the local distribution of particles has a shape of an ellipse (as expected for a diagonal diffusion tensor), while with $D_{\alpha_0P}$, this ellipse is tilted, and the tilt direction is determined by the sign of $D_{\alpha_0P}$. With $D_{\alpha_0P}$ positive (as for the $\alpha_0 = 30^\circ, E = 1$ MeV case) the ellipse tilts clockwise, and with $D_{\alpha_0P}$ negative ($\alpha_0 = 50^\circ, E = 3$ MeV), the ellipse tilts counterclockwise. These results are consistent with previous analytical results using Green functions [Albert and Young, 2005].
Appendix B

One-step error of the layer method

To calculate the one-step error of the layer method presented in Chapter 4, consider a 1D initial value problem,

\[ \frac{\partial f}{\partial t} = b(t, x) \frac{\partial f}{\partial x} + \frac{1}{2} \sigma^2(t, x) \frac{\partial^2 f}{\partial x^2}, \quad f(t_0, x) = g(x). \]  

(B.1)

With the layer method, we discretize \( t \) equidistantly into \( t_0, t_1, t_2, \ldots \) with time step \( h \), and the approximate solution of equation (B.1) at \( (t_{k+1}, x) \) is given by

\[ f_{k+1}(x) = \frac{1}{2} f \left[ t_k, x + b(t_{k+1}, x)h + \sigma(t_{k+1}, x)\sqrt{h} \right] \\
+ \frac{1}{2} f \left[ t_k, x + b(t_{k+1}, x)h - \sigma(t_{k+1}, x)\sqrt{h} \right]. \]  

(B.2)

To show the error of \( f_{k+1}(x) \), we expand \( f(t_{k+1} - h, x + bh \pm \sigma \sqrt{h}) \) using Taylor
The resulting equation is

\[ f(t_{k+1} - h, x + bh \pm \sigma \sqrt{h}) = f(t_{k+1}, x) - h \frac{\partial f}{\partial t} \]

\[ + (bh \pm \sigma \sqrt{h}) \frac{\partial f}{\partial x} + \frac{1}{2} (\sigma^2 h \pm 2b\sigma h^{3/2}) \frac{\partial^2 f}{\partial x^2} \]

\[ + (\mp \sigma h^{3/2}) \frac{\partial^2 f}{\partial t \partial x} + (\pm \sigma^3 h^{3/4}) \frac{1}{6} \frac{\partial^3 f}{\partial x^3} + O(h^2). \quad (B.3) \]

Inserting equation (B.3) into (B.2) yields

\[ f_{k+1}(x) = f(t_{k+1}, x) - h \left( \frac{\partial f}{\partial t} - b \frac{\partial f}{\partial x} - \frac{1}{2} \sigma^2 \frac{\partial^2 f}{\partial x^2} \right) + O(h^2). \quad (B.4) \]

Using equation (B.1), equation (B.4) becomes

\[ f_{k+1}(x) = f(t_{k+1}, x) + O(h^2), \quad (B.5) \]

which shows that the approximation \( f_{k+1}(x) \) differs from the true solution \( f(t_{k+1}, x) \) by a term proportional to \( h^2 \).

In practice, we use interpolations to obtain \( f(t_k, x + bh \pm \sigma \sqrt{h}) \) in equation (B.2) from fixed grid points, denoted as \( \tilde{f}(t_k, x + bh \pm \sigma \sqrt{h}) \). E.g., using linear interpolation, we have

\[ \tilde{f}(t_k, x + bh \pm \sigma \sqrt{h}) = f(t_k, x + bh \pm \sigma \sqrt{h}) + O(\Delta x^2). \quad (B.6) \]

Thus with \( \Delta x = c_x h \), where \( c_x \) is a constant, the error from interpolation is also \( O(h^2) \) and equation (B.5) is not changed. Overall the one-step error of the above layer method is \( O(h^2) \). Milstein [2002] shows that the global error (the error accumulated from time \( t_0 \) to \( t \)) of the layer method is \( O(h) \).
Appendix C

Relationship between finite difference methods and layer methods

Consider a 2D diffusion equation with constant diffusion coefficients and no cross diffusion,

\[
\frac{\partial f}{\partial t} = D_{xx} \frac{\partial^2 f}{\partial x^2} + D_{yy} \frac{\partial^2 f}{\partial y^2}.
\] (C.1)

According to the layer method described in Chapter 4, the updated value of \( f(x, y) \) after a timestep \( \Delta t \) is just the average of \( f \) at the four points \( f(x \pm L_x, y \pm L_y, t) \), where \( L_x \equiv \sqrt{2D_{xx}\Delta t} \) and \( L_y \equiv \sqrt{2D_{yy}\Delta t} \) according to the prescription in Section 4.2.2. Use bilinear interpolation on a regular grid with spacing \( (\Delta x, \Delta y) \), and take \( \Delta t \) small enough that \( L_x \leq \Delta x \) and \( L_y \leq \Delta y \) so that the points \( (x \pm L_x, y \pm L_y) \) lie within the neighboring grid cells.
Then the value of \( f \) at grid point \( i, j \) and time \( t_{n+1} \), \( f_{ij}^{n+1} \), is given by

\[
4f_{ij}^{n+1} = r_x r_y f_{i-1,j+1}^n + 2(1-r_x) r_y f_{ij+1}^n + r_x r_y f_{i+1,j+1}^n \\
+ 2r_x (1-r_y) f_{i-1,j}^n + 4(1-r_x)(1-r_y) f_{ij}^n + 2r_x (1-r_y) f_{i+1,j}^n \\
+ r_x r_y f_{i-1,j-1}^n + 2(1-r_x)r_y f_{ij-1}^n + r_x r_y f_{i+1,j-1}^n \tag{C.2}
\]

where \( r_x = L_x/\Delta x \) and \( r_y = L_y/\Delta y \). Note that \( f_{ij}^{n+1} \) is guaranteed to be positive since all the coefficients of \( f_{i\pm 1,j\pm 1}^n \) are non-negative.

On the other hand, the conventional explicit finite difference scheme for equation (C.1) can be written as

\[
4f_{ij}^{n+1} = 2c_x f_{ij+1}^n + 2c_y f_{i-1,j}^n + 4(1-c_x-c_y) f_{ij}^n + 2c_x f_{i+1,j}^n + 2c_y f_{ij-1}^n \tag{C.3}
\]

for timestep \( \Delta t \), where \( c_x = 2D_{xx}\Delta t/(\Delta x)^2 \) and \( c_y = 2D_{yy}\Delta t/(\Delta y)^2 \).

If \( r_x = c_x \) and \( r_y = c_y \), the two schemes come into close agreement. Then the difference between the two expressions for \( 4f_{ij}^{n+1} \) can be recognized as the finite difference expression for \( r_x r_y (\Delta x)^2 (\Delta y)^2 (\partial^4 f/\partial x^2 \partial y^2) \). The corresponding difference in \( \partial f/\partial t \) is

\[
D_{xx} D_{yy} \Delta t \frac{\partial^4 f}{\partial x^2 \partial y^2}. \tag{C.4}
\]

This extraneous term vanishes as \( \Delta t \to 0 \).

The conditions \( r_x = c_x \) and \( r_y = c_y \) are equivalent to \( 2D_{xx}\Delta t/(\Delta x)^2 = 1 \) and \( 2D_{yy}\Delta t/(\Delta y)^2 = 1 \), respectively. Then the combination \( 2D_{xx}\Delta t/(\Delta x)^2 + 2D_{yy}\Delta t/(\Delta y)^2 \) has the value 2, while the CFL stability criterion for the explicit scheme (for the original diffusion equation) requires it to be less than one. Thus for the simple case discussed above,
the layer method may be interpreted as an explicit scheme run at an unstably large timestep, but stabilized by the small extra term of $\mathcal{O}(\Delta t)$, which is the same order as the error of both methods. This is analogous to grid diffusivity terms added to finite difference schemes in the Lax method.
Bibliography


Albert, J. M., Evaluation of quasi-linear diffusion coefficients for EMIC waves in a multi-

Albert, J. M., Using quasi-linear diffusion to model acceleration and loss from wave-

Albert, J. M., Evaluation of quasi-linear diffusion coefficients for whistler mode waves


Albert, J. M., Efficient approximations of quasi-linear diffusion coefficients in the radiation


Albert, J. M., and S. L. Young, Multidimensional quasi-linear diffusion of radiation belt

Albright, B. J., D. Winske, D. S. Lemons, W. Daughton, and M. E. Jones, Quiet direct

Baker, D. N., S. Kanekal, J. B. Blake, B. Klecker, and G. Rostoker, Satellite anomalies
linked to electron increase in the magnetosphere, *Eos Trans. AGU*, 75(34), 401, 1994.


Horne, R. B., and R. M. Thorne, Potential waves for relativistic electron scattering and


Khazanov, G. V., and K. V. Gamayunov, Effect of oblique electromagnetic ion cyclotron waves on relativistic electron scattering: Combined Release and Radiation Ef-


Shprits, Y. Y., D. A. Subbotin, N. P. Meredith, and S. R. Elkington, Review of modeling of


