Development and Application of Stochastic Methods for Radiation Belt Simulations

by

Liheng Zheng

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

Approved, Thesis Committee:

Anthony A. Chan, Chair
Professor of Physics and Astronomy

Richard A. Wolf
Professor Emeritus of Physics and Astronomy

Brandon Dugan
Associate Professor of Earth Science

Houston, Texas

May, 2015
ABSTRACT

Development and Application of Stochastic Methods for Radiation Belt Simulations

by

Liheng Zheng

This thesis describes a method for modeling radiation belt electron diffusion, which solves the radiation belt Fokker-Planck equation using its equivalent stochastic differential equations, and presents applications of this method to investigating drift shell splitting effects on radiation belt electron phase space density.

The theory of the stochastic differential equation method of solving Fokker-Planck equations is formulated in this thesis, in the context of the radiation belt electron diffusion problem, and is generalized to curvilinear coordinates to enable calculation of the electron phase space density as a function of adiabatic invariants $M$, $K$ and $L$. Based on this theory, a three-dimensional radiation belt electron model in adiabatic invariant coordinates, named REM (for Radbelt Electron Model), is constructed and validated against both known results from other methods and spacecraft measurements. Mathematical derivations and the essential numerical algorithms that constitute REM are presented in this thesis.

As the only model to date that can solve the fully three-dimensional diffusion problem, REM is used to study the effects of drift shell splitting, which gives rise to $M$-$L$ and $K$-$L$ off-diagonal terms in the radiation belt diffusion tensor. REM simulation results suggest that drift shell splitting reduces outer radiation belt electron phase space density enhancements during electron injection events. Plots of the phase
space density sources, which are unique products of the stochastic differential equation method, and theoretical analysis further reveal that this reduction effect is caused by a change of the phase space location of the source to smaller $L$ shells, and has a limit corresponding to two-dimensional local diffusion on a curved surface in the $(M, K, L)$ phase space.
Acknowledgements

I owe my deepest gratitude to my advisor, Prof. Anthony Chan, for his guidance and support. He leads the research project with great enthusiasm and is always patient in advising me. Our inspiring discussions had constantly encouraged me to challenge myself in research, and honed this work toward mathematical rigorousness. It is my great honor and pleasure to work with him. I would like to thank Prof. Richard Wolf and Prof. Rui-Rui Du for their reviewing and comments on my research and progress. I am grateful to Prof. Brandon Dugan for his serving as my thesis committee member.

My sincere thanks go to Prof. Matthew Baring and Prof. Frank Toffoletto. This work has benefited a lot from their lectures. I would like to express my gratitude to Dr. Xin Tao, my former fellow student and now professor at University of Science and Technology of China, for his early instructions on the SDE method and his pioneering work that formed the starting point of this thesis. I thank all the research collaborators who had contributed to my thesis work. Among them, I am especially grateful to Dr. Paul O’Brien, Dr. Weichao Tu and Dr. Jay Albert, for their unselfish help and insightful thoughts that made this work possible.

I owe my appreciation to Ms. Umbelina Cantú and Ms. Rosa Almendarez for their excellent administrative work and wholehearted assistances to me. I feel lucky to have known Ms. Cantú since the very early years of my coming to Rice University. She has always been a great source of help and relief, and she is more like a friend, a family, in addition to an institute administrator.

Last but not least, I wish to name my closest friends in the United States who had made me forget the loneliness of studying seven thousand miles away from home. They are: Yong Cui, Tingting Gu, Wanding Zhou, Xia Wang, Zan Yang, Zhechao
“Merry” Ruan, Jian Yang and Yi Guo. I am very fortunate to have met my girlfriend and violinist Xinze Liu. The beautiful melodies that flew out from her violin aroused many ideas in this work. Finally, I am indebted to my parents Xiuzhen Wang and Xiang Zheng. They raised me and poured me with unconditional love, and had always been a strong support that backed me up through the journey of PhD. This thesis is devoted to them.
Contents

Abstract ii
List of Illustrations ix
List of Tables xi

1 Introduction 1
  1.1 Magnetosphere and radiation belts ............................. 1
  1.2 Electron adiabatic motion ....................................... 5
  1.3 Important magnetospheric plasma waves ......................... 10
  1.4 Kinetic description of the radiation belt electrons ............ 14
  1.5 Review of previous work ....................................... 20
  1.6 Thesis organization ............................................ 22

2 Solving Fokker-Planck Equations with the SDE Method 24
  2.1 The Wiener process and Itō stochastic calculus ................. 24
  2.2 Relation between the SDE and the Fokker-Planck equation ..... 28
  2.3 The SDE method from the viewpoint of transition probabilities:
    generalization to curvilinear coordinates ........................ 34
    2.3.1 Cartesian coordinates .................................... 35
    2.3.2 Curvilinear coordinates ................................. 38
    2.3.3 Conclusion ............................................ 42
  2.4 Solution of general diffusion equations: the Feynman-Kac formula 43

3 The Radbelt Electron Model 52
  3.1 Solving the radiation belt Fokker-Planck equation in REM ....... 52
3.1.1 Transformation of the Fokker-Planck equation and connection to SDEs ........................................ 52
3.1.2 Transformation matrix in dipole field ........................................ 56
3.1.3 Computational domain .................................................. 57
3.1.4 Boundary conditions ..................................................... 59
3.2 Numerical techniques ...................................................... 64
3.2.1 Overview ............................................................... 64
3.2.2 Matrix factorization ..................................................... 67
3.2.3 The adaptive time step Euler-Maruyama scheme ....................... 72
3.2.4 Estimation of statistical error .......................................... 75
3.3 PSD source plots .......................................................... 77
3.4 Summary and discussion .................................................... 78

4 REM Validations and Application to the October 2002 Storm .... 80
4.1 REM validations with known solutions .................................... 80
4.2 Application of REM to the October 2002 storm ......................... 82
4.3 Summary and discussion .................................................... 91

5 Effects of Drift Shell Splitting on Radiation Belt Electron Diffusion 93
5.1 Introduction .............................................................. 93
5.2 Simulations of an idealized MeV electron enhancement .................. 95
  5.2.1 Simulation setup ..................................................... 95
  5.2.2 Simulation results ................................................... 98
5.3 Theoretical analysis ...................................................... 104
5.4 Discussion and conclusions .............................................. 115
6 Summary and Future Work

6.1 Summary and conclusions .................................................. 118
6.2 Numerical improvements of REM ....................................... 121
6.3 Future studies ..................................................................... 124

A Divergence Operator and Diffusion Equation in Curvilinear Coordinates .................................................. 128

B An Example REM Solution Statistics Sheet .......................... 134

C Designing Initial and Boundary Conditions for the Radiation Belt Fokker-Planck Equation ......................... 137

D REM Simulations of the 8 October 2012 Storm Event Using Observed Data .................................................. 143

Bibliography ............................................................................. 149
# Illustrations

1.1 The Earth’s magnetosphere ........................................ 2
1.2 Changes of radiation belt electron fluxes in storm time .......... 4
1.3 Radiation belt structures ........................................... 6
1.4 Electron quasi-periodic motions ................................... 7
1.5 Drift shell splitting .................................................. 11
1.6 Magnetospheric plasma waves .................................... 12

2.1 A 2-D domain with a Neumann-type boundary .................... 49

3.1 Example computational domain in dipole field geometry .......... 58

4.1 REM 1-D radial diffusion tests .................................... 81
4.2 REM 2-D chorus-wave diffusion tests ............................. 83
4.3 Initial and boundary conditions of the Oct. 2002 storm simulation . 85
4.4 BAS chorus wave diffusion coefficients ........................... 87
4.5 Simulation results of the Oct. 2002 storm ........................ 88
4.6 1-D PSD source plot for the Oct 2002 storm simulation .......... 90

5.1 *O'Brien* [2015] drift shell splitting chorus wave diffusion coefficients . 96
5.2 The POES factors .................................................... 99
5.3 REM simulated PSD radial profiles with drift shell splitting ....... 100
5.4 PSD source plots for the $t = 17$ hour solutions at $M = 2000 \text{ MeV/G}$, $K = 0.01 \text{ G}^{1/2}\text{R}_E$ and $L = 4.4$ .............................................. 102
5.5 PSD source plots for the $t = 17$ hour solutions at $M = 2000 \text{ MeV/G}$, $K = 0.01 \text{ G}^{1/2}\text{R}_E$ and $L = 5.4$ .............................................. 103
5.6 PSD differences versus peak POES factors ........................................... 105
5.7 PSD ratios versus peak POES factors ................................................. 106
5.8 Schematic illustration of bounce-averaged diffusion surfaces ............... 109
5.9 PSD source plots for the $t = 17$ hour solutions at $M = 2000 \text{ MeV/G}$, $K = 0.01 \text{ G}^{1/2}\text{R}_E$ and $L = 5.4$ (2-D diffusion) ......................... 111
5.10 PSD source plots for the $t = 17$ hour solutions at $M = 2000 \text{ MeV/G}$, $K = 0.01 \text{ G}^{1/2}\text{R}_E$ and $L = 4.4$ (2-D diffusion) ......................... 112
5.11 REM simulated flux profiles in generalized $\alpha_0$ with drift shell splitting 114
6.1 REM simulated flux radial profiles .................................................... 127
A.1 The tangent- and reciprocal-basis vectors ....................................... 130
C.1 Fitted electron fluxes to the 8 Oct. 2012 storm ................................ 142
D.1 PSD radial profiles at $M = 1279 \text{ MeV/G}$, $K = 0.115 \text{ G}^{1/2}\text{R}_E$ .... 144
D.2 PSD source plots for the $t = 17$ hour solutions at $M = 1279 \text{ MeV/G}$, $K = 0.115 \text{ G}^{1/2}\text{R}_E$ and $L = 4.25$ ................................. 145
D.3 PSD radial profiles at $M = 3433 \text{ MeV/G}$, $K = 0.115 \text{ G}^{1/2}\text{R}_E$ .... 147
D.4 PSD source plots for the $t = 17$ hour solutions at $M = 3433 \text{ MeV/G}$, $K = 0.115 \text{ G}^{1/2}\text{R}_E$ and $L = 4.25$ ................................. 148
Tables

B.1 An example REM solution statistics sheet . . . . . . . . . . . 136
Chapter 1

Introduction

This dissertation describes the mathematical formulation and numerical techniques that constitute the Radbelt Electron Model (REM), a stochastic model of the Earth’s radiation belt electrons, and the application of this model to investigating the effects of an asymmetric geomagnetic field on radiation belt dynamics. The purpose of this research is to provide a new radiation belt electron model that is more general than existing models, and through which to advance understanding of physical processes that are responsible for changes of electron intensities in the radiation belts. The increase of these relativistic electrons is potentially hazardous to many space infrastructures that modern society heavily relies on, such as the navigation constellations on medium Earth orbits and the geostationary satellites [Baker et al., 1994, 1997], and to prospective human space expeditions that travel through the radiation belts to the Moon and Mars.

In this chapter, I will briefly introduce some background information, give an overview of previous work, and outline the organization of this dissertation.

1.1 Magnetosphere and radiation belts

The Earth is, among other things, a giant magnet. Its magnetic field, which is dominated by an intrinsic dipole approximately at the center of the Earth, is subjected continuously to a plasma flow (a flow of freely moving charged particles) directed
outward from the Sun, known as the solar wind [Parker, 1958]. The geomagnetic field diverts the solar wind, and in return is compressed in the dayside and stretched into a long tail in the nightside, as sketched in Figure 1.1; the boundary of their interactions is called the magnetopause. Inside the magnetopause is a region of space that is dominated by the geomagnetic field and contains the geomagnetically influenced charged particles, mainly protons and electrons. This region, together with the magnetized plasma within, is named the magnetosphere.

Figure 1.1: Schematic illustration of the Earth’s magnetosphere, including ring current, plasmasphere (cold, dense plasma near Earth), plasma sheet, and radiation belts. Shaded areas indicate the region of geomagnetically trapped particles. Figure adopted from http://galileo.ftecs.com/stone-diss/fig1-1.jpg in April 2015.

In addition to being a large container of charged particles, the magnetosphere is
a very dynamic entity in its own right. The Sun exerts its influence on the magnetosphere through the highly supersonic (approximately Mach 8, as compared to the magnetosonic wave speed) magnetized solar wind, and variations of the solar wind cause disturbances in the electric current systems in the magnetosphere [e.g., Kivelson and Russell, 1995]. These disturbances, which are observable on the ground as perturbations in the induced magnetic field, are known as the geomagnetic activity. To quantify the geomagnetic activity, ground perturbations recorded from a global network of magnetic observatories are averaged and categorized into an integral quasilogarithmic $Kp$ index [Bartels et al., 1939], with 0 the minimum, corresponding to the quietest magnetospheric conditions, and 9 the maximum, corresponding to the largest fluctuations.

Among the complicated magnetospheric current systems, the ring current that encircles the Earth at about 3 to 5 Earth radii ($R_E$) near the magnetic equator, as indicated in Figure 1.1, is of particular importance to the inner magnetosphere. Formed by low energy ($10 \sim 10^2$ keV) particles convecting from the nightside to the dayside, with protons circumventing the Earth from the dusk side and electrons from the dawn side, the ring current is effectively a westward circular current that induces a southward magnetic component near the magnetic equator on the ground. This induced magnetic field is indexed as $Dst$ (in unit of nT), whose value is usually negative, meaning a southward perturbation, and the magnitude is proportional to the total energy of the drifting particles in the ring current [Dessler and Parker, 1959]. During magnetic storms, which are large-scale geomagnetic events caused by abrupt increases in the solar wind speed or particle density, the ring current is strengthened as a result of enhanced magnetospheric convection, and the value of $Dst$ usually dips below $-100$ nT (see Figure 1.2(b)).
Figure 1.2: (a) Hourly averaged electron flux variation at geostationary Earth orbit (GEO) for 2–8 November 1993, measured by the LANL spacecraft 1984-129 for four high-energy channels and (b) $D_{st}$ variation for the same period. Figure reprinted from Kim and Chan [1997].
It was deep within the magnetosphere that the Earth’s radiation belts, consisting of energetic \((10^{-1} \sim 10 \text{ MeV})\) particles trapped by the geomagnetic field, were discovered by Van Allen et al. [1958] and Vernov et al. [1959] at the very beginning of the space age. Typically, the radiation belts are two toroidal regions in space [e.g., Baker et al., 2014], as schematically illustrated in Figure 1.3(A). The inner belt is relatively stable [Schulz and Lanzerotti, 1974]. It is formed by protons and electrons that are primarily born through the decay of albedo neutrons ejected from the upper atmospheric gas atoms by the colliding solar protons and cosmic-ray particles [e.g., Preszler et al., 1972]. The outer belt mainly consists of electrons, and is on the other hand very dynamic in response to geomagnetic activity. During magnetic storms, the outer belt electron fluxes can vary by several orders of magnitude in time scales less than a day (see Figure 1.2) [e.g., Kim and Chan, 1997; Reeves et al., 2003], and in some peculiar cases, even an entirely new belt can arise [Baker et al., 2013], as shown schematically in Figure 1.3(B). The sources of the outer belt electrons and the underlying physics involved in their dynamical processes are of major importance in radiation belt research, and are the focus of the research in this thesis.

1.2 Electron adiabatic motion

Electrons (and protons) trapped in the radiation belts display a hierarchy of three distinctive quasi-periodic motions with well separated periods [e.g., Roederer, 1970; Schulz and Lanzerotti, 1974]. An example of a charged particle trajectory is sketched in Figure 1.4. The most rapid motion in this hierarchy is a helical “cyclotron” motion about a magnetic field line, and the instantaneous center of the cyclotron orbit is named the guiding center. An average over cyclotron motion reveals a bounce motion of the guiding center between the conjugate mirror points along a guiding field line.
Similarly, an average over bounce motion reveals an eastward (westward for protons) drift motion of the bouncing trajectory around the Earth. Hence, the trajectory of a guiding center generates a shell, known as the drift shell, encircling the Earth. The phrase “quasi-periodic” indicates that the periods of these motions are not exactly fixed, but they vary slightly up and down along the trajectory of motion. For an electron with 1 MeV kinetic energy ($E$) on a guiding dipole field line that intersects the magnetic equatorial plane at 4 $R_E$, the three periods are in the orders of $10^{-3}$, 1, and $10^3$ seconds respectively.

The three quasi-periodic motions are associated with three action integrals that are invariant during the course of motion, as long as the forces (magnetic, electric, gravitational, etc.) seen by the particle remain virtually constant in time over the respective quasi-periodic orbit. They are hence called adiabatic invariants. In Hamiltonian mechanics [e.g., Landau and Lifshitz, 1976; Goldstein, 1980; Schulz and Lanzerotti, 1974], the action integral $J_i$ ($i = 1, 2, 3$) is defined as a path integral of the
canonical momentum $\mathbf{P}$ over a quasi-periodic orbit, i.e.,

$$ J_i = \oint_i \mathbf{P} \cdot d\mathbf{l} = \oint_i \left( \mathbf{p} - \frac{e}{c} \mathbf{A} \right) \cdot d\mathbf{l}, \quad (1.1) $$

in which $\mathbf{p}$ is the particle mechanical momentum, $e$ is the absolute value of electron charge, $c$ is the speed of light in vacuum, $\mathbf{A}$ is the magnetic vector potential, and $d\mathbf{l}$ is the infinitesimal displacement along the orbit. We note that Gaussian units are adopted in this thesis, except that electron-volt (eV) is used for energy. The variables canonically conjugate to $J_i$ are the phase angles $\varphi_i$ of the quasi-periodic motions; they are cyclic coordinates of the particle Hamiltonian if the forces felt by the particle do not vary on time scales of the respective quasi-periodic orbit [Landau and Lifshitz, 1976].

The action integral $J_1$ corresponding to cyclotron motion is

$$ J_1 = \frac{2\pi m_0 c}{e} M, \quad (1.2) $$

where

$$ M = \frac{p^2}{2m_0 B} \quad (1.3) $$
is, apart from constants, equal to $J_1$, and is more commonly referred to as the first adiabatic invariant. In the above equations, $m_0$ is the electron mass, and $p_\perp$ is the component of $p$ perpendicular to the magnetic field $B$ evaluated instantaneously at the guiding center. $M$ has the physical meaning that $M/\gamma$ is the relativistic magnetic moment of the particle gyration, where $\gamma$ is the Lorentz factor. With the aid of (1.3), the action integral $J_2$ associated with bounce motion can be expressed as

$$J_2 = \oint p_\parallel ds = \sqrt{8m_0MK},$$

where $p_\parallel$ is the component of $p$ parallel to $B$, $s$ is the arc-length along the guiding field line, and the quantity

$$K = \int_{s_m}^{s_m'} \sqrt{B_m - B(s)} ds$$

(1.5)

is commonly referred to as the second adiabatic invariant. Unlike $J_2$, $K$ is independent of the particle’s energy. $B_m$ in (1.5) is the magnetic intensity at the mirror points, and the integral is from one mirror point to its conjugate. The action integral of the drift motion can be represented by the magnetic flux $\Phi$ enclosed by the drift shell as

$$J_3 = \frac{e}{c} \Phi.$$  

(1.6)

Conventionally, it is more convenient to define the dimensionless $L$ value (after Roedderer [1970]) of a drift shell as the third adiabatic invariant, by specifying

$$L = 2\pi \left| \frac{\mu_E}{\Phi R_E} \right|,$$

(1.7)

in which $\mu_E \approx 0.31 \text{ GR}_E^3$ (Gauss, 1 G = $10^5$ nT) is the Earth’s magnetic moment.

If the geomagnetic field is adiabatically reduced to a dipole field, the $L$ value is equal to the radial distance $(r)$, measured in $R_E$, from the dipole to the equatorial intersection of the guiding field line, as given by the equation of a dipole field line

$$r = L R_E \sin^2 \theta,$$

(1.8)
where $\theta$ is the polar angle with respect to the dipole axis. In this geometry, the adiabatic invariants $M$ and $K$ can be written as

$$M = \frac{p^2 y^2 L^3 R_E^3}{2m_0 \mu_E}, \quad (1.9)$$

and

$$K = \sqrt{\frac{\mu_E}{L R_E}} \frac{Y(y)}{y}. \quad (1.10)$$

In equations (1.9) and (1.10), $y \equiv \sin \alpha_0$, where $\alpha_0$ is the angle subtended by $p$ and $B$ in the magnetic equatorial plane. This angle is known as the equatorial pitch angle, and it is related to the magnetic intensities $B_0$ in the equatorial plane ($s = 0$) and $B_m$ at the mirror point along the guiding field line, by

$$\sin^2 \alpha_0 = \frac{B_0}{B_m}, \quad (1.11)$$

as a result of conservation of $M$ (not necessarily in a dipole field). The function $Y$ in (1.10) is a geometric integral over the bounce orbit [Schulz, 1991, eq. (161)]

$$Y = \frac{1}{2L R_E} \int \cos \alpha ds = \int \cos \alpha ds \approx 2.760346 + 2.357194 y - 5.117540 y^{3/4}, \quad (1.12)$$

in which $\alpha$ is the local pitch angle. For completeness, it is worth mentioning the analytical approximate inverse function of $K$ [Schulz, 1991, eq. (132)], given by,

$$\frac{1}{y^2} \approx 1 + 1.35048 \kappa - 0.030425 \kappa^{4/3} + 0.10066 \kappa^{5/3} + \left(\frac{\kappa}{2.760346}\right)^2, \quad (1.13)$$

where

$$\kappa = \sqrt{\frac{L R_E}{\mu_E}} K. \quad (1.14)$$

Because the dipole field is axially symmetric, a drift shell can be simply obtained by revolving the portion of the guiding field line between mirror points around the
dipole axis, and therefore the drift shells of particles originally on the same guiding field line with different $K$ values coincide. As a result, $\alpha_0$ is a constant on a drift shell, and $L$ is independent of $\alpha_0$. This degeneracy of drift shells breaks when such axial symmetry is broken. In the geomagnetic field, drift shells generated by adiabatic motion of particles sharing a common guiding field line at one longitude do not coincide at other longitudes, if their $K$ values, or equivalently $\alpha_0$ values on the common field line, are different [Schulz and Lanzerotti, 1974]. In this situation, $L$ becomes dependent on $\alpha_0$, and $\alpha_0$ is no longer a constant along a drift trajectory. This phenomenon is known as drift shell splitting. Figure 1.5 gives examples of numerically calculated drift shells, seen as their cross sections in the noon-midnight meridian, in a model day-night asymmetric magnetic field [Roederer, 1967].

### 1.3 Important magnetospheric plasma waves

During magnetically active times, intense plasma waves are excited in the magnetosphere. In the basically collision-less magnetosphere, violations of adiabatic invariants are caused by resonance of electron quasi-periodic motions with these plasma waves at appropriate frequencies [Schulz and Lanzerotti, 1974]. Figure 1.6 schematically summarizes the spatial distribution of some magnetospheric plasma waves that are especially relevant to radiation belt electrons.

Ultra Low Frequency (ULF, $\sim$mHz, not shown in Figure 1.6) waves are hydromagnetic waves that are excited at the magnetopause in response to solar wind pressure fluctuations or velocity shear, and also internally by plasma instabilities [Kivelson and Russell, 1995]. ULF waves can resonate with electron drift motion, and cause radial diffusion across drift shells. Inward radial diffusion was originally thought to be the dominant mechanism for accelerating electrons in the radiation belts, as elec-
Figure 1.5: (a) Computed drift shell splitting for particles starting on common field lines in the noon meridian. Dots represent particles’ mirror points. Curves giving the position of mirror points for constant $\alpha_0$ are shown. (b) Similar to (a), but showing computed drift shell splitting for particles starting on common field lines in the midnight meridian. These figures are adopted from Roederer [1967].
Figure 1.6 : Schematic illustration of the spatial distribution of important waves in the inner magnetosphere, in relation to the plasmapause and the drift paths of ring current ions and electrons. Figure reprinted from Thorne [2010].
trons are radially transported from a source region at higher $L$ while conserving $M$ [e.g., Schulz and Lanzerotti, 1974; Elkington et al., 1999]. However, later observations revealed that electron phase space density (see next section) can peak around $L \sim 5$ during storm times [e.g., Chen et al., 2006, 2007a; Reeves et al., 2013], indicating a local acceleration source in the heart of the outer belt not explained by radial diffusion.

Kinetic waves in the frequency range of $1 \sim 10^4$ Hz can resonate with gyration or bounce motion or both, thus violating the first two adiabatic invariants. Among these, chorus waves are discrete coherent whistler mode (electron cyclotron mode) emissions that occur in two distinct bands above and below one-half the electron gyro-frequency [Tsurutani and Smith, 1974]. They are excited by cyclotron resonance with low energy plasma sheet electrons that are injected into the inner magnetosphere during enhanced convection [Hwang et al., 2007], and they are distributed along electron drift paths, as shown in Figure 1.6. On the nightside, chorus waves are most intense inside $L \sim 8$ and are mainly confined to magnetic latitudes below $15^\circ$, due to strong Landau damping during propagation towards higher latitude from the equatorial sources [Bortnik et al., 2007]. Dayside chorus, in contrast, is distributed over a broad range of latitudes and is strongest in the outer magnetosphere [Li et al., 2009]. Chorus waves play a dual role in radiation belts: they both accelerate electrons to the MeV energy range and scatter their pitch angle into the loss cone (a conical region in velocity space), within which the electrons ultimately precipitate into the atmosphere [Horne and Thorne, 1998].

A plasmaspheric hiss wave is an incoherent whistler mode emission mostly confined within the plasmasphere and dayside plasmaspheric plumes. It originates from a subset of chorus emissions that survive Landau damping during propagation toward
higher latitudes and are reflected into the plasmasphere, where the discrete chorus emissions merge together to form incoherent hiss [Bortnik et al., 2008]. Hiss waves are mainly responsible for pitch-angle scattering, which depletes electron distributions and forms the slot region between inner and outer radiation belts [Lyons and Thorne, 1973; Abel and Thorne, 1998].

Analogous to chorus, injected ring current protons during storm times can also excite discrete emissions, known as the electromagnetic ion cyclotron (EMIC) waves [Jordanova et al., 2008]. Favorable regions for this excitation include the overlap between the ring current and the plasmasphere and the dayside plumes [Morley et al., 2009; Pickett et al., 2010]. EMIC waves are circularly polarized in the opposite sense to electron gyration along the direction of propagation, and thus they are only resonant with MeV electrons via the Doppler effect. In consequence, they are potentially responsible for rapid loss of relativistic electrons during the main phase of a storm (see Figure 1.2) [Li et al., 2007], but not for electron acceleration.

1.4 Kinetic description of the radiation belt electrons

As a collection of a large number of particles, the radiation belt electrons are best described by the kinetic theory of statistical mechanics [e.g., Huang, 1987]. In mechanics, the state of a particle is fully described by its three configuration coordinates \( q \) and three momentum components \( p \), and is represented by a point in the six-dimensional (6-D) \( \mu \)-space spanned by \( q \) and \( p \). No loss of generality occurs by supposing that \( q \) are Cartesian, and so are \( p \). Similarly, the state of \( N \) particles is fully described by the \( 3N \) coordinates and the \( 3N \) momenta of all the particles, as represented by one point in the \( 6N \)-D \( \Gamma \)-space, and the \( \Gamma \)-space can thus be regarded as a concatenation of the \( N \) individual \( \mu \)-spaces. The motion of this representative point in \( \Gamma \)-space is
governed by the kinematics of the system, and the representative point is assigned a probability $F_N d^{3N}q d^{3N}p$ of appearing in a volume element $d^{3N}q d^{3N}p$ at the $\Gamma$-space position $(q_1, \ldots, q_N, p_1, \ldots, p_N)$. The Liouville theorem prescribes that the flow of the probability density $F_N$ is incompressible, so that $F_N$ is invariant along the system’s dynamical trajectory, i.e.,

$$\frac{d}{dt} F_N = \partial_t F_N + \dot{q}_i \partial_{q_i} F_N + \dot{p}_i \partial_{p_i} F_N = 0,$$

in which an over dot means a time derivative, and summation over repeated indices, from 1 to $3N$, is implied. The summation convention is henceforth assumed throughout this dissertation except when explicitly indicated otherwise.

If these $N$ particles have no interactions (collisions), $F_N$ is then equal to the product of the single-particle probability density function over the $N$ particles

$$F_N = f_1(1)f_1(2) \cdots f_1(N) = \prod_i f_1(i),$$

where $f_1(i)$ is a short-hand for $f_1(t, q_i, p_i)$, and $f_1$ is proportional to the particle phase space density (PSD) $f$, such that $f(t, q, p) d^3 q d^3 p$ gives the number of particles in the 6-D phase space element $d^3 q d^3 p$. Integrating the Liouville equation (1.15) over the $(N-1)$ $\mu$-spaces yields the BBGKY (Bogoliubov-Born-Green-Kirkwood-Yvon) hierarchy [Huang, 1987], whereby the $(N-1)$ particle degrees of freedom are integrated out, resulting in an equation for the PSD:

$$\frac{\partial}{\partial t} f + \dot{q}_i \frac{\partial}{\partial q_i} f + \dot{p}_i \frac{\partial}{\partial p_i} f = 0.$$

This equation is known as the Vlasov equation, and it describes the dynamics of collision-less plasma (here the summation over $i$ is from 1 to 3).

It is useful to change the coordinates of the phase space from $q$ and $p$ to the canonical variables $q$ and $P$, where the expression of $P$ has been given implicitly in
(1.1), and then the Vlasov equation becomes
\[
\frac{\partial}{\partial t}f + \dot{q}_i \frac{\partial}{\partial q_i} f + \dot{P}_i \frac{\partial}{\partial P_i} f = 0, \tag{1.18}
\]
where \( f(t, q, P)d^3q d^3p = f(t, \mathbf{q}, \mathbf{p})d^3q d^3p \), due to the unit Jacobian of this coordinate change. Applying to (1.18) the canonical equations of the single particle Hamiltonian \( H \),
\[
\dot{q}_i = \frac{\partial H}{\partial P_i}, \tag{1.19}
\]
\[
\dot{P}_i = -\frac{\partial H}{\partial q_i}, \tag{1.20}
\]
we obtain
\[
\frac{\partial}{\partial t}f + \frac{\partial H}{\partial P_i} \frac{\partial}{\partial q_i} f - \frac{\partial H}{\partial q_i} \frac{\partial}{\partial P_i} f = 0. \tag{1.21}
\]
Similarly, performing the coordinate transform from \( \mathbf{q} \) and \( \mathbf{P} \) to the action-angle variables \( \varphi_i \) and \( J_i \), which is a canonical transform and therefore also has a unit Jacobian [e.g., Landau and Lifshitz, 1976; Goldstein, 1980], and assuming that \( \varphi_i \) are cyclic coordinates of \( H \) (so that \( \partial H/\partial \varphi_i = 0 \)), the Vlasov equation becomes
\[
\frac{\partial}{\partial t}f + \frac{\partial H}{\partial J_i} \frac{\partial}{\partial \varphi_i} f = 0. \tag{1.22}
\]
Further, it is reasonable to assume in the radiation belts that the electrons are sufficiently phase mixed [Schulz and Lanzerotti, 1974], so that \( f \) is not a function of \( \varphi_i \), or we may take phase averages of \( f \) over \( \varphi_i \), and in either case the equation for the phase-averaged PSD \( \bar{f}(t, J_i) \) is simply
\[
\frac{\partial}{\partial t}\bar{f} = 0. \tag{1.23}
\]
This is to say, without violation of adiabatic invariants, the phase-averaged PSD of the radiation belt electrons is constant in time in the coordinates of the adiabatic invariants.
For plasma with inter-particle collisions, in which case $F_N$ depends on the 2, 3, \ldots, $N$-particle correlation functions, the BBGKY hierarchy results in a Boltzmann equation in the form [Rosenbluth et al., 1957]

$$
\frac{\partial}{\partial t} f + \dot{q}_i \frac{\partial}{\partial q_i} f + \dot{p}_i \frac{\partial}{\partial p_i} f = -\frac{\partial}{\partial p_i} (b_i f) + \frac{1}{2} \frac{\partial^2}{\partial p_i \partial p_j} (a_{ij} f),
$$

(1.24)

where $b_i \equiv \langle \Delta p_i / \Delta t \rangle$ and $a_{ij} \equiv \langle \Delta p_i \Delta p_j / \Delta t \rangle$ represent averaged frictional and diffusive transport coefficients due to collisions. Compared to equation (1.17), the collisional terms on the right-hand side hence cause a diffusion of PSD in the momentum space.

In general, we can introduce the functions $h_i$ such that, in Cartesian coordinates,

$$
h_i \equiv b_i - \frac{1}{2} \frac{\partial}{\partial p_j} a_{ij},
$$

(1.25)

and write equation (1.24) in the form

$$
\frac{\partial}{\partial t} f + \dot{q}_i \frac{\partial}{\partial q_i} f + \dot{p}_i \frac{\partial}{\partial p_i} f = -\frac{\partial}{\partial p_i} (h_i f) + \frac{1}{2} \frac{\partial}{\partial p_i} \left( a_{ij} \frac{\partial}{\partial p_j} f \right)
= -\frac{\partial}{\partial p_i} \left( h_i f - \frac{1}{2} a_{ij} \frac{\partial}{\partial p_j} f \right).
$$

(1.26)

From Appendix A, we identify that $-(1/2)a_{ij} \partial f / \partial p_j$ gives the diffusive flux of $f$, hence $(1/2)a_{ij}$ correspond to the diffusion coefficients. The quantities $h_i f$ are the advective flux of $f$ corresponding to dynamical friction, and $h_i$ are the momentum space analogue to components of an advective flow velocity.

However, in the outer radiation belt, inter-particle collisions are rare, whereas electrons “collide” with the electromagnetic field via resonant interactions with plasma waves as described in the last section. As such, we can formally include the electromagnetic field as a “particle” in the system, generalize the $\Gamma$-space to incorporate its “coordinates” and “momenta”, and repeat the above described BBGKY hierarchy. During this procedure, the degrees of freedom of the electromagnetic field are
integrated out, resulting in a diffusion of PSD in the momentum space, due to the stochastic nature of the wave-particle interactions. Further, in light of the convenience introduced by equation (1.23), we transform the coordinates into action-angle variables, and eventually obtain a Boltzmann equation in the form

$$\frac{\partial}{\partial t} \bar{f} = -\frac{\partial}{\partial J_i} (h_i \bar{f}) + \frac{1}{2} \frac{\partial}{\partial J_i} \left( a_{ij} \frac{\partial}{\partial J_j} \bar{f} \right) + \left( \frac{\partial \bar{f}}{\partial t} \right)_{\text{coll.}},$$

(1.27)

which is now a Fokker-Planck equation (diffusion equation). Note that, in this equation the coefficients $h_i$ and $a_{ij}$ are with respect to the $J_i$ coordinates and are different from those in equation (1.26), which are with respect to $p_i$. On the right-hand side of (1.27), the first two terms represent the diffusion caused by wave-particle interactions, and all other possible diffusion terms from inter-particle collisions are collectively represented in $\left( \frac{\partial \bar{f}}{\partial t} \right)_{\text{coll.}}$ (this term is usually a loss term and is often simply approximated by $-\bar{f}/\tau$ with $\tau$ a characteristic lifetime).

It is informative to note that (following Haerendel [1968]), if in a region of the phase space $\bar{f}$ is uniform in $J_i$ (i.e., $\partial \bar{f}/\partial J_i = 0$) and inter-particle collisions are absent, wave-particle interaction alone would not change $\bar{f}$ in this region, because the particle motions are still determined by the single-particle Hamiltonian, and the Vlasov equation

$$\frac{\partial}{\partial t} \bar{f} - \frac{\partial H}{\partial \varphi_i} \frac{\partial}{\partial J_i} \bar{f} = \frac{\partial}{\partial t} \bar{f} = 0$$

(1.28)

hence applies while treating the electromagnetic waves as an external field. Note that in (1.28) $\varphi_i$ are no longer cyclic coordinates of $H$ due to violation of the adiabatic invariants. Equations (1.27) and (1.28) then indicate that

$$\frac{\partial}{\partial J_i} h_i = 0,$$

(1.29)

which means the advective flow is incompressible in the $J_i$ space, so that it must flow across the region or circulate within. However, this conclusion is not restricted
to the considered region because it is derived from the nature of the wave-particle interaction and the collision-less plasma, and furthermore, in the $J_i$ space, there is no preferred direction for an incompressible flow because of spatial symmetry, hence the conjectured advective flow could not exist at all; in other words, $h_i = 0$ identically, and the effect of wave-particle interaction is a pure diffusion without advection. Thus, from (1.25), we have reached the conclusion that, in the radiation belts

$$b_i = \frac{1}{2} \frac{\partial}{\partial J_j} a_{ij}, \quad (1.30)$$

and the Fokker-Planck equation is

$$\frac{\partial}{\partial t} \bar{f} = \frac{\partial}{\partial J_i} \left( \bar{D}_{ij} \frac{\partial}{\partial J_j} \bar{f} \right), \quad (1.31)$$

with $D_{ij} \equiv (1/2)a_{ij}$ the tensor diffusion coefficients in the $J_i$ coordinates and inter-particle collisions neglected. We note that the above theoretical treatment derives the general form of the Fokker-Planck equation, but in practice the diffusion coefficients are not obtained from the BBGKY hierarchy. Rather, they are calculated by perturbation theory using the linearized Vlasov equation [e.g., *Kennel and Engelmann*, 1966; *Kaufman*, 1972; *Lyons*, 1974; *Brizard and Chan*, 2001, 2004; *Albert*, 1994, 1999], or are obtained from diffusively perturbed single-particle equations of motion [e.g., *Schulz and Lanzerotti*, 1974; *Schulz*, 1991; *Fei et al.*, 2006]. The corresponding theory is hence named quasi-linear diffusion theory.

To change into other coordinates $Q_i$, which may be computationally advantageous, the Fokker-Planck equation (1.31) transforms as

$$\frac{\partial}{\partial t} \bar{f} = \frac{1}{G} \frac{\partial}{\partial Q_i} \left( G \bar{D}_{ij} \frac{\partial}{\partial Q_j} \bar{f} \right). \quad (1.32)$$

In this equation, $G = \det(\partial J_i/\partial Q_j)$ is the Jacobian that arises from the transformation of the divergence operator from Cartesian coordinates into curvilinear coordi-
nates (see, e.g., D’haeseleer et al. [1991] and Appendix A), and $\tilde{D}_{ij}$ are the diffusion coefficients in the $Q_i$ coordinates, which are related to $D_{lm}$ by

$$
\tilde{D}_{ij} = \frac{\partial Q_i}{\partial J_l} D_{lm} \frac{\partial Q_j}{\partial J_m},
$$

(1.33)

(as explained in Appendix A).

1.5 Review of previous work

With the above theoretical preparation, understanding the behavior of the radiation belt particles is largely condensed to solving the radiation belt Fokker-Planck equation (1.32), with diffusion coefficients provided from various plasma waves, under certain initial and boundary conditions. Though conceptually simple, solving the Fokker-Planck equation is by no means an easy task, especially when off-diagonal components exist in a multi-dimensional diffusion tensor. Historically, the Fokker-Planck equation had been solved with increasing dimensions and components: from early 1-D radial diffusion (in $L$) [e.g., Frank, 1965] or elastic pitch-angle diffusion (in $\alpha_0$) [e.g., Roberts, 1969], to mixed energy and pitch-angle 2D diffusion [e.g., Albert and Young, 2005; Tao et al., 2008, 2009], and to the recent quasi-3D simulations that incorporate both radial and local energy pitch-angle diffusions [e.g., Albert et al., 2009; Subbotin et al., 2010; Tu et al., 2013; Glauert et al., 2014]. With every newly added diffusion component, new physics is introduced into the radiation belt Fokker-Planck equation, but consequently the equation becomes more and more difficult to solve, and the resultant phase space density solutions are accordingly changed.

Typically, solving the Fokker-Planck equation involves using a numerical grid and finite difference methods. However, unphysical negative numerical solutions occur when off-diagonal diffusion tensor components exist [Albert, 2009, 2013; Camporeale
et al., 2013a, b]. To overcome this in 2-D energy-pitch-angle diffusion simulations, Albert and Young [2005] globally diagonalized the diffusion tensor by coordinate transformation, and applied a standard finite-difference method thereafter. But in 3-D, where radial diffusion is included, such matrix diagonalization is not easily achieved. Instead, an operator-splitting technique is widely adopted, which uses two sets of grids, one for the radial diffusion and the other for the energy-pitch-angle diffusion. To communicate between grids, frequent high-accuracy interpolations are required. As a result, a high grid resolution is required in these methods [Subbotin et al., 2010]. Moreover, the two sets of grids imply exclusion of possible radial-energy and radial-pitch-angle mixed diffusion; in other words, the diffusion coefficients are at most a 2-D energy-pitch-angle tensor plus a 1-D radial diffusion coefficient, which is not fully 3-D.

Alternatively, it is useful to model radiation belt diffusion in the space of the adiabatic invariants. In adiabatic invariant coordinates, PSD is conserved under adiabatic processes, as shown by equation (1.23), and the model results can be mapped to particle fluxes for a variety of magnetic field models. Beutier and Boscher [1995] wrote the Salammbô model using one 3-D grid in the three action integrals $J_1$, $J_2$ and $J_3$. However, a box computational domain in the action-integral space transforms to a highly-extended region of complicated shape in the $(E, \alpha_0, L)$ coordinates, and vice versa. This can make setting boundary conditions difficult, since they are often observationally obtained, and thus conveniently specified, in the $(E, \alpha_0, L)$ coordinates. Later, the Salammbô model was converted to the operator-splitting technique in the $(E, \alpha_0, L)$ coordinates [Varotsou et al., 2008]. Subbotin and Shprits [2012] designed a new set of adiabatic invariant variables from mixtures of the action integrals, whose coordinate lines, by clever choice of a constant related to magnetic field, roughly
follow those of the \((E,\alpha_0, L)\) coordinates. Currently, their model is also not fully 3-D.

From an entirely different approach, Tao et al. [2008] demonstrated successful application of the stochastic differential equation (SDE) method [e.g., Øksendal, 1998] in solving the 2-D energy-pitch-angle Fokker-Planck equation. The SDE method obtains probabilistic PSD solutions by solving a set of Itô stochastic differential equations that are mathematically equivalent to the Fokker-Planck equation. This method avoids global matrix diagonalization or reliance on a grid; and this is advantageous compared to finite-difference methods in dealing with off-diagonal diffusion components and complex domain geometries. Later, Tao et al. [2009] developed their SDE model with the layer method, which is computationally more efficient compared to that in Tao et al. [2008] but relies on a grid.

1.6 Thesis organization

This thesis presents (i) the application of the SDE method to solving the fully 3-D Fokker-Planck equation in adiabatic invariant coordinates, which results in a model named the Radbelt Electron Model (REM), and (ii) the application of this model to investigating effects of drift shell splitting on radiation belt electrons. The thesis is organized as follows. In Chapter 2, we give a thorough description of the mathematical theories that constitute the foundations of REM. These theories, including the SDE method, are generally presented in Cartesian coordinates in the mathematical literature. However, the radiation belt Fokker-Planck equation is more commonly solved in curvilinear coordinates, and Chapter 2 gives this generalization. Specific considerations with regard to constructing REM and its key numerical techniques are described in Chapter 3. Chapter 4 gives comparisons of REM results with vari-
ous known Fokker-Planck equation solutions and with spacecraft observations, which serve as a validation of the model. In Chapter 5, we present an application of REM in solving the radiation belt diffusion problem with drift shell splitting effects, and a theoretical investigation of the simulation results that reveals more profound understanding of the phenomenon. Original research work in this thesis includes the generalization to curvilinear coordinates and treatment of the advection term in Chapter 2, most of Chapter 3, except that Subsection 3.2.4 borrows ideas from Kloeden and Platen [1992], the entirety of Chapters 4 and 5, and all of the appendices. We summarize our research and discuss possible future developments in Chapter 6.
Chapter 2

Solving Fokker-Planck Equations with the SDE Method

In this chapter we present enough of the stochastic differential equation (SDE) theory to enable us to apply the SDE method to solution of the radiation belt diffusion equations. For simplicity of derivation, these formulations are first developed in Cartesian coordinates in Sections 2.1 and 2.2, and then generalized into curvilinear coordinates in Section 2.3. Section 2.4 applies these formulations to solving the radiation belt Fokker-Planck equation. The majority of this chapter is published in Zheng et al. [2014].

2.1 The Wiener process and Itô stochastic calculus

Just as calculus deals with differentials, integrals and differential equations, stochastic calculus deals with these objects, but with stochastic processes included. A stochastic process, \( X_t \), is a time sequence that takes random values \( x_0, x_1, x_2, \ldots \) at time instances \( 0 \leq t_0 < t_1 < t_2 < \ldots \). Generally, the conditional probability of \( X_{t_{n+1}} \) taking a specific value \( x_{n+1} \) would depend on its current value \( X_{t_n} = x_n \) and the history of previous values \( X_{t_{n-1}} = x_{n-1}, \ldots, X_{t_1} = x_1, X_{t_0} = x_0 \). If, however, this conditional probability only depends on the current value but not the previous values, such a stochastic process is called a Markov process [Gardiner, 2004]. In this thesis we will consider only Markov processes, and we will assign capital letters to stochastic processes but use the corresponding lowercase letter to denote their values.
For simplicity, let us start from a 1-D time-homogeneous SDE of the form

\[ dX_t(\omega) = b(X_t(\omega)) \, dt + \sigma(X_t(\omega)) \, dW_t(\omega), \]  

(2.1)
in which \( X_t(\omega) \) is the unknown stochastic process and \( \omega \) is one sample in the sample space \( \Omega \). For example, if \( \Omega \) is a set of pollen grains undergoing random motions, \( \omega \) labels one grain among them, and the stochastic process \( X_t(\omega) \) is a time sequence of the spatial positions of that grain. The terminology “time-homogeneous” means the coefficients \( b(X_t) \) and \( \sigma(X_t) \) depend only on the value of the stochastic process, but not explicitly on \( t \). \( W_t(\omega) \) is a 1-D stochastic process called the Wiener process of \( \omega \). The Wiener process \( W_t \) is a Gaussian random process with initial value \( W_0 = 0 \) (with probability 1) and independent increments \( dW_t = W_{t+dt} - W_t \), such that

\[ \mathbb{E}[W_t] = 0, \quad \mathbb{E}[W_t^2] = t, \]  

(2.2)

where the symbol \( \mathbb{E} \) denotes expectation over \( \Omega \) [Øksendal, 1998; Kloeden and Platen, 1992]. For different \( \omega \)'s, the corresponding Wiener processes \( W_t(\omega) \) must be statistically independent.

The stochastic process \( X_t(\omega) \) can be regarded as a mapping from the variable set \( \{t, \omega\} \) to a real number \( x \). For each fixed \( t \), the mapping from \( \omega \) to \( X_t(\omega) \) defines a random variable; on the other hand, fixing \( \omega \) we can consider \( X_t(\omega) \) as a function of \( t \), which is called a path of \( X_t(\omega) \). It can be proved that such a function is continuous in \( t \) everywhere with probability 1, but nowhere differentiable [Øksendal, 1998; Gardiner, 2004]. With these concepts in mind, for brevity we will henceforth suppress the variable \( \omega \) except when that results in ambiguity.

The solution of the 1-D SDE (2.1) is formally represented by

\[ X_t = X_0 + \int_0^t b(X_s) \, ds + \int_0^t \sigma(X_s) \, dW_s. \]  

(2.3)
On the right-hand side of equation (2.3), the first integral is a Riemann integral over time, i.e., its value converges as we repeatedly divide the integration interval, no matter where we choose to evaluate the integrand within each subinterval. The second integral is a stochastic integral. In contrast to a Riemann integral, its value depends on the choices of $s$ within subintervals because of the non-differentiability of $W_s$; this implies that $dW_s$ has (with probability 1) infinite variation on every subinterval [Øksendal, 1998; Freidlin, 1985]. A widely used and well studied choice is to evaluate the integrand at the beginning of every subinterval (so its value does not depend on the future), and this defines the so-called Itô integral

$$\int_0^t \sigma(X_s) dW_s \equiv \sum_{n=0}^N \sigma(X_{s_n})(W_{s_{n+1}} - W_{s_n}), \quad \max(s_{n+1} - s_n) \to 0. \quad (2.4)$$

Other choices of $s$ define different types of stochastic integral [e.g., Kloeden and Platen, 1992]. We use the Itô integral because of its close relation to diffusion processes [Freidlin, 1985].

Aside from linearity, the Itô integral has the particular properties that [Øksendal, 1998; Kloeden and Platen, 1992], for $0 \leq S \leq T$ and arbitrary functions of time and stochastic processes $g(t, \omega)$ and $h(t, \omega)$:

$$\mathbb{E}\left[\int_S^T g dW_t\right] = 0, \quad (2.5)$$

$$\mathbb{E}\left[\int_S^T g dW_t \int_S^T h dW_t\right] = \int_S^T \mathbb{E}[gh]dt. \quad (2.6)$$

Property (2.6) reveals that the differential $(dW_t)^2$ is of the same infinitesimal order as $dt$ in the mean-square sense. Combining (2.5) and (2.6), we may also write

$$\mathbb{E}\left[\int_S^T g dW_{ti} \int_S^T h dW_{tj}\right] = \delta_{ij} \int_S^T \mathbb{E}[gh]dt, \quad (2.7)$$

in which $\delta_{ij}$ is the Kronecker delta. When $i \neq j$, $W_{ti}$ and $W_{tj}$ are independent Wiener processes, so that the left-hand side of (2.7) splits to the product of two expectations.
of each individual Itô integral, both equal to zero following (2.5). In the mean-square sense, we symbolically denote this $dW_{ti}dW_{ij} = \delta_{ij}dt$.

In multiple dimensions, the SDE (2.1) becomes a vector equation

$$d\mathbf{X}_t = \mathbf{b}(\mathbf{X}_t)dt + \sigma(\mathbf{X}_t)d\mathbf{W}_t,$$

(2.8)

where $\mathbf{X}_t$ is an $n$-D stochastic process, $\mathbf{b}$ is an $n$-D vector, $\sigma$ is an $n \times m$ matrix, and $\mathbf{W}_t$ is an $m$-D Wiener process with each of its components an independent 1-D Wiener process.

Now let us consider a function $f(\mathbf{x})$, with continuous second-order partial derivatives, acting on $\mathbf{X}_t$. Property (2.7) allows differentiation of $f(\mathbf{X}_t)$ in the mean-square sense, with the aid of the SDE (2.8), which yields

$$df(\mathbf{X}_t) = \frac{\partial f}{\partial x_i}dX_{ti} + \frac{1}{2} \frac{\partial^2 f}{\partial x_i \partial x_j}dX_{ti}dX_{tj}$$

$$= \frac{\partial f}{\partial x_i}(b_i dt + \sigma_{ij}dW_{tj}) + \frac{1}{2} \frac{\partial^2 f}{\partial x_i \partial x_j}\sigma_{ik}\sigma_{jk}dt.$$  (2.9)

Collecting infinitesimals of the same order and defining the differential operator

$$\mathcal{L}f(\mathbf{x}) \equiv \frac{1}{2}a_{ij}(\mathbf{x})\frac{\partial^2}{\partial x_i \partial x_j}f(\mathbf{x}) + b_i(\mathbf{x})\frac{\partial}{\partial x_i}f(\mathbf{x}),$$  (2.10)

with

$$a_{ij} \equiv \sigma_{ik}\sigma_{jk} = (\sigma \sigma^T)_{ij},$$  (2.11)

equation (2.9) gives rise to the following Itô formula [e.g., Øksendal, 1998], which describes the SDE of the new stochastic process $f(\mathbf{X}_t)$:

$$df(\mathbf{X}_t) = \mathcal{L}f(\mathbf{X}_t)dt + \sigma_{ij}(\mathbf{X}_t)\frac{\partial}{\partial x_i}f(\mathbf{X}_t)dW_{tj}.$$  (2.12)

Note that a matrix $\mathbf{a}$ given by (2.11) must be symmetric and positive semi-definite, meaning all its eigenvalues are non-negative and so is its determinant.
2.2 Relation between the SDE and the Fokker-Planck equation

Having introduced stochastic calculus, we establish in this section the relation between the SDE (2.8) and the corresponding Fokker-Planck equation, through manipulations of the Itô formula (2.12) and the differential operator $L$. This relation is developed in three steps: first, we show that a functional expectation of the stochastic paths, i.e., the expectation of a functional $f(X_t)$, is a solution of a partial differential equation (PDE) called the Kolmogorov backward equation; second, we show that the distribution function of the stochastic processes satisfies another PDE called the Kolmogorov forward equation, which is equivalent to the Fokker-Planck equation; and third, we show that, for pure diffusion processes (i.e., equation (1.30) holds), the distribution function can also be expressed as a functional expectation, which relates the Kolmogorov forward equation to the Kolmogorov backward equation, and provides the connection to the SDE (2.8). This relation, in consequence, provides a means of solving the radiation belt Fokker-Planck equation by numerically evaluating functional expectations of the stochastic paths. In this section we assume Cartesian coordinates; the generalization to curvilinear coordinates is made in the next section. Relaxation of the pure diffusion condition is discussed in Section 2.4.

**Step one: the Kolmogorov backward equation**

Integration of the Itô formula (2.12) over time yields

$$
 f(X_t) = f(\xi) + \int_0^t \mathcal{L}f(X_s)ds + \int_0^t \sigma_{ij}(X_s) \frac{\partial f}{\partial x_i} dW_{sj},
$$

(2.13)
in which $\xi \equiv x_0$ marks the initial position of the Itô process. Taking expectation values of both sides of (2.13) and using property (2.5) of the Itô integral gives

$$\mathbb{E}^\xi[f(X_t)] = f(\xi) + \int_0^t \mathbb{E}^\xi[Lf(X_s)]ds,$$

(2.14)

where the superscript of $\mathbb{E}^\xi$ denotes the common initial position of the stochastic processes. If we let $t$ approach 0 (from above), the integrand in equation (2.14) can be approximated by its value at $s = 0$ and $X_s$ thus takes the non-random initial value $\xi$. In this situation, $\mathbb{E}$ can be removed from the integrand, and (2.14) then gives another expression of $L$ in terms of the limit

$$L f(\xi) = \lim_{t \downarrow 0} \frac{\mathbb{E}^\xi[f(X_t)] - f(\xi)}{t}.$$ 

(2.15)

With this interpretation of $L$ and regarding $\mathbb{E}^\xi[f(X_s)]$ as a function of $\xi$, the integrand in equation (2.14) can be transformed as follows:

$$\mathbb{E}^\xi[Lf(X_s)] = \lim_{r \downarrow 0} \frac{\mathbb{E}^\xi[\mathbb{E}^{X_r}[f(X_r)]] - \mathbb{E}^\xi[f(X_s)]}{r}$$

$$= \lim_{r \downarrow 0} \frac{\mathbb{E}^\xi[\mathbb{E}^{X_r}[f(X_s)]] - \mathbb{E}^\xi[f(X_s)]}{r}$$

$$= L \mathbb{E}^\xi[f(X_s)],$$

(2.16)

where the interchange of “$X_s$” and “$X_r$” is permissible due to the history-independent (Markovian) property of the stochastic process described by our SDE [Øksendal, 1998]. Inserting (2.16) into (2.14) and taking a time derivative, we have

$$\frac{\partial}{\partial t} \mathbb{E}^\xi[f(X_t)] = L \mathbb{E}^\xi[f(X_t)].$$

(2.17)

Denoting the function $\mathbb{E}^\xi[f(X_t)]$ as $u(t, \xi)$, we have thus derived the PDE that the functional expectation of the stochastic processes satisfies. Together with the initial condition and zero boundary condition at infinity, they form an initial value problem:

$$\frac{\partial}{\partial t} u(t, \xi) = L u(t, \xi),$$

(2.18)
\[ u(0, \xi) = \mathbb{E}^\xi [f(\xi)] = f(\xi). \quad (2.19) \]

Equation (2.18) is the Kolmogorov backward equation associated with the Itô process (2.8).

From another point of view, consider the transition probability density \( p(0, x; t, y) \) of the Itô process \( X_t \), i.e., the probability density of a stochastic process traveling from \( x \) to \( y \) in time period \( t \) [Øksendal, 1998]. By definition, the expectation \( \mathbb{E}^\xi [f(X_t)] \) can be calculated from

\[ \mathbb{E}^\xi [f(X_t)] = \int_D f(\eta) p(0, \xi; t, \eta) d\eta, \quad (2.20) \]

where the integration is over the domain \( D \) in which \( f \) is defined. Applying this integral representation to the Kolmogorov backward equation (2.17), and noting that the function \( f \) is arbitrary, we obtain

\[ \frac{\partial}{\partial t} p(0, \xi; t, \eta) = \mathcal{L}_\xi p(0, \xi; t, \eta), \quad (2.21) \]

where the subscript of \( \mathcal{L} \) indicates the variable it is operating on. Hence, the transition probability density, as a function of the time period and the starting position, also satisfies the Kolmogorov backward equation.

**Step two: the Kolmogorov forward equation**

The integral representation of the relation (2.16) is

\[ \int_D p(0, \xi; t, \eta) \mathcal{L}_\eta f(\eta) d\eta = \mathcal{L}_\xi \int_D p(0, \xi; t, \eta) f(\eta) d\eta. \quad (2.22) \]

Applying this relation and (2.20) to the Kolmogorov backward equation (2.17) gives the integral equation

\[ \int_D f(\eta) \frac{\partial}{\partial t} p(0, \xi; t, \eta) d\eta = \int_D p(0, \xi; t, \eta) \mathcal{L}_\eta f(\eta) d\eta. \quad (2.23) \]
To remove the arbitrary function $f$, it is desirable to introduce the adjoint operator $L^*$ of $L$ [Øksendal, 1998; Arfken and Weber, 1995], defined as

$$L^* \phi(x) \equiv \frac{1}{2} \frac{\partial^2}{\partial x_i \partial x_j} [a_{ij}(x) \phi(x)] - \frac{\partial}{\partial x_i} [b_i(x) \phi(x)].$$  \hspace{1cm} (2.24)

For pure diffusion processes, $a_{ij}$ and $b_i$ are not independent, but are related by

$$b_i = \frac{1}{2} \frac{\partial a_{ij}}{\partial x_j}$$ \hspace{1cm} (2.25)

as components of rank two and rank one tensors, respectively. Even in more general cases, we can always split $b$ into a diffusive part and an advective part, such that $b = (1/2) \nabla \cdot a + h$ (cf. equations (1.25) and (1.30)). Using this relation, we have, for arbitrary functions $u$ and $v$ with continuous second-order partial derivatives, the identity (Green’s identity)

$$vL u - uL^* v = \frac{\partial}{\partial x_i} \left( \frac{1}{2} v a_{ij} \frac{\partial u}{\partial x_j} - \frac{1}{2} u a_{ij} \frac{\partial v}{\partial x_j} + uv h_i \right) = \frac{\partial}{\partial x_i} (u F_i^u - v F_i^v + uv h_i),$$ \hspace{1cm} (2.26)

in which $F_i^u = -(1/2) a_{ij} \partial u / \partial x_j$ and $F_i^v = -(1/2) a_{ij} \partial v / \partial x_j$ are the $i$-component of the diffusive fluxes of $u$ and $v$ respectively. Then it follows that the relation between adjoint operators

$$\int_D vL u \, dx = \int_D uL^* v \, dx$$ \hspace{1cm} (2.27)

is satisfied when the functions $u$ and $v$ vanish on the boundary. This is natural when the boundary is at infinity; but even if the domain $D$ is finite, (2.27) is still guaranteed by the condition that, on the boundary where $u$ or $v$ are nonzero, both the diffusive and advective fluxes normal to the boundary vanish, i.e., the fluxes cannot flow across that boundary. The latter is indeed the case in the radiation belt diffusion problem.
Under the assumption that $f$ satisfies the aforementioned conditions of $u$ and $v$, applying (2.27) to the right-hand side of (2.23) immediately gives

$$\int_D f(\eta) \frac{\partial}{\partial t} p(0, \xi; t, \eta) d\eta = \int_D f(\eta) L^*_\eta p(0, \xi; t, \eta) d\eta. \quad (2.28)$$

Again, since $f$ is arbitrary, we conclude that the transition probability density, as a function of the time period $t$ and the ending position $\eta$, satisfies the PDE

$$\frac{\partial}{\partial t} p(0, \xi; t, \eta) = L^*_\eta p(0, \xi; t, \eta). \quad (2.29)$$

This equation is called the Kolmogorov forward equation. Given an initial distribution function $\phi(0, \xi)$ of the diffusion process, which gives the number of stochastic processes initially in $d\xi$, $dN = \phi(0, \xi)d\xi$, the distribution function at a later time $t$ is calculated from

$$\phi(t, \eta) = \int_D \phi(0, \xi)p(0, \xi; t, \eta) d\xi. \quad (2.30)$$

Applying this integration to both sides of (2.29) then yields the Kolmogorov forward equation for the distribution function

$$\frac{\partial}{\partial t} \phi(t, \eta) = L^* \phi(t, \eta). \quad (2.31)$$

This equation describes the diffusion process as a PDE for the distribution function, and it is more commonly called the Fokker-Planck equation (cf. the form of the radiation belt Fokker-Planck equation (1.27)).

**Step three: the Kolmogorov backward equation for the distribution function**

Whereas the above derivations apply for the general cases that the advective “velocity” $h$ could be arbitrary, in this step, we temporarily restrict ourselves to pure diffusion
processes only, which means $h = 0$ and hence the relation (2.25) holds, and we show that, under this restriction, $\phi(t, \eta)$ also satisfies the Kolmogorov backward equation. Taking a particular $f(\eta) = \phi(0, \eta)$ in the integral equation (2.20) and reversing the derivations from (2.21) back to (2.17) gives

$$\frac{\partial}{\partial t} \mathbb{E}^{{\xi}}[\phi(0, X_t)] = \mathcal{L} \mathbb{E}^{{\xi}}[\phi(0, X_t)].$$

(2.32)

To see the meaning of $\mathbb{E}^{{\xi}}[\phi(0, X_t)]$, comparison between equations (2.20) and (2.30) suggests that we construct a time-backward stochastic process $Z^t_s$ with transition probability density $\pi(t, \eta; 0, \xi) = p(0, \xi; t, \eta)$, where the parameter $s$ increases from 0 to $t$ and time is given by $t - s$. Then, following (2.30), $\phi(t, \eta)$ is represented by

$$\phi(t, \eta) = \int_D \phi(0, \xi) \pi(t, \eta; 0, \xi) d\xi = \mathbb{E}^{{\eta}}[\phi(0, Z^t_t)].$$

(2.33)

However, in the absence of advection, it actually does not matter whether time is increasing or decreasing in the stochastic process. Therefore, going from $\eta$ to $\xi$ while time is decreasing is identical to going from $\eta$ to $\xi$ while time is increasing. This amounts to saying that, for pure diffusion processes, starting and ending positions commute in the transition probability density

$$p(0, \xi; t, \eta) = p(0, \eta; t, \xi),$$

(2.34)

and thus $Z^t_s$ is identical to $X_s$. Writing out the SDE for $Z^t_s$

$$dZ^t_s = b(Z^t_s) ds + \sigma(Z^t_s) dW_s$$

(2.35)

and comparing with (2.8) makes this statement more obvious. As a result, in equation (2.32)

$$\mathbb{E}^{{\xi}}[\phi(0, X_t)] = \mathbb{E}^{{\xi}}[\phi(0, Z^t_t)] = \phi(t, \xi).$$

(2.36)
Therefore the distribution function $\phi(t, \eta)$ is also a solution of the Kolmogorov backward equation

$$\frac{\partial}{\partial t} \phi(t, \eta) = \mathcal{L} \phi(t, \eta),$$

as claimed. In fact, the relation (2.34) implies that, given one of its Kolmogorov equations, it automatically satisfies the other.

The fact that $\phi(t, \eta)$ satisfies both the Kolmogorov forward equation and the Kolmogorov backward equation ensures that, given the Kolmogorov forward equation (2.31), we can always write it (by algebraic manipulations) in the Kolmogorov backward equation form (2.37), and from the latter, determine the corresponding SDE of the time-backward stochastic process of form (2.35) from coefficients of the equation, in which the matrix $\sigma$ is calculated by the factorization in (2.11). By equation (2.36), solution of the Kolmogorov forward equation is represented by the functional expectation of the stochastic processes, where the function of the stochastic paths is the initial distribution. This constitutes the basic idea of solving the Fokker-Planck equation in the SDE method in Cartesian coordinates.

### 2.3 The SDE method from the viewpoint of transition probabilities: generalization to curvilinear coordinates

The formulation of the SDE method in Section 2.2, which is derived directly from the Itô formula, is not easily generalized to curvilinear coordinates. To generalize, in this section we start with a reexamination of the formulation of the SDE method in Cartesian coordinates, but from another viewpoint that explicitly uses the Markovian property of the stochastic processes. This approach centralizes the role of the transition probability densities, and allows for a transition to curvilinear coordinates.
by obtaining the corresponding transition probability densities in these coordinates from their Cartesian counterparts. In this section, we consider only pure diffusion processes, and we use Latin letters to denote Cartesian coordinates and Greek letters to denote curvilinear coordinates.

2.3.1 Cartesian coordinates

From Section 2.2, we know that both the distribution function \( \phi(t, y) \) and the transition probability density \( p(0, x; t, y) \), as a function of \( t \) and \( y \), satisfy the Kolmogorov forward equation, i.e.,

\[
\frac{\partial}{\partial t} \phi(t, y) = L^* \phi(t, y),
\]

(2.38)

and

\[
\frac{\partial}{\partial t} p(0, x; t, y) = L_y^* p(0, x; t, y).
\]

(2.39)

For example, the radiation belt Fokker-Planck equation (1.31) in Cartesian coordinates

\[
\frac{\partial}{\partial t} \phi(t, y) = \frac{\partial}{\partial y_i} \left( D_{ij} \frac{\partial \phi}{\partial y_j} \right),
\]

(2.40)

can be written in the Kolmogorov forward equation form as

\[
\frac{\partial}{\partial t} \phi(t, y) = \frac{1}{2} \frac{\partial^2}{\partial y_i \partial y_j} (2D_{ij} \phi) - \frac{\partial}{\partial y_i} \left( \frac{\partial D_{ij}}{\partial y_j} \phi \right),
\]

(2.41)

in which \( \phi(t, y) \) is the distribution function with respect to the Cartesian coordinates, i.e., \( \phi d y \) gives the number of particles in the volume element \( d y \). Note that, comparing (2.41) with (2.24), it follows that, for the Fokker-Planck equation (2.40), the coefficients in \( L^* \) satisfies the relation given by (2.25). Returning to the general case, the purpose of the SDE method is to solve for \( \phi(t, y) \) in (2.38) with given initial condition.
For Markovian process, the probability of its going into a volume element at position \( y \) at time \( t + \Delta t \) depends only on its current position \( x \) at time \( t \), but not on the earlier history. This property further implies that the following Chapman-Kolmogorov equation \([Gardiner, 2004]\), which is the key equation of Markovian processes, must hold

\[
p(0, x; \tau, z) = \int p(0, x; t, y)p(t, y; \tau, z)dy, \quad (0 \leq t \leq \tau). \tag{2.42}
\]

Given the Kolmogorov forward equation, the Kolmogorov backward equation is implied by the Chapman-Kolmogorov equation. To see this, differentiating both sides of equation (2.42) with respect to \( t \) yields

\[
0 = \int p(t, y; \tau, z)\frac{\partial}{\partial t}p(0, x; t, y)dy + \int p(0, x; t, y)\frac{\partial}{\partial t}p(t, y; \tau, z)dy. \tag{2.43}
\]

Multiplying both sides of equation (2.39) with \( p(t, y; \tau, z) \) and integrating over \( y \), with the relation in equation (2.43), gives

\[
-\int p(0, x; t, y)\frac{\partial}{\partial t}p(t, y; \tau, z)dy = \int p(t, y; \tau, z)L_y^*p(0, x; t, y)dy. \tag{2.44}
\]

Next, we want to put \( p(0, x; t, y) \) on the right-hand side of (2.44) out of the differential operator. As we have seen in the previous section, appealing to the adjoint operators \( L_y \) and \( L_y^* \), we immediately obtain

\[
-\int p(0, x; t, y)\frac{\partial}{\partial t}p(t, y; \tau, z)dy = \int p(0, x; t, y)L_yp(t, y; \tau, z)dy. \tag{2.45}
\]

By the same token, because the function \( p(0, x; t, y) \) is arbitrary, we must have the differential equation

\[
-\frac{\partial}{\partial t}p(t, y; \tau, z) = L_y p(t, y; \tau, z). \tag{2.46}
\]

To make (2.46) more symmetric with (2.39), we denote the transition time interval \( s \equiv \tau - t \), and hence \( \partial / \partial s = -\partial / \partial t \). Then, equation (2.46) becomes

\[
\frac{\partial}{\partial s}p(\tau - s, y; \tau, z) = L_y p(\tau - s, y; \tau, z), \tag{2.47}
\]
which is the Kolmogorov backward equation for the transition probability density as a function of the time period $s$ and the starting position $y$.

As we have seen in the previous section, multiplying an arbitrary function $f(z)$ on both sides of equation (2.47) and integrating over $z$ gives the Kolmogorov backward equation for the functional expectation

$$\frac{\partial}{\partial s} \mathbb{E}^{s,y}[f(X_s)] = \mathcal{L} \mathbb{E}^{s,y}[f(X_s)], \quad (2.48)$$

where the superscripts of $\mathbb{E}$ indicate both traveling time period and the initial position of the ensemble of stochastic processes.

It has been shown in Section 2.2 that, for pure diffusion processes, $\phi(t, y)$ also satisfies the Kolmogorov backward equation, by introducing the time-backward stochastic process $X_s (=-Z^t_s)$ with transition probability density

$$\pi(t + \Delta t, y; t, x) = p(t, x; t + \Delta t, y), \quad (2.49)$$

and writing the integral representation as

$$\phi(t, y) = \int \phi(0, x) \pi(t, y; 0, x) dx = \mathbb{E}^y[\phi(0, X_t)], \quad (2.50)$$

in which the subscript in the time-backward stochastic process $X_t$ indicates the time period but not the ending time. Section 2.2 has argued that the time-backward stochastic process is in fact mathematically the same as the time-forward one, therefore the corresponding differential operator appearing in the equation for $\phi(t, y)$ would be identical to that for the time-forward stochastic processes in (2.48), i.e.,

$$\frac{\partial}{\partial t} \phi(t, y) = \mathcal{L} \phi(t, y). \quad (2.51)$$

With the operator $\mathcal{L}$ readily given by the adjoint of the operator in the Fokker-Planck equation (2.38), we can immediately write down the SDEs of the corresponding time-
backward stochastic process, and then the Fokker-Planck equation (2.38) is solved by evaluating the functional expectation in (2.50).

For the exemplified equation (2.40), direct expansion of the right-hand side yields its Kolmogorov backward equation form

$$\frac{\partial}{\partial t} \phi(t, y) = \frac{1}{2} (2D_{ij}) \frac{\partial^2 \phi}{\partial y_i \partial y_j} + \frac{\partial D_{ji}}{\partial y_j} \frac{\partial \phi}{\partial y_i}.$$  (2.52)

Compared with the differential operator in (2.41), we find that the differential operator here for the time-backward stochastic processes is indeed its adjoint, due to symmetry of the diffusion tensor $D_{ij} = D_{ji}$.

### 2.3.2 Curvilinear coordinates

In curvilinear coordinates, the situation becomes a little more complicated because of the appearance of the Jacobian determinant. Consider a transformation from Cartesian coordinates $y$ to curvilinear coordinates $\eta$. For the distribution function, we have

$$\phi(t, y) dy = \phi(t, y(\eta)) G(y; \eta) d\eta \equiv \psi(t, \eta) d\eta,$$  (2.53)

where $G(y; \eta) = \det(\partial y_i / \partial \eta_j)$ is the Jacobian of the transformation, and $\psi(t, \eta) = \phi(t, \eta) G(y; \eta)$ is the distribution function with respect to the curvilinear $\eta$ coordinates, i.e., $\psi d\eta$ gives the number of particles in a volume element $d\eta$.

To find the transformation of transition probability density to curvilinear coordinates, we look for the integral representation of the new distribution function $\psi(t, \eta)$. Starting from the Cartesian integral representation (2.30), we have

$$\phi(t, y(\eta)) = \int \phi(0, x(\xi)) p(0, x(\xi); t, y(\eta)) G(x; \xi) d\xi \equiv \int \psi(0, \xi) p(0, \xi; t, \eta) d\xi,$$  (2.54)
where $\psi(0, \xi) = \phi(0, x(\xi)) \ G(x; \xi)$ is the initial distribution in curvilinear coordinates, but $p(0, \xi; t, \eta)$ is still the transition probability density in Cartesian coordinates, although expressed as a function of $\xi$ and $\eta$. Multiplying both sides of (2.54) with $G(y; \eta)$, we obtain

$$\psi(t, \eta) = \int \psi(0, \xi) \ p(0, \xi; t, \eta) \ G(y; \eta) \ d\xi \equiv \int \psi(0, \xi) \ \tilde{p}(0, \xi; t, \eta) \ d\xi.$$

(2.55)

Therefore, we find that the transition probability density in curvilinear coordinates is

$$\tilde{p}(0, \xi; t, \eta) = p(0, \xi; t, \eta) G(y; \eta),$$

(2.56)
i.e., $\tilde{p}(0, \xi; t, \eta) d\eta$ gives the probability of a particle going from point $\xi$ into a volume element $d\eta$ at $\eta$ in curvilinear coordinates. Applying the transformation (2.56) to equation (2.42), it is straightforward to show that the Chapman-Kolmogorov equation retains the same form in curvilinear coordinates,

$$\tilde{p}(0, \xi; \tau, \zeta) = \int \tilde{p}(0, \xi; t, \eta) \ \tilde{p}(t, \eta; \tau, \zeta) \ d\eta.$$

(2.57)

We note that in equation (2.56), only the ending position is involved in the Jacobian but not the starting position, and that the Kolmogorov backward equation involves the starting position. Therefore, the transition probability density in curvilinear coordinates must satisfy the same Kolmogorov backward equation as in Cartesian coordinates, except for a coordinate transform for the differential operator $L$.

Applying this conclusion to the transition probability $\tilde{p}(t, \eta; \tau, \zeta)$, we have the Kolmogorov backward equation in the same form as equation (2.46):

$$-\frac{\partial}{\partial t} \tilde{p}(t, \eta; \tau, \zeta) = L_\eta \tilde{p}(t, \eta; \tau, \zeta).$$

(2.58)

Then, by a procedure similar to that illustrated in Cartesian coordinates from (2.43) to (2.46), but now multiplying on both sides of equation (2.58) by the transition
probability density $\tilde{p}(0, \xi; t, \eta)$ and integrating over $\eta$, and again using the adjoint relation between $L_\eta$ and $L^*_\eta$, we can prove that, given the Kolmogorov backward equation (2.58) and the Chapman-Kolmogorov equation (2.57), the transition probability density also satisfies the Kolmogorov forward equation in curvilinear coordinates

$$\frac{\partial}{\partial t} \tilde{p}(0, \xi; t, \eta) = L^*_\eta \tilde{p}(0, \xi; t, \eta). \quad (2.59)$$

Again, this has the same form as in Cartesian coordinates except for a coordinate transform in $L^*_\eta$. With the integration in (2.55), it follows that the Kolmogorov forward equation for the distribution function in curvilinear coordinates is

$$\frac{\partial}{\partial t} \psi(t, \eta) = L^* \psi(t, \eta), \quad (2.38')$$

which is an equation for the distribution function with respect to the $\eta$ coordinates. This equation is numbered as (2.38') to emphasize that it is the curvilinear analogue of equation (2.38).

We now apply these general results to the radiation belt Fokker-Planck equation. The radiation belt Fokker-Planck equation (2.40) in curvilinear coordinates $\eta$ is (cf. equation (1.32))

$$\frac{\partial}{\partial t} \phi(t, \eta) = \frac{1}{G} \frac{\partial}{\partial \eta_i} \left( G \tilde{D}_{ij} \frac{\partial \phi}{\partial \eta_j} \right), \quad (2.60)$$

where $G = G(y; \eta)$ and $\tilde{D}_{ij}$ are the diffusion coefficients in the $\eta$ coordinates. This can be rewritten in the Kolmogorov forward equation form

$$\frac{\partial}{\partial t} \psi(t, \eta) = \frac{1}{2} \frac{\partial^2}{\partial \eta_i \partial \eta_j} (2 \tilde{D}_{ij} \psi) - \frac{\partial}{\partial \eta_i} \left[ \frac{1}{G} \frac{\partial (G \tilde{D}_{ij})}{\partial \eta_j} \psi \right]. \quad (2.61)$$

Note that the distribution function with respect to the $\eta$ coordinates appears in equation (2.61); and the relation between coefficients becomes

$$b_i = \frac{1}{2G} \frac{\partial}{\partial \eta_j} (Ga_{ij}). \quad (2.62)$$
This is no surprise, because this is just how the divergence operator transforms from Cartesian to curvilinear coordinates (cf. Appendix A).

Returning to the general case, to solve the Fokker-Planck equation in curvilinear coordinates, we again devise the time-backward stochastic process with transition probability density \( \tilde{\pi}(t + \Delta t, \eta; t, \xi) \). By the relation (2.56), since \( \tilde{\pi} \) is a transition probability density, it must be related to the Cartesian counterpart by

\[
\tilde{\pi}(t + \Delta t, \eta; t, \xi) = \pi(t + \Delta t, y(\eta); t, x(\xi)) G(x; \xi).
\]

(2.63)

Therefore, by the definition of \( \pi \) (2.49), \( \tilde{\pi} \) is related to \( \tilde{p} \) by

\[
\tilde{\pi}(t + \Delta t, \eta; t, \xi) = p(t, x(\xi); t + \Delta t, y(\eta)) G(x; \xi)
\]

\[
= \tilde{p}(t, \xi; t + \Delta t, \eta) G^{-1}(y; \eta) G(x; \xi).
\]

(2.64)

Compared to (2.49), (2.64) loses its symmetry between both sides of the equation by having two extra Jacobian determinants on the right. This is because the transition process itself is not symmetric: the stochastic process starts from a point in space but goes into a volume element. Applying (2.64) to (2.55) yields

\[
\psi(t, \eta) = \int \psi(0, \xi) \tilde{\pi}(t, \eta; 0, \xi) G^{-1}(x; \xi) G(y; \eta) d\xi.
\]

(2.65)

which, by canceling the Jacobian in \( \psi \) on both sides, reduces to

\[
\phi(t, \eta) = \int \phi(0, \xi) \tilde{\pi}(t, \eta; 0, \xi) d\xi = \mathbb{E}^\eta[\phi(0, \tilde{X}_t)].
\]

(2.50’)

It is interesting to note that, even in curvilinear coordinates, the expectation representation is nevertheless constructed with the Cartesian distribution function \( \phi \), though \( \tilde{X}_t \) is the stochastic process in curvilinear coordinates.

As a transition probability density, \( \tilde{\pi} \) satisfies the Kolmogorov backward equation

\[
\frac{\partial}{\partial t} \tilde{\pi}(t, \eta; 0, \xi) = \mathcal{L}_\eta \tilde{\pi}(t, \eta; 0, \xi),
\]

(2.66)
which the functional expectation in (2.50') satisfies too. Therefore, in curvilinear
coordinates, $\phi(t, \eta)$ can still be evaluated by calculating the functional expectation
of the time-backward stochastic processes, whose SDEs are determined by writing
the Fokker-Planck equation into the Kolmogorov backward equation form. For the
example in (2.60), the Kolmogorov backward equation form is
\[
\frac{\partial}{\partial t} \phi(t, \eta) = \frac{1}{2} \left( 2 \tilde{D}_{ij} \right) \frac{\partial^2}{\partial \eta_i \partial \eta_j} \phi + \left[ \frac{1}{G} \frac{\partial (G \tilde{D}_{ij})}{\partial \eta_j} \right] \frac{\partial}{\partial \eta_i} \phi.
\] (2.67)
Note that it is $\phi$ that appears in this equation, rather than $\psi$ as in the Kolmogorov
forward form (2.61).

2.3.3 Conclusion

Transforming from Cartesian coordinates $y$ to curvilinear coordinates $\eta$, the distribu-
tion function transforms to $\psi(t, \eta) = \phi(t, y(\eta)) G(y; \eta)$, and equations (2.38) and
(2.51) transform to
\[
\frac{\partial}{\partial t} \psi(t, \eta) = \mathcal{L}^* \psi(t, \eta),
\] (2.38')
\[
\frac{\partial}{\partial t} \phi(t, \eta) = \mathcal{L} \phi(t, \eta),
\] (2.51')
with the diffusion coefficients in $\mathcal{L}^*$ and $\mathcal{L}$ changed accordingly (cf. equation (1.33)),
and the corresponding expectation representation is
\[
\mathbb{E}^\eta[\phi(0, \widetilde{X}_t)] = \phi(t, \eta),
\] (2.50')
where $\widetilde{X}_t$ is the time-backward stochastic process in $\eta$ coordinates. Note that, al-
though the function in the Fokker-Planck equation (2.38') becomes the new distribu-
tion function $\psi$, the function in the Kolmogorov backward equation (2.51') and the
expectation representation (2.50') remains $\phi$, the distribution function with respect
to Cartesian coordinates, except for a change of variables.
In mechanics, the role of Cartesian coordinates in phase space is played by canonical variables (as has been seen in Section 1.4), and transformations between canonical variables have unit Jacobian determinant [e.g., Goldstein, 1980]. Hence, we can choose just one distribution function for all sets of canonical variables. Consequently, regardless of what distribution function we start with in (2.38′), we always end up with the distribution function with respect to canonical variables when writing (2.51′) and (2.50′). Given that (2.51) and (2.50) have the same form as (2.51′) and (2.50′), we can thus apply the same SDE method for both Cartesian and curvilinear coordinates.

In conclusion, for pure diffusion processes, the Fokker-Planck equation can be solved by first writing it into the Kolmogorov backward equation form, then determining the SDEs of the corresponding time-backward stochastic processes from the differential operator in the Kolmogorov backward equation, and then evaluating the (Cartesian) distribution function by calculating the functional expectation of these stochastic processes about the initial condition.

2.4 Solution of general diffusion equations: the Feynman-Kac formula

In this section we consider three generalizations of the conclusions of the previous sections. For simplicity, we only formulate in Cartesian coordinates. Following the conclusions in Subsection 2.3.3, generalization of the discussion in this section to curvilinear coordinates is straightforward. This section assumes the same symbols as those in Section 2.2.

First, in time-inhomogeneous diffusion processes, the coefficients \( a_{ij}(t, \mathbf{x}) \) and \( b_i(t, \mathbf{x}) \) depend on both spatial coordinates and time \( t \). To construct the operator
\( \mathcal{L} \) with explicit time dependence from the Itô formula, the strategy is to treat time as one extra “stochastic” variable and construct a new Itô process in the extended space \( t \otimes \mathbb{R}^n \) as [Freidlin, 1985]

\[
Y_s = \begin{pmatrix} t - s \\ X_s \end{pmatrix}, \quad 0 \leq s \leq t, \tag{2.68}
\]

and the corresponding SDE is

\[
dY_s = \begin{pmatrix} -1 \\ b(t - s, X_s) \end{pmatrix} ds + \begin{pmatrix} 0 \cdots 0 \\ \sigma(t - s, X_s) \end{pmatrix} dW_s \\
= \tilde{b}(Y_s) ds + \tilde{\sigma}(Y_s) dW_s. \tag{2.69}
\]

Equation (2.69) shows that \( Y_s \) is a time-homogeneous Itô process, and thus all our previous conclusions apply. Repeating those procedures, it is straightforward to verify that the operators \( \mathcal{L} \) and \( \mathcal{L}^* \) associated with \( Y_s \) are indeed the time-dependent operators we wanted to construct. On the other hand, the projection of \( Y_s \) onto the \( \mathbb{R}^n \) space is a time-inhomogeneous Itô process

\[
dX_s^t = b(t - s, X_s^t) ds + \sigma(t - s, X_s^t) dW_s, \quad 0 \leq s \leq t, \tag{2.70}
\]

where \( s \) is the running time and \( t \) is the time period. Consequently, we can say that the time-dependent operators \( \mathcal{L} \) and \( \mathcal{L}^* \) are associated with this time-inhomogeneous Itô process, and the expectation representation from the previous subsection should have the SDE (2.8) replaced by (2.70).

Second, a general Kolmogorov backward equation might contain a source (loss) term characterized by \( c(t, \xi) \), such as

\[
\frac{\partial}{\partial t} u(t, \xi) = \mathcal{L} u(t, \xi) + c(t, \xi) u(t, \xi). \tag{2.71}
\]
Defining the operator
\[ \tilde{L} \equiv L + c(t, x), \]  
(2.72)
it can be verified that, to generate \( \tilde{L} \) instead of \( L \) in the Itô formula (2.12), we need only replace the function \( f \) therein with a functional \( F[c] \), regarded as a function of \( t \) and \( x \) for given \( c \), such that [Freidlin, 1985; Øksendal, 1998]
\[ F(t, x) \equiv f(x) \exp \left( \int_0^t c(t - s, x) \, ds \right). \]  
(2.73)

Then, by the same derivations as for the Kolmogorov backward equation (2.18), the solution of the general diffusion equation (2.71), with initial condition (2.19) and boundaries at infinity, is
\[ u(t, \xi) = \mathbb{E}^{\xi} \left[ f(X_t^s) \exp \left( \int_0^t c(t - s, X_s^t) \, ds \right) \right], \]  
(2.74)
\[ = \mathbb{E}[F^{t, \xi}] \]  
(2.75)
where the symbol \( X_t^s \) is a shorthand of \( X_s^t \), and \( F^{t, \xi} \) is the functional of the path \( X_s^t \), which is started from \( \xi \) and travels for time period \( t \). This formula is known as the Feynman-Kac formula. Observing that the Itô processes travel backward in time, the meaning of the Feynman-Kac formula is apparent: the solution at position \( \xi \) and time \( t \) is obtained by averaging the initial condition and the propagation information carried by the Itô processes along their paths.

The Feynman-Kac formula allows us to solve for the Fokker-Planck equation containing source or loss terms. Given a Fokker-Planck equation in the form
\[ \frac{\partial}{\partial t} \phi(t, \eta) = \mathcal{L}^* \phi(t, \eta) + c(t, \eta)\phi(t, \eta), \]  
(2.76)
from the discussion in Subsection 2.3.1, we can transform this equation by algebraic manipulations to the Kolmogorov backward equation form as
\[ \frac{\partial}{\partial t} \phi(t, \eta) = \mathcal{L} \phi(t, \eta) + c(t, \eta)\phi(t, \eta). \]  
(2.77)
Then, following equations (2.71) to (2.75), solution of the Fokker-Planck equation is represented by the expectation of the functional

\[ F^{t,\eta} = \phi(0, \mathbf{X}_t^t) \exp \left( \int_0^t c(t - s, \mathbf{X}_s^t) ds \right) . \]  

We note that, from (2.74), a functional expectation with respect to any arbitrary function \( f \) satisfies the Kolmogorov backward equation; however, from (2.78), only the functional expectation with respect to the initial condition solves the initial value problem with the Fokker-Planck equation (2.76).

The Feynman-Kac formula also allows for relaxation of the pure diffusion condition (2.25). As discussed in Section 2.2, given such a Fokker-Planck equation, we can always split the coefficients \( b_i \) into a diffusive part, denoted \( b_i^* \equiv (1/2)\partial a_{ij}/\partial \eta_j \), and an advective part \( h_i \), so that \( b_i = b_i^* + h_i \). Then, the Fokker-Planck equation can be transformed into the Kolmogorov backward equation form as follows:

\[
\frac{\partial}{\partial t} \phi(t, \eta) = \frac{1}{2} \frac{\partial^2}{\partial \eta_i \partial \eta_j} (a_{ij} \phi) - \frac{\partial}{\partial \eta_i} (b_i \phi) + \frac{1}{2} \frac{\partial^2}{\partial \eta_i \partial \eta_j} (a_{ij} \phi) - \frac{\partial}{\partial \eta_i} (b_i^* \phi) - \frac{\partial}{\partial \eta_i} (h_i \phi) = \frac{1}{2} a_{ij} \frac{\partial^2}{\partial \eta_i \partial \eta_j} \phi + \frac{1}{2} a_{ij} \frac{\partial^2}{\partial \eta_i \partial \eta_j} \phi + (b_i^* - h_i) \frac{\partial}{\partial \eta_i} \phi - \left( \frac{\partial h_i}{\partial \eta_i} \right) \phi, \tag{2.79}
\]

where in step (2.79) we have used the conclusions in Subsection 2.3.1. Comparing (2.80) with (2.71), we find that now the divergence of \( h \) serves as a “source or loss” in the Kolmogorov backward equation. The time-backward SDE corresponding to equation (2.80) is

\[
d\mathbf{X}_s^t = \left[ b^*(t - s, \mathbf{X}_s^t) - h(t - s, \mathbf{X}_s^t) \right] ds + \sigma(t - s, \mathbf{X}_s^t) dW_s, \tag{2.81}
\]

in which the minus sign before \( h \) is natural since the advection velocity must be reversed if time runs backward. The solution of the Fokker-Planck equation is thus
given by the expectation of 

\[ F^{t,n} = \phi(0, X^t_t) \exp \left( - \int_0^t \frac{\partial}{\partial \eta_i} h_i(t - s, X^t_s) ds \right), \tag{2.82} \]

with \( X^t_s \) integrated from (2.81). The relaxation of the pure diffusion condition implies that, given a general Fokker-Planck equation, we can always transform it into the corresponding Kolmogorov backward equation form. Therefore, solving the Kolmogorov backward equation is equivalent to solving the Fokker-Planck equation, and hence we need not explicitly distinguish between them.

Third, we consider boundary conditions of the diffusion problem. To implement a Dirichlet boundary condition, we introduce the “stopping time”, which is denoted by \( \tau \) henceforth. Given a domain \( D \), the stopping time is the time \( s \) until the Itô process first reaches the boundary and after which it remains at the same place. Such boundaries are denoted as \( \partial_1 D \), where the subscript 1 stands for a boundary condition of “the first kind”, also known as a Dirichlet boundary condition [Øksendal, 1998; Freidlin, 1985]. If \( \partial_1 D \) can never be reached, \( \tau \) is \( +\infty \).

Let us consider the initial value and Dirichlet boundary value problem:

\[ \frac{\partial}{\partial t} u(t, \xi) = \tilde{L} u(t, \xi), \tag{2.83} \]
\[ u(0, \xi) = f(\xi), \quad \xi \in D, \tag{2.84} \]
\[ u(t, \xi) = g(t, \xi), \quad \xi \in \partial_1 D, \tag{2.85} \]

in which \( g(t, \xi) \) is the Dirichlet boundary condition and where \( g(0, \xi) = f(\xi) \) for \( \xi \in \partial_1 D \). To solve this problem, contributions from the boundary condition is superposed with those from the initial condition, and the functional \( F^{t,\xi} \) is generalized as

\[ F^{t,\xi} = \mathbb{1}_{\tau \geq t} f(X^t_t) \exp \left( \int_0^t c(t - s, X^t_s) ds \right) \]
\[ + \mathbb{1}_{\tau < t} g(t - \tau, X^t_\tau) \exp \left( \int_0^\tau c(t - s, X^t_s) ds \right), \tag{2.86} \]
where the indicator $I_C$ has value 1 when condition $C$ is true, and 0 otherwise [Freidlin, 1985]. The meaning of (2.86) is simple: if the Itô process remains in $D$ when time $s$ runs out, apply the initial condition; if it stops on $\partial_1 D$ before that, apply the Dirichlet boundary condition. The solution of this problem is then given by the expectation of the new functional (2.86) as in equation (2.75).

A Neumann-type (the second kind) boundary $\partial_2 D$ is different from a Dirichlet-type boundary $\partial_1 D$. Instead of stopping on the boundary, the Itô process is reflected on $\partial_2 D$ [Freidlin, 1985; Øksendal, 1998; Costantini et al., 1998]. For a homogeneous Neumann boundary condition of the form

$$\gamma(\xi) \cdot \nabla u(t, \xi) = 0, \quad \xi \in \partial_2 D,$$

where $\gamma(\xi)$ is a continuous unit vector field pointing toward the interior of $D$, the Itô process undergoes a reflection about the direction of $\gamma$ at the place where it reaches $\partial_2 D$. In the SDE, this reflection is implemented by introducing an additional term involving $\gamma$ such that

$$dX^t_s = b(t-s, X^t_s)ds + \sigma(t-s, X^t_s)dW_s + \gamma(X^t_s)dk_s,$$

where $k_s$ is a non-decreasing 1-D stochastic variable, called local time, with initial value $k_0 = 0$. The increment $dk_s$ is nonzero only when $X^t_s$ reflects from $\partial_2 D$, and at other times it remains zero. In this sense, the vector $\gamma$ can be looked upon as a “local velocity” of the stochastic motion on $\partial_2 D$. Consider a 2-D domain as sketched in Figure 2.1, for example. Because $\gamma$ is not tangent to $\partial_2 D$, in the vicinity of a segment of the boundary we are able to construct a local curvilinear coordinate system, such that the $\tilde{x}$ coordinate lines are tangent to $\gamma$ on $\partial_2 D$, the boundary coincides with the $\tilde{y} = 0$ coordinate line, and the interior of the domain has $\tilde{y} > 0$. When $X^t_s$ reaches
Figure 2.1: A schematic illustration of a 2-D domain $D$ with a Neumann-type boundary $\partial_2 D$ and a local curvilinear coordinate system $(\tilde{x}, \tilde{y})$ constructed near the boundary. The $\tilde{x}$ coordinate lines are tangent to the $\gamma$ vectors on $\partial_2 D$, and the $\tilde{y} = 0$ coordinate line coincides with the boundary. The global coordinate is labelled $\xi$. 
$\partial_2 D$, we can decompose it into the $(\tilde{x}, \tilde{y})$ coordinates, and the local time $k_s$ increases to ensure that the $\tilde{y}$ component of $X_s^t$ is always greater than 0, while its $\tilde{x}$ component is left unaffected. When $X_s^t$ is away from $\partial_2 D$, $k_s$ then remains unchanged.

Next, we consider the initial value and mixed boundary condition problem, which is the type of problem that we encounter to solve the radiation belt Fokker-Planck equation in this thesis:

$$
\frac{\partial}{\partial t} u(t, \xi) = \tilde{L}u(t, \xi), \quad (2.89)
$$

$$
u(0, \xi) = f(\xi), \quad \xi \in D, \quad (2.90)
$$

$$
u(t, \xi) = g(t, \xi), \quad \xi \in \partial_1 D, \quad (2.91)
$$

$$
\gamma(\xi) \cdot \nabla u(t, \xi) = 0, \quad \xi \in \partial_2 D. \quad (2.92)
$$

The functional $F^{t, \xi}$ for this problem has the same form as (2.86), and the solution of the problem is still given by (2.75), except now $X_s^t$ must be expressed by (2.88).

Finally, for completeness of the theory, we mention here the situation involving the boundary condition of the third kind, specified by

$$
\gamma(\xi) \cdot \nabla u(t, \xi) - \lambda(t, \xi)u(t, \xi) = 0, \quad \xi \in \partial_3 D, \quad (2.93)
$$

where $\lambda(t, \xi)$ is a continuous and bounded function on the boundary $\partial_3 D$. In this case, for the initial value and mixed boundary condition problem with equations (2.89), (2.90), (2.91) and (2.93), the SDE for $X_s^t$ still has the form of (2.88), but now the functional must be modified as [Costantini et al., 1998]

$$
F^{t, \xi} = \mathbb{I}_{t \geq t} f(X_t^t) \exp \left( \int_0^t c(t - s, X_s^t)ds - \int_0^t \lambda(t - s, X_s^t)dk_s \right) + \mathbb{I}_{t < \tau} g(t - \tau, X_t^\tau) \exp \left( \int_0^\tau c(t - s, X_s^\tau)ds - \int_0^\tau \lambda(t - s, X_s^\tau)dk_s \right). \quad (2.94)
$$

Using the variational theory of PDEs (similar to that illustrated in Section 3.1.4), it can be shown that the radiation belt Fokker-Planck equation with nonzero dynamical
friction (advection) is related to this case. For the equation (2.80), the third kind boundary condition is

$$\frac{1}{2} \hat{n} \cdot a \cdot \nabla \phi(t, \eta) - \hat{n} \cdot h \phi(t, \eta) = 0, \quad \eta \in \partial_3 D,$$

(2.95)

in which $\hat{n}$ is the unit inward normal vector of $\partial_3 D$ (the $K = 0$ boundary). Therefore, the unit vector $\gamma$ and the function $\lambda$ are given respectively by

$$\gamma(\eta) = \frac{\hat{n} \cdot a}{|\hat{n} \cdot a|},$$

(2.96)

and

$$\lambda(t, \eta) = 2 \frac{\hat{n} \cdot h}{|\hat{n} \cdot a|},$$

(2.97)

and the “source or loss” term $c$ is given by $-\nabla \cdot h$ as shown in (2.82). Although solvable by the SDE method, the Fokker-Planck equation with nonzero dynamical friction is more commonly associated with the inner radiation belt particle diffusion, and is out of the scope of this thesis.
Chapter 3

The Radbelt Electron Model

Based on the SDE theory presented in Chapter 2, the 3-D Radbelt Electron Model (REM) for solving the radiation belt Fokker-Planck equation in adiabatic invariant coordinates is encoded with FORTRAN’03 and C, and parallelized using OpenMP. The model obtains a probabilistic solution by solving a set of Itô SDEs that are mathematically equivalent to the Fokker-Planck equation. In Section 3.1, the radiation belt Fokker-Planck equation is transformed into its corresponding SDEs, and special considerations on the boundary conditions are provided. Section 3.2 addresses the key numerical techniques that constitute the essence of REM, including matrix factorization, the adaptive Euler-Maruyama scheme of integrating SDEs, and an adaptive statistical error estimate. Section 3.3 describes the generation of the PSD source plots, which is a unique feature of the SDE method. A short summary of the SDE method is given in Section 3.4. A part of the content in this chapter is published in Zheng et al. [2014].

3.1 Solving the radiation belt Fokker-Planck equation in REM

3.1.1 Transformation of the Fokker-Planck equation and connection to SDEs

In Section 1.2 we have defined the adiabatic invariants for the radiation belt electron motion, and in Section 1.4 we have formulated the radiation belt Fokker-Planck
equation (1.31) as well as its transformation to other coordinates (1.32). Instead of the action integrals $J_i$, in practice it is customary to regard the PSD $\bar{f}$ as a function of the more convenient adiabatic invariants $M, K$ and $L$, whose expressions are given in (1.3), (1.5) and (1.7) respectively.

For computational purposes, $u \equiv \ln(M)$ is a better variable than $M$, for the following three reasons. First, in the typical energy range of electrons in the outer radiation belt, the value of $M$ spans four orders of magnitude, from a few to over ten thousand MeV/G. Second, changing to $\ln(M)$ sends the lower bound of the variable from 0 to $-\infty$, which prevents unphysical scenarios of stochastic processes traveling into “negative $M$” regions during numerical implementation. Third, in the phase space with the $\ln(M)$ coordinate, the family of constant energy surfaces have the same shape in a dipole magnetic field, which is not true using $M$; this feature makes the computational domain less irregular. Similarly, the numerical value of $K$ increases dramatically when $\alpha_0$ is small. However, unlike $M$, $K$ reaches zero for equatorially bouncing electrons. Therefore we use $v \equiv \ln(K + 1)$ as the second coordinate variable, with $K$ evaluated in the unit $G^{1/2}R_E$. The third coordinate is $L$.

After these transforms, the three adiabatic invariant coordinates $(u, v, L)$ are hence all dimensionless. An early stage of REM that works in the $(u, K, L)$ coordinates was published in Zheng et al. [2014]. In this section, the latest REM is presented.

Following equation (1.32), we write the Fokker-Planck equation in the coordinates $(u, v, L)$. The Jacobian determinant is

$$G = 8\sqrt{2}\pi^2 \frac{m_0^{3/2} \mu_E \exp(\frac{3}{2}u) \exp(v)}{R_E^2}. $$

The Fokker-Planck equation is transformed to the Kolmogorov backward equation form exactly as shown in the example equation (2.67), with $\eta$ replaced by $Q =$
\((Q_1, Q_2, Q_3) \equiv (u, v, L), \ G \) given by \((3.1)\), and \(\phi\) replaced by \(\tilde{f}\):

\[
\frac{\partial}{\partial t} \tilde{f}(t, Q) = \frac{1}{2} \left(2D_{Q_i Q_j}\right) \frac{\partial^2}{\partial Q_i \partial Q_j} \tilde{f} + \left[ \frac{1}{G} \frac{\partial (GD_{Q_i Q_j})}{\partial Q_j} \right] \frac{\partial}{\partial Q_i} \tilde{f}.
\tag{3.2}
\]

Here \(\tilde{f}\) is the PSD with respect to canonical variables, though expressed in the \((u, v, L)\) coordinates (cf. Section 1.4 and Subsection 2.3.3). Comparing the differential operator in \((3.2)\) with its definition \((2.10)\), we find that the coefficients \(a_{ij}\) and \(b_i\) are given by

\[
a = 2 \begin{pmatrix} D_{uu} & D_{uv} & D_{uL} \\ D_{uv} & D_{vv} & D_{vL} \\ D_{uL} & D_{vL} & D_{LL} \end{pmatrix},
\tag{3.3}
\]

and

\[
b_1 = \exp\left(-\frac{3}{2}u\right) \frac{\partial}{\partial u} \left[\exp\left(\frac{3}{2}u\right)D_{uu}\right]
+ \exp(-v) \frac{\partial}{\partial v} \left[\exp(v)D_{uv}\right] + L^2 \frac{\partial}{\partial L} \left(\frac{D_{uL}}{L^2}\right),
\tag{3.4}
\]

\[
b_2 = \exp\left(-\frac{3}{2}u\right) \frac{\partial}{\partial u} \left[\exp\left(\frac{3}{2}u\right)D_{uv}\right]
+ \exp(-v) \frac{\partial}{\partial v} \left[\exp(v)D_{vv}\right] + L^2 \frac{\partial}{\partial L} \left(\frac{D_{vL}}{L^2}\right),
\tag{3.5}
\]

\[
b_3 = \exp\left(-\frac{3}{2}u\right) \frac{\partial}{\partial u} \left[\exp\left(\frac{3}{2}u\right)D_{uL}\right]
+ \exp(-v) \frac{\partial}{\partial v} \left[\exp(v)D_{vL}\right] + L^2 \frac{\partial}{\partial L} \left(\frac{D_{LL}}{L^2}\right).
\tag{3.6}
\]

Following the results of Chapter 2, the 3-D SDE associated with the Kolmogorov backward equation \((3.2)\), in the form shown by \((2.88)\), has coefficients \(b\) given by expressions \((3.4)\)–\((3.6)\), and the matrix \(\sigma\) calculated from the decomposition of \(a\) in \((3.3)\) according to \((2.11)\), i.e.,

\[
\sigma \sigma^\top = a.
\tag{3.7}
\]
The reflection vector $\gamma$ will be addressed in Section 3.1.3 together with the boundary conditions. The solution of equation (3.2) is represented by

$$\bar{f}(t, Q) = E[F^{t,Q}], \quad (3.8)$$

with the functional $F^{t,Q}$ in the form as given by (2.86).

The matrix $\sigma$ is not uniquely determined by equation (3.7). However, according to Levi’s theorem [Freidlin, 1985], different $\sigma$’s satisfying (3.7) generate equivalent stochastic processes, and hence yield the same solution of the diffusion equation. To see this, we observe that, for any orthogonal matrix $U$, the new $\sigma$ obtained from $\tilde{\sigma} \equiv \sigma U$ still satisfies equation (3.7), and, when inserted into the SDE (2.88), $\tilde{\sigma}d\tilde{W}_s$ is equivalent to $\sigma d\tilde{W}_s$, where $\tilde{W}_s$ is merely the Wiener process $W_s$ after a rotation and perhaps a reflection of coordinate axes, which is still a Wiener process [Zhang, 1999].

The 3-D diffusion tensor $D_{Q_iQ_j}$ in the $(u,v,L)$ coordinates needs additional specification. Among its six independent components, the radial diffusion coefficient $D_{LL}$ may be provided by empirical models [e.g., Brautigam and Albert, 2000; Ozeke et al., 2012], or calculated using global MHD simulations [e.g., Fei et al., 2006; Tu et al., 2012]. The off-diagonal components $D_{uL}$ and $D_{vL}$ are related to drift shell splitting in asymmetric magnetic field [e.g., O’Brien, 2014, 2015] (cf. Section 1.2), which will be addressed in Chapter 5. In dipole field, they are both zero. The remaining components $D_{uu}$, $D_{uv}$ and $D_{vv}$ are related to energy-pitch-angle scattering of the electrons, and they can be transformed from the more commonly computed diffusion coefficients in the $\alpha_0$ and $p$ coordinates, by a transformation like (1.33). Written explicitly, they are

$$
\begin{pmatrix}
D_{uu} & D_{uv} \\
D_{uv} & D_{vv}
\end{pmatrix} =
\begin{pmatrix}
\frac{\partial u}{\partial \alpha_0} & \frac{\partial u}{\partial p} \\
\frac{\partial v}{\partial \alpha_0} & \frac{\partial v}{\partial p}
\end{pmatrix}
\begin{pmatrix}
D_{\alpha_0\alpha_0} & D_{\alpha_0p} \\
D_{\alpha_0p} & D_{pp}
\end{pmatrix}
\begin{pmatrix}
\frac{\partial u}{\partial \alpha_0} & \frac{\partial u}{\partial p} \\
\frac{\partial v}{\partial \alpha_0} & \frac{\partial v}{\partial p}
\end{pmatrix}.
$$

(3.9)
The partial derivatives on the right-hand side of equation (3.9) should be regarded as functions of \( u, v \) and \( L \), and are field-geometry dependent, so that in general particle tracings are required to evaluate these derivatives.

### 3.1.2 Transformation matrix in dipole field

For dipole field geometry, the components of the transformation matrix in equation (3.9) can be expressed analytically and are derived below. To express these derivatives as functions of \( u, v \) and \( L \) in dipole field, equations (1.9) and (1.10) must be inverted to yield \( y \) and \( p \) as functions of \( M, K \) and \( L \). For this purpose, we use the formula (1.13), and denote \( 1/y^2 \equiv z(v, L) \). The derivatives in (3.9) are then calculated as:

\[
\frac{\partial u}{\partial \alpha_0} = \frac{dy}{d\alpha_0} \frac{\partial u}{\partial y} = 2\sqrt{z - 1}, \quad (3.10)
\]

\[
\frac{\partial u}{\partial p} = -\frac{2}{p} \frac{\mu_E}{R_E} \frac{z(z - 1)}{L} \frac{\exp(-v)}{\exp(u)}, \quad (3.11)
\]

\[
\frac{\partial v}{\partial \alpha_0} = -2 \frac{\mu_E}{R_E} \frac{T^{-1/2}(y)}{y^2} \sqrt{\frac{z(z - 1)}{L}} \exp(-v), \quad (3.12)
\]

\[
\frac{\partial v}{\partial p} = 0. \quad (3.13)
\]

In the derivation of equation (3.12), we have used the differential relation \([Schulz, 1991, eq. (159)]\)

\[
\frac{d}{dy} \frac{Y(y)}{y} = -\frac{2T(y)}{y^2}, \quad (3.14)
\]

where

\[
T(y) = \frac{1}{4LR_E} \oint \frac{ds}{\cos \alpha} \quad (3.15)
\]

is the normalized dipole bounce period, which is well approximated by

\[
T(y) \approx 1.380173 - 0.639693 \, y^{3/4}. \quad (3.16)
\]
We note that, because $K$ is independent of particle energy, $\partial v/\partial p$ is always zero, even in general magnetic fields, as long as the field is slowly varying (compared to a bounce period) and the particles experience no external force other than the magnetic force [Roederer, 1970].

### 3.1.3 Computational domain

The ability to deal with complicated boundary geometries gives the SDE method advantages in solving the electron Fokker-Planck equation in adiabatic invariant coordinates. The computational domain is determined by selecting a range of electron kinetic energy $E$, usually from a few hundred keV to several MeV, $\alpha_0$ from 90° to the local bounce loss-cone angle, and a range of $L$. These boundaries are subsequently transformed into the $(u,v,L)$ space. The boundaries transformed using dipole field geometry are illustrated in Figure 3.1. The upper panel of Figure 3.1 shows a 3-D view of the computational domain, a skewed hexahedron, with colored contours at four fixed $L$ values. Visible on the right-hand side of the hexahedron (maximum $u$) is the curved surface of constant maximum energy, which has the same shape as the constant minimum energy surface (invisible from this angle) due to the dipole field geometry. The front surface is the plane of the $v = 0$ boundary corresponding to $\alpha_0 = 90°$. The top surface is the upper-$L$ plane. The maximum $v$ surface (corresponding to dipole bounce loss-cone angles) and the lower-$L$ plane are not visible from this point of view. The lower panel of Figure 3.1 shows the projection of the colored constant-$L$ contours in the $u$-$v$ plane. As $L$ increases, the maximum $v$ value increases as a result of the diminishing loss cone, and the contour moves as a whole toward the larger $u$ region.
Figure 3.1: An example computational domain in the $(u,v,L)$ space in a dipole magnetic field (upper panel), defined by $E$ from 0.1 to 10.0 MeV, $\alpha_0$ from bounce loss cone to 90°, and $L$ from 2.0 to 8.0. The lower panel shows the projections in the $u$-$v$ plane of the colored contours in the upper panel, labeled by their $L$ values.
3.1.4 Boundary conditions

Boundary conditions are assigned as follows. Dirichlet boundary conditions are specified on the surfaces of constant maximum energy ($E_{\text{max}}$), constant minimum energy ($E_{\text{min}}$), maximum $L$ ($L_{\text{max}}$), minimum $L$ ($L_{\text{min}}$) and loss-cone $v$ ($v_{\text{max}}$). In particular, we assume negligible electron flux at the maximum energy and the loss cone, therefore the $E_{\text{max}}$ and $v_{\text{max}}$ boundary values are zero. In other words, Itô processes are lost upon reaching the $E_{\text{max}}$ or $v_{\text{max}}$ boundaries. These boundary conditions are apparent from the nature of the diffusion problem, though the loss cone could be nonempty in the strong diffusion limit.

It is, however, not so apparent what form the boundary condition takes for the $v = 0$ boundary. To answer this question, we examine the Fokker-Planck equation in light of the variational theory of PDEs. Starting from equation (1.32) and assuming that the spatial and temporal variables are separable, i.e., $\tilde{f}(t, Q_i) = \Psi_n(t) \Phi_n(Q_i)$ with $n$ indexing the $n$-th eigenmode, equation (1.32) can be separated into (with no summation over repeated $n$ in (3.17), (3.18) and (3.20))

$$\frac{d \Psi_n}{dt} = -\lambda_n \Psi_n,$$  \hspace{1cm} (3.17)

and

$$\frac{\partial}{\partial Q_i} \left( GD_{Q_i, Q_j} \frac{\partial \Phi_n}{\partial Q_j} \right) + \lambda_n G \Phi_n = 0,$$  \hspace{1cm} (3.18)

where $\lambda_n$ is the eigenvalue (physically, the temporal decay rate of the $n$-th eigenmode).

For each $n$, equation (3.18) is the Euler-Lagrange equation of the functional

$$\mathcal{I}[\Phi_n] = \int_D \mathcal{F}(\Phi_n, \Phi_n, Q_i) dQ_1 dQ_2 dQ_3,$$  \hspace{1cm} (3.19)

in which $\Phi_{n,i} \equiv \partial \Phi_n / \partial Q_i$, and the integrand function $\mathcal{F}$ is

$$\mathcal{F}(\Phi_{n,i}, \Phi_n, Q_i) = G \left( D_{Q_i, Q_j} \Phi_{n,j} \Phi_{n,j} - \frac{1}{2} \lambda_n \Phi_n^2 \right).$$  \hspace{1cm} (3.20)
Whereas \( \bar{f} \), and hence \( \Phi_n \), are often held fixed on the domain boundaries, they are free to change on the \( Q_2 = 0 \) boundary. As a result, to make the functional \( \mathcal{I}[\Phi_n] \) stationary against variations \( \delta \Phi_n \), in addition to the Euler-Lagrange equation, we must also require the condition [e.g., Mathews and Walker, 1970]
\[
\frac{\partial \mathcal{F}}{\partial \Phi_n, 2} \bigg|_{Q_2=0} = 0. \tag{3.21}
\]
Applied to (3.20), equation (3.21) then gives a Neumann boundary condition, written explicitly in terms of \( \bar{f} \) and \( (u,v,L) \),
\[
\left( D_{uv} \frac{\partial \bar{f}}{\partial u} + D_{vv} \frac{\partial \bar{f}}{\partial v} + D_{vL} \frac{\partial \bar{f}}{\partial L} \right) \bigg|_{v=0} = 0. \tag{3.22}
\]
In the SDE language, this boundary condition indicates that, on the \( v = 0 \) boundary, the stochastic processes are reflected locally about the unit vector \( \gamma(Q) \propto (D_{uv}, D_{vv}, D_{vL})^\top \). From another point of view, the left-hand side of (3.22) is just the negative of the \( v \)-component of the PSD diffusive flux (in a curvilinear coordinate system). Therefore, the physical meaning of (3.22) is that the PSD is not allowed to flow across the \( v = 0 \) surface, which is natural since the region of “negative \( v \)” is physically meaningless. This conclusion confirms the criteria of adjoint operators discussed in Section 2.2 (equations (2.26)–(2.27)).

**Non-degeneracy condition for Neumann boundary condition (3.22)**

Neumann-type boundaries require that the diffusion tensor must be non-degenerate in the direction of \( \gamma \) on \( \partial_2 D \), which means the following quadratic form inequality must always hold [Freidlin, 1985]
\[
\gamma^\top(Q) a(t, Q) \gamma(Q) > 0, \quad Q \in \partial_2 D. \tag{3.23}
\]
For the \( v = 0 \) boundary with boundary condition (3.22) to be non-degenerate, we now prove that the sufficient and necessary condition is \( D_{vv} > 0 \) on the boundary.
For convenience of discussion, we will use numeric indices 1, 2, 3, instead of $u, v, L$, as the subscripts of components in this proof. The unnormalized reflection vector

$$\tilde{\gamma} \equiv (D_{12}, D_{22}, D_{32})^\top \quad (3.24)$$

coincides with the second column of the diffusion tensor $\mathbf{D}$. Because $\mathbf{D}$ is symmetric, the quadratic form of $\mathbf{D}$ with its own column vector is equal to a diagonal component of the matrix product $\mathbf{D}^3$. In this case, it is

$$\tilde{\gamma}^\top \mathbf{D} \tilde{\gamma} = (\mathbf{D}^3)_{22}. \quad (3.25)$$

A real symmetric matrix can be diagonalized via orthogonal transformations, i.e.,

$$\mathbf{D} = \mathbf{P} \Lambda \mathbf{P}^\top,$$

where $\Lambda$ is a real diagonal matrix of $\mathbf{D}$’s eigenvalues, and $\mathbf{P}$ is an orthogonal matrix composed of $\mathbf{D}$’s orthogonalized eigenvectors. Consequently,

$$\mathbf{D}^3 = \mathbf{P} \Lambda^3 \mathbf{P}^\top. \quad (3.26)$$

Therefore the diagonal components of $\mathbf{D}$ and $\mathbf{D}^3$ can be expressed respectively by

$$D_{ii} = \Lambda_k P_{ik}^2, \quad (3.27)$$

and

$$\left(\mathbf{D}^3\right)_{ii} = \Lambda_k^3 P_{ik}^2, \quad (3.28)$$

where $\Lambda_k$ is the $k$-th diagonal component of $\Lambda$, i.e., the $k$-th eigenvalue of $\mathbf{D}$. Diffusion processes require that the diffusion tensor must be at least positive semi-definite (this can be seen from equation (2.11) and the discussion thereafter), which implies $\Lambda_k$ must be non-negative. Hence, from equations (3.27) and (3.28), it follows that there is always $(\mathbf{D}^3)_{ii} \geq 0$; and the only situation for $(\mathbf{D}^3)_{ii} = 0$ is when at least either $P_{ik}^2$ or $\Lambda_k$ is zero for each $k$, which is also the situation for $D_{ii} = 0$. In other words,
$D_{ii} > 0$ is a sufficient and necessary condition for $(D^3)_{ii} > 0$. Relating this conclusion to equation (3.25), equation (3.3), and the non-degeneracy requirement of Neumann boundary conditions (3.23), we see that the sufficient and necessary condition for the boundary condition (3.22) to be non-degenerate is $D_{22}$, or $D_{vv}$, being greater than zero on the boundary.

**Proof that $\frac{\partial K}{\partial \alpha} \bigg|_{\alpha=\pi/2} = 0$ in a general magnetic field**

We next prove that $\partial K/\partial \alpha_0$ vanishes at $K = 0$ in general magnetic fields, and so must the transformation matrix component $\partial v/\partial \alpha_0$. This then simplifies the Neumann boundary condition (3.22). Taylor expanding the magnetic field strength $B(s)$ around its equatorial minimum $B_0$ along the field line gives

$$B(s) = B_0 + \frac{1}{2} \left( \frac{\partial^2 B}{\partial s^2} \right)_0 s^2 + \mathcal{O}(s^4),$$

(3.29)

where $s$ is the field line arc length, and $(\partial^2 B/\partial s^2)_0$ is a field-geometry dependent quantity, which we assume is nonzero. Note that, up to the second order term in (3.29), for any near equatorially bouncing particles (meaning $s_m$ is small), the field is symmetric about the minimum point regardless of the specific field geometry, and so are the two mirror point positions. By equation (1.11) and truncating (3.29) to the second order term, we can express the mirror point field intensity as

$$B_m = \frac{B_0}{y^2} = B_0 + \frac{1}{2} \left( \frac{\partial^2 B}{\partial s^2} \right)_0 s_m^2.$$  

(3.30)

Introducing the auxiliary variables

$$\Delta \equiv B_m - B_0 = B_0 \left( \frac{1}{y^2} - 1 \right),$$

(3.31)

and

$$\Gamma \equiv \frac{1}{2} \left( \frac{\partial^2 B}{\partial s^2} \right)_0,$$  

(3.32)
solving (3.30) for $s_m$ gives

$$s_m = \sqrt{\frac{\Delta}{\Gamma}}, \quad (3.33)$$

and

$$B_m - B(s) = \Delta - \Gamma s^2. \quad (3.34)$$

Now, according to (1.5), the adiabatic invariant $K$ can be worked out explicitly, for $y \to 1^-$,

$$K = 2 \int_0^{s_m} \sqrt{B_m - B(s)} \, ds$$

$$= 2 \left[ \frac{s}{2} \sqrt{\Delta - \Gamma s^2} + \frac{\Delta}{2 \sqrt{\Gamma}} \arcsin \left( s \sqrt{\frac{\Gamma}{\Delta}} \right) \right]_0^\Delta$$

$$= \frac{\pi}{2} \frac{\Delta}{\sqrt{\Gamma}} = \frac{\pi B_0}{2 \sqrt{\Gamma}} \left( \frac{1}{y^2} - 1 \right). \quad (3.35)$$

The particular field geometry enters the expression of $K$ through $\sqrt{\Gamma}$ in (3.35). Taking the derivative

$$\left. \frac{\partial K}{\partial y} \right|_{y=1} = -\frac{\pi B_0}{\sqrt{\Gamma}}, \quad (3.36)$$

which is a finite and field-geometry dependent quantity. But since

$$\frac{dy}{d\alpha_0} = \cos \alpha_0, \quad (3.37)$$

we always have

$$\left. \frac{\partial K}{\partial \alpha_0} \right|_{\alpha_0=\pi/2} = \left( \frac{\cos \alpha_0}{\partial y} \frac{\partial K}{\partial \alpha_0} \right)_{\alpha_0=\pi/2} = 0, \quad (3.38)$$

for all magnetic field lines with a nonzero $(\partial^2 B/\partial s^2)_0$, as claimed above.

At $v = 0$, since both $\partial v / \partial \alpha_0$ and $\partial v / \partial p$ are zero, the transformation matrix in (3.9) is singular; and assuming that $D_{\alpha \alpha_0}$ is finite, the transformed $D_{uv}$ and $D_{vv}$ are both zero there. ($D_{vL}$ would also be zero at $v = 0$ in an asymmetric field, due to this singularity. See equation (5.1) in Chapter 5.) This is an artifact of calculating the
$u$-$v$ diffusion coefficients from the $\alpha_0$-$p$ coefficients. To circumvent this artifact, we invoke small but nonzero bounce-resonance effects in the equatorial plane (as argued physically by Roberts [1969]) to give a nonzero $D_{vv}$. Then, the boundary condition (3.22) reduces to
\begin{equation}
\frac{\partial \bar{f}}{\partial v} \bigg|_{v=0} = 0,
\end{equation}
and the reflection vector $\gamma$ is now simply $(0, 1, 0)^T$, i.e., the normal to the $v = 0$ boundary. We note that the Salammbô code [Beutier and Boscher, 1995] used (without explanation) the Neumann boundary condition $\partial \bar{f}/\partial J_2 = 0$ at $J_2 = 0$, which is equivalent to (3.39). Lastly, we emphasize that the variational method of determining the Neumann boundary condition and the discussion of the non-degeneracy condition are general; they apply to the Fokker-Planck equation in any other coordinate variables.

### 3.2 Numerical techniques

#### 3.2.1 Overview

Implementation of the SDE method starts from numerically integrating the SDE (2.88) in time. In the interior of the domain, the SDE is discretized using the Euler-Maruyama scheme [Kloeden and Platen, 1992]:
\begin{equation}
X_{n+1} = X_n + b(t - s_n, X_n)\Delta s + \sigma(t - s_n, X_n)\Delta W_n,
\end{equation}
where the time step $\Delta s = s_{n+1} - s_n$, and $X_n$ is the approximation of $X_s'$ at time $s_n$. The increment of the Wiener process $\Delta W_n$ is generated by
\begin{equation}
\Delta W_n = \sqrt{\Delta s} \begin{pmatrix} \mathcal{N}_1 \\ \mathcal{N}_2 \\ \mathcal{N}_3 \end{pmatrix},
\end{equation}
with $\mathcal{N}_i$ normally distributed random variables.
where \( N_1, N_2 \) and \( N_3 \) are three standard Gaussian random numbers each with zero mean and unit variance, which are transformed from three independent uniform deviations via the Box-Muller algorithm \([\text{Press et al.}, 1992]\). The uniform deviations are generated by the Dynamic Creator of Mersenne Twisters \([\text{Matsumoto and Nishimura}, 1998, 2000]\), a fast and sophisticated multiple-stream pseudo-random number generator based on Mersenne primes, which ensures their mutual statistical independence. Mutual independence between pseudo-random number streams is important for implementing code parallelization for all Monte-Carlo-like methods, because otherwise the numerical sampling will be biased.

Calculation of the \( \sigma \) matrix in (3.40) is based on equation (3.7). According to the discussion after equation (3.7), the non-uniqueness of \( \sigma \) gives us freedom to choose the most economic method to accomplish the factorization, and most commonly, it is Cholesky decomposition. However, some special cases of the radiation belt diffusion problem prevent straightforward application of a Cholesky decomposition. Subsection 3.2.2 gives a description of the numerical techniques of REM to solve equation (3.7).

Numerical integration of (3.40) presents the question of how to determine the time step size. In the mathematical literature, the Euler-Maruyama scheme is most thoroughly studied and usually implemented using equidistant time steps \([\text{e.g., Kloeden and Platen}, 1992]\). For evaluation of functional expectations like ours, where only the statistical distribution of random walks matters but not their individual paths, the convergence of the Euler-Maruyama scheme is of order 1.0, meaning that the mean error of the scheme is proportional to the first power of \( \Delta s \) \([\text{Kloeden and Platen}, 1992]\). But for the radiation belt diffusion problem, the diffusion coefficients differ by orders of magnitude across the phase space, which render the fixed step-size Euler-Maruyama scheme both slow and inaccurate. In Subsection 3.2.3, we present
an adaptive time step-size method of REM, which is based on an estimate of an analogous “mean-free-path” of the stochastic motions. The adaptive method improves both the speed and accuracy of the Euler-Maruyama scheme in practice, but whether it increases the order of convergence remains an open question.

Stochastic paths are nowhere differentiable. Therefore, integration of (3.40) only provides an approximation to the stochastic path at the time discretization steps $s_n$. An approximation of the entire path over the continuous time interval $0 \leq s \leq t$ is obtained by linear interpolations between $X_n$ and $X_{n+1}$, i.e., connecting them by straight line segments. On Dirichlet-type boundaries, the stopping position of an Itô process is then determined by the intersection between the boundary and this linear-interpolated path. The accuracy of this stopping strategy is of order $0.5$ [Gobet, 2001]. Sophisticated order 1.0 schemes are also available [e.g., Gobet, 2001], but their implementations often require much more complicated computation near the boundary. On the Neumann-type boundary, we use the Symmetrized Euler Scheme (order 1.0) [Bossy et al., 2004] to reflect the Itô processes.

The functional expectation $\mathbb{E}[F_{t,Q}]$ is evaluated by the arithmetic mean of the functional values obtained from various stochastic paths. In practice, only a finite number of Itô processes can be simulated, which introduces statistical fluctuations to the result. The statistical error is proportional to $1/\sqrt{N}$ with sufficiently large total number of stochastic paths $N$, and has zero expectation. Subsection 3.2.4 explains how the statistical error, in the form of a confidence interval, is estimated in REM, so that $N$ is truncated at the right value to both meet a prescribed error tolerance and reduce computational effort. In this sense, REM is also adaptive in the number of stochastic paths.

Consequently, the error of the SDE method comes from two parts: one is the
systematic error intrinsic to the numerical schemes, which can be directly reduced by using smaller step size (but above the round-off error limit); the other is the statistical error from the finite number of stochastic paths. For the systematic error, schemes of higher order than Euler can be adopted to improve the convergence, but with a penalty of more intricate computations. For the statistical error, variance reduction techniques, which use cleverer but more expensive ways to estimate expectation than the arithmetic mean, are possible mitigations [Kloeden and Platen, 1992]. These prospective improvements are discussed in the last chapter.

### 3.2.2 Matrix factorization

A key step in the SDE method of solving diffusion equations is the calculation of the $\sigma$ matrix, or in mathematical terms, the factorization of a known real symmetric matrix $A$ into

$$ A = LL^\top. \quad (3.42) $$

This calculation is done at every step of numerically integrating an SDE, thus it is extremely heavily performed (more than $10^6$ times for one solution point with physical time period one day in REM). Therefore, in numerically implementing (3.42), algorithm speed is an important consideration, if not the first. In REM, the factorization routines (Cholesky and spectral, see below) involved in solving (3.42) are employed from LAPACK [Anderson et al., 1999].

If the real symmetric matrix $A$ is positive definite, which means all its eigenvalues are positive, Cholesky decomposition [e.g., Press et al., 1992] is in general the simplest and fastest method of choice. Briefly, Cholesky decomposition results in a unique lower-triangular (or upper-triangular if one so chooses) factorization matrix $L$, where
the entries of $\mathbf{L}$ are calculated by (without summation convention)

$$L_{ii} = \left( A_{ii} - \sum_{k=1}^{i-1} L_{ik}^2 \right)^{1/2},$$

(3.43)

and

$$L_{ji} = \frac{1}{L_{ii}} \left( A_{ij} - \sum_{k=1}^{i-1} L_{ik} L_{jk} \right), \quad j > i.$$  

(3.44)

In diffusion problems, however, the matrix $\mathbf{A}$ is allowed to be positive semi-definite, so that one or more of its eigenvalues can be zero. Even worse, $\mathbf{A}$ could be indefinite by having small negative eigenvalues, which should have been zero or positive, because of numerical errors from calculating the matrix. These cases are readily indicated by failure of Cholesky decomposition in numerical implementation. One remedy is to make a spectral factorization of $\mathbf{A}$, i.e., to decompose it into a diagonal matrix $\mathbf{\Lambda}$ of its eigenvalues and an orthogonal matrix $\mathbf{Q}$ of the corresponding eigenvectors

$$\mathbf{A} = \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^T.$$  

(3.45)

Then the matrix square-root can be calculated for $\mathbf{\Lambda}$, which yields a diagonal matrix with elements (without summation convention)

$$V_{ii} = \Lambda_{ii}^{1/2},$$

(3.46)

provided that all small negative eigenvalues are corrected to zero. Combining (3.45) and (3.46), the matrix $\mathbf{A}$ is then successfully factorized into

$$\mathbf{A} = (\mathbf{Q} \mathbf{V}) (\mathbf{Q} \mathbf{V})^T,$$  

(3.47)

and this product in return guarantees positive semi-definiteness of $\mathbf{A}$. Unlike Cholesky decomposition, numerically diagonalizing a matrix involves a potentially large number of coordinate rotations (this is known as the Jacobi method) and is more expensive
[e.g., Press et al., 1992]. If we know a priori that $\mathbf{A}$ is positive semi-definite, pivoted Cholesky decomposition [e.g., Lucas, 2004] could also be a candidate, and may be significantly faster than spectral factorization when the matrix is large enough. However, in REM, the diffusion tensor is only $3 \times 3$; and further, to distinguish between positive semi-definiteness and indefiniteness one needs to take extra effort first, e.g., calculating its determinant. For these reasons, the pivoted Cholesky decomposition method is not adopted in REM for dealing with positive semi-definite matrices.

There is one situation of positive semi-definiteness that we should pay attention to and can make use of: the semi-definiteness caused by $\mathbf{A}$’s zero diagonal elements. A corollary of positive semi-definiteness is that the determinants of all $\mathbf{A}$’s principal minors are non-negative. Suppose that $A_{ii} = 0$; taking an arbitrary $j \neq i$, we have

$$A_{ii}A_{jj} - A_{ij}^2 \geq 0 \Rightarrow A_{ij} = 0.$$  

(3.48)

Therefore, running $j$ through $\mathbf{A}$’s dimensions (except $i$), it is proved that the $i$-th column and $i$-th row of $\mathbf{A}$ are all zeros. Suppose that the $n \times n$ matrix $\mathbf{A}$ has $m$ zero diagonal entries, it is computationally wasteful to directly spectral-factorize such a matrix, because the rank of this problem is actually $n - m$. Rather, let us consider a special case that $\mathbf{A}$’s first $n - m$ rows are nonzero; and we denote $\mathbf{A}$’s leading $(n - m) \times (n - m)$ principal minor as $\mathbf{A}^\ast$. Then, it is apparent that the factorization of $\mathbf{A}$ can be devised in such a manner

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}^\ast & \mathbf{0}_{(n-m)\times m} \\ \mathbf{0}_{m\times(n-m)} & \mathbf{O}_{m \times m} \end{pmatrix} = \begin{pmatrix} \mathbf{L}^\ast & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{L}^\ast & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = \mathbf{LL}^\top,$$

(3.49)

where the factorization of $\mathbf{A}^\ast = \mathbf{L}^\ast \mathbf{L}^\ast\top$ can be either Cholesky if it is positive definite, or spectral if it is not.
In the more general case where \( A \)'s nonzero rows are not aligned from the 1st to the \((n - m)\)-th, we can always switch the nonzero rows with the zero rows, and switch the columns in the same manner, so that the transformed matrix \( \tilde{A} \) has the same form as in (3.49), and the orders of the nonzero rows and columns are kept unchanged. In matrix representation, these switches are accomplished by the following elementary matrix

\[
E(ij) = \begin{pmatrix}
I_{i-1} & 0 & 1 \\
0 & I_{j-i-1} & 1 \\
1 & 0 & I_{n-j}
\end{pmatrix}, \quad (j > i), \quad (3.50)
\]

where \( I_k \) is the identity matrix of rank \( k \). \( E(ij) \) has the effect that, when left-multiplied to a matrix \( A \), it exchanges \( A \)'s \( i,j \) rows; and when right-multiplied, it exchanges \( A \)'s \( i,j \) columns. Additionally,

\[
E(ij) = E(ij)^\top = E(ij)^{-1}. \quad (3.51)
\]

Therefore, the matrix \( \tilde{A} \) can be obtained after a series of such elementary transformations, by

\[
\tilde{A} = \begin{pmatrix}
\tilde{A}^* & 0 \\
0 & 0
\end{pmatrix} = E_p^\top \cdots E_1^\top A E_1 \cdots E_p = P^\top A P, \quad (3.52)
\]

in which the so-called pivoting matrix \( P = E_1 \cdots E_p \) is orthogonal, as seen from (3.51). The factorization of \( \tilde{A} \) then follows straightforwardly as that in (3.49), and gives

\[
\tilde{A} = \tilde{L} \tilde{L}^\top. \quad (3.53)
\]

Consequently, \( A \) is factorized by

\[
A = P \tilde{A} P^\top = (P\tilde{L})(P\tilde{L})^\top. \quad (3.54)
\]
In numerical practice, however, we need neither to really perform the transformations in (3.52) nor to calculate $P$. $\tilde{A}^*$ is effectively achieved by deleting all zero rows and columns from $A$, and $\tilde{L}^*$ as well as $\tilde{L}$ are obtained as those in (3.49). The only problem is how we can construct $L$ from $L = P\tilde{L}$ without recording $P$. From equation (3.52), the effect of left-multiplying $P^\top$ is to shift all $A$’s “in-between” zero rows after the nonzero rows. Since $P = (P^\top)^{-1}$, left-multiplying $P$ on $\tilde{L}$ would therefore shift its zero rows, which are all after the leading principal minor $\tilde{L}^*$, back to their original positions according to $A$. Consequently in numerical realization, only $\tilde{L}^*$ needs to be recorded. $L$ is readily obtained by inserting into $\tilde{L}^*$ the zero rows deleted from $A$ to their original positions, and then padding zero columns to the right to make the matrix square. For instance, consider the following $3 \times 3$ matrix

$$A = \begin{pmatrix} A_{11} & 0 & A_{13} \\ 0 & 0 & 0 \\ A_{13} & 0 & A_{33} \end{pmatrix}.$$  \tag{3.55}$$

Its principal minor after deleting the zero rows and columns is

$$\tilde{A}^* = \begin{pmatrix} A_{11} & A_{13} \\ A_{13} & A_{33} \end{pmatrix},$$  \tag{3.56}$$

and factorization of $\tilde{A}^*$ yields

$$\tilde{L}^* = \begin{pmatrix} L_{11}^* & L_{12}^* \\ L_{21}^* & L_{22}^* \end{pmatrix}.$$  \tag{3.57}$$

Then, the matrix $L$ is

$$L = \begin{pmatrix} L_{11}^* & L_{12}^* & 0 \\ 0 & 0 & 0 \\ L_{21}^* & L_{22}^* & 0 \end{pmatrix}.$$  \tag{3.58}$$
Of course, one may also insert another pair of $P^TP$ in (3.54), and get

$$A = (P\tilde{L}P^T)(P\tilde{L}P^T)^T.$$  

(3.59)

This means in the construction of $L$, in addition to inserting back all the zero rows, one also inserts back all the zero columns, so that the nonzero entries of $L$ are all at the same positions as the nonzero entries in $A$. Although formally more coherent, computationally this construction requires more operations than that in (3.54), therefore (3.54) is adopted in REM.

As a byproduct, this matrix factorization technique may also be useful in interpolation of a positive definite or semi-definite tensor field, for example the diffusion tensor field. Direct component-wise interpolation does not preserve definiteness. However, a component-wise interpolation of $L$ as calculated in this subsection guarantees definiteness of $A$ through the product $A = LL^T$.

### 3.2.3 The adaptive time step Euler-Maruyama scheme

Because the radiation belt diffusion problem has highly variable diffusion coefficients, it is numerically inefficient to use a fixed time step Euler-Maruyama scheme. In regions with very little diffusion, the simulated stochastic motion (3.40) takes tiny spatial steps, which is a waste of computational effort. In contrast, in regions with very large diffusion, the stochastic motion makes huge strides that may be greater than the variation length scale of the diffusion coefficients, and this becomes a source of numerical error. Therefore, an adaptive time step scheme, which adjusts the time step size according to local diffusion coefficients, is desirable.

In solving ordinary differential equations, an adaptive method often contains an error estimator, and the step size is repeatedly adjusted until the estimated one-step
error is acceptable compared to a prescribed tolerance. However, for SDE solvers, this adaptive paradigm is not easily realized because of the stochastic term. Instead, we ask this question: given a time increment $dt$, what is the mean spatial step size $dS$ of the stochastic motion as described by the SDE (2.8). If we have this knowledge in numerical integration, we can prescribe a fixed $dS$ and adjust $dt$ accordingly, so that the aforementioned numerical problem is avoided.

For this purpose, we reexamine the coefficient matrix $a$ given in equation (2.11). Following equation (3.47), we can decompose $a$ as

$$ a = \sigma \sigma^T = (QV)(QV)^T, \quad (3.60) $$

where $Q$ and $V$ are defined as in (3.45) and (3.46). This equation, however, does not mean that $\sigma = QV$. In fact, we can always insert an arbitrary orthogonal matrix pair $U^T U$, and have $\sigma \sigma^T = (QVU^T)(QVU^T)^T$. Thus in general,

$$ \sigma = QVU^T. \quad (3.61) $$

Equation (3.61) is called the singular value decomposition of $\sigma$, and $V_{ii}$ are its singular values. With this relation, we have the corollary

$$ \sigma^T \sigma = U \Lambda U^T, \quad (3.62) $$

which will be used later.

In time $dt$, the distance that a stochastic motion has travelled can be calculated as follows. First,

$$ dX_t^2 = b^2 dt^2 + 2b^T \sigma dW_t + dW_t^T \sigma^T \sigma dW_t $$

$$ = b^2 dt^2 + 2b^T \sigma dW_t + dW_t^T U \Lambda U^T dW_t \quad (3.63) $$

$$ = b^2 dt^2 + 2b^T \sigma dW_t + d\tilde{W}_t^\top \Lambda d\tilde{W}_t, \quad (3.64) $$
where in step (3.63), we have applied the corollary (3.62), and in step (3.64), according to Levi’s theorem, $d\tilde{W}_i$ is an equivalent infinitesimal Wiener process to $dW_t$ (cf. Subsection 3.1.1). Now we can calculate the expectation of $dX_t^2$, by (with explicit summation)

$$
\mathbb{E}[dX_t^2] = \mathbb{E}\left[\sum_i \Lambda_{ii} d\tilde{W}_i^2 + b^2 dt^2\right] + b^2 dt^2
$$

$$
= \sum_i \Lambda_{ii} \mathbb{E}[d\tilde{W}_i^2] + b^2 dt^2
$$

$$
= \sum_i \Lambda_{ii} dt + b^2 dt^2
$$

$$
\approx \text{tr}(a) dt,
$$

where in steps (3.65) and (3.66) we have used the property (2.2) of Wiener processes, and in step (3.67) the second order term is neglected and $\text{tr}(a)$ is the trace of the matrix $a$. Therefore, the root-mean-square distance travelled by a stochastic motion in time $dt$ is

$$
dS = \sqrt{\mathbb{E}[dX_t^2]} = \sqrt{\text{tr}(a)} dt.
$$

In numerical integration of (3.40), the time step size $\Delta s$ is not infinitesimal, so the spatial step caused by $b\Delta s$ cannot be omitted. In practice, we prescribe a desired root-mean-square spatial step size $\Delta S$, which is sufficiently small compared to the length scale of diffusion coefficient variation, then calculate the corresponding time step sizes both from $\Delta S = |b| \Delta s$ and from equation (3.68), with $a$ related to the in situ diffusion tensor by (3.3). The smaller of them is chosen as the actual time step size. Finally, this time step size is further bounded by a maximum of $10^{-2}$ day, roughly an electron drift period, and a minimum of $5 \times 10^{-7}$ day, comparable to a gyration period. Appendix B gives statistics of an example REM result that shows the effect of this adaptive scheme.
3.2.4 Estimation of statistical error

A simple and reliable way to estimate the expectation value of a random sequence is to calculate its arithmetic mean, especially when the statistical distribution of the sequence is unknown. In REM, the unknown functional expectation \( \mathbb{E}[F_{t,Q}] \) is thus estimated by the arithmetic mean of \( N \) simulated functional values as

\[
\hat{\mathbb{E}}[F_{t,Q}] = \frac{1}{N} \sum_{i=1}^{N} F_{t,Q}^i,
\]

(3.69)

where \( F_{t,Q}^i \) is the \( i \)-th simulated functional value, and we use a hat to indicate the sample statistical values (i.e., quantities evaluated using samples). According to the law of large numbers [e.g., Li and Zhou, 1999b], \( \hat{\mathbb{E}}[F_{t,Q}] \) approaches the real value \( \mathbb{E}[F_{t,Q}] \) asymptotically as \( N \) approaches \( +\infty \). However this is not enough. As a practical numerical method, we also want to know how much the sample mean differs from the expectation given a finite \( N \), and conversely how large an \( N \) is enough for a prescribed difference. The evaluation of confidence intervals answers these questions.

Let us consider \( m \) independent identically distributed Gaussian random variables \( X_1, X_2, \ldots, X_m \), whose expectation is \( \mu \), sample mean is \( \hat{\mu}_m \), and variance is unknown. From statistical theory, the random variable

\[
T_m = \frac{\hat{\mu}_m - \mu}{\sqrt{\frac{\hat{\sigma}^2_m}{m}}}
\]

(3.70)

satisfies the Student t-distribution with \((m - 1)\) degrees of freedom [e.g., Li and Zhou, 1999b; Kloeden and Platen, 1992], in which \( \hat{\sigma}^2_m \) is the sample variance, defined by

\[
\hat{\sigma}^2_m = \frac{1}{m - 1} \sum_{i=1}^{m} (X_i - \hat{\mu}_m)^2.
\]

(3.71)

(We note in passing that numerical evaluation of \( \hat{\sigma}^2_m \) does not use this exact formula, due to its round-off error [Press et al., 1992].) From knowledge of the t-distribution,
it follows that $\mu$ has a probability $(1-\alpha)$ of lying in the interval $(\hat{\mu}_m - \Delta\hat{\mu}, \hat{\mu}_m + \Delta\hat{\mu})$, where the probability is known as the significance level, the interval is called the confidence interval, $\Delta\hat{\mu}$ is given by

$$\Delta\hat{\mu} = t_{1-\alpha,m-1} \sqrt{\frac{\hat{\sigma}^2}{m}}, \quad (3.72)$$

and the factor $t_{1-\alpha,m-1}$, which depends on $\alpha$ and $m$, can be found in statistical tables.

In REM, the statistical distribution of $F_{t_i}^t Q$ is unknown, therefore the above theory cannot be applied directly. However, the central limit theorem [e.g., Li and Zhou, 1999b] provides that, for an arbitrary independent identically distributed random sequence $Y_1, Y_2, \ldots, Y_n$, with expectation $E[Y]$ and variance $V[Y]$, the random variable

$$Z_n = \frac{1}{n} \sum_{i=1}^{n} Y_i, \quad (3.73)$$

which is the arithmetic mean of the sequence, asymptotically obeys a Gaussian distribution with expectation $E[Y]$ and variance $V[Y]/n$, as $n$ approaches $+\infty$. Therefore in REM, we divide the simulations into $m$ batches, with each batch containing a number $n$ of simulations, so that the total number of simulations is $N = mn$. Although the statistical distribution of $F_{t_i}^t Q$ is unknown, the batch averages

$$\hat{A}_{n,j} = \frac{1}{n} \sum_{i=1}^{n} F_{t_i}^t Q_{j,i} \quad (3.74)$$

of the $m$ batches $j = 1, 2, \ldots, m$ are independent and approximately Gaussian for large enough $n$, and hence allow application of the t-test as described above. The arithmetic mean of $\hat{A}_{n,j}$

$$E[\hat{A}_{n,j}] = \frac{1}{m} \sum_{j=1}^{m} \hat{A}_{n,j} = \frac{1}{mn} \sum_{j=1}^{m} \sum_{i=1}^{n} F_{t_i}^t Q = \frac{1}{N} \sum_{i=1}^{N} F_{t_i}^t Q, \quad (3.75)$$

gives $E[F_{t_i}^t Q]$, and the confidence interval of $E[F_{t_i}^t Q]$ can readily be calculated from the asymptotic Gaussian sequence $\hat{A}_{n,j}$ according to (3.70) through (3.72).
In the REM implementation, \( n \) is a user-specified parameter, and \( m \) is increased adaptively, with the confidence interval \( \Delta \hat{E}[\hat{A}_{n,j}] \) (e.g., with the significance level \( \alpha = 0.1 \)) monitored after each \( m \) increment. When it is found that the relative error \( \Delta \hat{E}[\hat{A}_{n,j}]/\hat{E}[\hat{A}_{n,j}] \) satisfies a prescribed tolerance (e.g., 20\%), the solution process is then stopped. See Appendix B for an example of the solution statistics.

### 3.3 PSD source plots

In addition to solving the Fokker-Planck equation, the SDE method can also provide insights into the diffusion process by generating PSD source plots. In the context of electron PSD, the functional value \( F^{t,Q} \) of a stochastic path is the amount of PSD transferred along this path from the time-backward ending point \( q \) to the time-backward starting point \( Q \). Suppose that in an ensemble of \( N \) stochastic motions starting from \( Q \), \( k \) of them end in a phase space element \( \Delta q \) around \( q \). Then the PSD contribution from \( \Delta q \) to \( Q \) is given by

\[
\Delta \bar{f}(t, Q) = \lim_{N \to +\infty} \frac{1}{N} \sum_{i=1}^{k} F^{t,Q}_i.
\] (3.76)

Summing over all elements of the phase space, the sum of \( \Delta \bar{f}(t, Q) \) yields the solution \( \bar{f}(t, Q) \). Therefore, the distribution of \( \Delta \bar{f}(t, Q) \) is the distribution of the sources of the PSD at \( Q \). This is additional information provided by the SDE method in solving the Fokker-Planck equation.

In the numerical realization, the 3-D phase space is partitioned into small bins in the \((E, \alpha_0, L)\) coordinates, because the computational domain is essentially a cube in these coordinates, by definition. This implies a coordinate transform from \( Q \) to \((E, \alpha_0, L)\) for the time-backward ending positions. Each bin is a numerical approximation of \( \Delta q \), and \( \Delta \bar{f}(t, Q) \) is estimated by summing over the functional values of
the stochastic motions that stop in the bin, as in (3.76), though \(N\) and \(k\) are actually finite numbers.

To view such a 3-D distribution by eye, we need to project it into 2-D spaces. Mathematically, this projection is an integration of the \(\Delta \tilde{f}(t, Q)\) distribution along one dimension of the phase space, and there are three such projections for the 3-D phase space. In REM, these integrations are carried out along the \(E, \alpha_0\) and \(L\) dimensions, which result in three 2-D PSD source plots. The 2-D source distributions can be further integrated once again to yield 1-D source distributions. Usually, the 1-D distribution is only plotted for the \(L\) dimension to give an \(L\)-shell distribution of the PSD sources. Examples of the source plots will be shown in Chapters 4 and 5.

3.4 Summary and discussion

SDE theory, which relates a diffusion problem to stochastic processes, and thereby facilitates solutions of the diffusion equation from functional expectations of the stochastic processes, is applied to solving the radiation belt Fokker-Planck equation. Compared to finite-difference-based methods, the SDE method has several advantages when applied to radiation belt simulations. First and foremost, the SDE method does not need a numerical grid. This allows the method to deal with boundary geometries with great complexity, which facilitates the use of electron adiabatic invariant coordinates. Consequently, adiabatic variations are completely separated from non-adiabatic diffusion in the computations, and yet adiabatic losses due to the change of loss cone can still be easily modeled (as a time-varying boundary). The absence of a grid also makes the method more efficient for obtaining solutions at a limited number of phase space positions, which is often the case with spacecraft observations, since the method only samples the phase-space regions that contribute to the solutions.
Even in the opposite situation, where simulations are extended into larger regions, the method is parallelized very efficiently to achieve the necessary coverage. Second, the SDE method can solve fully multi-dimensional diffusion problems. Hence it provides a computational tool to assess diffusion effects of the electron Shabansky orbits and drift-shell splitting, which manifest themselves in the \( M-L \) and \( K-L \) components of the diffusion tensor. Last but not least, because solutions in the SDE method are obtained by summing over contributions from stochastic processes and then taking the average, the method is very robust. In particular, it can tolerate ranges of PSD of several orders of magnitude, and it never gives a negative PSD.

Based on SDE theory, we have formulated the SDE representation of the electron Fokker-Planck equation in the adiabatic invariant coordinates \((u, v, L)\), where \( u \equiv \ln(M) \) and \( v \equiv \ln(K + 1) \). We have also mathematically clarified the form of the Neumann boundary condition at \( v = 0 \) (or \( \alpha_0 = 90^\circ \)), which has often been simply implemented as \( \partial \bar{f} / \partial \alpha_0 = 0 \) (this form is only correct when off-diagonal diffusion components are zero on the boundary). A fully 3-D numerical code in adiabatic invariant coordinates (named REM) has been constructed and the key numerical techniques are presented. Validity of the model and its numerical techniques will be discussed in the next chapter.
Chapter 4

REM Validations and Application to the October 2002 Storm

This chapter provides comparisons of REM results with known Fokker-Planck equation solutions (Section 4.1), and application to a real magnetic storm event simulation, in which the model results are compared with observational data (Section 4.2). Together, they form a validation of correctness and effectiveness of the SDE method and the numerical schemes. Section 4.3 briefly summarizes this chapter. A part of the content in this chapter is published in Zheng et al. [2014].

4.1 REM validations with known solutions

REM has been tested against known solutions for 1-D radial diffusion (by setting \( u-v \) diffusion coefficients to zero) and for 2-D chorus-wave diffusion (by setting the radial diffusion coefficient to zero). In the radial diffusion test, we solve the steady state diffusion equation

\[
0 = L^2 \frac{\partial}{\partial L} \left( \frac{D_{LL}}{L^2} \frac{\partial \tilde{f}}{\partial L} \right) \tag{4.1}
\]

with \( D_{LL} \propto L^n \) for \( n = 2, 6, 10 \), a linear initial distribution of \( \tilde{f} \) from 0 at \( L = 2 \) to 1 at \( L = 9 \), and fixed-value Dirichlet boundary conditions. Figure 4.1 shows the comparison of the SDE results with analytic solutions of equation (4.1). For \( n = 2 \), equation (4.1) is degenerate, and the solution remains the same as the initial condition, as shown in Figure 4.1 by the black straight line. The SDE results are
clearly in agreement with the analytic solutions.

Figure 4.1: SDE solutions (discrete symbols) and analytic solutions (solid curves) of equation (4.1) for $n = 2, 6, 10$ (see text). The solid black line coincides with the initial condition.

In the chorus-wave diffusion test, we compare the SDE solutions with those from Albert and Young [2005], using the same initial and boundary conditions and the same diffusion coefficients. Albert and Young [2005] solved the $\alpha_0$-$p$ electron Fokker-Planck equation at $L = 4.5$, with off-diagonal diffusion components, by a finite difference method with matrix diagonalization. To compare, we reproduce the fluxes for $E = 0.5$ MeV and $E = 2.0$ MeV electrons after 0.1 day and 1.0 day, with $\alpha_0$ ranging from $6^\circ$ to $88^\circ$ with $2^\circ$ spacing (see Figure 4.2). Results from an early version of REM (Figure 4.2(a) and (b)), which does not use the adaptive methods described in Subsections 3.2.2 and 3.2.3, and results from the latest REM (Figure 4.2(c) and (d)) are shown in the figure, to illustrate the effects of the adaptive methods. Statistical
variance is visible in the 2.0 MeV and 1.0 day case in Figure 4.2(b), where the SDE solution fluctuates around its mean value. Comparing panel (d) with (b), it is clear that the statistical fluctuation is greatly reduced by the adaptive methods. Within numerical errors associated with each of the methods, they yield solutions in very good agreement.

4.2 Application of REM to the October 2002 storm

This section describes an application of an early version of REM that uses the $(u, K, L)$ coordinates. The main difference between the $(u, K, L)$ and $(u, v, L)$ coordinates are in their numerical resolutions in different parts of the computational domain, with the latter coordinates giving more balanced resolution through all $\alpha_0$’s. However, the SDE method and the physical meaning of the results are not affected by the change of coordinates.

Motivated by observations of PSD as a function of adiabatic invariants made using NASA’s Polar satellite, the Los Alamos National Laboratory geosynchronous (LANL GEO) satellites and the Global Positioning System (GPS) constellation [Chen et al., 2007b; Koller et al., 2007], we apply REM to a simplified simulation of the PSD increase at GPS orbit during a moderate high-speed stream (HSS) storm with minimum $Dst \sim -70$ nT from 15 October (Day 288) to 20 October (Day 293) 2002. Chen et al. [2007b] gave a thorough description of the observational data during this storm. In that work, for $M = 2083$ MeV/G and $K = 0.03$ G$^{1/2}$R$_E$, they concluded that during the recovery phase of this storm chorus wave acceleration played an essential role in replenishing the outer belt and generating the PSD peak observable at GEO.

Figure 4.3 summarizes the initial condition and the outer-$L$ boundary condition
Figure 4.2: Comparisons between solutions from the SDE method (solid lines) and the Albert and Young [2005] method (dashed lines) for 0.5 MeV (a and c) and 2.0 MeV (b and d) electron fluxes after $t = 0.1$ day (blue) and $t = 1.0$ day (red) diffusion. Solid black lines show the initial condition. Panels (a) and (b) are calculated using an early version of REM with fixed step-size Euler-Maruyama scheme and fixed total number of stochastic simulations. Panels (c) and (d) are from the latest REM, within which the 90% confidence intervals are plotted for each solution point.
in our simulation. To set up an initial 3-D PSD distribution, we fit the PSD data at a list of $M$’s (167, 462, 1051, 2083 MeV/G) and $K$’s (0.005, 0.01, 0.03, 0.1, 0.3 G$^{1/2}$/RE) in three $L$ groups, corresponding to GPS, GEO and Polar orbits respectively. The top panel of the figure illustrates the fitted initial PSD radial profile at the four pairs of $M$ and $K$ that we simulate. On the outer-$L$ boundary at geosynchronous distance ($L = 6$), we simplify the observed PSD temporal variations as exponential increases from the beginning to the end of the storm, as shown in the bottom panel of the figure. On the inner-$L$ boundary ($L = 2$), PSD is kept the same as its initial value over time. To make the boundary conditions consistent, the exponential increase is also carried onto the $E_{min}$ boundary in such a way that the increase rate diminishes as $L$ decreases from 6 to 2. Thus, on the $E_{min}$ boundary, at $L = 6$, the increase is at the full rate; whereas at $L = 2$, the PSD is unchanged.

During the storm, the $Kp$ index varied only a small amount about an average value of about 3; for simplicity, we fix $Kp = 3$ in assigning the diffusion coefficients. $D_L$ uses the Brautigam and Albert [2000] magnetic radial diffusion coefficient (henceforth denoted $D_M^{LL}[B&A]$), because in the outer belt it dominates over the electric field counterpart ($D_E^{LL}[B&A]$). Chorus wave diffusion coefficients in the $(u, K, L)$ coordinates are transformed from the British Antarctic Survey (BAS) drift and bounce averaged diffusion matrix, by the method described in Subsection 3.1.1 and assuming dipole field geometry. The BAS diffusion matrix was calculated using the PADIE code [Glauert and Horne, 2005] and plasma wave observations from the CRRES spacecraft [Meredith et al., 2001, 2003] as described in Varotsou et al. [2005]. Varotsou et al. [2005] considered the effects of equatorial chorus, defined as chorus wave power within $\pm 15^\circ$ of the magnetic equator. In this paper we have included wave power up to $\pm 30^\circ$ to include the effects of strong mid-latitude chorus that is observed on the dayside.
Figure 4.3: Initial condition (top) and the outer-$L$ boundary condition (bottom) used in the October 2002 HSS storm simulation. In the top panel, markers represent satellite data at the beginning of 15 October (Day 288) from GPS ($L \approx 4$), GEO ($L \approx 6$) and Polar ($L \approx 8$). Marker colors identify the corresponding $M$ and $K$ values to the fitted curves. The dashed line at $L = 6$ denotes the location of the outer-$L$ boundary. In the bottom panel, GEO observations are shown in colored dots. Simplified exponential increases are drawn as straight lines.
Further, we have extrapolated for energies between 3 and 4 MeV. Plots of the transformed chorus wave diffusion rates are shown in Figure 4.4. Inward of geosynchronous orbit, the diffusion rates grow monotonically with $L$. In contrast to the $(\alpha_0, p)$ coordinates, where the off-diagonal term often changes sign, the transformed $D_{uK}$ values are all negative throughout the entire $L$ range of the BAS diffusion matrix for this $Kp$ level.

Model simulations are performed to obtain PSD at four phase-space positions at $L = 4$, with $(M, K)$ values of $(1051, 0.005)$, $(2083, 0.005)$, $(1051, 0.3)$ and $(2083, 0.1)$. In a dipole field, these phase-space positions correspond respectively to energies (in MeV) and $\alpha_0$’s of $(1.85, 81.1^\circ)$, $(2.78, 81.1^\circ)$, $(3.26, 37.6^\circ)$ and $(3.47, 54.4^\circ)$. At each position, three independent simulations are carried out: one with radial diffusion only, one with chorus wave diffusion only, and one with the two mechanisms combined. The results are given in Figure 4.5.

In the top row (a and b) of Figure 4.5, where $\alpha_0$’s are near $90^\circ$, chorus waves have very little effect on the PSD variation, as seen from the almost flat “chorus only” curves. This could also be appreciated from the close alignment between the “radial only” and the “combined” curves. As argued in Horne et al. [2005] and Shprits [2009], this is because energy diffusion by chorus waves is less effective at large pitch angles (i.e. $\sim 80^\circ$). In panels (a) and (b) radial diffusion curves largely follow the observed data points. Thus, the equatorial PSD increases in this HSS storm are consistent with mainly radial diffusion acting on the particles.

In the bottom row (c and d), where $\alpha_0$’s are in the range of peak chorus wave effectiveness ($30^\circ \sim 60^\circ$, according to Horne et al. [2005]), chorus wave effects are more significant. Nevertheless, chorus wave diffusion alone is insufficient to explain the observed PSD increases. Radial diffusion alone, which reproduces the mild increase
Figure 4.4: Normalized $u$-$K$ diffusion rates (day$^{-1}$) transformed from the BAS drift and bounce averaged chorus wave diffusion matrix at three $L$ values. $D_{uu}$ is dimensionless since $u$ is dimensionless. $D_{uK}$ and $D_{KK}$ are normalized against $K = 1$ G$^{1/2}$R$_E$. The uniform blue color in each sign($D_{uK}$) panel indicates negative regions of $D_{uK}$. Positive regions would have been red.
Figure 4.5: Comparison of the simulated PSD temporal changes with GPS observations at $L = 4$. In each of the panels, observed data are marked with black crosses. Black straight lines indicate the driving boundary conditions at $L = 6$ at each specified $M$ and $K$. Note that the ordinate scales are different in these panels.
of PSD (less than one order of magnitude) in the near equatorial cases, overestimates the observed increases here. As a result, the PSD variations of the combined diffusion are even more exaggerated. Apart from the simplified inputs to this simulation, we suggest two possible reasons for this overestimation. First, $D_{LL}[B&A]$ is derived for equatorially bouncing electrons, whereas radial diffusion rates caused by ULF waves are $\alpha_0$-dependent. From the estimate given in Schulz and Lanzerotti [1974], it follows that the radial diffusion rate in the $\alpha_0$ range of $35^\circ \sim 55^\circ$ is only about $0.2 \sim 0.4$ times that at $90^\circ$. Therefore, using $D_{LL}[B&A]$ will result in exaggerated radial diffusion at small $\alpha_0$’s. Second, $L = 4$ is likely on the edge of the plasmasphere for this moderate storm, and electron loss from electromagnetic ion cyclotron (EMIC) waves and plasmaspheric hiss waves are not included in our simulation. At this $L$, effects of magnetopause shadowing or drift-shell splitting are expected to be negligible.

In addition to solving the diffusion equation, the SDE method also provides insight into the diffusion process that is not easily obtained in other methods. Take for example the solutions at $M = 1051$ MeV/G, $K = 0.3 \ G^{1/2}R_E$ and $t = 5$ day (Day 293 in Figure 4.5c), where solutions of “radial only” and “combined” diffusions are well separated. By recording the diffusion stopping times of the time-backward stochastic processes (as in the example illustrated in Appendix B), we find that the electrons take about 1.2 days to diffuse from GEO to GPS orbit in our simulation. On the other hand, by binning the stopping position distribution of the time-backward stochastic processes, we are able to determine the PSD contribution from different phase-space regions to the solution; in other words, to determine a phase-space distribution of the energetic electron sources. In Figure 4.6, we illustrate these “source distributions” by plotting radial profiles of the source PSD contribution in the “radial only” and the “combined” simulations. The radial profile in the “chorus only” simulation is trivial,
Figure 4.6: Radial distributions of the PSD contribution for the solution at $M = 1051 \text{ MeV/G}$, $K = 0.3 \text{ G}^{1/2}R_E$, and $L = 4$ at $t = 5$ day, from the “radial only” (square) and the “combined” (circle) simulations. The contributions are binned in $L$ with bin size 0.4. The first and the last bins are cumulative; thus, the first marker represents all contributions from $L < 2.0$, and the last represents all contributions from $L \geq 6.0$. The interior markers reside at the center of each bin, and give the value in that bin.

since it is simply a spike at $L = 4$.

As illustrated in the figure, at the end of the storm, at GPS orbit, the majority of the PSD comes from the outer-$L$ boundary in both simulations. Radial diffusion thus plays an indispensable role in energizing all these electrons. However, the outer-$L$ boundary contributes about twice the amount of PSD in the “combined” simulation as in the “radial only” simulation, demonstrating that with local acceleration turned on, the outer-$L$ boundary becomes more efficient in providing source electrons. In a synergistic manner, more low-energy electrons are accelerated by chorus waves while diffusing inward; these particles could not have been energized as much with radial
diffusion only. A bump is clearly observed in the “combined” curve between \( L = 3.8 \) and 5.4, whereas the “radial only” curve remains almost flat in this range. This difference indicates the regions where chorus wave acceleration is most important for the electron population at \( L = 4 \). We note that, from Figure 4.4, although chorus wave diffusion is stronger in larger \( L \) regions, its effect on the \( L = 4 \) solution is nevertheless radially localized. This is due to both the length of diffusion time and the shape of the initial PSD distribution.

### 4.3 Summary and discussion

REM has been successfully tested against known solutions of 1-D radial diffusion and 2-D chorus wave diffusion with off-diagonal components. REM has also been applied to simulating the PSD increase observed at GPS locations during the October 2002 high-speed stream storm. It has been argued in Chen et al. [2007b] that chorus wave acceleration was active during this storm, and was responsible for the PSD peak observed on geosynchronous orbit. For near-equatorial electrons (Figure 4.5(a) and (b)), the model gives reasonably accurate simulations of the observed GPS increase, consistent with radial diffusion from the outer-\( L \) boundary and weak chorus wave diffusion at large equatorial pitch-angles. At smaller equatorial pitch-angles (Figure 4.5(c) and (d)) radial diffusion, and consequently the combined diffusion, overestimates the increase of PSD at GPS locations. Possible explanations could be the neglect of equatorial pitch-angle dependence in the radial diffusion coefficients, or the neglect of electron loss from pitch-angle scattering by hiss and EMIC waves in the model.

The SDE method allows us to extract information such as the electron diffusion time and PSD source distribution, by analyzing the statistics of the stochastic paths.
Our model reveals the intrinsic differences between radial and combined diffusion from a new point of view, by plotting the radial distribution of the source PSD. In the case exemplified in this work, although both diffusion scenarios have the majority of the PSD coming from the outer boundary, with chorus waves present the efficiency of the outer boundary in providing seed electron populations is increased by a factor of two. Chorus waves also generate a peak to the PSD source distribution well within the outer radiation belt, indicating the $L$ ranges in which they are of particular interest.
Chapter 5

Effects of Drift Shell Splitting on Radiation Belt Electron Diffusion

We begin with a brief preview of this chapter. Drift shell splitting in the presence of pitch-angle scattering breaks all three adiabatic invariants of radiation belt electron motion, and produces new diffusion terms that fully populate the diffusion tensor in the Fokker-Planck equation. The REM solves such a Fokker-Planck equation, and is used to investigate the phase space density sources. Simulation results and theoretical arguments suggest that drift shell splitting changes the phase space location of the source to smaller $L$ shells, which typically reduces outer zone phase space density enhancements, and this reduction has a limit corresponding to two-dimensional local diffusion on a curved surface in the phase space.

In Section 5.1, we give a brief introduction to the drift shell splitting problem. In Section 5.2, we show the REM simulations results with drift shell splitting effects in an idealized model storm, which allows more control on the simulation conditions than a realistic data-driven simulation. In Section 5.3, we present theoretical arguments to explain the simulation results. Discussion and conclusions are given in Section 5.4.

5.1 Introduction

Radiation belt dynamics can be simulated by solving a Fokker-Planck equation in a set of phase space coordinates with a given diffusion tensor. In an axially symmetric magnetic field, such as a dipole field, particle drift shells do not depend on equato-
rial pitch-angle \(\alpha_0\). Thus, \(\alpha_0\) is a preserved quantity along a drift trajectory, and the variables \(\alpha_0\), the magnitude of the mechanical momentum \(p\), and Roederer’s \(L\) \cite{Roederer:1970} are useful phase space coordinates. When axial symmetry is broken, as in the geomagnetic field, \(\alpha_0\) is no longer constant for the entire drift shell. This necessitates the use of adiabatic invariants, e.g., \(J_1\), \(J_2\) and \(J_3\), or equivalently and more conventionally \(M\), \(K\) and \(L\), as coordinates of the phase space. Moreover, \(L\) becomes dependent on \(\alpha_0\), i.e., drift shells would split for an ensemble of originally colocated particles with different \(\alpha_0\), which gives rise to the terminology “drift shell splitting” (cf. Section 1.2).

The dependence of \(L\) on \(\alpha_0\) introduces new components to the diffusion tensor. This occurs as one transforms the 2-D bounce-averaged diffusion tensor from the \((\alpha_0, p, L)\) coordinates to the \((M, K, L)\) coordinates on a guiding field line (cf. equation (1.33)):

\[
\begin{pmatrix}
D_{MM} & D_{MK} & D_{ML}^* \\
D_{MK} & D_{KK} & D_{KL}^* \\
D_{ML}^* & D_{KL}^* & D_{LL}^*
\end{pmatrix}
= \begin{pmatrix}
\frac{\partial M}{\partial \alpha_0} & \frac{\partial M}{\partial p} & 0 \\
\frac{\partial K}{\partial \alpha_0} & 0 & 0 \\
\frac{\partial L}{\partial \alpha_0} & 0 & 1
\end{pmatrix}
\begin{pmatrix}
D_{\alpha_0 \alpha_0} & D_{\alpha_0 p} & 0 \\
D_{\alpha_0 p} & D_{pp} & 0 \\
0 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
\frac{\partial M}{\partial \alpha_0} & \frac{\partial K}{\partial \alpha_0} & \frac{\partial L}{\partial \alpha_0} \\
\frac{\partial M}{\partial p} & 0 & 0 \\
0 & 0 & 1
\end{pmatrix}.
\]

(5.1)

Here the starred components are generated by drift shell splitting. With vanishing \(\partial L/\partial \alpha_0\) in the Jacobian matrix, \(D_{ML}^*\), \(D_{KL}^*\) and \(D_{LL}^*\) also vanish, leaving the left-hand side tensor still a \(2 \times 2\) block. Based on this transformation, \(O’Brien\) [2014, 2015] calculated the bounce-and-drift-averaged chorus wave diffusion tensor using a statistical chorus wave model (as in \(Shprits\ et\ al.\ [2011]\)), the \(Sheeley\ et\ al.\ [2001]\) plasma density model, and realistic geomagnetic field models. These geomagnetic field models include the International Geomagnetic Reference Field \(\text{[Finlay et al., 2010]}\), which provides effects of the Earth’s internal magnetic multiples (though negligible in the
outer belt), and T89 model [Tyaganenko, 1989] that provides the external magnetic field asymmetry. In these calculations, field-aligned chorus wave propagation was assumed. The values of these starred components for 0.1 MeV electrons in the midnight meridian are plotted with the un-starred components (combined into the same dimensions) in Figure 5.1. From comparisons in the figure, the magnitudes of the starred diffusion components are most significant in the midnight $\alpha_0$ range 40° to 80°. As energy increases from $\sim$ 100 keV into the MeV range, the relative magnitude of the starred components decreases compared to the un-starred counterparts [O’Brien, 2015]. This chapter investigates the effect of drift shell splitting on radiation belt electron diffusion using the diffusion tensor calculated by O’Brien [2015].

5.2 Simulations of an idealized MeV electron enhancement

5.2.1 Simulation setup

In this study, the simulation domain is defined by electron kinetic energy $E$ from 100 keV to 10 MeV, $K$ from the bounce loss cone to zero or the drift loss cone, and $L$ from 2 to 8. The mapping between $E$ and adiabatic invariants is provided via particle tracing in the geomagnetic field model.

To examine drift shell splitting effects on electron PSD, we have made simulations of a model storm motivated by the 8 October 2012 electron acceleration event [Tu et al., 2014]. We emphasize that our purpose is not to precisely reproduce that event; rather, it is to evaluate the magnitude of the drift shell splitting terms on the electron PSD using a representative idealized simulation. Initial and boundary conditions are obtained from time-dependent fits to Van Allen Probes electron flux data at $L = 4$
Figure 5.1: Comparisons of the drift-shell-splitting diffusion components with the non-drift-shell-splitting counterparts, namely, (a) $D_{ML}^* \text{ versus } (D_{MM}D_{ULF}^{ULF})^{1/2}$ ("$D_{ML}^*$" is indicated by ‘$D_{(S)}^{(ML)}$’ in the figure and note the typo "$D_{MK}$"), (b) $D_{KL}^* \text{ versus } (D_{KK}D_{ULF}^{ULF})^{1/2}$, and (c) $D_{LL}^* \text{ versus } D_{ULF}^{ULF}$, where $D_{ULF}^{ULF}$ are the ULF diffusion coefficients obtained from Ozeke et al. [2014]. All diffusion coefficients are evaluated for the 0.1 MeV electrons and $Kp = 4$ on guiding field lines in the midnight meridian, with distances of their equatorial crossings indicated by $R$. These figures are reprinted from O’Brien [2015].
during the 8 October 2012 storm. The fits are made with an approximate analytic solution to the Fokker-Planck equation, assuming the lowest eigenmode pitch-angle diffusion and steady-state radial diffusion (see Appendix C for details). On 8 October, the observed $\sim 100$ keV (“seed”) electron flux increased exponentially for about 7 hours and then stayed steady; this provides the low energy boundary condition for our simulation. We set the onset of this seed electron injection as the initial time of this study.

In contrast to a simulation in a dipole field, in a geomagnetic field a new phase space boundary arises due to electron Shabansky orbits [Shabansky, 1971; Öztürk and Wolf, 2007], on which the electron drift trajectories bifurcate into either of the two magnetic field minima before local noon, and join together afterwards. The Roederer $L$ for electrons on these orbits is not defined [Ukhorskiy et al., 2011], and hence they are excluded from the 3-D adiabatic invariant phase space. Further, these electrons are quasi-trapped; they drift from a few to a few hundred periods depending on the orbit before escaping into the bounce or drift loss cone [Ukhorskiy et al., 2011]. In REM, for simplicity we assign a 50 drift-period characteristic e-folding lifetime for the electrons on the boundary between the stably trapped and the quasi-trapped regions. Locations of this boundary (in $K$ and $L$) are determined from particle tracing in the geomagnetic field model.

This study adopts Ozeke et al. [2014] ULF wave diffusion coefficients ($D_{LL}^{ULF}$) to invoke radial diffusion at constant $M$ and $K$, and the aforementioned O’Brien [2015] drift-shell-splitting chorus wave diffusion coefficients (drift average of $D_{ij}$ in the left-hand side of equation (5.1)), so that the Fokker-Planck equation solved is

$$\frac{\partial \bar{f}}{\partial t} = \frac{1}{G} \frac{\partial}{\partial Q_i} \left( G(D_{ij}) \frac{\partial \bar{f}}{\partial Q_j} \right) + L^2 \frac{\partial}{\partial L} \left( \frac{1}{L^2} D_{LL}^{ULF} \frac{\partial \bar{f}}{\partial L} \right),$$

(5.2)

where $\bar{f}$ is the phase-averaged PSD, $G$ is the Jacobian determinant for the coordinate
transform from \((J_1, J_2, J_3)\) to \((M, K, L)\) as represented collectively by the \(Q_i\)’s, and summation over repeated indices is implied. The Ozeke et al. [2014] radial diffusion coefficients are derived from statistical representations of compressional magnetic field wave power and azimuthal electric field wave power in the ULF band, and are analytically approximated as functions of \(L\) and \(Kp\). As a simple but reasonable storm-time value, we assume constant \(Kp = 4\), which also results in significant asymmetry in the T89 magnetic field [O’Brien, 2015]. To resemble the 8 October 2012 storm, we multiply statistically obtained chorus wave amplitudes by an event-specific factor inferred from POES spacecraft data [e.g., Li et al., 2013; Chen et al., 2014]. In our simulation, this modeled POES factor has a radial profile that peaks at \(L = 4\) with a height of 10 (see Figure 5.2).

5.2.2 Simulation results

Figure 5.3 shows the radial profile of PSD solutions at \(M = 2000\) MeV/G and \(K = 0.01\) G\(^{1/2}\)R\(_E\) at 9 and 17 hours after the initial time, with comparisons in which drift shell splitting is turned on or off by retaining or removing the starred components in equation (5.1). In both cases, strong local acceleration produces PSD peaks in the heart of the radiation belt. Inside \(L = 4\), the “on” and “off” solutions are very close because of small asymmetry in the inner geomagnetic field. Beyond \(L = 4\), the “on” solutions are all lower than the “off” solutions, with the largest difference appearing around the PSD peaks.

To better investigate the diffusion processes, we use REM to generate PSD source plots for a given phase space solution point (cf. Section 3.3). These plots show the PSD contribution from various phase space locations to that solution point. Figure 5.4 shows source plots for the “off” (a–c) and “on” (d–f) solutions at the PSD peaks.
Figure 5.2: Real-time magnetic local time averaged POES factors for every half hour from 23:00UT, 08 to 16:00UT, 09 October 2012 (blue dashed lines), their time average during this period (red solid line), and the model POES factor with a peak height of 10 at $L = 4$ (black solid line). Model POES factors of a different peak height (namely 4, 7, and 13) have the same peak location and width as the one shown here.
Figure 5.3: REM simulated PSD radial profiles with drift shell splitting off (blue) and on (red) for $M = 2000 \text{ MeV/G}$ and $K = 0.010 \text{ G}^{1/2}R_E$ at $t = 9$ hour (dashed) and $t = 17$ hour (solid). The vertical bar at each data point indicates the estimated 90% confidence interval of the solution.


\((L = 4.4, \, t = 17\ \text{hour})\) in Figure 5.3. From left to right in each row, the three panels are projections of the 3-D PSD source distribution along \(E\), \(\alpha_0\), and \(L\) dimensions respectively. Each pixel in these panels represents a 2-D phase space element, whose color represents the amount of PSD it contributes to the solution point (indicated by the cross). In these plots and Sections 5.2 and 5.3, the definition of \(\alpha_0\) on a drift shell is generalized using the mirror point magnetic strength \(B_m(K)\) compared to that of the equatorially-bouncing drift shell with the same \(L\) according to the formula

\[
\sin^2 \alpha_0(K) \bigg|_L = \left. \frac{B_m(0)}{B_m(K)} \right|_L.
\]  

This generalized \(\alpha_0\) is hence drift invariant and reduces to equatorial pitch-angle if the field is reduced adiabatically to a symmetric field.

A striking feature of Figure 5.4 is that, in both “off” and “on” solutions, the majority of the PSD comes from a narrow range of \(L\) around 4.4 near the low energy boundary, suggesting that local acceleration together with the seed population play a decisive role in electron energization, and the diffusion is approximately 2-D \(\alpha_0-E\) diffusion. The seed electron source region in the “on” case (d and e) is broader in \(L\) than in the “off” case (a and b). The low energy sources in the “off” case (b) occur roughly at \(L = 4.4\), whereas in the “on” case (e) they occur at a lower \(L\) around 4.2. The \(\alpha_0\) distribution of the sources in both cases (a and d) has two blobs, one between 25° and 70° corresponding to energization from the nightside chorus, and the other below 20° from the dayside chorus.

The prominence of the drift shell splitting effects depends on the relative importance of chorus wave diffusion. The latter is a function of the peak height of the POES factors. In the large POES factor limit chorus wave scattering dominates, so that the diffusion is asymptotically 2-D; and in the opposite limit the diffusion approaches 1-D radial diffusion (see Figure 5.5), and the drift shell splitting effect vanishes. Figure
Figure 5.4: PSD source plots for the $t = 17$ hour solutions at $M = 2000$ MeV/G, $K = 0.01 \, G^{1/2}R_E$ and $L = 4.4$ (white cross), with drift shell splitting off (a–c) and on (d–f). White dashed lines are projections of the 3-D constant-$M$-and-$K$ curve, or “radial diffusion curve” through the cross, calculated with dipole field geometry. The lines of dark blue pixels within $L = 6 \sim 8$ and $\alpha_0 = 40^\circ \sim 90^\circ$ in (a) and (d) indicate locations of the Shabansky orbit boundary.
Figure 5.5: PSD source plots, in the same format as Figure 5.4, for the $t = 17$ hour solutions at $M = 2000 \text{ MeV/G}$, $K = 0.01 \text{ G}^{1/2} R_E$ and $L = 5.4$, calculated with uniform POES factor 1. The Shabansky orbit boundary is seen as the line of green pixels in (a) and (d). Panels (b) and (e) show that the energization mechanism here is primarily radial diffusion, with the majority of PSD contributed by the Shabansky orbit boundary.
5.6 plots the difference of PSD solutions (“off” minus “on”) at $L = 4.4$ and $t = 17$ hour versus a range of peak POES factors. Regardless of the POES factor value, all the PSD differences are positive. As the POES factor increases, the PSD difference increases monotonically but approaches a 2-D diffusion limit as indicated by the dashed line, which is the difference calculated with peak POES factor 13 but without ULF wave diffusion. (The “ro” points are referred to in Section 5.4.) The ratio of PSD solutions (“off” over “on”) is shown in Figure 5.7. Comparing Figure 5.6 with Figure 5.7 shows that the difference of PSD has a much simpler dependence on POES factor than the ratio of PSD. In the next section we use analytic theory to investigate the changes in PSD due to drift shell splitting and to interpret the simulation results.

5.3 Theoretical analysis

Symbolically, equation (5.1) may be written as $D = GDG^\top$, where different fonts of the diffusion tensors denote their different coordinates. By construction, the diffusion tensor $D$ has two positive (or zero) eigenvalues, whose corresponding eigenvectors are the semi-axes of the elliptic contour of the diffusion Green function [cf. Albert, 2009] in the $\alpha_0$-$p$ plane, and a third trivially zero eigenvalue with the eigenvector in the $L$ direction. Arranging eigenvalues in descending order, this corresponds to a diagonalized factorization of $D$ such that

$$D = \begin{pmatrix} \hat{u} & \hat{v} & \hat{l} \end{pmatrix} \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & 0 \end{pmatrix} \begin{pmatrix} \hat{u}^\top \\ \hat{v}^\top \\ \hat{l}^\top \end{pmatrix}, \quad (\lambda_1 \geq \lambda_2 \geq 0), \quad (5.4)$$

in which $\hat{u}$ and $\hat{v}$ are orthogonal unit eigenvectors corresponding to eigenvalues $\lambda_1$ and $\lambda_2$, and $\hat{l}$ is the unit vector along the $L$ axis. In mathematical terms, the continuous one-to-one mapping of a coordinate transform makes the two spaces homeomorphic
Figure 5.6: Difference in PSD solutions at $M = 2000$ MeV/G, $K = 0.01$ G$^{1/2}$RE, $L = 4.4$ and $t = 17$ hour versus the peak height of the POES factor. In the legend, “off” or “on” indicates drift shell splitting is turned off or on; “ro” means only $D_{LL}^{*}$ is included and not $D_{ML}^{*}$ or $D_{KL}^{*}$; and “2D” means the simulation is done with a peak POES factor 13 and no ULF wave diffusion. At “peak” POES factor 1, the POES factor is actually uniformly 1 with no peak.
Figure 5.7: Ratio between PSD solutions at $M = 2000$ MeV/G, $K = 0.01$ G$^{1/2}$RE, $L = 4.4$ and $t = 17$ hour versus the peak height of the POES factor. In the legend, “off” or “on” indicates drift shell splitting is turned off or on, and “2D” means the simulation is made with a peak POES factor 13 and no ULF wave diffusion. At POES factor 1, the PSD ratio is very close to 1, indicating that the drift shell splitting effect is very weak. As the peak POES factor increases, both local acceleration and the drift shell splitting effect increase, and their competition results in a non-monotonic relation between the PSD ratio and the strength of chorus wave diffusion. At a moderate peak POES factor 4, the ratio reaches a local maximum where “on” is nearly one order of magnitude smaller than “off”.
[e.g., Zorich, 2004], meaning that the topology of a geometrical object in one space is unchanged after the coordinate transform into another. In this sense, though the \((M, K, L)\) space does not have a metric, we are still able to discuss a geometrical object in this space, because it is topologically identical to that in the \((J_1, J_2, J_3)\) space, and the latter is a Euclidean space of the generalized momenta. For the topological discussion in this section, we understand the coordinate variables as dimensionless as being scaled by constant values bearing their units.

According to Sylvester’s law of inertia [e.g., Norman, 1986], the congruence transform in (5.1) conserves the numbers of positive and zero eigenvalues of the diffusion tensors on both sides. Therefore, although fully populated, the left-hand side tensor \(D\) is nonetheless singular due to the zero eigenvalue. At a given point in the \((M, K, L)\) space, the bounce-averaged diffusion is hence in a local 2-D plane containing the first two eigenvectors of \(D\). In general, it is difficult to analytically calculate these eigenvectors to obtain the orientation of the plane. However, since the plane is just the local \(\alpha_0-p\) plane after a coordinate transform, the transformed \(\hat{u}\) and \(\hat{v}\) vectors, denoted as \(\xi\) and \(\eta\) respectively, must also lie in this plane, where \(\xi\) and \(\eta\) are readily given by the matrix equation

\[
\begin{pmatrix}
\xi \\
\eta
\end{pmatrix} = G
\begin{pmatrix}
\hat{u} \\
\hat{v}
\end{pmatrix}.
\]

(5.5)

Thus, the normal of the local 2-D diffusion plane points in the direction of

\[
n = \xi \times \eta = \frac{\partial M}{\partial p} \left( \frac{\partial L}{\partial \alpha_0} \hat{k} - \frac{\partial K}{\partial \alpha_0} \hat{l} \right),
\]

(5.6)

where \(\hat{k}\) is the unit vector along the \(K\) axis. This is also the direction of the third eigenvector of \(D\), corresponding to its zero eigenvalue.

In equation (5.6), \(\partial M/\partial p = 2M/p\) is positive, and \(\partial L/\partial \alpha_0\) and \(\partial K/\partial \alpha_0\) are field geometric quantities, with \(\partial K/\partial \alpha_0\) negative except at \(\alpha_0 = \pi/2\) where it van-
ishes. Thus the direction of \( n \) is solely determined by the geometry of the guiding field line. For a symmetric field, because \( \partial L/\partial \alpha_0 \) is identically zero, \( n \) is uniformly in the \( \hat{l} \) direction and the diffusion is confined to an \( M-K \) plane. For the asymmetric geomagnetic field, \( n \) tilts in the \( \hat{k} \) direction by an amount that depends on the sign and relative magnitude of \( \partial L/\partial \alpha_0 \) compared to \( \partial K/\partial \alpha_0 \). In particular, \( \partial L/\partial \alpha_0 \) is typically positive for guiding field lines in the nightside and negative for those in the dayside (cf. Figure 1 in O'Brien [2015]), and \( \partial K/\partial \alpha_0 \) decreases without bound with decreasing \( \alpha_0 \) (in dipole field, \( \partial K/\partial \alpha_0 \sim -\cos \alpha_0/\sin^2 \alpha_0 \)). At small \( K \), the \( \hat{k} \) component of \( n \) dominates because of vanishing \( \partial K/\partial \alpha_0 \) as \( K \) approaches 0 (cf. equation (3.38)); but as \( K \) increases, since \( \partial L/\partial \alpha_0 \) cannot be boundless, the \( \hat{l} \) component quickly becomes significant, and eventually dominates. For all \( M \) and \( K \) values on the guiding field line, all of these local planes thus form a curved surface in phase space, with its normal direction at every point given by \( n \). Figure 5.8 schematically illustrates such a surface for a guiding field line in the nightside. On this guiding field line, the 2-D bounce-averaged diffusion takes place in this surface (\( \Sigma^* \)) rather than in an \( M-K \) plane (\( \Sigma \)) as in a symmetric field. For a representative point stochastically scattering in \( \Sigma^* \), its \( L \) value thus changes due to the geometry of the surface.

Different guiding field lines have different \( \Sigma^* \) diffusion surfaces, so a fixed phase space point (such as point \( P \) in Figure 5.8) resides on a family of \( \Sigma^* \) surfaces corresponding to all the guiding field lines on the drift shell. At this point, though bounce-averaged diffusion on each guiding field line is confined to its individual surface, drift-averaged diffusion, as a superposition of the former, effectively takes place in a phase space volume. This means that the diffusion tensor loses its singularity after drift averaging [O’Brien, 2015]. In this sense, strict 2-D drift-averaged diffusion does not occur in an asymmetric geomagnetic field even without drift-resonant radial
Figure 5.8: Schematic illustration of a bounce-averaged diffusion surface for a nightside guiding field line in a symmetric magnetic field (the $\Sigma$ plane), and in an asymmetric geomagnetic field (the $\Sigma^*$ surface), with a common $L$ value at the phase space point $P$. The intersection of $\Sigma$ and $\Sigma^*$ through $P$ is a straight line parallel with the $M$ axis. Shaded areas indicate the phase space regions of the seed electrons (see text). The $\Sigma^*$ surface for a dayside guiding field line would be tilted upward in $L$ with increasing $K$, due to negative $\partial L/\partial \alpha_0$. 
diffusion.

Because wave activity is not evenly distributed in magnetic local time, the bounce-averaged diffusion in each $\Sigma^*$ surface contributes differently to the drift average. Outside the plasmapause, the diffusion rate of nightside chorus typically exceeds that of dayside chorus by orders of magnitude except at small $\alpha_0$’s [Li et al., 2007]. Consequently, except at large $K$, the drift-averaged diffusion would primarily take place in the $\Sigma^*$ surfaces of nightside guiding field lines. At large $K$, all the bounce-averaged diffusion surfaces are close to an $M$-$K$ plane anyway due to the dominant $\hat{i}$ component in the normal, and so is their drift average. As a result, the drift-averaged chorus wave diffusion in an asymmetric geomagnetic field is still approximately 2-D, around a surface in the phase space similar to that of a nightside guiding field line.

This approximate 2-D drift-averaged diffusion is indeed observable in the PSD source plots. Figure 5.9 shows source plots for solutions at $M = 2000 \text{ MeV}/G$, $K = 0.01 G^{1/2} R_E$, $L = 5.4$ and $t = 17$ hour from the simulations of Figure 5.6 that are labeled “2D”. In the “off” case (Figure 5.9 a–c), PSD sources align exactly in the constant $L = 5.4$ plane; whereas in the “on” case (d–f), PSD sources are distributed near a surface that tilts toward smaller $L$ in smaller $\alpha_0$ and smaller $E$. This is consistent with the theoretically predicted 2-D diffusion near the $\Sigma^*$ surface in Figure 5.8, where PSD sources contributing to the point $P$ from smaller $M$ and larger $K$ regions have smaller $L$ values than that of $P$, due to the geometry of the surface. The same phenomenon is also observed in the source plots for the solution with the same $M$, $K$ and $t$ but at $L = 4.4$, as shown in Figure 5.10. In this case, the PSD sources are mainly from the low-energy boundary instead of the interior phase space. PSD contributions from the portion of the $\Sigma^*$ surface that is at smaller $K$ and larger $L$ to the solution point are visible in panels (d) and (e).
Figure 5.9: PSD source plots, in the same format as Figure 5.4, for the $t = 17$ hour solutions at $M = 2000$ MeV/G, $K = 0.01 \ G^{1/2}/R_E$ and $L = 5.4$, calculated with the 2-D diffusion setup (see text).
Figure 5.10: PSD source plots, in the same format as Figure 5.4, for the $t = 17$ hour solutions at $M = 2000$ MeV/G, $K = 0.01$ $G^{1/2}R_E$ and $L = 4.4$, calculated with the 2-D setup. In the “off” case (a–c), PSD sources are distributed in the $L = 4.4$ plane, mainly near the low energy boundary. In the “on” case (d–f), PSD sources are distributed near a surface that, at smaller $\alpha_0$’s (larger $K$) has $L$ value smaller than that of the solution point (cross), and at larger $\alpha_0$’s (smaller $K$) has $L$ value larger than that of the cross, as predicted in the theoretical explanation. Integration of the source distribution in (e) along the $E$ dimension shows that it is centered at $L = 4.1$. 
On the basis of this discussion, we can better understand the results in Section 5.2. In Figure 5.4, the solution point is at smaller $K$ than the range of strong chorus wave scattering, which occurs for generalized $\alpha_0$ between about $10^\circ$ to $70^\circ$. By the geometry of the surfaces in Figure 5.8, the seed electrons for point $P$ are located at a smaller $L$ in the $\Sigma^*$ surface than those in the $\Sigma$ plane (as indicated by the shaded areas); this is why the seed electrons in the “on” case are from smaller $L$ than those in the “off” case. Furthermore, drift shell splitting makes the drift average of $D$ nonsingular, so that the seed electron regions in the “on” case are spread more in $L$. The PSD of the seed electrons typically has a positive gradient in $L$ at fixed $M$ and $K$, so it follows that local acceleration of the seed electrons from smaller $L$ provides less PSD. The lower PSD solutions with drift shell splitting in Figure 5.3 are consistent with this mechanism. One caveat is that a difference in chorus wave diffusion rates on $\Sigma^*$ and $\Sigma$ may also have an effect, and that should be investigated in future work.

As the point $P$ moves toward larger $K$ into the range of strong chorus wave scattering, the $\Sigma^*$ surface is raised in $L$ so that its intersection with the $\Sigma$ plane goes through the shaded areas, and the shaded areas are closer in $L$. Moreover, the shape of $\Sigma^*$ at large $K$ is close to that of $\Sigma$. Therefore, the drift shell splitting effect becomes negligible at large $K$, as illustrated in Figure 5.11.

In Figure 5.6, when chorus wave diffusion is decreased by the decreasing peak POES factor, the source regions in Figure 5.8 are moved closer to $P$ along the $\Sigma^*$ surface or the $\Sigma$ plane, so that they are less separated in $L$. Moreover, radial diffusion, which is then relatively stronger, would move these source regions in $L$ and further spread them. As a result, the source regions will increasingly overlap, and the drift shell splitting effect diminishes. Because both the chorus wave acceleration effect and the drift shell splitting effect depend on the POES factors, the PSD differences in
Figure 5.11: REM simulated flux profiles in generalized $\alpha_0$ with drift shell splitting off (blue) and on (red) for $E = 2.0$ MeV and $L = 4.25$ at $t = 9$ hour (dashed) and $t = 17$ hour (solid), calculated with a POES factor profile that peaks at $L = 5$ with a height 7. At $\alpha_0 \lesssim 40^\circ$, the two 17 hour solutions are effectively indistinguishable, suggesting that the drift shell splitting effect vanishes as $K$ becomes large.
Figure 5.6 reveal a simpler relation to POES factor than the PSD ratios in Figure 5.7; in the latter, this relation is complicated by the competition between the two effects.

5.4 Discussion and conclusions

Early drift shell splitting studies [e.g., Fälthammar and Walt, 1969; Roederer and Schulz, 1969; Schulz, 1972; Schulz and Lanzerotti, 1974, III.7–8] focused mainly on anomalous radial diffusion caused by elastic pitch-angle scattering in a distorted geomagnetic field. Since energy $E$ is conserved in this process, the corresponding anomalous radial diffusion coefficient, as estimated by drift average of the $L$-displacement at constant $\alpha_0$ and $E$ [e.g., Schulz, 1972], must be expressed in the $(\alpha_0, p, L)$ coordinates, and it is not the same as the $\langle D_{LL}^* \rangle$ obtained from a drift average of equation (5.1). In these early studies, the drift-resonant radial diffusion was derived assuming impulsive variations of the geomagnetic field and its induced electric field [Kellogg, 1959; Parker, 1960], which played a comparable role to ULF waves in the present work, and the pitch-angle diffusion rate $(D_{\alpha_0\alpha_0})$ was assumed longitudinally uniform due to theoretical simplicity and the lack of in situ observations. Roederer [1968] (also Schulz and Lanzerotti [1974, pp.106–107]) pointed out that, under such a diffusion scheme, the energy-conserving anomalous radial diffusion provided extra acceleration for electrons diffusing into the radiation belts from an external source. Given that a typical electron PSD radial profile has a positive $L$-gradient at fixed $M$ and $K$ and a negative gradient at fixed $E$, electrons have a favored inward diffusion at constant $M$ and $K$ and a favored outward diffusion at constant $E$. The consequence is that a fraction of the electrons would have repeated opportunities to gain energy by the energizing inward diffusion, and thus an enhanced PSD increase should occur.
Roederer’s prediction seems contrary to the results of the current study. However, this apparent contradiction is caused by the different longitudinal distributions of pitch-angle diffusion rate in the two schemes. In the picture of the phase-space surfaces of Figure 5.8, with longitudinally uniform pitch-angle diffusion rate, bounce-averaged diffusion in each $\Sigma^*$ surface of a drift shell contributes equally to the drift-averaged diffusion, so that drift shell splitting causes no systematic $L$-shell dislocation of source regions. In contrast, in our diffusion scheme, the pitch-angle scattering rate due to chorus waves is much greater on the nightside than the dayside, and Roederer’s energization mechanism becomes secondary compared to the reduction of the source electron abundance. Since the chorus wave distribution is statistically derived from observations, we would expect that a reduced PSD enhancement is more realistic.

As a practical consideration, for the radiation belt models that cannot deal with the $\langle D^*_{ML} \rangle$ and $\langle D^*_{KL} \rangle$ components, we have tested the effects of including $\langle D^*_{LL} \rangle$ alone as the drift shell splitting diffusion coefficient. In Figure 5.6, the PSD solutions obtained with only $\langle D^*_{LL} \rangle$ but not $\langle D^*_{ML} \rangle$ and $\langle D^*_{KL} \rangle$ (labelled “ro” for “radial only”) are compared to those with the full set of drift shell splitting diffusion coefficients turned “on” or “off”. At POES factor 1, compared to “off”, the “ro” solution is even further separated from the “on” solution that incorporates the full set of starred diffusion components. This is expected since adding $\langle D^*_{LL} \rangle$ alone is no different from increasing radial diffusion at constant $M$ and $K$. At larger peak POES factors, where chorus wave acceleration dominates the amount of PSD increase, the strengthened radial diffusion has very little effect on PSD.

In summary, we have shown, from simulations and theoretical arguments, that drift shell splitting decreases outer radiation belt electron PSD enhancements at small to intermediate $K$. This conclusion is derived from the basic geomagnetic field con-
figuration that $\partial L/\partial \alpha_0 > 0$ on the nightside, the dominance of nightside over dayside chorus in resonant scattering, and the typically positive seed electron PSD radial gradient. The magnitude of the decrease depends on both the separation of the PSD source regions in $L$, which is determined by the asymmetry of the magnetic field, and the PSD radial gradient of the seed electrons. In our simulations, we assumed a T89 magnetic field model with $Kp = 4$ and steady-state radial diffusion in the initial and boundary conditions, and found a nearly one-order-of-magnitude decrease in PSD enhancement. During magnetic storms both the field asymmetry and the source PSD radial gradient could be larger than in our model, thus leading to a stronger drift shell splitting effect.
Chapter 6

Summary and Future Work

6.1 Summary and conclusions

This dissertation presents research work on statistical modeling of the radiation belt electron diffusion in adiabatic invariant coordinates, and the application of this method to investigating effects of drift shell splitting. Apart from the background knowledge and review of previous work given in Chapter 1, the main results of the dissertation are summarized as follows.

In Chapter 2, we introduced the general theory relating a diffusion process to individual stochastic motions whose equation of motion is a stochastic differential equation (SDE), and provided the connection between the Fokker-Planck equation and its equivalent SDEs. For simplicity of formulation, the SDE theory was first presented in Cartesian coordinates. However, for radiation belt electrons, diffusion in phase space is more often described in curvilinear coordinates for physical convenience, and so the SDE theory was subsequently generalized to curvilinear coordinates. An interesting conclusion then followed that, though the coordinates had been changed from Cartesian to curvilinear, the distribution function satisfying the Kolmogorov backward equation (and the expectation representation) is still with respect to Cartesian coordinates. For radiation belt electron diffusion, the role of Cartesian coordinates in phase space is played by canonical variables. This not only illustrates a special property of canonical variables in mechanics, but also allows us to continue to use
the \((q,p)\) phase space density (PSD), even when we transform the Fokker-Planck equation to other coordinates.

Given the theoretical foundation of Chapter 2, Chapter 3 described the implementation of the SDE method to the particular radiation belt Fokker-Planck equation to be solved throughout this thesis. In particular, the question of the Neumann boundary condition form at vanishing second adiabatic invariant, which had been unclear in the literature for a long time, was mathematically settled from the variational theory of partial differential equations (PDEs). Implementation of the SDE method presented three major numerical problems: (i) efficiently decomposing the diffusion tensor into the product of a matrix and its transpose to set the coefficients for the SDEs, (ii) stochastically integrating the SDE in time, and (iii) evaluating the expectation of the functional values obtained from integration of the SDEs. Optimized numerical techniques for these tasks were developed in Chapter 3 in the context of the radiation belt diffusion problem.

Compared to the finite-difference PDE solvers, the SDE method has several numerical and theoretical advantages. From the numerical viewpoint, the SDE method does not rely on a computational grid, and thus it is capable of dealing with complicated boundary geometries, which facilitates the use of adiabatic invariants as coordinates. More importantly, finite-difference methods usually require global matrix diagonalization by coordinate transforms to retain positivity of their PSD solutions, which is generally unachievable for matrix dimensions higher than two, but the SDE method does not have this problem. Solution positivity is intrinsically guaranteed by the method, and the SDE method is the only candidate to date that can solve the Fokker-Planck equation with a fully three-dimensional (3-D) diffusion tensor, such as that with the drift shell splitting effect. From the theoretical viewpoint, the SDE
method gives additional information on the diffusion process while solving the Fokker-Planck equation, which allows the generation of PSD source plots. As a unique and powerful tool, the source plots provide deep insights into the origins of the energetic electrons, which are a major focus of radiation belt research. The main disadvantage of the SDE method is in its computing speed. But this can be ameliorated with development and application of high performance computing equipment and large scale parallelization.

The radiation belt model based on the SDE method is named REM (for Radbelt Electron Model). The validity of REM was successfully demonstrated in Chapter 4 by comparing its solutions both with known solutions from other methods and with spacecraft observations during a magnetic storm event. Effectiveness of the numerical algorithms presented in Chapter 3 in improving accuracy and reducing statistical fluctuation was also demonstrated in the model tests.

In Chapter 5, REM was used to investigate drift shell splitting effects on radiation belt electron diffusion, which had previously been limited to approximate theoretical treatments, mainly due to the lack of applicable simulation tools. REM simulation results of an idealized electron injection event suggested that, with drift shell splitting, the energetic electron PSD enhancement would be reduced compared to that without. This is in contrast to the increased acceleration prediction made in Schulz and Lanzerotti [1974], but that analysis considered only the additional radial diffusion caused by drift shell splitting. PSD source plots and theoretical analysis further revealed that the electron source regions were moved to lower $L$-shells by the drift shell splitting effect. The lowered $L$-shell source regions had fewer seed electrons because of the positive radial gradient of the hundred-keV electron PSD, and hence reduced the PSD enhancement. In the 2-D $a_0E$ diffusion limit, electron diffusion with drift
shell splitting takes place on a curved surface in phase space due to the greater chorus wave scattering rates on the nightside, rather than on a constant-$L$ plane without drift shell splitting, and this caused the lowered $L$-shell of the source.

6.2 Numerical improvements of REM

Both computational speed and numerical accuracy can be improved for REM. For improved computational speed, the current REM employs OpenMP parallelization, which is relatively simple to implement in coding, but is limited to use of only one computation node of a cluster. For instance, on the Rice cluster DAVinCI, the parallelized REM using all the 12 processors of a node runs about 6 times faster than that using a single processor (serial execution). However, considering the large number of stochastic paths to be simulated for each solution point, more powerful speed-up measures, such as MPI parallelization, that allow the use of multiple nodes simultaneously, or code implementation on advanced graphics processing units (GPU), should be considered for future development, and promise to reduce REM computational time by orders of magnitude.

For increased numerical accuracy, improvements can be made in the SDE integration scheme and in the reduction of statistical variance in evaluating expectations. The current SDE integration scheme, the Euler-Maruyama scheme (cf. Section 3.2.1), is order 1 in the weak convergence sense, meaning that when only the statistical distribution of the stochastic paths matters but not the individual path, the expectation of the scheme error is proportional to the first power of the time step size [Kloeden and Platen, 1992]. Higher order schemes, which are devised by Taylor expanding the stochastic process $X_t$ (consider a 1-D example here) to higher infinitesimal orders, such as the terms involving $dW_t dW_s$, $dt dW_s$, $dt ds$, ..., can be adopted by REM.
Time integration of these higher order terms gives corrections to the first order term, and yields higher order numerical schemes. Higher order schemes are systematically analyzed in *Kloeden and Platen* [1992].

However, caution should be used when employing higher order schemes. First, even for an order 2 weak scheme, the higher order terms in the Taylor expansion contain the first and the second order derivatives of the SDE coefficients $b$ and $\sigma$ (as in equation (2.1)). Evaluating these derivatives would not be a problem if they were analytically given, but unfortunately this is not the case in REM. Numerically calculating these derivatives for the highly variable radiation belt diffusion coefficients, which must also be carried out on an unevenly partitioned grid in REM, provides a source of numerical error that might well negate the benefit brought by the higher order scheme itself. Second, the multiple Itô integrals of the higher order terms are not easy to evaluate. Even worse, in REM the SDEs are 3-D rather than 1-D. This means the multiple Itô integrals would also involve components of the Wiener process from different dimensions, which turn out to be more difficult to evaluate [*Kloeden and Platen*, 1992]. Nonetheless, if these difficulties can be overcome in REM, not only would the code accuracy improve, but also the integration time step may be lengthened, which might make the code faster.

Reducing statistical variance in the distribution of the functional values $F^t, Q$ could be another improvement of REM. In some extreme solutions, such as the one illustrated in Figure D.4, among the $10^6$ simulated stochastic paths, only a few tens of them were from the low-energy boundary, but contributed the majority of the solution. From the pixel colors in Figure D.4, it is seen that the functional values of a stochastic path from the low-energy boundary is orders of magnitude higher than that from the interior of the phase space. The large variance in the distribution makes
the estimation of its expectation value inaccurate by the arithmetic mean, and this statistical error only decreases with $N^{-1/2}$ as $N$ (total number of stochastic paths) approaches $+\infty$. Therefore, it is desirable to reduce the statistical variance in the functional value distribution while keeping its expectation unchanged; techniques for this purpose are called variance reduction methods.

One variance reduction method, the measure transformation method [Kloeden and Platen, 1992], has been tried on REM. This method uses a transformation named after Girsanov that changes the underlying probability measure of the stochastic motions, and results in an additional advection term in the transformed SDE [Kloeden and Platen, 1992; Øksendal, 1998]. Consequently, in addition to scattering randomly in phase space due to the diffusion coefficients, the transformed stochastic motions also advect to the regions of greater contribution (in a time-backward sense), and the problem illustrated in the above paragraph would then be mitigated. To do this, we need an initial guess on which parts of the phase space are more important; in other words, an initial guess of the solution of the Fokker-Planck equation. This initial guess is used as a generation function for the Girsanov transformation. Then, the transformed SDEs are integrated in REM, and the expectation of the transformed functional values are evaluated with a correction factor (named the Radon-Nikodym derivative). This expectation is theoretically equal to the untransformed expectation, and has a reduced variance depending on the proximity of the initial guess to the true solution. (An initial guess identical to the true solution reduces the variance to zero.)

However, numerical experiments of the measure transformation method, using the initial condition as the generation function, were not successful in REM. It not only failed to yield the correct PSD solutions, but also disabled the generation of PSD source plots, since the statistics of the transformed stochastic motions were
different from the untransformed ones. Reasons for this failure are not clear, possibly involving the error introduced by numerically calculating derivatives of the initial condition during the Girsanov transformation. To verify and overcome this, the use of an analytic initial guess, such as the approximate solution given in Appendix C, might be a good next step. Other variance reduction methods are also described in Kloeden and Platen [1992]. If any of the variance reduction methods is applicable to REM, it may potentially reduce by orders of magnitude the number of stochastic path simulations, hence greatly improving both the accuracy and speed of the code.

Nevertheless, even without these numerical improvements, the current version of REM as shown is still a useful, powerful, and in some ways unique radiation belt electron model that can be employed in many future studies. This is the topic of the next section.

6.3 Future studies

One apparent future application of REM is the data-driven event simulation, with spacecraft observed initial and boundary conditions, and event-specific magnetic field models, diffusion coefficients, and possible loss terms. In particular, the SDE method can easily adopt time-varying boundary locations, which enables the use of spacecraft apogee observations as a natural outer-$L$ boundary condition, and hence avoids the extrapolation of observational data (to potentially unrealistic values). The time-varying boundaries would be very difficult for a finite-difference-based method to incorporate. The PSD source plots make REM simulations more informative than the other radiation belt electron models, by revealing explicitly the dominant diffusion processes and the sources of the energetic electrons. Systematic investigation of the PSD source plots of various events can produce knowledge on the radiation belt
diffusion from an entirely new viewpoint.

From the PSD source plots in the drift shell splitting simulations, we have found that, when radial diffusion dominates, the Shabansky orbit boundary can contribute the majority of the electron PSD (cf. Figure 5.5). This is the first time in a fully 3-D radiation belt diffusion simulation that the Shabansky orbits have been considered. A thorough understanding of the role of the Shabansky orbits might be a topic of future study. In this thesis, we have treated Shabansky orbits as exterior to the phase space domain, and assigned on the Shabansky orbit boundary a reasonable yet still arbitrary characteristic lifetime for electrons, mainly for simplicity. Although electrons on Shabansky orbits do not have the third invariant defined due to their unclosed drift shells, their states of motion are still describable using a set of parameters, e.g., the first two adiabatic invariants plus the equatorial crossing distance in the midnight meridian, as in Ukhorskiy et al. [2011]. The stochastic motion of Shabansky orbit electrons in the parameter space consists of both diffusive and ballistic components [Öztürk and Wolf, 2007; Ukhorskiy et al., 2011], but it can still be described by SDEs, and thus in principle these electrons could still be incorporated in radiation belt simulations, provided a coordinate transform across the Shabansky orbit boundary from the three adiabatic invariants to the selected set of parameters and vice versa is available. This might be a better way of dealing with the Shabansky orbits.

The effects of localized chorus wave diffusion, which are quantified by the POES factors, could be another topic of future work. POES factors modulate the relative strength of chorus wave diffusion, and the spatial distribution of POES factors affects the spatial distribution of electron local acceleration rates, which could cause new spatial structures in the radiation belts. For example, in the idealized storm event simulation in Chapter 5, if the modeled POES factor profile peaks at $L = 5$ instead
of 4 in Figure 5.2, i.e., roughly the position of the secondary peak in the observations, the simulated electron fluxes would feature a strong local acceleration centered around $L \sim 5$, as shown in Figure 6.1. In Figure 6.1, the initial electron flux profile (black curve) peaks around $L \sim 3$, which resembles the quiet time outer belt, and is referred to in Baker et al. [2013] as the storage ring since it stores the energetic electrons before the beginning of strong acceleration. After 9 hours, in either the case with drift shell splitting turned off or on, a strong local acceleration takes place beyond $L \sim 4$, which generates a new belt of energetic electrons, with a flux trough at about $L \sim 4$ separating it from the storage ring. This could be a possible scenario for the formation of the “three-belt structure” as reported in Baker et al. [2013]. Once again, the PSD increases in the drift shell splitting on case are lower than those in the off case, as we have discussed in Chapter 5.

The research on drift shell splitting effects in this dissertation marks a starting point to more thorough investigations. The drift shell splitting effects associated with electron diffusion caused by various plasma waves other than chorus, such as plasmaspheric hiss and EMIC waves, are still unknown and have yet to be studied. Dependences of the drift shell splitting effects on wave models, initial and boundary conditions, and magnetic activity levels are also possible future work. With REM, the first fully 3-D radiation belt electron diffusion model, and the unprecedented in situ observation from the Van Allen Probes mission [Lanzerotti, 2013], we expect to see advances toward better understanding of the physics in the radiation belts.
Figure 6.1: REM simulated flux radial profiles with drift shell splitting off (blue) and on (red) for $E = 2.0$ MeV and $K = 0.010 \ G^{1/2} R_E$ electrons at $t = 9$ hour (dashed) and $t = 17$ hour (solid), calculated with a POES factor profile that peaks at $L = 5$ with a height 7. The other setups are the same as the idealized simulations in Chapter 5.
Appendix A

Divergence Operator and Diffusion Equation in Curvilinear Coordinates

This appendix provides explanation of the divergence operator and the diffusion equation in curvilinear coordinates from the perspective of Gauss and Stokes theorems. In such an approach, only a minimal reliance on tensor analysis in curvilinear coordinates is required. For an alternative, more conventional approach using the metric tensors, see, e.g., D’haeseleer et al. [1991].

In this appendix, we use subscbbted vectors to indicate tangent-basis vectors, and superscbbibed vectors to indicate reciprocal-basis vectors. The tangent-basis vectors are the vectors along the (curved) coordinate lines at some spatial point \( P \), defined through the expansion of an infinitesimal displacement \( d\mathbf{r} \) at \( P \),

\[
d\mathbf{r} = \frac{\partial \mathbf{r}}{\partial \xi^i} d\xi^i = d\xi^i \mathbf{e}_i,
\]

(A.1)

with

\[
\mathbf{e}_i \equiv \frac{\partial \mathbf{r}}{\partial \xi^i}.
\]

(A.2)

Therefore, \( d\xi^i \mathbf{e}_i \) (no summation) gives the infinitesimal increment along the \( \xi^i \) coordinate line in the \( \mathbf{e}_i \) direction. A vector can be expanded on the basis \( \{ \mathbf{e}_i \} \), yielding

\[
\mathbf{a} = a^i \mathbf{e}_i,
\]

(A.3)

where \( a^i \) are the contravariant components of \( \mathbf{a} \). To define inner product, we have to introduce another basis, reciprocal to the tangent-basis, that satisfy

\[
\mathbf{e}^i \cdot \mathbf{e}_j = \delta_j^i,
\]

(A.4)
where $\delta^i_j$ is the Kronecker delta. This basis $\{\epsilon^i\}$, naturally, is named the reciprocal-basis. Then, we can calculate the inner product between $\epsilon^i$ and $dr$ and, using (A.1) and (A.4), get

$$d\mathbf{r} \cdot \epsilon^i = d\xi^i \epsilon_j \cdot \epsilon^i = d\xi^i.$$  \hfill (A.5)

From another point of view, the increment $d\xi^i$ is the projection of $d\mathbf{r}$ in the direction of the steepest increase of $\xi^i$, in other words, in the direction of $\nabla \xi^i$, thus,

$$d\xi^i = d\mathbf{r} \cdot \nabla \xi^i.$$  \hfill (A.6)

Comparing (A.5) and (A.6), we find that the reciprocal-basis vector $\epsilon^i$ is the gradient of the coordinate $\xi^i$, i.e.,

$$\epsilon^i = \nabla \xi^i.$$  \hfill (A.7)

In a curvilinear coordinate system, the gradient $\nabla \xi^i$ is the normal vector of the hypersurface defined by constant $\xi^i$. A two-dimensional (2D) example of such geometry of tangent- and reciprocal-basis vectors is illustrated in Figure A.1. A vector can also be expanded on the basis $\{\epsilon^i\}$

$$\mathbf{b} = b_i \epsilon^i,$$  \hfill (A.8)

where $b_i$ are the covariant components of $\mathbf{b}$. Then, the inner product between $\mathbf{a}$ and $\mathbf{b}$ is calculated by

$$\mathbf{a} \cdot \mathbf{b} = a^i b_j \epsilon_i \cdot \epsilon^j = a^i b_i.$$  \hfill (A.9)

Specifically, in Cartesian coordinates, the tangent- and reciprocal-basis are identical. They are denoted as $\{\epsilon_i\}$ and $\{\epsilon^i\}$ respectively in this appendix.

A diffusion equation can be obtained from the continuity equation of a distribution function $f$ and the corresponding diffusive flux $\mathbf{F}$, i.e.,

$$\frac{\partial f}{\partial t} + \nabla \cdot \mathbf{F} = 0,$$  \hfill (A.10)
Figure A.1: A 2D illustration of the tangent- and reciprocal-basis vectors. $\xi^1$ and $\xi^2$ are curvilinear coordinate lines intersecting at point $P$. $\epsilon_1$ and $\epsilon_2$ are the tangent-basis vectors, and $\epsilon^1$ and $\epsilon^2$ are the reciprocal-basis vectors.

where $\mathbf{F}$ is proportional to the steepest descent of $f$ by tensor diffusion coefficients $\mathbf{D}$

$$\mathbf{F} = -\mathbf{D} \cdot \nabla f. \quad (A.11)$$

Therefore, symbolically, the diffusion equation is written as

$$\frac{\partial f}{\partial t} = \nabla \cdot (\mathbf{D} \cdot \nabla f). \quad (A.12)$$

In Cartesian coordinates $\{x^i\}$, this equation is expressed as

$$\frac{\partial}{\partial t} f(t, \mathbf{x}) = \frac{\partial}{\partial x^i} \left( D_{x}^{ij} \frac{\partial f}{\partial x^j} \right), \quad (A.13)$$

where $D_{x}^{ij}$ are the contravariant components of $\mathbf{D}$ in these coordinates.

In curvilinear coordinates $\{\xi^i\}$, the gradient and divergence operators are written
as follows. Consider an infinitesimal increment of $f$ over $dr$,

$$df = d\xi^i \frac{\partial f}{\partial \xi^i} = d\xi^i \epsilon_i \cdot \epsilon^j \frac{\partial f}{\partial \xi^j} = dr \cdot \nabla f.$$  \hspace{1cm} (A.14)

By the last equality, we have

$$\nabla f = \epsilon^i \frac{\partial f}{\partial \xi^i}.$$  \hspace{1cm} (A.15)

Using the metric tensor, D’haeseleer et al. [1991] proves that the divergence operator in three-dimensional curvilinear coordinates has the form

$$\nabla \cdot \mathbf{g} = \frac{1}{G} \frac{\partial}{\partial \xi^i} \left( G g^i_{\xi} \right),$$  \hspace{1cm} (A.16)

where $G = \det(\partial x^i/\partial \xi^j)$ is the Jacobian determinant, and $g^i_{\xi}$ are the contravariant components of $\mathbf{g}$ in $\{\xi^i\}$ coordinates. Alternatively, without appealing to the metric tensor, the expression (A.16) can be verified in $n$ dimensions using the Gauss and generalized Stokes theorems, in the following manner. Gauss theorem states that

$$\int_{\Omega} \nabla \cdot \mathbf{g} \, d\Omega = \int_{\partial \Omega} \mathbf{g} \cdot \mathbf{\hat{n}} \, d\sigma,$$  \hspace{1cm} (A.17)

in which $\partial \Omega$ is the closed boundary of domain $\Omega$, and $\mathbf{\hat{n}} \, d\sigma$ is an oriented hypersurface element on $\partial \Omega$ with $\mathbf{\hat{n}}$ its outward unit normal. Let us now check that the expression (A.16) gives the correct Gauss formula in Cartesian coordinates. Integrating $\nabla \cdot \mathbf{g}$ over space yields

$$\int_{\Omega} \nabla \cdot \mathbf{g} \, dx = \int_{\Omega} \frac{1}{G} \frac{\partial}{\partial \xi^i} \left( G g^i_{\xi} \right) G \, d\xi = \int_{\Omega} \frac{\partial}{\partial \xi^i} \left( G g^i_{\xi} \right) \, d\xi,$$  \hspace{1cm} (A.18)

where $dx$ is the abbreviation for the spatial volume element, in terms of exterior differential forms [e.g., Spivak, 1968; Li and Zhou, 1999a], it is $dx^1 \wedge dx^2 \wedge \cdots \wedge dx^k$; and so is $d\xi$. The wedge product $\wedge$ is like an ordinary product, but changes sign when the order of differentials is switched, e.g., $dx^1 \wedge dx^2 = -dx^2 \wedge dx^1$. By the generalized
Stokes theorem (the fundamental theorem of calculus in more than one dimension \[e.g., \text{Spivak, 1968; Li and Zhou, 1999a; Zorich, 2004}\]), the right-hand side of (A.18) is given by

\[
\int_{\Omega} \frac{\partial}{\partial \xi^i} (Gg^i_i) \, d\xi^1 \wedge \cdots \wedge d\xi^k = \oint_{\partial\Omega} (-1)^{i-1} Gg^i_i \, d\xi^1 \wedge \cdots \wedge d\xi^i \wedge \cdots \wedge d\xi^k, \tag{A.19}
\]

where, following the notation of \text{Zorich [2004]}, the \(\hat{\cdot}\) stands over a quantity that is to be omitted. The Jacobian \(G\) can be expanded in minors along the \(i\)-th column as

\[
G = \sum_j (-1)^{i+j} \frac{\partial x^j}{\partial \xi^i} \frac{\partial (x^1, \ldots, \hat{x}^j, \ldots, x^k)}{\partial (\xi^1, \ldots, \hat{\xi}^i, \ldots, \xi^k)}, \tag{A.20}
\]

in which the summation is only over \(j\) as has been explicitly indicated. The right-most factor in (A.20) is a \((k-1)\)-dimensional Jacobian, it transforms the differential forms in \(\{\xi^i\}\) to those in \(\{x^i\}\), by

\[
\frac{\partial (x^1, \ldots, \hat{x}^j, \ldots, x^k)}{\partial (\xi^1, \ldots, \hat{\xi}^i, \ldots, \xi^k)} \, d\xi^1 \wedge \cdots \wedge d\xi^i \wedge \cdots \wedge d\xi^k = dx^1 \wedge \cdots \wedge \hat{dx}^j \wedge \cdots \wedge dx^k. \tag{A.21}
\]

Additionally, expanding \(g\) in Cartesian coordinates gives

\[
g = (g^i_\xi e_i \cdot e^j) \, e_j = g^i_\xi \frac{\partial x^j}{\partial \xi^i} \, e_j, \tag{A.22}
\]

therefore,

\[
g^j_x = g^i_\xi \frac{\partial x^j}{\partial \xi^i}. \tag{A.23}
\]

Inserting (A.20) into (A.19) and using (A.21) and (A.23) gives

\[
\oint_{\partial\Omega} (-1)^{i-1} Gg^i_i \, d\xi^1 \wedge \cdots \wedge d\xi^i \wedge \cdots \wedge d\xi^k = \oint_{\partial\Omega} (-1)^{j-1} g^j_x \, dx^1 \wedge \cdots \wedge \hat{dx}^j \wedge \cdots \wedge dx^k. \tag{A.24}
\]

In Cartesian coordinates, projections of a hypersurface element \(d\sigma\) in the \(e^j\) direction are given by \[cf. Zorich, 2004, section 13.2.4\]

\[
\hat{n}_j d\sigma = (-1)^{j-1} dx^1 \wedge \cdots \wedge \hat{dx}^j \wedge \cdots \wedge dx^k. \tag{A.25}
\]
Therefore, combining (A.25), (A.24) and (A.18), we find that

$$\int_{\Omega} \nabla \cdot \mathbf{g} \, dx = \oint_{\partial \Omega} g^i_j \hat{n}_j \, d\sigma,$$  \hspace{1cm} (A.26)

which gives the correct Gauss formula in the \( \{x^i\} \) coordinates and completes our verification of the curvilinear coordinate divergence operator (A.16).

Now, returning to our symbolic diffusion equation (A.12), and applying the gradient and divergence operators (A.15) and (A.16), the diffusion equation in the \( \{\xi^i\} \) coordinates becomes

$$\frac{\partial}{\partial t} f(t, \xi) = \frac{1}{G} \frac{\partial}{\partial \xi^i} \left( G D^{ij}_\xi \frac{\partial f}{\partial \xi^j} \right),$$  \hspace{1cm} (A.27)

where \( D^{ij}_\xi \) are the contravariant components of \( D \) on the \( \{\epsilon_i\} \) basis. They are related to the Cartesian components by

$$D_{\xi} = D^{ij}_\xi \epsilon_i \epsilon_j = D^{ij}_x \left( \epsilon_i \cdot \epsilon^l \right) \left( \epsilon_j \cdot \epsilon^m \right) \epsilon_l \epsilon_m$$  \hspace{1cm} (A.28)

Therefore we have

$$D^{lm}_\xi = \frac{\partial \xi^l_x}{\partial x^i} D^{ij}_x \frac{\partial \xi^m_x}{\partial x^j},$$  \hspace{1cm} (A.29)

which constitutes a congruence transform of a matrix from the \( \{x^i\} \) to the \( \{\xi^i\} \) coordinates.
Appendix B

An Example REM Solution Statistics Sheet

This appendix illustrates an example REM solution statistics sheet taken from a simulation using Ozeke et al. [2014] ULF radial diffusion and O’Brien [2015] drift shell splitting chorus wave diffusion coefficients. The statistics listed in Table B.1 are taken from the solution at $M = 1212$ MeV/$G$, $K = 0.01$ G$^{1/2}$R$_E$, $L = 4.25$ and $t = 17$ hour (0.71 day) in the idealized simulation described in Chapter 5 of this thesis. The specified root-mean-square spatial step size is 0.01, the significance level for estimating the confidence interval is 90%, and the solution relative error tolerance is 10%. Each batch contains 2400 simulations of stochastic paths. The resultant relative error after 60 batches is 9.9%, where the solution process is stopped, and the maximum spatial step size that occurred is 0.71.

In Table B.1, $\tau$ denotes the mean stopping time (in days) of all reachable Dirichlet boundaries, and $\tau_i$ specifies the mean stopping time of the $i$-th piece of surface of the boundary, which is specified in the code, e.g., $i = 1$ stands for the $E_{min}$ boundary. The upper two rows in the table give the batch-wise means (Mean) and standard deviations (StDev) over the total 60 batches; and the lower rows list the path-wise means obtained within each batch, with the batch number indicated in the first column. For the stopping times, in the upper rows a “−1” indicates that the corresponding boundary piece is not reached, so that mean and standard deviation are not defined; and in the lower rows this is indicated by zeros.

From the mean spatial step size 0.01, it is evident (from column 4 of Table B.1)
that the adaptive Euler-Maruyama scheme (cf. Subsection 3.2.3) has successfully maintained the prescribed root-mean-square step size by adaptively varying the time step sizes, although statistical fluctuation made the biggest space step (0.71) nearly two orders of magnitude larger than the specified root-mean-square value (0.01).

The standard deviation of the PSD result (in unit of $c/cm/MeV^3$) is about a half of the mean. Considering that the batch PSD averages ($\hat{A}_{n,j}$ in Subsection 3.2.4) have a distribution that is lower-bounded by zero, a standard deviation that is comparable with the mean indicates that the actual distribution of $\hat{A}_{n,j}$ does not well approximate a Gaussian, which further indicates that the number of simulations in each batch is not quite enough to apply the central limit theorem in this case. However, this only affects our estimation of an accurate confidence interval, and it does not affect the expectation result. In some extreme cases, the Gaussian distribution is so well approximated that the standard deviation is smaller than the mean by orders of magnitude, so that the result can reach a relative error lower than 1% with only a few batches. Whereas at the other extreme, the standard deviation could be one order of magnitude larger than the mean, and REM exhausts its 1000 batch limit before reaching the error tolerance. Both of these extremes are rare in practice.
Table B.1: An example REM solution statistics sheet. See the text in this appendix for its explanation.

<table>
<thead>
<tr>
<th>PSD</th>
<th>time step</th>
<th>spatial step</th>
<th>$\tau$</th>
<th>$\tau_1$</th>
<th>$\tau_2$</th>
<th>$\tau_3$</th>
<th>$\tau_4$</th>
<th>$\tau_5$</th>
<th>$\tau_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>2.299876E-08</td>
<td>0.000097</td>
<td>0.011129</td>
<td>0.501865</td>
<td>0.651163</td>
<td>0.500192</td>
<td>-1.000000</td>
<td>-1.000000</td>
<td>0.587493</td>
</tr>
<tr>
<td>StDev</td>
<td>1.052553E-08</td>
<td>0.000005</td>
<td>0.000134</td>
<td>0.010980</td>
<td>0.032244</td>
<td>0.010931</td>
<td>-1.000000</td>
<td>-1.000000</td>
<td>0.083350</td>
</tr>
<tr>
<td>1</td>
<td>2.101352E-08</td>
<td>0.000100</td>
<td>0.011183</td>
<td>0.501129</td>
<td>0.000000</td>
<td>0.500052</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.699416</td>
</tr>
<tr>
<td>2</td>
<td>2.060861E-08</td>
<td>0.000105</td>
<td>0.010966</td>
<td>0.514427</td>
<td>0.000000</td>
<td>0.513508</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>3</td>
<td>2.311559E-08</td>
<td>0.000091</td>
<td>0.011123</td>
<td>0.515323</td>
<td>0.000000</td>
<td>0.512825</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>4</td>
<td>2.149226E-08</td>
<td>0.000103</td>
<td>0.011091</td>
<td>0.508406</td>
<td>0.000000</td>
<td>0.506237</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>5</td>
<td>2.051684E-08</td>
<td>0.000095</td>
<td>0.011128</td>
<td>0.516372</td>
<td>0.000000</td>
<td>0.515074</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>6</td>
<td>4.860785E-08</td>
<td>0.000102</td>
<td>0.010940</td>
<td>0.492762</td>
<td>0.673963</td>
<td>0.491084</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.434169</td>
</tr>
<tr>
<td>7</td>
<td>9.567874E-08</td>
<td>0.000091</td>
<td>0.011140</td>
<td>0.512563</td>
<td>0.628363</td>
<td>0.511439</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>8</td>
<td>2.087146E-08</td>
<td>0.000096</td>
<td>0.011314</td>
<td>0.487075</td>
<td>0.000000</td>
<td>0.484507</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>9</td>
<td>2.049250E-08</td>
<td>0.000092</td>
<td>0.011169</td>
<td>0.486782</td>
<td>0.000000</td>
<td>0.485425</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>10</td>
<td>2.024433E-08</td>
<td>0.000087</td>
<td>0.011325</td>
<td>0.500778</td>
<td>0.000000</td>
<td>0.499921</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>60</td>
<td>2.056399E-08</td>
<td>0.000091</td>
<td>0.011185</td>
<td>0.509467</td>
<td>0.000000</td>
<td>0.508313</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.648108</td>
</tr>
</tbody>
</table>
Appendix C

Designing Initial and Boundary Conditions for the Radiation Belt Fokker-Planck Equation

A realistic initial condition for the radiation belt Fokker-Planck equation should be a solution of the equation at that instant of time, since the initial condition itself is a result of the radiation belt dynamical evolution from an earlier time. In this appendix we seek, under certain assumptions, an approximate special solution to the Fokker-Planck equation by separation of variables, for use in assigning initial and boundary conditions.

We consider bimodal radial and pitch-angle diffusion, which approximately conserves the quantity \( \zeta \equiv M/y^2 \) [Walt, 1970], where \( y \equiv \sin \alpha_0 \). In dipole field, the corresponding Fokker-Planck equation is [Schulz and Lanzerotti, 1974, p110]

\[
\frac{\partial \tilde{f}}{\partial t} = L^{5/2} \frac{\partial}{\partial L} \left[ L^{-5/2} D_{LL} \frac{\partial \tilde{f}}{\partial L} \right]_{x,\zeta} + \frac{1}{xT(y)} \frac{\partial}{\partial x} \left[ xT(y)D_{xx} \frac{\partial \tilde{f}}{\partial x} \right]_{\zeta,L},
\]  
(C.1)

where \( x \equiv \cos \alpha_0 \) and \( T(y) \) is the dipole bounce function defined in (3.15). A special solution of (C.1) can be obtained by assuming that the radial diffusion has reached a steady state, so that

\[
L^{5/2} \frac{\partial}{\partial L} \left[ L^{-5/2} D_{LL} \frac{\partial a}{\partial L} \right]_{x,\zeta} = 0,
\]  
(C.2)

equation (C.1) becomes

\[
\frac{\partial g}{\partial t} = \frac{1}{xT(y)} \frac{\partial}{\partial x} \left[ xT(y)D_{xx} \frac{\partial g}{\partial x} \right]_{\zeta,L},
\]  
(C.3)
and \( \tilde{f}(t, \zeta, x, L) \) is given by

\[
\tilde{f}(t, \zeta, x, L) = a(\zeta, L) g(t, x) = a(\zeta, L) \sum_n g_n(x) \exp(-\lambda_n t), \tag{C.4}
\]

where \( g_n(x) \) and \( \lambda_n \) are the eigenfunctions and eigenvalues of (C.3). For \( x \)-independent \( D_{xx} \) and constant loss cone cosine \( x_c \), \( g_n(x) \) is proportional to the zero order Bessel function \( g_n(x) \propto J_0(\kappa_n x/x_c) \), where \( \kappa_n \) are defined by \( J_0(\kappa_n) = 0 \) [Roberts, 1969].

At fixed \( \zeta \), with the form \( D_{LL} = D_0 L^\nu \) and the boundary conditions

\[
\begin{cases}
  a(\zeta, L_1) = 0 \quad & (1 < L_1 < L_2, \ c'(\zeta) > 0), \\
  a(\zeta, L_2) = c'(\zeta)
\end{cases} \tag{C.5}
\]

in which \( L_1 \) is the radial position of the radiation belt slot and \( L_2 \) an arbitrary outer position, the solution of (C.2) is

\[
a(\zeta, L) = \frac{c'(\zeta)}{L_1^{7/2-\nu} - L_2^{7/2-\nu}} (L_1^{7/2-\nu} - L^{7/2-\nu})
\equiv c(\zeta) \left[ 1 - \left( \frac{L}{L_1} \right)^{7/2-\nu} \right], \quad (L \geq L_1, \ c(\zeta) > 0), \tag{C.6}
\]

with \( c(\zeta) \) a free function to be determined from the electron energy spectrum. Note that in integrating for (C.6) we have tacitly assumed \( \nu > 7/2 \). For the purpose of designing a simple initial condition, we assume that, at \( t = 0 \), \( \tilde{f} \) stays at the lowest eigenmode of pitch-angle diffusion, which is arguably reasonable between storms [O’Brien et al., 2014], so that

\[
\tilde{f}(0, \zeta, x, L) = c(\zeta) \left[ 1 - \left( \frac{L}{L_1} \right)^{7/2-\nu} \right] g_0(x). \tag{C.7}
\]

To find a simple form for \( g_0(x) \), note that the lowest eigenfunction \( J_0(\kappa_0 x/x_c) \) resembles the function \( 1 - (x/x_c)^2 \) [Schulz and Lanzerotti, 1974, pp.162–166]. In the outer radiation belt, we may further assume that \( x_c \simeq 1 \), and hence use the simple form

\[
g_0(x) \propto 1 - x^2. \tag{C.8}
\]
On the other hand, in dipole field particle momentum $p$ is drift invariant, so we can phase average the particle flux $j$ and obtain

$$\bar{j}(t, E, x, L) = p^2 \bar{f}(t, E, x, L) = \frac{2m_0\mu_E\zeta}{L^3 R_E^3} \bar{f}(t, E, x, L), \quad (C.9)$$

where $m_0$ is the particle rest mass, $\mu_E$ is the Earth’s dipole moment, and $R_E$ is the Earth’s radius. Assuming an exponential energy spectrum with e-folding energy $E_0$, the form of $\bar{j}$ is

$$\bar{j}(t, E, x, L) = \bar{j}(t, 0, x, L) \exp(-E/E_0), \quad (C.10)$$

with kinetic energy

$$E = \mathcal{E}_0(\gamma - 1) = \mathcal{E}_0 \left( \sqrt{\frac{2\mu_E\zeta}{\mathcal{E}_0 R_E^3 L^3}} + 1 - 1 \right), \quad (C.11)$$

a function of $\zeta$ and $L$, where $\gamma$ is the Lorentz factor and $\mathcal{E}_0$ is the rest energy.

Conservation of the first and second adiabatic invariants by radial diffusion implies that [Schulz and Lanzerotti, 1974, p131]

$$\left( \frac{\partial \ln \bar{j}}{\partial \ln p} \right)_{L,y} = \text{const.} \quad (C.12)$$

For highly relativistic electrons ($p^2 c^2 \gg \mathcal{E}_0$), this restriction prescribes a relation for the e-folding energy that $E_0 L^{3/2} = \text{const}$, which is observationally evident (e.g., Paolini et al. [1968] report $E_0 L^{1.3} \simeq \text{const}$). With this relation, $\exp(-E/E_0)$ is only a weak function of $L$, and independent of $L$ for highly relativistic electrons. Combining (C.7) through (C.11), we have

$$\bar{f}(0, \zeta, x, L) = c(\zeta) \left[ 1 - \left( \frac{L}{L_1} \right)^{7/2-\nu} \right] (1 - x^2)$$

$$= \frac{R_E^3}{2m_0\mu_E} \frac{1}{\zeta} \exp \left( -\frac{E}{E_0} \right) L^3 \bar{j}(0, E = 0, x, L). \quad (C.13)$$

Therefore, comparison of variables between the two equations in (C.13) yields

$$\bar{j}(0, E = 0, x, L) = \frac{\Gamma}{L^3} \left[ 1 - \left( \frac{L}{L_1} \right)^{7/2-\nu} \right] (1 - x^2), \quad (C.14)$$
which has a peak at \([(\nu - 1/2)/3]^{2/(2\nu - 7)}L_1\), and

\[
c(\zeta) \simeq \frac{R^2 \Gamma}{2m_0 \mu E} \frac{1}{\zeta} \exp \left[ - \frac{E_0 L^{3/2}}{E'_0} \left( \frac{2 \mu E \zeta}{\mathcal{E}_0 R^2 L^3} + 1 - 1 \right) \right]. \tag{C.15}
\]

In these expressions, \(\Gamma\) is an arbitrary constant that scales the asymptotic “zero energy” equatorial flux amplitude

\[
\bar{j}(0, E = 0, x = 0, L) = \frac{\Gamma}{L^3}, \quad (L \gg L_1), \tag{C.16}
\]

and \(E'_0 = E_0 L^{3/2}\) a constant scaling the e-folding energy.

For Dirichlet boundary conditions, which must reduce to \(\bar{f}(0, \zeta, x, L)\) on boundaries at \(t = 0\), \(\bar{f}(t, \zeta, x, L)\) can be obtained by varying the two constants \(\Gamma\) and \(E'_0\) with time. Cast in more conventional variables, we have

\[
\bar{f}(t, E, x, L) = \frac{\Gamma(t)}{L^3} \left[ 1 - \left( \frac{L}{L_1} \right)^{\gamma/2 - \nu} \right] \frac{c^2}{E(E + 2E_0)} \exp \left[ - \frac{E L^{3/2}}{E'_0(t)} \right] (1 - x^2), \tag{C.17}
\]

where \(c\) is the speed of light, and the values of \(\Gamma(t)\) and \(E'_0(t)\) are determined from fitting to observational data at different \(t\). Note that, comparing with (C.4), we have absorbed the \(\exp(-\lambda_0 t)\) term into \(\Gamma(t)\). Exponential variations of the particle flux with time further restricts the functional forms of \(\Gamma(t)\) and \(E'_0(t)\). From equations (C.10) and (C.14), we have

\[
\bar{j}(t, E, x = 0, L) = \frac{\Gamma(t)}{L^3} \left[ 1 - \left( \frac{L}{L_1} \right)^{\gamma/2 - \nu} \right] \exp \left[ - \frac{E}{E'_0(t)} L^{3/2} \right], \tag{C.18}
\]

so that the time derivative

\[
\frac{\partial}{\partial t} \ln \bar{j}(t, E, x = 0, L) = \frac{1}{\Gamma} \frac{d\Gamma}{dt} - E L^{3/2} \frac{d}{dt} \left( \frac{1}{E'_0} \right) \tag{C.19}
\]

should be piecewise constant in \(t\). Therefore, the functional forms of \(\Gamma(t)\) and \(E'_0(t)\) are

\[
\Gamma(t) = \Gamma(0) \exp(\lambda t), \tag{C.20}
\]

which have a peak at \([(\nu - 1/2)/3]^{2/(2\nu - 7)}L_1\), and
\[
\frac{1}{E_0'(t)} = \beta t + \frac{1}{E_0'(0)},
\]  
(C.21)

where \( \lambda \) and \( \beta \) are piecewise constant parameters to be determined from fitting to observational data. Figure C.1 illustrates the fitted electron fluxes for the 8 October 2012 storm, as calculated from (C.17) with \( \nu = 6 \), corresponding to \( K_p = 4 \) in the Ozeke et al. [2014] ULF wave diffusion coefficients, and the slot region position \( L_1 = 2.5 \).

The solution (C.17) has the typical characteristics of a quiet time PSD: it has a positive \( L \)-gradient at fixed \( M \) and \( K \), and a negative \( L \)-gradient at fixed \( E \) and \( x \) beyond a peak. On the low energy boundary, equation (C.17) physically assumes that the relaxation time for the seed electrons to radially diffuse to a steady state is much shorter than the characteristic time scale of the injection. On the Shabansky orbit boundary, where a 50 drift-period \( (T_d) \) electron lifetime is assumed, the boundary condition \( (\bar{h}) \) is multiplied by the factor as

\[
\bar{h}(t, E, x, L) \bigg|_{boundary} = \bar{f}(t, E, x, L) \exp \left( -\frac{t}{50T_d} \right) \bigg|_{boundary},
\]  
(C.22)

in which \( T_d \) is also a function of \( E \), \( x \) and \( L \).
Figure C.1: (top) Fitted omni-directional electron fluxes on the low-energy boundary ($E = 100$ keV) versus $L$ at different times and (bottom) fitted uni-directional fluxes at $L = 4$ and $\alpha_0 = 50^\circ$ versus time for a range of electron energies. Here $t = 0$ corresponds to 02:30UT 8 October 2012 (the onset of the low-energy electron injection). The bottom panel shows that the low-energy electron injection is energy dependent, and high-energy ($> 1.5$ MeV) electron fluxes decrease during the low-energy electron injection.
Appendix D

REM Simulations of the 8 October 2012 Storm Event Using Observed Data

The REM simulations presented in Chapter 5 were motivated by the 8 October 2012 radiation belt electron event, but they were designed primarily to assess the effect of drift shell splitting on PSD changes, rather than performing an event simulation. In this appendix, we present a preliminary attempt to model the 8 October 2012 event more closely and compare to measurements.

The simulation setups are similar to those described in Section 5.2, except that here: (i) initial and boundary conditions are specified using Van Allen Probes MagEIS and REPT measurements of electron PSD [Baker et al., 2012], starting from 23:00UT 8 October, and ending at 16:00UT 9 October 2012; (ii) POES factors are the time-averaged profile (red line in Figure 5.2) instead of the model profile; and (iii) a constant $K_p = 6$ is adopted for the O’Brien [2015] and Ozeke et al. [2014] diffusion coefficients, which is closer to the observed $K_p$ values during the storm. The simulations are carried out with drift shell splitting diffusion components turned off and on. The simulation results are summarized in the following.

Figure D.1 gives a comparison of the observed PSD radial profiles and the REM simulation results at $M = 1279$ MeV/G and $K = 0.115$ G$^{1/2}$R$_E$. The simulated PSD profile with drift shell splitting turned on is consistently lower than the profile with that turned off. The “on” profile is closer to the observed PSD, both in the magnitude and the position of the PSD peak, whereas the “off” profile is higher and peaks at a
Figure D.1: (Top) Van Allen Probes observed PSD temporal variations at $M = 1279$ MeV/G, $K = 0.115 G^{1/2}_{RE}$, from 06:00UT 8 October to 16:00UT 9 October 2012, and (bottom) REM simulated PSD at the same $M$ and $K$, starting from 23:00UT 8 October 2012 (black curve, approximately corresponding to curve 5 in the top panel) and ending at 16:00UT 9 October 2012, with drift shell splitting turned off (blue) and on (red). Note that the axis ranges are different in these panels. Top panel is reprinted from Tu et al. [2014].
larger $L$.

The PSD source plots for both the “off” and “on” solutions at $L = 4.25$ are shown in Figure D.2. Apparently, less PSD is contributed from the low-energy boundary

![Figure D.2](image)

Figure D.2: PSD source plots, in the same format as Figure 5.4, for the $t = 17$ hour solutions at $M = 1279$ MeV/G, $K = 0.115 \, G^{1/2} R_E$ and $L = 4.25$. Drift shell splitting effects are turned off in (a)–(c), and on in (d)–(f).

in the “on” case (d–f); this is consistent with the theoretical arguments of Section 5.3. Note that in panels (a) and (d), the boundary of Shabansky orbits is as close as $L \approx 5$ near generalized $\alpha_0 \approx 90^\circ$, as a result of enhanced day-night asymmetry in T89 at $Kp = 6$, and a cusp exists along this boundary. To the smaller $\alpha_0$ side of the cusp, the boundary is not caused by Shabansky orbits but by open drift shells that
intersect the magnetopause.

Figure D.3 gives a comparison of the observed PSD radial profiles and the REM simulation results corresponding to Figure D.1 at $M = 3433$ MeV/G and $K = 0.115 \, G^{1/2}\text{R}_E$. Each of the REM solution points in this figure is not obtained from one single simulation, but is an average of five PSD solutions from one hour earlier, with half-hour increments, to one hour later than the nominal ending time. This is because the phase space point of the PSD solution is too far from the low-energy boundary, which is the major source of PSD contributions, and thus the intrinsic statistics of the SDE method become low. Again, we observe that the solutions with drift shell splitting on is consistently lower than the off solutions. The “on” solutions agree better with observed data in most $L$ ranges than the “off” solutions.

Figure D.4 shows the PSD source plots for the “off” and “on” solutions at $M = 3433$ MeV/G, $K = 0.115 \, G^{1/2}\text{R}_E$ and $L = 4.25$. We again see that in the “on” case, the low-energy boundary contributes less PSD to the solution point than that in the “off” case. Compared with the source plots for a lower $M$ in Figure D.2, the contributing pixels (phase space elements) in Figure D.4 are much more sparse. But the majority of the PSD sources are nonetheless from low energies, indicating the importance of the seed electron population to the radiation belt PSD enhancement even at the multi-MeV energies.
Figure D.3: (Top) Van Allen Probes observed PSD temporal variations at $M = 3433 \text{ MeV/G}$, $K = 0.115 \text{ G}^{1/2}R_E$, from 18:22UT 8 October to 9 October 2012, and (bottom) REM simulated PSD at the same $M$ and $K$, starting from 23:00UT 8 October 2012 (black curve, approximately corresponding to the cyan curve in the top panel) and ending at 16:00UT 9 October 2012 (corresponding to “Late Times” in the top panel), with drift shell splitting turned off (blue) and on (red). Note that the axis ranges are different in these panels. The top panel is reprinted from Reeves et al. [2013].
Figure D.4: PSD source plots, in the same format as Figure 5.4, for the $t = 17$ hour solutions at $M = 3433$ MeV/G, $K = 0.115 \ G^{\frac{1}{2}}R_E$ and $L = 4.25$, as in Figure D.3.
Bibliography

Abel, B., and R. M. Thorne, Electron scattering loss in Earth’s inner magnetosphere:  


Albert, J. M., N. P. Meredith, and R. B. Horne, Three-dimensional diffusion simula-
tion of outer radiation belt electrons during the 9 October 1990 magnetic storm, 


Baker, D., et al., The Relativistic Electron-Proton Telescope (REPT) instrument on 
board the Radiation Belt Storm Probes (RBSP) spacecraft: Characterization of 
earth’s radiation belt high-energy particle populations, *Space Sci. Rev.*, 179(1-4), 

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*(35), 401, 
1994.

Baker, D. N., et al., Recurrent geomagnetic storms and relativistic electron enhance-
ments in the outer magnetosphere: ISTP coordinated measurements, *J. Geophys. 

Baker, D. N., et al., A long-lived relativistic electron storage ring embedded in Earth’s 
outer Van Allen belt, *Science, 340*(6129), 186–190, doi:10.1126/science.1233518, 
2013.

Baker, D. N., et al., Gradual diffusion and punctuated phase space density enhance-


Chen, Y., G. D. Reeves, R. H. W. Friedel, and G. S. Cunningham, Global time-dependent chorus maps from low-Earth-orbit electron precipitation and Van Allen


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


Varotsou, A., D. Boscher, S. Bourdarie, R. B. Horne, N. P. Meredith, S. A. Glauert,
and R. H. Friedel, Three-dimensional test simulations of the outer radiation belt

Vernov, S., A. Chudakov, E. Gorchakov, J. Logachev, and P. Vakulov, Study of the


