

**A STUDY OF
THE GUIDING CENTER
APPROXIMATION**

A Dissertation
Presented to
The Faculty of the Department of Physics
The College of William and Mary in Virginia

In Partial Fulfillment
Of the Requirements for the Degree of
Doctor of Philosophy

by
Qun Yao
January 1993

APPROVAL SHEET

This dissertation is submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy



Qun Yao

Approved, January 1993



Allen H. Boozer



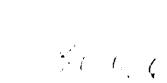
Franz L. Gross



Chi-Kwong Li



Eugene R. Tracy



George M. Vahala

To my wife Jing

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	viii
LIST OF FIGURES	x
ABSTRACT	xi
CHAPTER 1. INTRODUCTION	2
Appendix 1-A	8
References	10
CHAPTER 2. COORDINATE SYSTEMS	11
2-1 General Coordinates	11
2-2 Hamiltonian Representation of the Magnetic Field	17
2-3 Magnetic Coordinates	23
References	29
CHAPTER 3. HAMILTONIAN MECHANICS	30
3-1 Hamilton Equations of Motion	30
3-2 Canonical Transformation	32
3-3 Hamiltonian in Noncanonical Coordinates	37
3-4 The Canonical Perturbation Theory	41
3-5 Lie Transformation	51

3-6 Adiabatic Invariance	55
3-7 Motion Near the Resonance	61
3-8 Numerical Example	67
References	70
CHAPTER 4. THE HIGHER ORDER DRIFT HAMILTONIAN	
IN A CURL FREE FIELD	72
4-1 Introduction	72
4-2 The Magnetic Coordinates and the Drift Hamiltonian	77
4-3 The Exact Hamiltonian in the Guiding Center Coordinates	82
4-4 Higher Order Drift Hamiltonian	85
4-5 Conclusions	92
Appendix 4-A Second Order Hamiltonians	96
Appendix 4-B Derivation of Equation (4-4.4)	98
Appendix 4-C Proof of the First order Coordination to the Standard Drift	
Hamiltonian	100
Appendix 4-D The Hamiltonian Near the Magnetic Axis	103
References	107
CHAPTER 5. THE HIGHER ORDER DRIFT HAMILTONIAN	
IN A FULL FIELD	108
5-1 The Exact Hamiltonian in Boozer Coordinates	108
5-2 Higher Order Drift Hamiltonian	114

5-3 Conclusions	128
Appendix 5-A The Poisson Brackets for Darboux Transformation in Sec. 5-2	135
Appendix 5-B Proof of Our Definition for the Guiding Center	139
References	141
CHAPTER 6. A NUMERICAL STUDY OF THE DRIFT APPROXIMATION . . .	142
6-1 The Hamiltonians	142
6-2 The Numerical Results	148
6-3 Conclusions	154
References	158
CHAPTER 7 SUMMARY	159
BIBLIOGRAPHY	162

ACKNOWLEDGEMENTS

First, I would like to thank all of the members of the William and Mary physics department. The professors, staff and fellow graduate students have created such a wonderful working environment. Everyone has been friendly and helpful. During the last four years, I have felt as if I have been living in a big family. I shall never forget the support from the department when I and many other Chinese students were experiencing the painful tragedy of Tiananmen Square in June of 1989.

I would like to express my gratitude to Professor Allen H. Boozer, my thesis advisor, for his guidance and thoughtfulness. I consider myself very fortunate to have been his student. He has always been friendly and concerned about the welfare of his students. His sharpness and ability to immediately see to the core of a problem has been a constant inspiration to me. The knowledge I have gained from him goes far beyond the discipline of physics.

I would also like thank Professor Franz L. Gross, Professor Chi-Kwong Li, Professor Eugene R. Tracy and Professor George M. Vahala for serving on my thesis committee and Professor John B. Delos for his warm hearted help. The puzzle games at the Delos's parties were always challenging and fun.

I would like to thank Alastair Neil and Rick Weening for helping me with the computer in the early stages of this project. I would also like to thank Allan Greer, John Shaw, and A. B. Wakely for helping me with my English while writing of this dissertation.

To the members of the volleyball team, especially Douglas Baker, Carl Carlson, Chree Haas and Maciek Sasinowski, "thanks", it was a fun experience.

Dr. David Anderson and his colleagues at the Torsatron/Stellarator Laboratory of the University of Wisconsin, Madison, were gracious hosts and provided a stimulating environment during my visit in the summer of 1991.

During the summer of 1992 and the following frequent visits at the MIT Plasma Fusion Center, I had many stimulating discussions, especially with Dr. Chi-Tien Hsu. My special thanks to Dr. Dieter J. Sigmar, without whom my visit at PFC would have been impossible. During my visits, Dieter provided all the help I needed and made my stay at PFC an enjoyable experience.

I am deeply indebted to my mother. She unselfishly devoted her life to her children. The love she has given me I can never repay. I wish my father could have lived to see this day. Dad, I miss you very much.

None of this would have been accomplished without my beloved wife and colleague, Jing Gao. Her love, support, understanding and her belief in me are constant encouragement. I owe thanks to her most of all.

This work has been supported by the U. S. Department of Energy under Grant No. DE-FG05-84ER53176.

LIST OF FIGURES

Figure	Page
2.1 The Cartesian coordinates and the spherical coordinates	12
2.2 Toroidal coordinates	21
2.3. The magnetic coordinates	25
3.1 Motion of a two dimensional integrable Hamiltonian system in phase space	45
3.2 Surface of section for the field line Hamiltonian	69
6.1 The magnetic surface of a quasihelically symmetric stellarator	147
6.2 Orbits of the standard drift Hamiltonian in a quasihelically symmetric field	149
6.3 Trajectory of the high order drift Hamiltonian	151
6.4 Locations of close orbits	153
6.5 The change of μ_{sp} in time	155
6.6 Trajectory of the exact Hamiltonian	156

ABSTRACT

A study of the relationship between the trajectories of the exact and drift Hamiltonians for charged particles in toroidal magnetic fields is presented.

In a magnetic field, a charged particle moves very rapidly on a nearly circular orbit about a magnetic field line. The guiding center moves along the field line with a parallel velocity, and drifts slowly from one field line to another with a drift velocity. The separated time scales between the fast circular motion and the slow drift motion implies the existence of an adiabatic invariant, which effectively reduces the degrees of freedom from three to two. The guiding center motion can be found by the drift equations first derived by Alfvén using a Taylor expansion. The validity of this derivation only holds for a time $1/(\epsilon\omega)$ with ω the cyclotron frequency and $\epsilon \sim \rho/a$ the ratio of the gyroradius to plasma size. But in fusion plasmas the shortest time of interest for following particles is $1/(\epsilon^2\omega)$, and for α -particles one must follow their trajectories for times as long as $1/(\epsilon^4\omega)$.

Giving the exact Hamiltonian in magnetic coordinates, we perform canonical transformations and Taylor expansions. We show the exact Hamiltonian

$$H_e(\theta_g, \mu; \theta, p_\theta; \varphi, p_\varphi) = H_d(\mu; \theta, p_\theta; \varphi, p_\varphi) + \epsilon H_I(\mu; \theta, p_\theta; \varphi, p_\varphi) + o(\epsilon^2).$$

with H_d the standard drift Hamiltonian and H_I the first order correction of the standard drift Hamiltonian. H_e depends on both $|B|$ and the metric tensor of the magnetic coordinates. H_d depends only on $|B|$. H_I depends on the metric tensor of the magnetic coordinates.

A hypothetical field that has exact quasi-helical symmetry is used to simplify the comparison between H_d , $K_d = H_d + \epsilon H_I$ and H_e . In quasi-helical symmetry the field strength has a symmetry but the magnetic surfaces do not, and

$$H_e(\theta_g, \mu; \theta, p_\theta; \varphi, p_\varphi) = H_d(\mu; \theta, p_\theta; p_\varphi) + \epsilon H_I(\mu; \theta, p_\theta; \varphi, p_\varphi) + o(\epsilon^2).$$

The standard drift Hamiltonian $H_d(\mu; \theta, p_\theta; p_\varphi)$ has three invariants: H_d , μ , and p_φ . The system is completely integrable and the drift trajectories are perfectly confined. The exact Hamiltonian has only one invariant, the energy H_e , which is insufficient to insure the confinement of the exact trajectories. The drift Hamiltonian with first order correction has two invariants: K_d and μ , and can be treated as a two dimensional integrable system with a small perturbation which is proportional to the ratio of the gyroradius to system size. The numerical results show that the difference in phase space structure between the exact and standard drift Hamiltonians can be predicted by the drift Hamiltonian with the first order correction.

**A STUDY OF
THE GUIDING CENTER
APPROXIMATION**

CHAPTER 1

INTRODUCTION

The subject of this dissertation is the motion of charged particles in a toroidal electromagnetic field. The motion of the particle is governed by the equation

$$m \frac{d\mathbf{v}}{dt} = e(\mathbf{v} \times \mathbf{B} + \mathbf{E}), \quad (1-1)$$

with m the mass and e the charge of the particle. Although this equation appears simple, it is a very difficult task even for super computers such the Cray to integrate particle trajectories over a time scale of interest. Fortunately in most cases of interest there exists an adiabatic invariant. This invariant arises from the rapidity of the circular motion of a charged particle about a magnetic field line in comparison to the slow drift motion of a particle from one field line to another. The theory of adiabatic invariants implies that the exact trajectory of the particle can be approximated by following the center of the circle about which the particle gyrates. The lowest order equations of motion for the guiding center were given by Alfvén in 1940, and are known to plasma physicists as the drift equations.¹ Today almost all the trajectory calculations in fusion plasmas are based on this approximation. The Alfvén drift equations, which were derived using the lowest order terms of a Taylor expansion, are only guaranteed to be valid for a time which is short compared to $(\epsilon\omega)^{-1}$, with $\omega = eB/m$ (the gyrofrequency) and $\epsilon \sim \rho/a$ (the ratio of the gyroradius to the scale length on which the field changes). In fusion plasma the time that one wishes to follow particle orbits can be much longer than the time over which Alfvén's derivation of the drift equations is valid. Despite the obvious importance, there has been little study of the accuracy with which the drift trajectories approximate the exact particle

trajectories.

To have a general ideal about the motion of a charged particle in a given electromagnetic field, we start with a simple case of a uniform time-independent electromagnetic field. In Cartesian coordinates (x, y, z) we let

$$\mathbf{B} = B\hat{z}, \quad (1-2a)$$

$$\mathbf{E} = E_{||}\hat{z} + E_{\perp}\hat{y} \quad (1-2b)$$

By solving the equation of motion of Eq. (1-1), we obtain

$$\mathbf{v} = \left(v_{||0} + \frac{e}{m} E_{||} t \right) \hat{z} + \frac{E_{\perp}}{B} \hat{x} + v_{\perp} (\hat{x} \cos \omega t - \hat{y} \sin \omega t), \quad (1-3a)$$

$$\mathbf{x} = \hat{z} \left(v_{||0} t + \frac{1}{2} \frac{e}{m} E_{||} t^2 \right) + \hat{x} \frac{E_{\perp}}{B} t + \rho (\hat{x} \sin \omega t + \hat{y} \cos \omega t), \quad (1-3b)$$

with the initial conditions $\mathbf{x}(t = 0) = \rho \hat{y}$ and $\mathbf{v}(t = 0) = v_{||0}\hat{z} + v_{\perp}\hat{x}$. In Eq. (1-3) the gyrofrequency or cyclotron frequency is

$$\omega = \frac{eB}{m}, \quad (1-4)$$

and the gyroradius or Larmor radius is

$$\rho = \frac{v_{\perp}}{\omega} = \frac{mv_{\perp}}{|e|B}. \quad (1-5)$$

The motion contains two parts: the circular motion about the magnetic field line and the guiding center (the center of the circular motion) motion. The guiding center moves along the field line at the parallel velocity and slowly drifts cross field lines with a drift velocity,

$$\mathbf{v}_d = \frac{\mathbf{E} \times \mathbf{B}}{B^2}. \quad (1-6)$$

The position of the guiding center is

$$X = x + \frac{mv \times B}{eB^2}, \quad (1-7)$$

with x the position of the particle and v its velocity. For most cases of interest the gyroradius is small and the distinction between the location of the particle and its guiding center is irrelevant. On the other hand, to integrate along an exact trajectory is much more difficult than along its guiding center.

The parallel motion is accelerated by the force along the magnetic field line,

$$m \frac{dv_{||}}{dt} = F_{||} = \frac{l}{B} \mathbf{F} \cdot \mathbf{B}. \quad (1-8)$$

The force perpendicular to the field line causes the guiding center to drift across the field. The direction of this drift is not in the direction of the force, but rather in the direction that is perpendicular to both the force and the field line,

$$\mathbf{v}_d = \frac{\mathbf{F} \times \mathbf{B}}{eB^2}. \quad (1-9)$$

Now we consider a nonuniform magnetic field which satisfies

$$\epsilon \sim \frac{\rho}{B} \nabla B \ll 1. \quad (1-10)$$

Eq. (1-10) means that the scale length on which the field changes is much larger than the Larmor radius. By doing a Taylor expansion over the parameter ϵ , we can find that the parallel force due to the variation of the field can be approximated by (see Appendix 1-A)

$$F_{||} \approx -\frac{l}{2} e \rho v_{\perp} \nabla_{||} B. \quad (1-11)$$

We define the magnetic moment of the gyrating particle to be

$$\mu = \frac{mv_{\perp}^2}{2B}. \quad (1-12)$$

Putting Eqs. (1-11) and (1-12) into Eq. (1-8), we obtain

$$m \frac{dv_{\parallel}}{dt} = -\mu \nabla_{\parallel} B. \quad (1-13)$$

We now write the parallel velocity as $v_{\parallel} = ds/dt$ and multiply both sides of Eq. (1-13) by v_{\parallel} , we have

$$\frac{d}{dt} \left(\frac{1}{2} mv_{\parallel}^2 \right) = -\mu \frac{\partial B}{\partial s} \frac{ds}{dt} = -\mu \frac{dB}{dt}. \quad (1-14)$$

The particle energy is conserved, i.e.

$$\frac{d}{dt} \left(\frac{1}{2} mv_{\parallel}^2 + \frac{1}{2} mv_{\perp}^2 \right) = \frac{d}{dt} \left(\frac{1}{2} mv_{\parallel}^2 + \mu B \right) = 0. \quad (1-15)$$

Combining Eqs. (1-14) and (1-15), we have

$$\frac{d}{dt} \mu = 0. \quad (1-16)$$

Eq. (1-16) is one of the most important conservation laws of plasma physics. It should be pointed out that the magnetic moment μ is exactly conserved only in a uniform field. Generally the magnetic moment μ is an adiabatic invariant. Adiabaticity is an important concept in Hamiltonian mechanics, which we will discuss in more detail in Chapter 3.

When the field is nonuniform, the charged particle can also "feel" forces that are

perpendicular to the magnetic field (see Appendix 1-A). These forces cause the guiding center to drift across the field according to Eq. (1-9). The drift velocity, to the lowest order in ϵ , is

$$\mathbf{v}_d = \frac{\mathbf{E} \times \mathbf{B}}{B^2} + \frac{v_\perp^2}{2\omega} \frac{\mathbf{B} \times \nabla B}{B^2} + \frac{mv_\parallel^2}{e} \frac{\mathbf{B} \times \boldsymbol{\kappa}}{B^2}, \quad (1-17)$$

with $\boldsymbol{\kappa} = (\hat{\mathbf{b}} \cdot \nabla) \hat{\mathbf{b}}$, ($\hat{\mathbf{b}} = \mathbf{B}/B$), the curvature of the field line. Eq. (1-17) was given by Alfvén, and is called the drift equation. The first term in Eq. (1-17) is called the electric field drift, which is independent of mass and charge. The second term is called grad-B drift and the third term is called curvature drift. The last two terms depend on both mass and charge.

To complete the equations of motion for the guiding center, we have

$$\mathbf{v}_g = v_\parallel \hat{\mathbf{b}} + \mathbf{v}_d, \quad (1-18)$$

$$m \frac{dv_\parallel}{dt} = -\mu \nabla_\parallel B - e \nabla_\parallel \phi, \quad (1-19)$$

with ϕ the electric potential.

The equations of (1-18) and (1-19) have a defect of not satisfying Liouville's theorem. A more systematic way is to derive the equations of motion from the drift Hamiltonian ²

$$H_d = \frac{1}{2} mv_\parallel^2 + \mu B + e\phi. \quad (1-20)$$

The biggest advantage of the Hamiltonian approach is the conservation of phase space, which is very important in the numerical study of the long term behavior for a chaotic system.

In Chapters 2 and 3 a background knowledge of magnetic coordinates and

Hamiltonian mechanics is provided. Chapters 4, 5 and 6 each contain a research topic. In chapter 4, the first order correction to the standard drift Hamiltonian is derived in a vacuum magnetic field. This formulation is extended into a time-independent full electromagnetic field with plasma equilibrium in chapter 5. In chapter 6 the numerical study of the effect of the first order correction is presented. Finally, in chapter 7, a summary is given. Applications and directions for further research are discussed.

Appendix 1-A: Forces Due to the Change of Magnetic Field

In this appendix we give a simple derivation of the forces due to the change of magnetic field, which follows from Chen.³

We first consider the case that ∇B is perpendicular to B . We assume the magnetic field is

$$\mathbf{B} = B(y)\hat{z}. \quad (1\text{-A.1})$$

We Taylor expand the field at $y = 0$,

$$\mathbf{B} \equiv \hat{z}\left(B_0 + y \frac{\partial B}{\partial y}\right). \quad (1\text{-A.2})$$

The force due to the magnetic field is

$$\mathbf{F} = e\mathbf{v} \times \mathbf{B} \equiv ev_{\perp}(\hat{x}\cos\omega t - \hat{y}\sin\omega t) \times \hat{z}\left(B_0 + y \frac{\partial B}{\partial y}\right). \quad (1\text{-A.3})$$

Inserting $y = \rho \cos \omega$ into Eq. (1-A.3) and averaging over the gyromotion, we obtain

$$\langle \mathbf{F} \rangle \equiv -\frac{1}{2}e\rho v_{\perp} \frac{\partial B}{\partial y} \hat{y} = -\frac{1}{2} \frac{ev_{\perp}^2}{\omega} \nabla B. \quad (1\text{-A.4})$$

Now we consider the case that there is a gradient in B along the field. We assume that the field is symmetric about the z-axis and its Taylor expansion is

$$\mathbf{B} = \left(B + z \frac{\partial B_z}{\partial z}\right)\hat{z} + x \frac{\partial B_x}{\partial x} \hat{x} + y \frac{\partial B_y}{\partial y} \hat{y}. \quad (1\text{-A.5})$$

The force is

$$\begin{aligned} \mathbf{F} = e\mathbf{v} \times \mathbf{B} &\equiv \hat{x} \left[-ev_{\perp} \sin \omega t \left(B + z \frac{\partial B_z}{\partial z} \right) - ev_{\parallel} y \frac{\partial B_y}{\partial y} \right] \\ &+ \hat{y} \left[ev_{\parallel} x \frac{\partial B_x}{\partial x} - ev_{\perp} \cos \omega t \left(B + z \frac{\partial B_z}{\partial z} \right) \right] \\ &+ \hat{z} \left[ev_{\perp} \cos \omega t y \frac{\partial B_y}{\partial y} + ev_{\perp} \sin \omega t x \frac{\partial B_x}{\partial x} \right]. \end{aligned} \quad (1-A.6)$$

Because $\nabla \cdot \mathbf{B} = 0$, we also have

$$\frac{\partial B_z}{\partial z} = -\frac{\partial B_x}{\partial x} - \frac{\partial B_y}{\partial y}. \quad (1-A.7)$$

Combining Eqs. (1-A.6), (1-3b) and (1-A.7) and averaging over gyromotion, we obtain

$$F_{\parallel} = \langle F \rangle \equiv -\frac{1}{2} e\rho v_{\perp} \frac{\partial B_z}{\partial z} = -\frac{1}{2} e\rho v_{\perp} \nabla_{\parallel} B. \quad (1-A.8)$$

When the field is bended, the particle also "feels" a centrifugal force,

$$\mathbf{F}_{curl} = -mv_{\parallel}^2 \boldsymbol{\kappa}, \quad (1-A.9)$$

with $\boldsymbol{\kappa}$ the curvature of the field line.

References

- 1 H. Alfvén, Ark. Mat. Astron. Fys. **27A**, 1 (1940)
- 2 J. B. Taylor, Phys. Fluids **7**, 767 (1964)
- 3 F. F. Chen, *Introduction to Plasma Physics* (Plenum Press, New York, 1974)

CHAPTER 2

COORDINATE SYSTEMS

In this chapter, some knowledge of coordinate systems, which follows most closely to that of Boozer,¹ is reviewed. The confinement time of a fusion plasma must be much longer than a collision time, so fusion plasma must be close to a Maxwellian. Magnetic field lines can confine a near-Maxwellian plasma only if they lie in a closed surface. Topology implies that the only shape possible for such a surface is toroidal. A good coordinate system should represent this toroidal configuration in a simple manner. The so called magnetic flux coordinates satisfy such a requirement. In the flux coordinates, one variable gives the surface and the other two give the location on this surface. In Sec. 2-1 we review the properties of a general coordinate system. In Sec. 2-2 we discuss the Hamiltonian description of magnetic field. Finally in Sec. 2-3, we present the magnetic coordinates and a special magnetic coordinate system, Boozer coordinates.

2-1 General Coordinates

In this section we review the non-orthogonal curvilinear coordinate system in the three-dimensional Euclidean space.

An arbitrary point \mathbf{x} in an Euclidean three-dimensional space can be given by its Cartesian coordinates $\mathbf{x} = (x^1, x^2, x^3)$, i.e.

$$\mathbf{x} = x^1 \mathbf{e}_1 + x^2 \mathbf{e}_2 + x^3 \mathbf{e}_3, \quad (2-1.1)$$

with e_1 , e_2 and e_3 three orthogonal unit vectors obeying the right hand rule. Let (ξ^1, ξ^2, ξ^3) be three arbitrary coordinates. We assume the two coordinate systems are related by three well-behaved functions,

$$x^i = x^i(\xi^1, \xi^2, \xi^3) \quad (i = 1, 2, 3). \quad (2-1.2)$$

The three functions can be inverted to give the ξ^i in terms of the x^i ,

$$\xi^i = \xi^i(x^1, x^2, x^3) \quad (i = 1, 2, 3). \quad (2-1.3)$$

A simple example is the relation between the Cartesian coordinates (x, y, z) and the spherical coordinates (r, θ, φ) ,

$$x = r \sin \theta \cos \varphi, \quad y = r \sin \theta \sin \varphi \quad \text{and} \quad z = r \cos \theta,$$

or

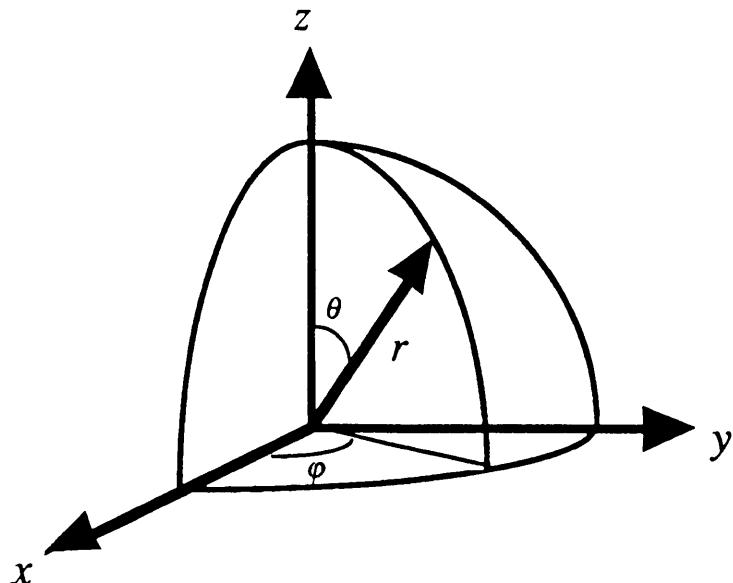


Figure 2.1 The Cartesian coordinates (x, y, z) and the spherical coordinates (r, θ, φ) .

$$r = \sqrt{x^2 + y^2 + z^2}, \quad \theta = \cos^{-1}\left(z/\sqrt{x^2 + y^2 + z^2}\right) \text{ and } \varphi = \cos^{-1}\left(x/\sqrt{x^2 + y^2}\right).$$

One must bear in mind that although Eqs. (2-1.2) and (2-1.3) in principle exist for well-behaved coordinate systems, it is generally very difficult to construct one from the other one by inversion. Even if one can do so, the inverse functions are often too complicated to be used in actual calculations. Hence we assume only Eq. (2-1.2), $x^i = x^i(\xi^1, \xi^2, \xi^3)$, is known.

Associated with a set of arbitrary coordinates (ξ^1, ξ^2, ξ^3) , there are two kinds of basis vectors, $\partial \mathbf{x} / \partial \xi^i$ and $\nabla \xi^i$ for $i = 1, 2, 3$. As long as the vectors $\partial \mathbf{x} / \partial \xi^1$, $\partial \mathbf{x} / \partial \xi^2$ and $\partial \mathbf{x} / \partial \xi^3$ (or $\nabla \xi^1$, $\nabla \xi^2$ and $\nabla \xi^3$) do not lie in one plane (same condition for Eqs. (2-1.2) and (2-1.3) existing), i.e.

$$\left(\frac{\partial \mathbf{x}}{\partial \xi^1} \times \frac{\partial \mathbf{x}}{\partial \xi^2} \right) \cdot \frac{\partial \mathbf{x}}{\partial \xi^3} \neq 0 \quad \text{or} \quad (\nabla \xi^1 \times \nabla \xi^2) \cdot \nabla \xi^3 \neq 0,$$

any vector can be expressed in these two basis. These two sets of basis vectors are related by the orthogonality relations and dual relations of general coordinates. The orthogonality relations hold for any coordinates (ξ^1, ξ^2, ξ^3) , and are

$$\frac{\partial \mathbf{x}}{\partial \xi^i} \cdot \nabla \xi^j = \delta_i^j. \tag{2-1.4}$$

The dual relations are

$$\nabla \xi^i = \frac{1}{2J} \epsilon^{ijk} \frac{\partial \mathbf{x}}{\partial \xi^j} \times \frac{\partial \mathbf{x}}{\partial \xi^k}, \tag{2-1.5}$$

and

$$\frac{\partial \mathbf{x}}{\partial \xi^i} = \frac{J}{2} \epsilon_{ijk} \nabla \xi^j \times \nabla \xi^k. \quad (2-1.6)$$

The convention of summing over the same index has been used here and will be used in the remainder of the dissertation. In Eqs. (2-1.5) and (2-1.6), ϵ^{ijk} and ϵ_{ijk} are the fully antisymmetric tensors, J is the Jacobian

$$J = \left(\frac{\partial \mathbf{x}}{\partial \xi^1} \times \frac{\partial \mathbf{x}}{\partial \xi^2} \right) \cdot \frac{\partial \mathbf{x}}{\partial \xi^3} = \frac{1}{(\nabla \xi^1 \times \nabla \xi^2) \cdot \nabla \xi^3}. \quad (2-1.7)$$

The dual relations hold only in those kinds of space which have a Jacobian. Given Eq. (2-1.2) the basis vectors $\partial \mathbf{x} / \partial \xi^i$ can be very easily calculated. To find the basis vectors $\nabla \xi^i$, one should use the dual relation, Eq. (2-1.5), instead of trying to invert Eq. (2-1.2).

A vector \mathbf{A} has a contravariant representation

$$\mathbf{A} = A^i (\xi^1, \xi^2, \xi^3) \frac{\partial \mathbf{x}}{\partial \xi^i}, \quad (2-1.8)$$

with

$$A^i = \mathbf{A} \cdot \nabla \xi^i, \quad (2-1.9)$$

and it also has a covariant representation

$$\mathbf{A} = A_i (\xi^1, \xi^2, \xi^3) \nabla \xi^i, \quad (2-1.10)$$

with

$$A_i = \mathbf{A} \cdot \frac{\partial \mathbf{x}}{\partial \xi^i}. \quad (2-1.11)$$

The components of a vector in two different representations are linked with the metric tensor of the coordinates g^{ij} and g_{ij} ,

$$g^{ij} = \nabla \xi^i \cdot \nabla \xi^j \quad \text{and} \quad g_{ij} = \frac{\partial \mathbf{x}}{\partial \xi^i} \cdot \frac{\partial \mathbf{x}}{\partial \xi^j}, \quad (2-1.12)$$

$$A^i = g^{ij} A_j \quad \text{and} \quad A_i = g_{ij} A^j. \quad (2-1.13)$$

The two metric tensors are related by

$$g^{ij} g_{jk} = \delta_k^i. \quad (2-1.14)$$

It can also be shown that the square of the Jacobian is the determinant of the metric tensor,

$$J^2 = \|g^{ij}\|^{-1} = \|g_{ij}\|. \quad (2-1.15)$$

Now we discuss the calculus in the general coordinates (ξ^1, ξ^2, ξ^3) . The gradient of a scalar function $f(\xi^1, \xi^2, \xi^3)$ is very simple,

$$\nabla f = \frac{\partial f}{\partial \xi^i} \nabla \xi^i, \quad (2-1.16)$$

which is a vector in the covariant representation. To calculate a curl of a vector function, one should use the covariant representation. Let the vector be $A(\xi^1, \xi^2, \xi^3) = A_i \nabla \xi^i$, then

$$\nabla \times A = \nabla A_j \times \nabla \xi^j = \frac{\partial A_j}{\partial \xi^i} \nabla \xi^i \times \nabla \xi^j,$$

where the identities $\nabla \times (sV) = \nabla s \times V + s \nabla \times V$ and $\nabla \times (\nabla s) = 0$ are used for arbitrary vector V and scalar s . Therefore the curl of a vector is in the contravariant representation,

$$\nabla \times A = \frac{\partial A_j}{\partial \xi^i} \nabla \xi^i \times \nabla \xi^j \quad \text{or} \quad \nabla \times A = \frac{\epsilon^{ijk}}{J} \frac{\partial A_k}{\partial \xi^j} \frac{\partial \mathbf{x}}{\partial \xi^i}. \quad (2-1.17)$$

To calculate the divergence of a vector function, however, one should use the contravariant representation. Let the vector be $\mathbf{A} = A^i (\partial \mathbf{x} / \partial \xi^i)$, we have

$$\nabla \cdot \mathbf{A} = \frac{\partial \mathbf{x}}{\partial \xi^i} \cdot \nabla A^i + A^i \nabla \cdot \left(\frac{\partial \mathbf{x}}{\partial \xi^i} \right).$$

The first term is trivial,

$$\frac{\partial \mathbf{x}}{\partial \xi^i} \cdot \nabla A^i = \frac{\partial \mathbf{x}}{\partial \xi^i} \cdot \nabla \xi^j \frac{\partial A^i}{\partial \xi^j} = \frac{\partial A^i}{\partial \xi^j} \delta_i^j = \frac{\partial A^i}{\partial \xi^i}.$$

To calculate the second term, we use Eq. (2-1.6) and

$$\nabla \cdot (\nabla \xi^j \times \nabla \xi^k) = \nabla \cdot [\nabla \times (\xi^j \nabla \xi^k)] = 0,$$

then

$$A^i \nabla \cdot \left(\frac{\partial \mathbf{x}}{\partial \xi^i} \right) = A^i \nabla \cdot \left(\frac{J}{2} \epsilon_{ijk} \nabla \xi^j \times \nabla \xi^k \right) = \frac{A^i}{J} \frac{\partial J}{\partial \xi^i}.$$

Thus we obtain the divergence of a vector function

$$\nabla \cdot \mathbf{A} = \frac{1}{J} \frac{\partial (JA^i)}{\partial \xi^i}. \quad (2-1.18)$$

The integrals in three-dimensional space include three types: line, area and volume. A line integral along a curve c is $\int_c \mathbf{A} \cdot d\mathbf{x}$. A curve can be described in the coordinates (ξ^1, ξ^2, ξ^3) by holding two of the three coordinates, say ξ^2 and ξ^3 , fixed. Therefore

to the trajectories of a time-dependent Hamiltonian with one degree of freedom.² Therefore, the structure of a magnetic field can be more conveniently studied by using a scalar (the field line Hamiltonian) rather than using a vector (the magnetic field \mathbf{B} or its vector potential \mathbf{A}).

The reason that the field line has a Hamiltonian is the divergence-free property of the magnetic field. Any divergence-free field has a so called canonical representation

$$\mathbf{B}(\psi, \theta, \varphi) = \nabla\psi \times \nabla\theta + \nabla\varphi \times \nabla\chi(\psi, \theta, \varphi). \quad (2-2.1)$$

In Eq. (2-2.1), (ψ, θ, φ) are some general non-orthogonal curvilinear coordinates in the three-dimensional Euclidean space. The function $\chi(\psi, \theta, \varphi)$ is the field line Hamiltonian, θ is the general coordinate, ψ is the conjugate momentum and φ is the time-like parameter. We first prove the existence of the canonical representation, and then show that the function $\chi(\psi, \theta, \varphi)$ is the field line Hamiltonian.

An arbitrary globally divergence-free field has a well behaved vector potential, i.e.

$$\forall \nabla \cdot \mathbf{B}(\mathbf{x}) = 0, \quad \exists \mathbf{A}(\mathbf{x}), \quad \text{s.t.} \quad \mathbf{B}(\mathbf{x}) = \nabla \times \mathbf{A}(\mathbf{x}).$$

To prove the existence of the canonical representation, we give the divergence-free field in arbitrary coordinates $\mathbf{B}(\xi^1, \xi^2, \xi^3)$, and its vector potential in the covariant representation

$$\mathbf{A}(\xi^1, \xi^2, \xi^3) = A_1(\xi^1, \xi^2, \xi^3)\nabla\xi^1 + A_2(\xi^1, \xi^2, \xi^3)\nabla\xi^2 + A_3(\xi^1, \xi^2, \xi^3)\nabla\xi^3. \quad (2-2.2)$$

A gauge can be freely chosen for the vector potential. If \mathbf{A} satisfies $\mathbf{B} = \nabla \times \mathbf{A}$, then $\mathbf{A}' = \mathbf{A} + \nabla G$ also does, with G a well behaved scalar function. We choose the function $G(\xi^1, \xi^2, \xi^3)$ that satisfies $\partial G / \partial \xi^1 = -A_1$. The new vector potential is

$$\mathbf{A}' = \left(A_2 + \frac{\partial G}{\partial \xi^2} \right) \nabla \xi^2 + \left(A_3 + \frac{\partial G}{\partial \xi^3} \right) \nabla \xi^3.$$

Letting $\theta = \xi^2$, $\varphi = \xi^3$, and choosing

$$\psi = A_2 + \partial G / \partial \xi^2, \quad (2-2.3a)$$

$$-\chi = A_3 + \partial G / \partial \xi^3, \quad (2-2.3b)$$

we have

$$\mathbf{A}' = \psi \nabla \theta - \chi \nabla \varphi. \quad (2-2.4)$$

We thus obtain the canonical representation for the divergence-free field of Eq. (2-2.1) by taking curl of Eq. (2-2.4).

Now we prove that the function $\chi(\psi, \theta, \varphi)$ is the field line Hamiltonian, i.e. we show that field lines in the spatial coordinates (ψ, θ, φ) satisfy Hamilton's equations of motion

$$\frac{d\psi}{d\varphi} = -\frac{\partial \chi}{\partial \theta} \quad \text{and} \quad \frac{d\theta}{d\varphi} = \frac{\partial \chi}{\partial \psi}. \quad (2-2.5)$$

A magnetic field line is defined to be a solution of the differential equation

$$\frac{dx}{d\lambda} = \mathbf{B}(x), \quad (2-2.6)$$

with \mathbf{B} the magnetic field and λ a parameter which describes positions along the field line (λ is the differential distance along the field line divided by the field strength B). For any well behaved function $f(x)$, we have

$$\frac{df}{d\lambda} = \frac{dx}{d\lambda} \cdot \nabla f = \mathbf{B} \cdot \nabla f. \quad (2-2.7)$$

We assume $\mathbf{B} \cdot \nabla \varphi$ is non-zero in the region of interest for simplicity. Thus

$$\frac{d\psi}{d\varphi} = \frac{d\psi/d\lambda}{d\varphi/d\lambda} = \frac{\mathbf{B} \cdot \nabla \psi}{\mathbf{B} \cdot \nabla \varphi}. \quad (2-2.8)$$

Using the canonical representation for \mathbf{B} of Eq. (2-2.1),

$$\mathbf{B} \cdot \nabla \psi = (\nabla \varphi \times \nabla \chi) \cdot \nabla \psi = (\nabla \varphi \times \nabla \theta) \cdot \nabla \psi = -\mathbf{B} \cdot \nabla \varphi \frac{\partial \chi}{\partial \theta}. \quad (2-2.9)$$

We hence prove that $d\psi/d\varphi = -\partial \chi/\partial \theta$. Similarly we have

$$\frac{d\theta}{d\varphi} = \frac{d\theta/d\lambda}{d\varphi/d\lambda} = \frac{\mathbf{B} \cdot \nabla \theta}{\mathbf{B} \cdot \nabla \varphi} = \frac{\partial \chi}{\partial \psi}.$$

It is important to be aware that any two of the four potentials ψ, θ, φ and χ can be arbitrarily chosen. For example, when we prove Eq. (2-2.1) we choose θ and φ arbitrarily and define ψ and χ according to Eqs. (2-2.3a) and (2-2.3b). This arbitrariness is exactly the freedom of choosing the gauge or the freedom of canonical transformation. Suppose two sets of potentials $(\psi, \theta, \varphi, \chi)$ and $(\bar{\psi}, \bar{\theta}, \bar{\varphi}, \bar{\chi})$ are related by a generating function $S(\bar{\psi}, \theta, \varphi, \bar{\chi})$,

$$\psi = \frac{\partial S}{\partial \theta}, \quad -\chi = \frac{\partial S}{\partial \varphi}, \quad \bar{\theta} = \frac{\partial S}{\partial \bar{\psi}} \quad \text{and} \quad \bar{\varphi} = -\frac{\partial S}{\partial \bar{\chi}}. \quad (2-2.10)$$

then if $(\psi, \theta, \varphi, \chi)$ give a canonical representation of the \mathbf{B} -field,

$$\mathbf{B} = \nabla \psi \times \nabla \theta + \nabla \varphi \times \nabla \chi,$$

the other set $(\bar{\psi}, \bar{\theta}, \bar{\varphi}, \bar{\chi})$ also does,

$$\mathbf{B} = \nabla \bar{\psi} \times \nabla \bar{\theta} + \nabla \bar{\varphi} \times \nabla \bar{\chi}.$$

Now we turn to the subject of how to obtain the field line Hamiltonian $\chi(\psi, \theta, \varphi)$ and the transformation equation $x(\psi, \theta, \varphi)$ for a given magnetic field $B(x)$.^{3, 4} Let us start with the toroidal coordinates (r, θ, φ) (see Fig. 2.2), i.e. in the cylindrical coordinates

$$x(r, \theta, \varphi) = R(r, \theta)\hat{R}(\varphi) + Z(r, \theta)\hat{Z}, \quad (2-2.11)$$

with

$$R = R_0 + r \cos(2\pi\theta) \quad \text{and} \quad Z = -r \sin(2\pi\theta). \quad (2-2.12)$$

With Eq. (2-2.1), we have

$$\mathbf{B} \cdot \nabla \varphi = (\nabla \psi \times \nabla \theta) \cdot \nabla \varphi = (\nabla r \times \nabla \theta) \cdot \nabla \varphi \frac{\partial \psi}{\partial r}. \quad (2-2.13)$$

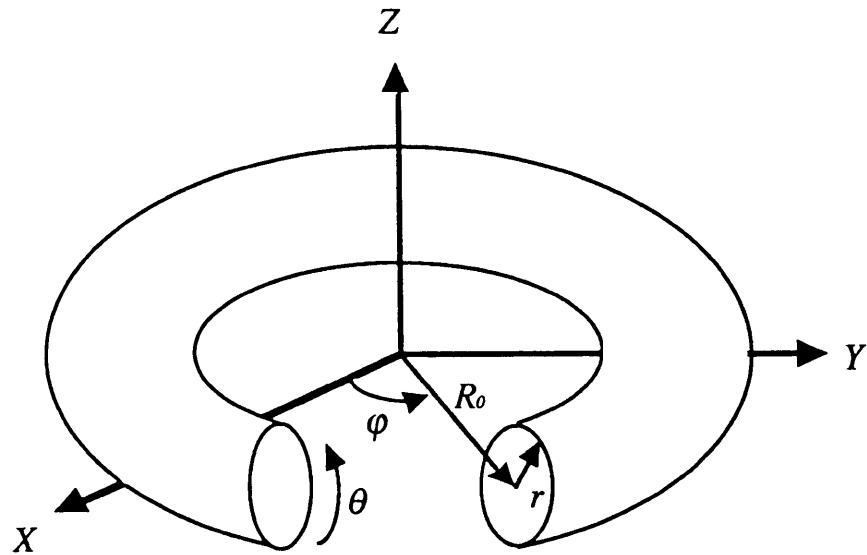


Figure 2.2 Toroidal coordinates. The constant R_0 is the major radius. The minor radius is r . The poloidal angle is θ and the toroidal angle is φ . The period of θ and φ are of unity instead of the conventional 2π .

From the dual relation of Eq. (2-1.5) and the Jacobian Eq. (2-1.7)

$$\nabla \varphi = [(\nabla r \times \nabla \theta) \cdot \nabla \varphi] \left(\frac{\partial \mathbf{x}}{\partial r} \times \frac{\partial \mathbf{x}}{\partial \theta} \right). \quad (2-2.14)$$

Thus we obtain the differential equation

$$\frac{\partial r}{\partial \psi} = \left[\mathbf{B} \cdot \left(\frac{\partial \mathbf{x}}{\partial r} \times \frac{\partial \mathbf{x}}{\partial \theta} \right) \right]^{-1}. \quad (2-2.15)$$

The right-hand side of Eq. (2-2.15) is a known function of (r, θ, φ) . Similarly the differential equation for the field line Hamiltonian can be obtained,

$$\frac{\partial \chi}{\partial \psi} = \frac{\mathbf{B} \cdot [(\partial \mathbf{x}/\partial \varphi) \times (\partial \mathbf{x}/\partial r)]}{\mathbf{B} \cdot [(\partial \mathbf{x}/\partial r) \times (\partial \mathbf{x}/\partial \theta)]}, \quad (2-2.16)$$

with the right-hand side a known function of (r, θ, φ) . Integrating Eqs. (2-2.15) and (2-2.16) at some fixed angle (θ_0, φ_0) with the boundary conditions $\psi = 0$ and $\chi = \text{constant}$ at $r = 0$, we can find numerically

$$\mathbf{x}(\psi, \theta_0, \varphi_0) = \mathbf{x}[r(\psi, \theta_0, \varphi_0), \theta_0, \varphi_0], \quad (2-2.17)$$

$$\chi(\psi, \theta_0, \varphi_0) = \chi[r(\psi, \theta_0, \varphi_0), \theta_0, \varphi_0]. \quad (2-2.18)$$

Since both $\mathbf{x}(\psi, \theta, \varphi)$ and $\chi(\psi, \theta, \varphi)$ are periodic in θ and φ , their Fourier series

$$\mathbf{x}(\psi, \theta, \varphi) = \sum_{m,n} x_{mn}(\psi) e^{2\pi i(n\varphi - m\theta)} \quad (2-2.19)$$

and

$$\chi(\psi, \theta, \varphi) = \sum_{m,n} \chi_{mn}(\psi) e^{2\pi i(n\varphi - m\theta)} \quad (2-2.20)$$

can be constructed from $x(\psi, \theta_0, \varphi_0)$ and $\chi(\psi, \theta_0, \varphi_0)$ at a finite number of choice of angles (θ_0, φ_0) . Thus the field line Hamiltonian and the transformation equation can be obtained numerically for a given magnetic field.

It is a great advantage to study the magnetic field using the field Hamiltonian. The topological structure can be investigated by powerful tools of Hamiltonian mechanics, such as Liouville's theorem, Noether's theorem, KAM theory, etc. We will give an example in Chapter 3.

2-3 Magnetic Coordinates

In the canonical representation of the magnetic field of Eq. (2-1.1), the field line Hamiltonian is a time-dependent Hamiltonian with one degree of freedom or equivalently a time-independent Hamiltonian with two degrees of freedom. If this Hamiltonian is integrable, then it can be transformed into action-angle variables so that the Hamiltonian only depends on the action. The magnetic field lines lie on the nested two dimensional tori. The surface of the torus is called the magnetic surface. The magnetic coordinates are the action-angle variables of the field line Hamiltonian.⁵

The existence of magnetic coordinates is equivalent to the existence of magnetic surfaces. In the language of mathematics, magnetic coordinates exist if, and only if, there is a well behaved function $f(\mathbf{x})$ that satisfies

$$\mathbf{B} \cdot \nabla f(\mathbf{x}) = 0, \quad (2-3.1)$$

with $\nabla f(\mathbf{x}) \neq 0$ except along isolated curves. The surfaces of $f(\mathbf{x}) = \text{const.}$ form the magnetic surfaces. This condition of Eq. (2-3.1) is met when plasma equilibrium exists. When the plasma pressure is balanced by the electromagnetic force, $\nabla P = \mathbf{j} \times \mathbf{B}$, the pressure distribution function $P(\mathbf{x})$ satisfies $\mathbf{B} \cdot \nabla P = 0$. The pressure surfaces coincide

with the magnetic surfaces. Plasma confinement requires that the pressure surfaces be bounded. The so-called hair theorem of topology tells us that the pressure surface must form a torus and that the magnetic field must have a toroidal configuration.

In magnetic coordinates (ψ, θ, φ) , the canonical representation of the field becomes

$$\mathbf{B}(\psi, \theta, \varphi) = \nabla \psi \times \nabla \theta + \nabla \varphi \times \nabla \chi(\psi). \quad (2-3.2)$$

The field line Hamiltonian χ depends on the action ψ alone. We define the rotational transform ι as the angular frequency of θ ,

$$\iota(\psi) = \frac{\partial \chi}{\partial \psi}. \quad (2-3.3)$$

It can be trivially shown that

$$\mathbf{B} \cdot \nabla \psi = 0 \quad \text{and} \quad \mathbf{B} \cdot \nabla(\theta - \iota \varphi) = 0,$$

which imply that the magnetic field lines satisfy

$$\psi = \psi_0 \quad \text{and} \quad \theta = \theta_0 + \iota(\psi) \varphi, \quad (2-3.4)$$

with ψ_0 and θ_0 constants ($\psi = \psi_0$ and $\theta = \theta_0$ at $\varphi = 0$). These can be explained in terms of Hamiltonian language. The Hamiltonian χ depends only on the action ψ . ψ is a constant of motion and θ increases at a fixed rate $d\theta/d\varphi = \iota(\psi)$.

The rotational transform ι is an important quantity. When ι is an irrational number, magnetic field lines never close on themselves. A single field line can cover the magnetic surface densely. When ι is a rational number, magnetic field lines close on themselves and form a rational surface. A rational surface is unstable and can be destroyed by a small perturbation. We will discuss more about the effect of the perturbation in Chapter 3. The physical interpretation of ψ, θ, φ and χ is given in Fig. 2.3. The toroidal flux inside a

magnetic surface is ψ . The poloidal angle is θ and the toroidal angle is φ . The poloidal flux outside a magnetic surface is $-\chi$. The period of the poloidal angle θ and the period of the toroidal angle φ are unity instead of the conventional 2π . We will first show that the contravariant representation of Eq. (2-3.2) exists when there is a function $f(x) \neq 0$ such that $\mathbf{B} \cdot \nabla f = 0$. We then prove that ψ is the toroidal flux and $-\chi$ is the poloidal flux.

We start with the coordinates (f, ξ, φ) and the general form of the contravariant representation of the magnetic field

$$\mathbf{B} = a \nabla f \times \nabla \xi + b \nabla \xi \times \nabla \varphi + c \nabla \varphi \times \nabla f, \quad (2-3.5)$$

where f satisfies Eq. (2-3.1), ξ is an arbitrary poloidal angle, φ is the toroidal angle and three functions a , b and c depend on (f, ξ, φ) . $\mathbf{B} \cdot \nabla f = 0$ implies $b = 0$, and

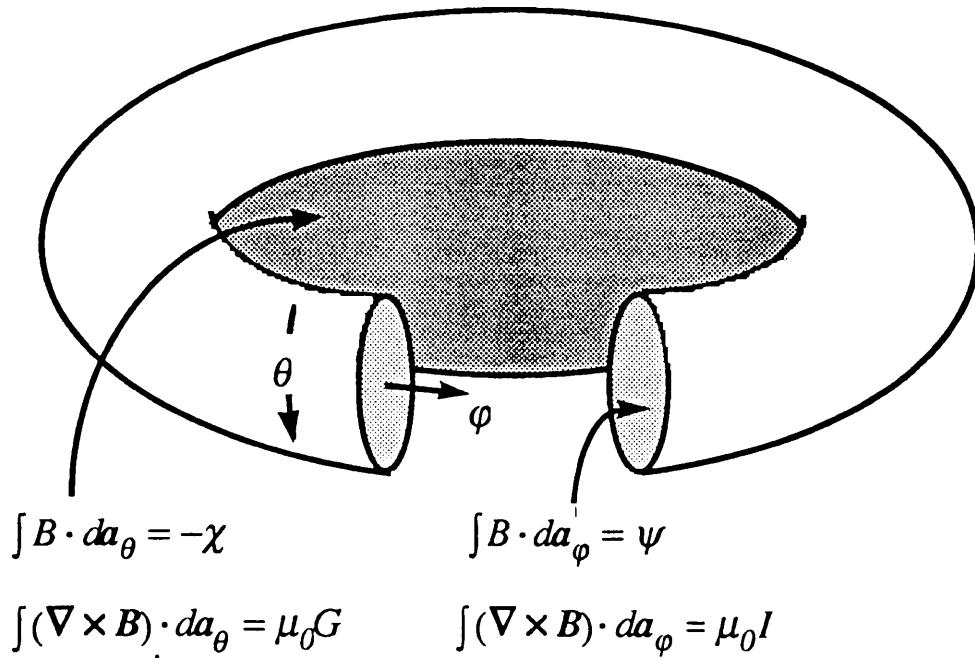


Figure 2.3. The magnetic coordinates. The toroidal angle φ and the poloidal angle θ are normalized to period unity. The toroidal flux ψ , the poloidal flux χ , the toroidal current I and the poloidal current G are defined by integrals over the two enclosed areas of the torus.

$$\nabla \cdot \mathbf{B} = (\nabla f \times \nabla \xi) \cdot \nabla \varphi \left(\frac{\partial a}{\partial \varphi} + \frac{\partial c}{\partial \xi} \right) = 0$$

implies

$$\frac{\partial a}{\partial \varphi} + \frac{\partial c}{\partial \xi} = 0. \quad (2-3.6)$$

The general solution of Eq. (2-3.6) is

$$a = a_0(f) \left(1 + \frac{\partial g}{\partial \xi} \right), \quad (2-3.7a)$$

$$c = c_0(f) - a_0(f) \frac{\partial g}{\partial \varphi}, \quad (2-3.7b)$$

with g a well behaved function of (f, ξ, φ) . We let $\theta = \xi + g$, $d\psi/df = a_0(f)$ and $d\chi/df = c_0(f)$, and we obtain Eq. (2-3.2). One should observe that in magnetic coordinates (ψ, θ, φ) one of the angles, say φ , can still be freely chosen. This freedom can be used to simplify the Jacobian of the coordinates.^{6, 7} And it can also be used to obtain a simple covariant representation of the magnetic field.

It is relatively easy to show that ψ and $-\chi$ are the toroidal and poloidal flux of the magnetic field. Using Eq. (2-1.21), the toroidal flux enclosed by a magnetic surface is

$$\iint \mathbf{B} \cdot d\mathbf{a}_\varphi = \iint J \mathbf{B} \cdot \nabla \varphi d\psi d\theta, \quad (2-3.9)$$

with $J = I/(\nabla \psi \times \nabla \theta) \cdot \nabla \varphi$. From the contravariant representation of the field Eq. (2-3.2), $\mathbf{B} \cdot \nabla \varphi = (\nabla \psi \times \nabla \theta) \cdot \nabla \varphi$. Therefore

$$\iint \mathbf{B} \cdot d\mathbf{a}_\varphi = \oint d\theta \int_0^\psi d\psi = \psi. \quad (2-3.10)$$

Similarly we obtain the toroidal flux outside a magnetic surface

$$\iint \mathbf{B} \cdot d\mathbf{a}_\theta = \oint d\varphi \int_0^\psi J \mathbf{B} \cdot \nabla \theta \, d\psi = \int_{\psi_c}^\psi \frac{d\chi}{d\psi} \, d\psi = - \int_0^\chi d\chi = -\chi, \quad (2-3.11)$$

with ψ_c the toroidal flux at the major axis ($R = 0$).

As we mentioned earlier, magnetic coordinates are arbitrary in one function of position, $\omega(\psi, \theta, \varphi)$. If $(\psi, \theta_1, \varphi_1)$ is one set of magnetic coordinates then $(\psi, \theta_2, \varphi_2)$ is another set if $\varphi_2 = \varphi_1 + \omega$ and $\theta_2 = \theta_1 + \iota(\psi)\omega$ with $\iota(\psi)$ the rotational transform defined in Eq. (2-3.3). Boozer used this freedom of $\omega(\psi, \theta, \varphi)$ to simplify the covariant representation of the magnetic field² to

$$\mathbf{B} = \mu_0 [G(\psi) \nabla \varphi + I(\psi) \nabla \theta + \beta_* (\psi, \theta, \varphi) \nabla \psi]. \quad (2-3.12)$$

The magnetic coordinates satisfying both Eqs. (2-3.2) and (2-3.12) are called Boozer coordinates.

The last term in Eq. (2-3.12) may be confusing. It seems that the magnetic field \mathbf{B} has a component in $\nabla \psi$ direction, which contradicts $\mathbf{B} \cdot \nabla \psi = 0$ following from Eq. (2-3.2). The fact is that the \mathbf{B} -field has no component in $\nabla \psi$ direction and Eq. (2-3.12) is correct. Because $\nabla \theta$ and $\nabla \varphi$ are not orthogonal to $\nabla \psi$, therefore the sum of all three terms in Eq. (2-3.12) has zero component in $\nabla \psi$ direction.

In the covariant representation of the field, Eq. (2-3.12), μ_0 is the permeability of free space, and the function β_* is closely related to the Pfirsch-Schlüter current. The physical interpretation of I and G is also given in Fig. 2.3. G is the poloidal current in Amperes outside a magnetic surface, and I is the toroidal current in Amperes enclosed by a magnetic surface. These can be proven by using $\nabla \times \mathbf{B} = \mu_0 \mathbf{j}$ and the procedures are similar to the proof that ψ and $-\chi$ are the magnetic fluxes.

Boozer coordinates exist if there is a function $\psi(x)$ that satisfies

$$\mathbf{B} \cdot \nabla \psi = 0, \quad (2-3.13a)$$

and

$$(\nabla \times \mathbf{B}) \cdot \nabla \psi = 0. \quad (2-3.13b)$$

When plasma equilibrium exists, $\nabla P = \mathbf{j} \times \mathbf{B}$, but $\nabla \times \mathbf{B} = \mu_0 \mathbf{j}$, thus Eqs. (2-3.13a) and (2-3.13b) are satisfied with the magnetic surface coincident with the pressure surface. To prove this, we first choose the coordinates $(\psi, \theta_0, \varphi_0)$ that satisfy Eq. (2-3.13a), hence

$$\mathbf{B} = \nabla \psi \times \nabla \theta_0 + \nabla \varphi_0 \times \nabla \chi(\psi).$$

In the coordinates $(\psi, \theta_0, \varphi_0)$, the magnetic field also has its covariant representation

$$\mathbf{B} = \mu_0(a \nabla \varphi_0 + b \nabla \theta_0 + c \nabla \psi), \quad (2-3.14)$$

with a, b and c being some functions of $(\psi, \theta_0, \varphi_0)$. Eq. (2-3.13b) implies that

$$\frac{\partial a}{\partial \theta_0} - \frac{\partial b}{\partial \varphi_0} = 0. \quad (2-3.15)$$

The solution of Eq. (2-3.15) are

$$a = G(\psi) + [G(\psi) + \iota(\psi)I(\psi)] \frac{\partial}{\partial \varphi_0} \omega(\psi, \theta_0, \varphi_0), \quad (2-3.16a)$$

$$b = I(\psi) + [G(\psi) + \iota(\psi)I(\psi)] \frac{\partial}{\partial \theta_0} \omega(\psi, \theta_0, \varphi_0). \quad (2-3.16b)$$

By choosing the new angle variables to be $\theta = \theta_0 + \iota \omega$ and $\varphi = \varphi_0 + \omega$, we obtain Eq. (2-3.12).

References

- 1 A. H. Boozer, Lecture notes on *Plasma Physics* (1991)
- 2 A. H. Boozer, Phys. Fluids **26**, 1288 (1983)
- 3 A. H. Boozer, Phys. Fluids **24**, 1999 (1981); **25**, 520 (1982)
- 4 G. Kuo-Petraovic, A. H. Boozer, J. A. Rome and R. H. Fowler, J. Comp. Phys. **51**, 261 (1983)
- 5 D. Kruskal and R. M. Kulsrud, Phys. Fluids **1**, 265 (1958)
- 6 S. Hamada, Nucl. Fusion **2**, 23 (1962)
- 7 J. M. Greene and J. L. Johnson, Phys. Fluids **5**, 510 (1962)

CHAPTER 3

HAMILTONIAN MECHANICS

In this chapter we review briefly the general theory of Hamiltonian mechanics, which mostly follows from Lichtenberg and Lieberman.¹ We first give the Hamilton equations of motion in Sec. 3-1. In Sec. 3-2 we discuss the canonical transformation. In Sec. 3-3, a noncanonical method is discussed. The canonical perturbation theory is presented in Sec. 3-4. In Sec. 3-5, the Lie transformation is introduced. In Sec. 3-6, a very important concept, adiabatic invariance, is discussed. In Sec. 3-7, a qualitative discussion of the motion near the resonance is presented. In the last section, Sec. 3-8, an numerical example is given

3-1 Hamilton Equations of Motion

Suppose the system we consider is of dimension N . The generalized coordinates of the system are \mathbf{q} , or q^i ($i = 1, 2, \dots, N$). The Lagrangian function of the system is

$$L(\mathbf{q}, \dot{\mathbf{q}}, t) = T(\dot{\mathbf{q}}) - U(\mathbf{q}, t), \quad (3-1.1)$$

in which $\dot{\mathbf{q}} = dq/dt$, T is the kinetic energy and U is the potential energy. The equations of motion in terms of the Lagrangian function can be derived from the variational principle

$$\delta \int_{t_1}^{t_2} L(\mathbf{q}, \dot{\mathbf{q}}, t) dt = 0. \quad (3-1.2)$$

The equations of motion are

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^i} \right) - \frac{\partial L}{\partial q^i} = 0, \quad (3-1.3)$$

with $i = 1, 2, \dots, N$. Eq. (3-1.3) normally contains N second order differential equations. Therefore $2N$ independent variables, or initial conditions, are needed to determine the motion of the system. Another way to describe the system is to use the Hamiltonian formalism. First we define the canonical momenta of \mathbf{q} : $\mathbf{p} = (p_1, p_2, \dots, p_N)$,

$$p_i = \frac{\partial L}{\partial \dot{q}^i}. \quad (3-1.4)$$

The Hamiltonian of the system is defined by a Legendre transformation,

$$H(\mathbf{p}, \mathbf{q}, t) = \dot{q}^i p_i - L(\mathbf{q}, \dot{\mathbf{q}}, t), \quad (3-1.5)$$

with $\dot{\mathbf{q}}$ as functions of (\mathbf{q}, \mathbf{p}) . Notice here the canonical coordinates, q^i , always have contravariant indices and the canonical momenta p_i always have covariant indices.

For simplicity, we only discuss the time-independent Hamiltonian system. For the time-dependent Hamiltonian, we will later in section 3-2, introduce the *extended phase space* in which a time-dependent Hamiltonian is converted into a time-independent Hamiltonian. In the remaining parts of this chapter we will assume the system is time-independent, unless we state the contrary.

Hamilton's equations for a time-independent Hamiltonian system are

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q^i}, \quad (3-1.6a)$$

$$\frac{dq^i}{dt} = \frac{\partial H}{\partial p_i} \quad (3-1.6b)$$

and

$$\frac{dH}{dt} = 0. \quad (3-1.7)$$

The motion of a point $\mathbf{q}(q^1, \dots, q^N)$ in the N -dimensional configuration space can also be viewed as a point $(\mathbf{p}, \mathbf{q}) = (p_1, \dots, p_N; q^1, \dots, q^N)$ in the $2N$ -dimensional phase space. The phase space has three important properties. 1) The trajectories in the phase space never intersect. The motion of the system is uniquely determined by the $2N$ initial conditions $(\mathbf{p}_0, \mathbf{q}_0)$. If two trajectories intersected at a certain time, then they would have the same values of \mathbf{p} and \mathbf{q} , and their motion before and after that time would be identical. If the Hamiltonian is independent of time, the trajectories in the phase space are independent of time and can not cross in the phase space. 2) Any closed curve in phase space C_1 will be transformed into another closed curve C_2 after a time τ , trajectories bounded by the curve C_1 will be bounded by C_2 . This property follows from the first one. 3) Liouville's Theorem: if the density distribution of system points in phase space is $\rho(\mathbf{p}, \mathbf{q}, t)$, then one can prove that the density function satisfies the continuity equation

$$\frac{d\rho}{dt} = \frac{\partial\rho}{\partial t} + \frac{d\mathbf{p}}{dt} \cdot \frac{\partial\rho}{\partial\mathbf{p}} + \frac{d\mathbf{q}}{dt} \cdot \frac{\partial\rho}{\partial\mathbf{q}} = 0. \quad (3-1.8)$$

Eq. (3-1.8) implies that the phase space flow, or the Hamiltonian flow, is incompressible.

3-2 Canonical Transformation

In this section we will first state the canonical transformation method. Then we introduce the concept of the extended phase space.

A. Canonical Transformation

One of the important techniques in solving a Hamiltonian problem is to make a certain canonical transformation such that in the new canonical coordinates the Hamiltonian is independent of some of the coordinates. The Hamiltonian problem is easier to solve in the new coordinates because some of the symmetries become apparent. Suppose a Hamiltonian of N degrees of freedom only depends on M generalized coordinates, (q^1, \dots, q^M) , with $M < N$. We denote those canonical momenta conjugate to the coordinates (q^{M+1}, \dots, q^N) as (c_{M+1}, \dots, c_N) . The Hamiltonian of the system is $H(p_1, \dots, p_M; q^1, \dots, q^M; c_{M+1}, \dots, c_N)$. Since

$$\frac{dc_l}{dt} = -\frac{\partial H}{\partial q^l} = 0 \quad (l = M+1, \dots, N),$$

the canonical momenta (c_{M+1}, \dots, c_N) are isolating constants of motion and can be considered as some constant parameters. The Hamiltonian problem of N degrees of freedom is reduced to a problem of M degrees of freedom.

The canonical transformation is a transformation from a set of old canonical variables (p, q) and the old Hamiltonian $H(p, q)$ to another set of new canonical variables (P, Q) and the new Hamiltonian $K(P, Q)$ such that the form of Hamilton's equations of motion is preserved, i.e. in the new canonical coordinates,

$$\frac{dP}{dt} = -\frac{\partial K}{\partial Q}, \quad \frac{dQ}{dt} = \frac{\partial K}{\partial P}. \quad (3-2.1)$$

A canonical transformation can be obtained by using a mixed variable generating function, which depends on both the old and the new variables. We state without proof the following results.² There are four types of generating function, which we denote as $F_1(q, Q)$, $F_2(q, P)$, $F_3(p, Q)$ and $F_4(p, P)$. For F_1 , we have

$$p_i = \frac{\partial F_1}{\partial q^i}, \quad P_i = -\frac{\partial F_1}{\partial Q^i}. \quad (3-2.2)$$

For F_2 ,

$$p_i = \frac{\partial F_2}{\partial q^i}, \quad Q^i = \frac{\partial F_2}{\partial P_i}. \quad (3-2.3)$$

For F_3 ,

$$q^i = -\frac{\partial F_3}{\partial p_i}, \quad P_i = -\frac{\partial F_3}{\partial Q^i}. \quad (3-2.4)$$

For F_4 ,

$$q^i = -\frac{\partial F_4}{\partial p_i}, \quad Q^i = \frac{\partial F_4}{\partial P_i}. \quad (3-2.5)$$

For all the four types of generating function, the relation between the old Hamiltonian $H(\mathbf{p}, \mathbf{q})$ and the new Hamiltonian $K(\mathbf{P}, \mathbf{Q})$ is

$$K(\mathbf{P}, \mathbf{Q}) = H(p(\mathbf{P}, \mathbf{Q}), q(\mathbf{P}, \mathbf{Q})). \quad (3-2.6)$$

In practice a generating function can be any combination of the four types.

Another way to define a canonical transformation is to use the Poisson bracket. Let (\mathbf{p}, \mathbf{q}) be a set of canonical coordinates and f and g be two well behaved functions of (\mathbf{p}, \mathbf{q}) . The Poisson bracket of f and g is

$$\{f, g\}_{(\mathbf{p}, \mathbf{q})} = \frac{\partial f}{\partial q^i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q^i}. \quad (3-2.7)$$

Obviously, for the canonical coordinates (\mathbf{p}, \mathbf{q}) , we have

$$\{q^i, p_j\}_{(\mathbf{p}, \mathbf{q})} = \delta_j^i, \quad \{q^i, q^j\}_{(\mathbf{p}, \mathbf{q})} = 0 \quad \text{and} \quad \{p_i, p_j\}_{(\mathbf{p}, \mathbf{q})} = 0. \quad (3-2.8)$$

A transformation from (\mathbf{p}, \mathbf{q}) to (\mathbf{P}, \mathbf{Q}) is canonical, if and only if the Poisson brackets are

invariant under this transformation, i.e.

$$\{f, g\}_{(P, Q)} = \{f, g\}_{(p, q)}, \quad (3-2.9)$$

or

$$\{Q^i, P_j\}_{(p, q)} = \delta_j^i, \quad \{Q^i, Q^j\}_{(p, q)} = 0 \quad \text{and} \quad \{P_i, P_j\}_{(p, q)} = 0. \quad (3-2.10)$$

We will drop the subscripts on the Poisson brackets since they are invariants under all canonical coordinates. Here we state the algebraic properties of the Poisson bracket without proof:

$$\{f, f\} = 0; \quad (3-2.11a)$$

$$\{f, g\} = -\{g, f\}; \quad (3-2.11b)$$

$$\{af + bg, h\} = a\{f, h\} + b\{g, h\}, \quad (3-2.11c)$$

with a and b constants;

$$\{fg, h\} = \{f, h\}g + f\{g, h\}; \quad (3-2.11d)$$

and

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0. \quad (3-2.11e)$$

Eq. (3-2.11e) is also called *Jacobi's identity*.

The advantage of using the mixed variable generating function is that one only needs to deal with a scalar function instead of a vector. The disadvantage is that one has to perform a function inversion to obtain the desired form of coordinate transformation, i.e.

$p_i = p_i(P, Q)$ and $q^i = q^i(P, Q)$, or $P_i = P_i(p, q)$ and $Q^i = Q^i(p, q)$, because the generating function depends on both the old and the new variables. The alternative procedure which avoids using the mixed variable generating function is the *Lie transformation*. However, the Lie transformation is a continuous transformation. Most coordinate transformations in Hamiltonian perturbation theory are continuous and the Lie transformation is a very powerful tool in perturbation theory. We will discuss Lie transformation in section 3-6.

B. Extended Phase Space

We now introduce the concept of extended phase space. Let the Hamiltonian be explicitly dependent on the time, $H(p, q, t)$. In the extended phase space we take the time t and the Hamiltonian H as two canonical variables. By using a generating function

$$F_2 = \sum_{i=1}^N P_i q^i + P_{N+1} t, \quad (3-2.12)$$

with

$$P_i = p_i, \quad i = 1, \dots, N \quad \text{and} \quad P_{N+1} = -H, \quad (3-2.13a)$$

one can find trivially that

$$Q^i = q^i, \quad i = 1, \dots, N \quad \text{and} \quad Q^{N+1} = t, \quad (3-2.13b)$$

and the new Hamiltonian in the extended phase space is

$$K(P, Q) = H(p, q, t) + P_{N+1}. \quad (3-2.13c)$$

There is a "time" like parameter τ in the $(2N+2)$ -dimensional extended phase space. The Hamilton equations of motion in the extended phase space are

$$\frac{dP_i}{d\tau} = -\frac{\partial K}{\partial Q^i}, \quad \frac{dQ^i}{d\tau} = \frac{\partial K}{\partial P_i}. \quad (3-2.14)$$

The new Hamiltonian is explicitly independent of the "time" τ . With $i = N + 1$, Eq. (3-2.14) yields $\iota(\tau) = \tau$ and $K = \text{constant}$. The motion of a system with a time-dependent Hamiltonian is equivalent to that of a time-independent Hamiltonian with an additional degree of freedom.

3-3 Hamiltonian in Noncanonical Coordinates

Contrary to the conventional canonical Hamiltonian method, there is noncanonical Hamiltonian method. The canonical Hamiltonian formulations are simple because the Poisson brackets of the canonical variables are simple. The requirement of using canonical coordinates sometimes hampers the Hamiltonian treatment of the problems, like guiding center motion. However the nice features of the Hamiltonian method, such as preservation of phase volume, only dealing with scalar functions and etc., are not based on using canonical coordinates. The noncanonical Hamiltonian method is developed in a number of places in the literature.^{3, 4, 5} In this section we briefly summarize the noncanonical Hamiltonian method and Darboux's theorem.

We will use z , or z^i , with $i = 1, \dots, 2N$, only for noncanonical coordinates of a Hamiltonian system having N degrees of freedom. We will only use p and q for the canonical coordinates.

If the Hamiltonian is given in a set of noncanonical coordinates, $H(z)$ with $z = (z^1, \dots, z^{2N})$, and the relation between the noncanonical variables and canonical variables is assumed known, i.e. $z^i = z^i(p, q)$, then the equations of motion in the noncanonical variables are

$$\frac{dz^i}{dt} = \frac{\partial z^i}{\partial q^l} \frac{dq^l}{dt} + \frac{\partial z^i}{\partial p_l} \frac{dp_l}{dt} = \frac{\partial z^i}{\partial q^l} \frac{\partial H}{\partial p_l} - \frac{\partial z^i}{\partial p_l} \frac{\partial H}{\partial q^l} = \left(\frac{\partial z^i}{\partial q^l} \frac{\partial z^j}{\partial p_l} - \frac{\partial z^i}{\partial p_l} \frac{\partial z^j}{\partial q^l} \right) \frac{\partial H}{\partial z^j}.$$

Thus the equation of motion is

$$\frac{dz^i}{dt} = J^{ij} \frac{\partial H}{\partial z^j}, \quad (3-3.1)$$

with J^{ij} the Poisson bracket

$$J^{ij} = \{z^i, z^j\} = \frac{\partial z^i}{\partial q^l} \frac{\partial z^j}{\partial p_l} - \frac{\partial z^i}{\partial p_l} \frac{\partial z^j}{\partial q^l}. \quad (3-3.2)$$

Eq. (3-3.1) can also be written in a vector form

$$\frac{d\mathbf{z}}{dt} = \mathbf{J} \cdot \frac{\partial H}{\partial \mathbf{z}}, \quad (3-3.1')$$

where the Poisson tensor is

$$\mathbf{J} = \begin{pmatrix} 0 & J^{ij} \\ -J^{ij} & 0 \end{pmatrix}. \quad (3-3.3)$$

The Poisson bracket in Eq. (3-3.2) is, according to the assumption $z^i = z^i(\mathbf{p}, \mathbf{q})$, a function of \mathbf{p} and \mathbf{q} , i.e. $J^{ij} = J^{ij}(\mathbf{p}, \mathbf{q})$.

In practice, the relation between \mathbf{z} and \mathbf{p}, \mathbf{q} is generally known as $\mathbf{p} = \mathbf{p}(\mathbf{z})$ and

$\mathbf{q} = \mathbf{q}(z)$. The most desirable form of the Poisson bracket J^{ij} is one in which it is a function of z , i.e. $J^{ij} = J^{ij}(z)$ instead of $J^{ij} = J^{ij}(\mathbf{p}, \mathbf{q})$. The formulation will become very messy if one tries to invert functions. It is necessary to introduce the action form and Lagrange bracket. The action form is defined as

$$\rho_i = p_l \frac{\partial q^l}{\partial z^i}. \quad (3-3.4)$$

The Lagrange bracket is defined as

$$\omega_{ij} = \frac{\partial p_l}{\partial z^i} \frac{\partial q^l}{\partial z^j} - \frac{\partial q^l}{\partial z^i} \frac{\partial p_l}{\partial z^j} = \frac{\partial \rho_j}{\partial z^i} - \frac{\partial \rho_i}{\partial z^j}, \quad (3-3.5)$$

which is a function of the noncanonical variables z . The Lagrange bracket we use here differs from the usual definition by a sign. One can easily prove that

$$\omega_{ij} J^{jk} = \delta_i^k. \quad (3-3.6)$$

The easy way to calculate the Poisson bracket as a function of the noncanonical variables is first to construct the Lagrange tensor with the Lagrange brackets and then invert the Lagrange tensor. The matrix inversion maybe tedious, but it is still much easier than trying to invert the functions.

Now we turn to the subject of Darboux's theorem. One of the important techniques in solving a canonical Hamiltonian problem is to make certain canonical transformation such that in the new canonical coordinates the Hamiltonian is independent of some of the

coordinates. The Hamiltonian problem is easier to solve in the new coordinates because some of the constants of the motion become apparent. In an arbitrary noncanonical coordinate system the same technique can not be applied easily because the Poisson brackets generally couple the motion of different variables. Suppose the Hamiltonian is independent of a certain variable z^k , the equations of motion are

$$\frac{dz^i}{dt} = J^{ij} \frac{\partial H}{\partial z^j}.$$

In order to make z^i a constant of motion, the Poisson brackets must satisfy $J^{ij} = 0$ for all $j \neq k$. In noncanonical coordinates, the coordinate transformation has to meet the requirements of both the Hamiltonian and the Poisson tensor in order to make the symmetry of a system apparent.

Darboux's theorem says that one can choose one arbitrary function of any coordinates, say $q = q(z)$, which can also be one of the coordinates, then construct other coordinates p and $Z^i (i = 3, \dots, 2N)$ as functions of z by solving the following equations,

$$\{q, p\} = c,$$

$$\{Z^i, q\} = 0,$$

$$\{Z^i, p\} = 0,$$

(3-3.7)

where c is a constant. In these new coordinates $(p, q; Z^3, \dots, Z^{2N})$ the motion of p and q is

decoupled from Z^i . The Poisson tensor is

$$\mathbf{J} = \begin{pmatrix} 0 & c & 0 \\ -c & 0 & \{Z^i, Z^j\} \\ 0 & \{Z^i, Z^j\} \end{pmatrix}. \quad (3-3.8)$$

If the Hamiltonian in the new coordinates is independent of q , then p is a constant of motion.

3-4 The Canonical Perturbation Theory

A. Introduction

If an N -dimensional system has N independent symmetries, then the Hamiltonian is said to be integrable. When a Hamiltonian of a system with N degrees of freedom can be separated into N independent Hamiltonians in a set of canonical coordinates, one for each degree of freedom, we say the original Hamiltonian is separable. A separable Hamiltonian is always integrable. A system of one degree of freedom is always integrable. Consider an arbitrary one degree of freedom Hamiltonian,

$$H(p, q) = E, \quad (3-4.1)$$

E is a constant of motion. The canonical momentum p is a function of q alone,

$$p = p(q; E), \quad (3-4.2)$$

with E a constant parameter. The solution, i.e. p and q as functions of time, can be obtained from one of Hamilton's equations of motion, Eq. (3-1.6b),

$$\frac{dq}{dt} = \frac{\partial H}{\partial p}.$$

Thus

$$t = \int_{q_0}^q \frac{dq}{\partial H / \partial p}, \quad (3-4.3)$$

with $\partial H / \partial p = \left. \frac{\partial H(p, q)}{\partial p} \right|_{p=p(q; E)}$. Thus a Hamiltonian of one degree of freedom is always integrable.

For an N -degree of freedom completely separable Hamiltonian,

$$H(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^N H_i(p_i, q^i). \quad (3-4.4)$$

The solution is

$$t = \int_{q_0^i}^{q^i} \frac{dq^i}{\partial H_i / \partial p_i}. \quad (3-4.5)$$

For a one-dimensional oscillatory system, we can define the action integral J ,

$$J = \frac{1}{2\pi} \oint p dq, \quad (3-4.6)$$

with the integral taking over one cycle of the oscillation. The canonical coordinate conjugate to the action J is an angle-like variable, usually denoted as θ . The canonical pair (J, θ) is called the *action-angle variables*. The Hamiltonian in the action-angle variables is a function of the action J alone,

$$H = H(J), \quad (3-4.7)$$

and J is a constant of motion. The change of θ with time is also a constant

$$\frac{d\theta}{dt} = \omega(J) = \frac{\partial H}{\partial J}. \quad (3-4.8)$$

For an N -dimensional integrable oscillatory system, the action-angle variables can be defined in a similar way. The concept of the action is important not only because it is the canonical momentum in action-angle variables but because it is also an *adiabatic invariant* in many nearly integrable systems. The subject of adiabatic invariance will be discussed in Sec. 3-6.

In reality almost all systems are non-integrable.⁶ However many systems of interest are near-integrable. By near-integrable we mean that the system is very close to an integrable system. In Hamiltonian language, the Hamiltonian of a near-integrable system, H , is

$$H = H_0 + H', \quad (3-4.9)$$

with H_0 an integrable Hamiltonian and H' very small compared to H_0 , $|H'/H_0| \ll 1$. Such a near-integrable system can be "solved" by perturbation methods. The perturbative solution is usually an asymptotic approximation, and is valid only for a finite time.

We first consider a periodic near-integrable Hamiltonian system with two degrees of freedom,

$$H(J_1, J_2; \theta^1, \theta^2) = H_0(J_1, J_2) + \varepsilon H'(J_1, J_2; \theta^1, \theta^2), \quad (3-4.10)$$

where (J, θ) are action-angle variables of the integrable part, H_0 . The perturbation part, H' , is much smaller than H_0 , and is periodic in the θ 's. In Eq. (3-4.10) and in the remainder of this dissertation, the dimensionless parameter ε serves as a small parameter for keeping track of ordering, and we will set $\varepsilon = 1$ at the end of the calculation.

We first consider the motion of the unperturbed part H_0 in which J_1 and J_2 are constants of motion. The trajectories that have the same J_1 and J_2 lie on a single 2-dimensional invariant torus in the phase space. Trajectories of the same energy and different J_1 and J_2 form a infinite number of nested invariant tori (see Fig. 3.1a). Two angle variables θ^1 and θ^2 change at frequencies

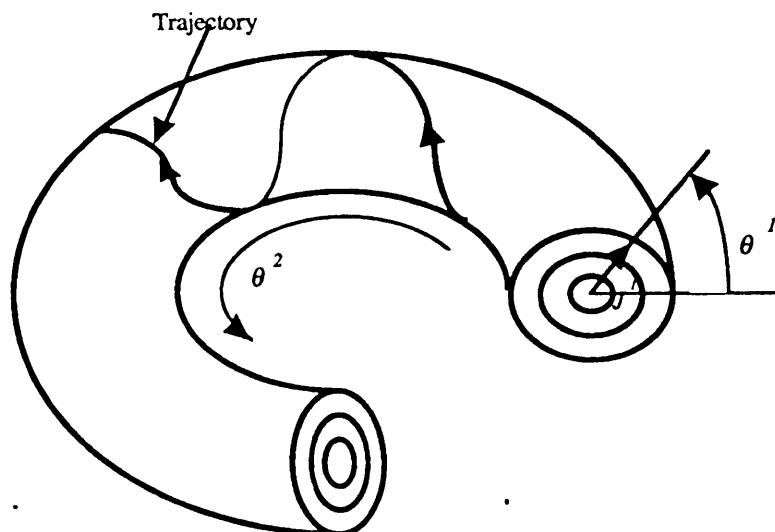
$$\omega_1 = \frac{\partial H_0}{\partial J_1}, \text{ and } \omega_2 = \frac{\partial H_0}{\partial J_2}. \quad (3-4.11)$$

If ω_1/ω_2 is an irrational number, a single trajectory densely covers the surface of the torus. We call such surfaces *irrational surfaces*. If ω_1/ω_2 is a rational number, $\omega_1/\omega_2 = m/n$ with m and n integers, the trajectory bites its own tail after θ^1 repeats m times and θ^2 repeats n times. We call these surfaces *rational surfaces*, or *resonant surfaces*. The surface of section (SOS), or Poincaré surface of section, of (J_1, θ^1) is shown in Fig. 3.1b. For the resonant surface $\omega_1/\omega_2 = m/n$, each trajectory hits the SOS n times before it repeats itself. The curve on the SOS is called period n curve. The rational surfaces are dense but of zero measure in the phase space.

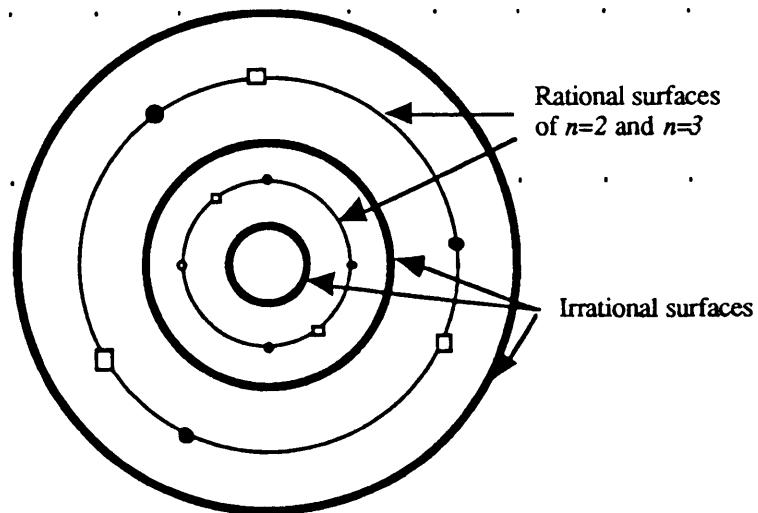
Now we add a very small perturbation term H' . Since H' is periodic in the θ 's, we can Fourier expand H' ,

$$H' = \sum_{n_1, n_2} H'_{n_1, n_2}(J_1, J_2) \exp[2\pi i(n_1\theta^1 - n_2\theta^2)]. \quad (3-4.12)$$

Each of those Fourier terms, H'_{n_1, n_2} , destroys the resonant surface of $\omega_1/\omega_2 = n_1/n_2$. On the SOS of (J_1, θ^1) , the period n_2 curve becomes n_2 fixed O points and n_2 fixed X points. The trajectories starting near the X points are stochastic and they occupy the region near the separatrix. The trajectories near the O points are regular, and their intersections with the SOS form the invariant curves around the O points (see Fig. 3.2). The width of the regular island is of order $|H'_{n_1, n_2}|^{1/2}$.



(a)



(b)

Figure 3.1 Motion of a two dimensional integrable Hamiltonian system in phase space. (a) Nested invariant tori. (b) Poincaré surface of section of (J_1, θ') .

The KAM theorem⁷ states that when the perturbation is small enough most of the irrational invariant surfaces still exist, they are merely distorted rather than destroyed. Although the motion near the separatrix is stochastic, the chaotic motion is constrained by nearby invariant surfaces, the KAM surfaces, and the motion is not globally stochastic. If the perturbation is large enough so that the islands of different resonant layers overlap with each other, the KAM surfaces are destroyed. The motion then becomes globally stochastic.

A system with three or more degrees of freedom has all the properties of a 2-dimensional system. However, the motion is more complicated because the system has more degrees of freedom. Even when the perturbation is very small and primary resonant layers do not overlap, the stochasticity is still a global phenomenon because of *Arnold diffusion*.⁸ For a system with two degrees of freedom, the 3-dimensional energy volume (4-dimensional phase space with energy as constant of motion) is divided into a set of isolated regions by the 2-dimensional KAM surfaces. When a system has more than two degrees of freedom, the N -dimensional KAM surfaces can no longer separate the $(2N - 1)$ -dimensional energy volume into any isolated regions. That is analogous to the case of a 3-dimensional space which can be disconnected by a 2-dimensional surface but not by a 1-dimensional curve. Because the resonance surfaces of the unperturbed Hamiltonian are dense in the phase space, a single trajectory can "diffuse" along the resonance and pass by any point in phase space arbitrarily close in a finite amount of time. The thin chaotic layers are all connected together and form a single network, called *Arnold web*. There is an essential difference between strong stochasticity and Arnold diffusion. Strong stochasticity appears when the perturbation becomes large enough that the different chaotic layers overlap each other, and the stochasticity diffuses across the layers. The Arnold diffusion always exists no matter how small the perturbation is, and the stochasticity diffuses along the thin stochastic Arnold web. The time scale of Arnold diffusion is much longer than that of strong stochasticity.

B. Classical Perturbation Method^{9, 10}

We first consider an oscillatory Hamiltonian with one degree of freedom. Suppose the

Hamiltonian is in near action-angle form,

$$H = H_0(J) + \varepsilon H_1(J, \theta) + \varepsilon^2 H_2(J, \theta) + \dots, \quad (3-4.13)$$

with the perturbations in Fourier series, i.e. $H_1 = H_{10}(J) + \sum_{n \neq 0} H_{1n}(J) \exp(in\theta)$ and etc..

The solution of H_0 is

$$J = J_0, \quad (3-4.14a)$$

$$\theta = \theta_0 + \omega t, \quad (3-4.14b)$$

with $\omega = \partial H_0 / \partial J$. In Eq. (3-4.14), ω , J_0 and θ_0 are constants of t . To solve the motion of H perturbatively, we apply the Poincaré-von Zeipel procedure. We need to find a generating function $S(\bar{J}, \theta)$ to transform (J, θ) to $(\bar{J}, \bar{\theta})$, so that in the new canonical coordinates the new Hamiltonian \bar{H} depends on \bar{J} only, $\bar{H} = \bar{H}(\bar{J})$. Since the perturbation is small, the difference between the old and new variables must be small. We choose the generating function as a power series in ε ,

$$S(\bar{J}, \theta) = \bar{J}\theta + \varepsilon S_1(\bar{J}, \theta) + \dots. \quad (3-4.15)$$

In Eq. (3-4.15), the zero order is an identity transformation and $|S_1/S| \sim o(\varepsilon)$. Following Eq. (3-2.3), we have

$$J = \bar{J} + \varepsilon \frac{\partial S_1(\bar{J}, \theta)}{\partial \theta} + \dots, \quad (3-4.16a)$$

$$\bar{\theta} = \theta + \varepsilon \frac{\partial S_1(\bar{J}, \theta)}{\partial \bar{J}} + \dots. \quad (3-4.16b)$$

To obtain the new Hamiltonian with Eq. (3-2.6), we have to invert Eq. (3-4.16). To order ε , the inversion can be easily done,

$$J = \bar{J} + \varepsilon \frac{\partial S_I(\bar{J}, \bar{\theta})}{\partial \bar{\theta}} + \dots, \quad (3-4.17a)$$

$$\theta = \bar{\theta} - \varepsilon \frac{\partial S_I(\bar{J}, \bar{\theta})}{\partial \bar{J}} + \dots \quad (3-4.17b)$$

with $\partial S_I(\bar{J}, \bar{\theta})/\partial \bar{\theta} = \partial S_I(\bar{J}, \theta)/\partial \theta|_{\theta=\bar{\theta}}$ and $\partial S_I(\bar{J}, \bar{\theta})/\partial \bar{J} = \partial S_I(\bar{J}, \theta)/\partial J|_{\theta=\bar{\theta}}$.

From Eq. (3-2.6),

$$\bar{H}(\bar{J}, \bar{\theta}) = H(J(\bar{J}, \bar{\theta}), \theta(\bar{J}, \bar{\theta})) = \bar{H}_0 + \varepsilon \bar{H}_I + \dots \quad (3-4.18)$$

We expand the H of Eq. (3-4.18) in power series in ε with Eq. (3-4.17),

$$H_0(J(\bar{J}, \bar{\theta})) = H_0(\bar{J}) + \varepsilon \frac{\partial H_0(\bar{J})}{\partial \bar{J}} \frac{\partial S_I(\bar{J}, \bar{\theta})}{\partial \bar{\theta}} + \dots, \quad (3-4.19a)$$

$$\varepsilon H_I(J(\bar{J}, \bar{\theta}), \theta(\bar{J}, \bar{\theta})) = \varepsilon H_I(\bar{J}, \bar{\theta}) + \dots \quad (3-4.19b)$$

Therefore we obtain the new Hamiltonian in the zero order

$$\bar{H}_0 = H_0(\bar{J}), \quad (3-4.20a)$$

and in the first order

$$\bar{H}_I = H_I(\bar{J}, \bar{\theta}) + \omega(\bar{J}) \frac{\partial S_I(\bar{J}, \bar{\theta})}{\partial \bar{\theta}}, \quad (3-4.20b)$$

with $\omega = \partial H_0(\bar{J})/\partial \bar{J}$.

We choose the proper S_I so that it cancels the oscillatory part of H_I . Thus S_I satisfies

$$\omega(\bar{J}) \frac{\partial S_I(\bar{J}, \bar{\theta})}{\partial \bar{\theta}} = -\{H_I(\bar{J}, \bar{\theta})\} = -\sum_{n \neq 0} H_{In}(\bar{J}) e^{in\bar{\theta}}, \quad (3-4.21)$$

and

$$S_I(\bar{J}, \bar{\theta}) = \sum_n S_{In}(\bar{J}) e^{in\bar{\theta}}, \quad (3-4.22)$$

with S_{I0} a constant and $S_{In} = -H_{In}/(in\omega)$ for $n \neq 0$. One should observe that the frequency ω cannot be zero. The first order of the new Hamiltonian is

$$\bar{H}_I = \langle H_I \rangle = H_{I0}(\bar{J}). \quad (3-4.23)$$

In Eqs. (3-4.21) and (3-4.23), $\langle \rangle$ denotes the average part and $\{ \}$ denotes the oscillatory part, i.e. for an arbitrary function f ,

$$\langle f \rangle = \frac{1}{2\pi} \int_0^{2\pi} f d\theta, \quad (3-4.23a)$$

$$\{f\} = f - \langle f \rangle. \quad (3-4.23b)$$

Putting Eqs. (3-4.20a) and (3-4.23) together, we have

$$\bar{H} = H(\bar{J}) + \varepsilon \langle H_I(\bar{J}, \bar{\theta}) \rangle + o(\varepsilon^2). \quad (3-4.24)$$

The Poincaré-von Zeipel procedure can be used to arbitrary order in ε as long as the frequency ω is not zero. The expansion converges because all one-dimensional systems are integrable, but the algebra becomes very tedious because of the mixed variable generating function. For higher order perturbation, it is better to use the Lie transformation method, which we will present in section 3-5.

Now we discuss the perturbation method for a system of two or more degrees of

freedom. Like the zero frequency case in the one-dimensional system, there are resonances in a system of two or more degrees of freedom, where the expansions fail to converge. The failure is not because of the Poincaré-von Zeipel procedure itself, it is because of the dynamic structure of the system. Any perturbation procedure is a one to one coordinate transformation, and it can not change the phase space topological structure. We will come back to this point again later.

Let the Hamiltonian of the system be

$$H(\mathbf{J}, \boldsymbol{\theta}) = H_0(\mathbf{J}) + \varepsilon H_I(\mathbf{J}, \boldsymbol{\theta}), \quad (3-4.25)$$

with \mathbf{J} and $\boldsymbol{\theta}$ the N -dimensional action-angle variables of H_0 , and H_I a multi-periodic function of the angles. The Fourier series of the perturbation is

$$H_I = \sum_{\mathbf{m}} H_{I\mathbf{m}}(\mathbf{J}) \exp(i\mathbf{m} \cdot \boldsymbol{\theta}), \quad (3-4.26)$$

in which $\mathbf{m} \cdot \boldsymbol{\theta} = \sum_{i=1}^N m_i \theta^i$. We apply a similar procedure as that used for the one-dimensional system by introducing the near-identity generating function,

$$S = \bar{\mathbf{J}} \cdot \boldsymbol{\theta} + \varepsilon \sum_{\mathbf{m}} S_{I\mathbf{m}}(\bar{\mathbf{J}}) \exp(i\mathbf{m} \cdot \boldsymbol{\theta}) + o(\varepsilon^2). \quad (3-4.27)$$

After obtaining the old variables as functions of the new variables and expanding the Hamiltonian, we find to order ε the new Hamiltonian

$$\bar{H} = H_0(\bar{\mathbf{J}}) + \varepsilon \langle H_I(\bar{\mathbf{J}}, \bar{\boldsymbol{\theta}}) \rangle + o(\varepsilon^2), \quad (3-4.28)$$

and the generating function

$$S(\bar{\mathbf{J}}, \bar{\boldsymbol{\theta}}) = \bar{\mathbf{J}} \cdot \bar{\boldsymbol{\theta}} + \varepsilon i \sum_{\mathbf{m} \neq 0} \frac{H_{I\mathbf{m}}(\bar{\mathbf{J}})}{\mathbf{m} \cdot \boldsymbol{\omega}(\bar{\mathbf{J}})} \exp(i\mathbf{m} \cdot \bar{\boldsymbol{\theta}}) + o(\varepsilon^2), \quad (3-4.29)$$

with

$$\omega(\bar{J}) = \frac{\partial H_0(\bar{J})}{\partial \bar{J}}. \quad (3-4.30)$$

Clearly when near the resonance, i.e. $\mathbf{m} \cdot \boldsymbol{\omega} = 0$, the expansion fails to converge. The resonance condition is precisely that of the resonant surface for the unperturbed Hamiltonian H_0 . The phase space structure near the resonance will be discussed in Sec. 3-7.

3-5 Lie Transformation

When a mixed variable generating function is used, the coordinate transformation is in a mixed form, and so is the relation between the old and the new Hamiltonian. In order to obtain the desired form, one has to invert functions. This complicates the procedure. In practice it is very inconvenient to use mixed variable generating function to calculate high order perturbation. Since the coordinate transformation in perturbation theory is always continuous, the *Lie transformation* can be used, which avoids using the mixed variable generating function. Since Hori¹¹ and Garrido¹² first introduced the Lie transformation to Hamiltonian perturbation theory, the method has been improved by Deprit¹³, Dewar¹⁴, Howland¹⁵ and etc.. In this section we review the Lie transformation.

The basic idea of Lie transformation is to use a generating function which depends on a parameter ε . For convenience we denote x for the old variable and X for the new variable. The coordinate transformation is characterized by the differential equations

$$\frac{dX}{d\varepsilon} = \{X, w\}, \quad (3-5.1)$$

with the initial condition $X(\varepsilon = 0) = x$ and w the *Lie generating function*. In Eq. (3-5.1), the Lie generating function w serves as a Hamiltonian and ε serves as a time-like

parameter. The solution of Eq. (3-5.1), $X(x, \varepsilon)$, is the transformation of the coordinates. Corresponding to the coordinate transformation, a function is transformed by the *evolution operator* T . Let $H(x)$ be the old Hamiltonian in the old coordinates and $K(X)$ be the transformed Hamiltonian in the new coordinates, then

$$H(x) = T K(X) = K(X(x, \varepsilon)). \quad (3-5.2)$$

Another way to think Eq. (3-5.2) is that the coordinates are changed instead of the function, i.e.

$$X = T x. \quad (3-5.3)$$

To simplify Eq. (3-5.1), we define the *Lie operator* L as

$$L = \{w, \quad\} . \quad (3-5.4)$$

Rewriting Eq. (3-5.1) with Eqs. (3-5.4) and (3-5.3), we have

$$\frac{dX}{d\varepsilon} = -LX = -LTx = \frac{dT}{d\varepsilon} x,$$

which gives the equation of evolution operator

$$\frac{dT}{d\varepsilon} = -LT. \quad (3-5.5)$$

The solution of Eq. (3-5.5) is

$$T = \exp\left[-\int_0^\varepsilon L(\varepsilon') d\varepsilon'\right]. \quad (3-5.6)$$

The new Hamiltonian in new coordinates is

$$K(X) = T^{-1} H(x) = H(x(X, \varepsilon)), \quad (3-5.7)$$

with T^{-1} the inverse operator of T , i.e. $T^{-1} T = T T^{-1} = I$.

Using the Deprit perturbation method¹³, we expand the Lie generating function w , the Lie operator L , the evolution operator T , the Hamiltonian H and new Hamiltonian K as power series in ϵ ,

$$w = \sum_{n=0}^{\infty} \epsilon^n w_n, \quad (3-5.8a)$$

$$L = \sum_{n=0}^{\infty} \epsilon^n L_n, \quad (3-5.8b)$$

$$T = \sum_{n=0}^{\infty} \epsilon^n T_n, \quad (3-5.8c)$$

$$H = \sum_{n=0}^{\infty} \epsilon^n H_n, \quad (3-5.8d)$$

$$K = \sum_{n=0}^{\infty} \epsilon^n K_n, \quad (3-5.8e)$$

where

$$L_n = \{w_n, \quad\}. \quad (3-5.8f)$$

Putting Eqs. (3-5.8b) and (3-5.8c) into Eq. (3-5.5), We have

$$T_0 = I \quad \text{and} \quad T_n = -\frac{1}{n} \sum_{m=0}^{n-1} T_m L_{n-m} \quad (n > 0). \quad (3-5.9)$$

Similarly we can also obtain

$$T_0^{-1} = I \quad \text{and} \quad T_n^{-1} = \frac{1}{n} \sum_{m=0}^{n-1} L_{n-m} T_m^{-1} \quad (n > 0). \quad (3-5.10)$$

To calculate the transformation of the Hamiltonian, we take derivatives of Eq. (3-5.2) and then insert Eq. (3-5.5) and the series expansions. We obtain

$$K_0 = H_0 \quad \text{and} \quad L_n H_0 = n(K_n - H_n) - \sum_{m=1}^{n-1} (L_{n-m} K_m + m T_{n-m}^{-1} H_m). \quad (3-5.11)$$

To write Eqs. (3-5.9), (3-5.10) and (3-5.11) out explicitly to second order, we have

$$T_0 = T_0^{-1} = I, \quad (3-5.12a)$$

$$T_I = -T_I^{-1} = -L_I, \quad (3-5.12b)$$

$$T_2 = \frac{1}{2}(-L_2 + L_I^2), \quad (3-5.12c)$$

$$T_2^{-1} = \frac{1}{2}(L_2 + L_I^2). \quad (3-5.12d)$$

$$K_0 = H_0, \quad (3-5.13a)$$

$$L_I H_0 = K_I - H_I, \quad (3-5.13b)$$

$$L_2 H_0 = 2(K_2 - H_2) - L_I(K_I + H_I). \quad (3-5.13c)$$

The Lie generating function can be found so that it cancels out the oscillatory part of the equations. To the first order, the results of the Lie transformation and that of the Poincaré-von Zeipel method are the same.

3-6 Adiabatic Invariance

The adiabatic invariance is a very important concept in physics, which is constructed from an asymptotic series. The use of adiabatic invariants can greatly simplify the study of a near-integrable system both analytically and numerically.

Following Bender and Orszag,¹⁶ we say a series $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ is asymptotic to a function $f(x)$, if for every N the remainder $\delta_N(x)$ after $(N - 1)$ terms of the series is much smaller than the last retained term as $x \rightarrow x_0$:

$$\delta_N(x) \equiv f(x) - \sum_{n=0}^N a_n(x - x_0)^n \ll (x - x_0)^N \quad (x \rightarrow x_0). \quad (3-6.1)$$

A series need not be convergent to be asymptotic. Most asymptotic series are not convergent. Let us compare convergent and asymptotic:

Convergent: $\delta_N(x) \equiv \sum_{n=N+1}^{\infty} a_n(x - x_0)^n \rightarrow 0, \quad N \rightarrow \infty, \quad x \text{ fixed};$

Asymptotic: $\delta_N(x) \ll (x - x_0)^N, \quad x \rightarrow x_0, \quad N \text{ fixed.}$

For a convergent series, $\delta_N(x)$ goes to zero as $N \rightarrow \infty$ for some fixed $|x - x_0|$. On the other hand, if the series is asymptotic, then $\delta_N(x)$ goes to zero faster than $(x - x_0)^N$ as $x \rightarrow x_0$, but need not go to zero as $N \rightarrow \infty$ for fixed $|x - x_0|$. Convergence is an absolute concept. One can prove that a series converges without knowing the function to which it converges. However, asymptoticity is a relative property of the expansion coefficients and

the function to which the series is asymptotic.

It can be shown that motion of a slow time-dependent Hamiltonian

$$H(p, q, t) = \frac{1}{2m} p^2 + \frac{1}{2} m\omega^2(\varepsilon t) q^2, \quad (3-6.2)$$

with $d\omega/dt \sim \varepsilon\omega$, has an asymptotic series expansion.^{17, 18, 19} An adiabatic invariant can be constructed

$$J \approx \oint p \, dq, \quad (3-6.3)$$

which is conserved for a time scale shorter than $t_c \sim (I/\omega)\exp(c/\varepsilon)$, with c a positive constant. For a multiply periodic system, the exponentially small change in the adiabatic invariant is the consequence of the resonant terms that modify the phase space, and ultimately destroy the conservation of the adiabatic invariant for ε not sufficiently small in a relative short time.

The condition for adiabatic approximation is that the system has well separated time scales for different variables. Suppose the Hamiltonian of the system has the form

$$H(J, \theta; p, q) = H_0(J; p, q) + \varepsilon H_1(J, \theta; p, q) + \varepsilon^2 H_2(J, \theta; p, q) + \dots, \quad (3-6.4)$$

where the action-angle variables (J, θ) are for the fast degree of freedom, and the remaining (p, q) are for the slow degrees of freedom. The zero order Hamiltonian H_0 is independent of the fast angle variable θ , and time scale of the fast variable is of order ε^{-1} faster than the slow variables. We need to find a canonical transformation so that the new Hamiltonian is independent of the fast angle θ . To avoid using the mixed variable generating function, we use Lie transformation method and start with Eqs. (3-5.12) and (3-5.13). Unlike the ordinary perturbation, which we discussed in early in Sec. 3-4 and Sec. 3-5, the Lie operator here contains both fast part and slow part, i.e.

$$L = L_f + \varepsilon L_s, \quad (3-6.5)$$

with the fast part

$$L_f = \{w, \quad\} = \frac{\partial w}{\partial \theta} \frac{\partial}{\partial J} - \frac{\partial w}{\partial J} \frac{\partial}{\partial \theta}, \quad (3-6.6a)$$

and the slow part

$$L_s = \{w, \quad\}_s = \frac{\partial w}{\partial q} \cdot \frac{\partial}{\partial p} - \frac{\partial w}{\partial p} \cdot \frac{\partial}{\partial q}. \quad (3-6.6b)$$

We now need to solve Eqs. (3-5.13) and find the Lie generating function. Because of the different ordering in time scales we need to expand L_n (or w_n) and K_n as power series in ε ,

$$w_n = \sum_{l=0}^{\infty} \varepsilon^l w_{nl}, \quad (3-6.7a)$$

$$K_n = \sum_{l=0}^{\infty} \varepsilon^l K_{nl}. \quad (3-6.7b)$$

To the first order in ε , Eq.(3-5.13b) becomes

$$L_{I0f} H_0 + \varepsilon (L_{IIf} + L_{I0s}) H_0 = K_{I0} - H_I + \varepsilon K_{II}.$$

Thus

$$\{w_{I0}, H_0\}_f = K_{I0} - H_I, \quad (3-6.8a)$$

$$\{w_{II}, H_0\}_f + \{w_{I0}, H_0\}_s = K_{II}. \quad (3-6.8b)$$

By letting the generating function to cancel the oscillatory part, we find

$$K_{I0} = \langle H_I \rangle, \quad (3-6.9a)$$

$$w_{I0} = -\frac{I}{\omega_0} \int \{H_I\} d\theta, \quad (3-6.9b)$$

$$K_{II} = \left\langle \{w_{I0}, H_0\}_s \right\rangle, \quad (3-6.9c)$$

$$w_{II} = -\frac{I}{\omega_0} \int d\theta \left\{ \{w_{I0}, H_0\}_s \right\}, \quad (3-6.9d)$$

where $\omega_0 = \partial H_0 / \partial J$.

Solving Eq.(3-5.13c) to zeroth order, we have

$$L_{20f} H_0 = 2(K_{20} - H_2) - L_{I0f}(K_{I0} + H_I). \quad (3-6.10)$$

Therefore,

$$K_{20} = \langle H_2 \rangle + \frac{I}{2} \left\langle \{w_{I0}, (K_{I0} + H_I)\}_f \right\rangle, \quad (3-6.11a)$$

$$w_{20} = -\frac{I}{\omega_0} \int d\theta \left\{ 2H_2 + \left\{ w_{I0}, (K_{I0} + H_I) \right\}_f \right\}. \quad (3-6.11b)$$

Because w_{I0} is oscillatory and H_0 is independent of θ , we have

$$K_{II} = \langle \{w_{I0}, H_0\}_s \rangle = 0.$$

Similarly,

$$\langle \{w_{I0}, (K_{I0} + H_I)\}_f \rangle = \langle \{w_{I0}, 2\langle H_I \rangle\}_f \rangle + \langle \{w_{I0}, \{H_I\}\}_f \rangle = \langle \{w_{I0}, \{H_I\}\}_f \rangle.$$

Therefore

$$K_{20} = \langle H_2 \rangle + \frac{1}{2} \langle \{w_{I0}, \{H_I\}\}_f \rangle. \quad (3-6.12)$$

One should be careful that in the above equations $\{\cdot\}$ sometimes means average and sometimes means Poisson bracket. It means average when there is one function inside, i.e. $\{g\}$; it means Poisson bracket when there are two functions inside, i.e. $\{f, g\}$.

Combining the above results, we have, to order ε^2 ,

$$\begin{aligned} K &= K_0 + \varepsilon K_{I0} + \varepsilon^2 (K_{20} + K_{II}) \\ &= H_0 + \varepsilon \langle H_I \rangle + \varepsilon^2 \left(\langle H_2 \rangle + \frac{1}{2} \langle \{w_{I0}, \{H_I\}\}_f \rangle \right). \end{aligned} \quad (3-6.13)$$

And the relation between the old canonical coordinates $x_{old} = (J, \theta; p, q)_{old}$ and the new ones $x_{new} = (J, \theta; p, q)_{new}$ is

$$\mathbf{x}_{\text{old}} = T^{-1} \mathbf{x}_{\text{new}}, \quad (3-6.14)$$

where

$$T^{-1} = I + \epsilon L_{I0f} + \epsilon^2 \left(L_{IIf} + L_{I0s} + \frac{1}{2} L_{20f} + \frac{1}{2} L_{I0f}^2 \right). \quad (3-6.15)$$

The new Hamiltonian is "independent" of the fast angle, and therefore the action is an adiabatic invariant. The change of the adiabatic invariant is caused by the resonance between the fast and slow variables. Let us assume that the slow variables are also in action-angle form, i.e.

$$H = H(J_f, \theta_f; J_s, \theta_s), \quad (3-6.16)$$

where the subscript f and s indicate the fast and the slow variables. We expand the Hamiltonian in Fourier series

$$H(J_f, \theta_f; J_s, \theta_s) = \sum_{k,n} H_{k,n}(J_f, J_s) \exp[i(k\theta_f + \mathbf{n} \cdot \boldsymbol{\theta}_s)], \quad (3-6.17)$$

with $\mathbf{n} = (n_1, \dots, n_l)$. There is a mathematical theorem that the Fourier series of an analytic function is exponentially convergent, which means

$$|H_{k,n}| < \alpha \exp[-(c_0|k| + c_1|n_1| + \dots + c_l|n_l|)], \quad (3-6.18)$$

with c_0, c_1, \dots, c_l and α positive constants. As $d\theta_f/dt \sim \epsilon^{-1} d\theta_s/dt$, the resonance happens when the n 's are at least of order $1/\epsilon$. Thus the leading order resonant terms are of order $\exp(-c/\epsilon)$, with c a positive constant, and it takes an exponentially long time for the adiabatic invariant to break down. When the fast and slow time scales become closer, i.e. ϵ become larger, the adiabatic invariant becomes less well conserved. The adiabatic invariant disappears when ϵ is large enough that resonant islands of different layers

overlap each other.

When carrying out the "adiabatic" perturbation, the slow variables need not be in canonical coordinates. The Hamiltonian can be in the form of

$$H(J, \theta; z) = H_0(J; z) + \varepsilon H_I(J, \theta; z) + \dots \quad (3-6.19)$$

In Eq. (3-6.19), (J, θ) are the fast variables in action-angle form and z are the slow variables in noncanonical coordinates. The Poisson brackets between the coordinates satisfy

$$\{\theta, J\} = I, \quad (3-6.20a)$$

$$\{\theta, z\} = 0, \quad (3-6.20b)$$

$$\{J, z\} = 0. \quad (3-6.20c)$$

When transforming the coordinates, two conditions must be satisfied simultaneously: the new Hamiltonian is to be independent of the fast angle θ and the Poisson brackets of the new coordinates are to be in the same form as Eqs. (3-6.20). The second condition requires the transformation to be symplectic, which preserves the Poisson brackets, and the Lie transformation is such a transformation. The procedure is similar to what we have done in the case when the slow variables are canonical. The results to order of ε are

$$K = H_0(J; z) + \varepsilon \langle H_I(J, \theta; z) \rangle_\theta, \quad (3-6.21a)$$

$$T = I - \varepsilon L_{I0f} \quad \text{and} \quad T^{-1} = I + \varepsilon L_{I0f}. \quad (3-6.21b)$$

3-7 Motion Near the Resonance

In this section we will present a qualitative discussion of the motion near the

resonance. As we discussed in early sections, motion near the resonant surfaces is very sensitive to perturbations. A very small perturbation can destroy the rational surface of the unperturbed Hamiltonian and form a stochastic layer near the separatrices.

Let us start with a Hamiltonian with two degrees of freedom,

$$H = H_0(J_1, J_2) + \varepsilon H_{r,s} \cos(r\theta^1 - s\theta^2) + \varepsilon H_{u,v} \cos(u\theta^1 - v\theta^2). \quad (3-7.1)$$

For simplicity, we only include two Fourier terms, with $H_{r,s}$ and $H_{u,v}$ constants, in the perturbation, and we also assume that the integers r, s and u, v have no common divisor. We investigate the motion near one of the resonances, say

$$r\omega_1 - s\omega_2 = 0, \quad (3-7.2)$$

with r, s integers, ω_1 and ω_2 the unperturbed frequencies

$$\omega_1 = \frac{\partial H_0}{\partial J_1} \quad \text{and} \quad \omega_2 = \frac{\partial H_0}{\partial J_2}. \quad (3-7.3)$$

We first turn off the other Fourier term by setting $H_{u,v} = 0$. The motion is regular, but the phase space topology is changed. To see this, we proceed a canonical transformation $(J_1, J_2, \theta^1, \theta^2) \rightarrow (I_1, I_2, \varphi^1, \varphi^2)$, using the generating function

$$S(\theta^1, \theta^2; I_1, I_2) = (r\theta^1 - s\theta^2)I_1 + \theta^2 I_2. \quad (3-7.4)$$

From Eq. (3-2.3),

$$J_1 = \frac{\partial S}{\partial \theta^1} = rI_1, \quad (3-7.5a)$$

$$J_2 = \frac{\partial S}{\partial \theta^2} = I_2 - sI_1, \quad (3-7.5b)$$

$$\varphi^I = \frac{\partial S}{\partial I_I} = r\theta^I - s\theta^2, \quad (3-7.5c)$$

$$\varphi^2 = \frac{\partial S}{\partial I_2} = \theta^2. \quad (3-7.5d)$$

The new Hamiltonian is

$$K = H_0(J_I(I_I, I_2), J_2(I_I, I_2)) + \varepsilon H_{r,s} \cos(\varphi^I). \quad (3-7.6)$$

The action I_2 is a constant of motion and the Hamiltonian is integrable. Furthermore if we expand H_0 near the resonant surface of $r\omega_I - s\omega_2 = 0$, (J_{I0}, J_{20}) , we can rewrite the Hamiltonian of motion

$$\mathcal{H} = \frac{1}{2} G p^2 + \varepsilon H_{r,s} \cos \phi, \quad (3-7.7)$$

with

$$G(J_{I0}, J_{20}) = \frac{1}{r^2} \left(r \frac{\partial}{\partial J_I} - s \frac{\partial}{\partial J_2} \right)^2 H_0(J_{I0}, J_{20}), \quad (3-7.8a)$$

$$p = J_I - J_{I0}, \quad (3-7.8b)$$

$$\phi = \varphi^I = r\theta^I - s\theta^2, \quad (3-7.8c)$$

$$\mathcal{H} = K - H_0(J_{I0}, J_{20}) + o((J_I - J_{I0})^3). \quad (3-7.8d)$$

The Hamiltonian of Eq. (3-7.7) describes the motion of a pendulum in the phase space of (p, ϕ) . The (p, ϕ) phase space is separated into rotational and librational regions by a curve called the separatrix. There are two fixed points, $(0,0)$ and $(0,\pi)$, meaning at these two points

$$\frac{dp}{dt} = -\frac{\partial \mathcal{H}}{\partial \phi} = 0 \quad \text{and} \quad \frac{d\phi}{dt} = \frac{\partial \mathcal{H}}{\partial p} = 0. \quad (3-7.9)$$

One of the fixed points is unstable, which is located at the intersections of the separatrices. If a very small perturbation is given, the motion can change to either rotation or libration. The unstable fixed point is also called an *X* point or hyperbolic point. The other fixed point is stable, and located at the center of the librational region. If a small perturbation is given, the motion will change into a libration around this fixed point. In the region very close to the stable fixed point, the motion is just like a harmonic oscillator and the angular frequency of ϕ is $(|GH_{r,s}|)^{1/2}$. The stable fixed point is also called an *O* point or elliptic point. On the separatrix, the maximum change of the momentum, or the half-width of the separatrix, is $p_{max} = 2(|H_{r,s}/G|)^{1/2}$.

Now let us visualize the motion in the phase space of (J, θ) instead of (p, ϕ) . This can be done by using one of the Poincaré surfaces of section, say (J_1, θ^I) . Inserting Eqs. (3-7.8b) and (3-7.8c) into the Hamiltonian of Eq. (3-7.7), we have

$$\mathcal{H} = \frac{1}{2}G(J_1 - J_{10})^2 + \epsilon H_{r,s} \cos(r\theta^I - s\theta^2). \quad (3-7.10)$$

On the SOS of (J_1, θ^I) the angle θ^2 is a constant. The curves on the SOS must satisfy Eq. (3-7.10), meaning the curves are $\mathcal{H}(J_1, \theta^I) = \text{constant}$. The SOS is like the phase space of a pendulum. There are r stable fixed points and r unstable fixed points evenly distributed on the period r curve of the unperturbed Hamiltonian H_0 . The motion near the separatrices is particularly sensitive to any perturbation. A very small "kick" can change the motion from rotational to librational, or *vise versa*. The half-width of the sparatrices is $2(|\epsilon H_{r,s}/G|)^{1/2}$.

Now we switch on the other Fourier term, $H_{u,v} \neq 0$, and examine the motion near the resonance of $r\omega_1 - s\omega_2 = 0$. We perform the same canonical transformation as Eqs. (3-

7.4) and (3-7.5). The new Hamiltonian is

$$K = H_0(I_1, I_2) + \varepsilon H_{r,s} \cos(\varphi^1) + \varepsilon H_{u,v} \cos\left[\frac{u}{r}\varphi^1 + \left(\frac{us}{r} - v\right)\varphi^2\right]. \quad (3-7.11)$$

Near the resonance, the angle $\varphi^1 = r\theta^1 - s\theta^2$ changes much slower than the angle $\varphi^2 = \theta^2$, we can use the adiabatic approximation to "average" over φ^2 . With Eq. (3-6.13), we obtain the the Hamiltonian to order ε

$$\bar{K} = H_0(I_1, I_2) + \varepsilon H_{r,s} \cos(\varphi^1). \quad (3-7.12)$$

Thus the motion is like a pendulum. The adiabatic approximation is valid except near the separatrices. Near the separatrices the last term in Eq. (3-7.11) acts like a kicking force. The direction and the magnitude of the "kick" depend on the change of φ^2 , which is almost random. These random kicks make the motion near the separatrices extremely irregular. In fact the motion is stochastic near the separatrices. The same structure appears near the resonance of $u\omega_1 - v\omega_2 = 0$. So when the perturbation is small, most of the invariant surfaces of the unperturbed Hamiltonian are distorted rather than destroyed. Only the perturbations which resonate with the unperturbed motion destroy those resonant surfaces. On a Poincaré surface of section, a period n resonant curve becomes n stable fixed points and n unstable fixed points evenly distributed. Near each stable fixed point there is a region (called island) in which motion is regular. The half-width of the islands is of order $\mathcal{O}(\varepsilon^{1/2})$, say $2(H_{r,s}/G)^{1/2}$ for the resonant surface of $r\omega_1 - s\omega_2 = 0$. The unstable fixed points are linked by separatrices which surround the regular islands. The motion around separatrices are chaotic. Because the perturbation is small, the islands of different resonant surfaces are small and do not overlap each others. The distorted invariant surfaces (also called KAM surfaces) in between separate the stochastic layers of different resonant surfaces. The stochastic motion is confined in a finite and isolated region.

Now we examine what happens when the perturbations ($H_{r,s}$ and $H_{u,v}$) become bigger. We still use the Hamiltonian of Eq. (3-7.1). The first thing one notices is that the

island widths of the primary resonances become bigger. This is easy to understand because the island width is of order $(H_{r,s})^{1/2}$ or $(H_{u,v})^{1/2}$. The next thing one notices is that some more islands appear at resonant surfaces of H_0 other than those satisfying $r\omega_1 - s\omega_2 = 0$ and $u\omega_1 - v\omega_2 = 0$. These new islands are called secondary resonance islands. To understand how these secondary resonances appear we go back to Eq. (3-5.13), i.e.

$$K_0 = H_0, \quad (3-7.13a)$$

$$\omega_1 \frac{\partial w_1}{\partial \theta^1} + \omega_2 \frac{\partial w_1}{\partial \theta^2} = K_1 - H_{r,s} \cos(r\theta^1 - s\theta^2) - H_{u,v} \cos(u\theta^1 - v\theta^2), \quad (3-7.13b)$$

$$\begin{aligned} \omega_1 \frac{\partial w_2}{\partial \theta^1} + \omega_2 \frac{\partial w_2}{\partial \theta^2} \\ = 2K_2 - \left\{ w_1, [K_1 + H_{r,s} \cos(r\theta^1 - s\theta^2) + H_{u,v} \cos(u\theta^1 - v\theta^2)] \right\}. \end{aligned} \quad (3-7.13c)$$

From Eq. (3-7.13b),

$$K_1 = \langle H_{r,s} \cos(r\theta^1 - s\theta^2) + H_{u,v} \cos(u\theta^1 - v\theta^2) \rangle = 0 \quad (3-7.14)$$

and

$$w_1 = w_{r,s} \sin(r\theta^1 - s\theta^2) + w_{u,v} \sin(u\theta^1 - v\theta^2), \quad (3-7.15a)$$

with

$$w_{r,s} = -\frac{H_{r,s}}{r\omega_1 - s\omega_2} \quad \text{and} \quad w_{u,v} = -\frac{H_{u,v}}{u\omega_1 - v\omega_2}. \quad (3-7.15b)$$

Therefore, to order ε , the secular terms come only at the primary resonances $r\omega_1 - s\omega_2 = 0$ and $u\omega_1 - v\omega_2 = 0$. Solving Eq. (3-7.13c), we obtain

$$\begin{aligned} K_2 &= \frac{1}{2} \left\langle \left\{ w_I, \left[H_{r,s} \cos(r\theta^I - s\theta^2) + H_{u,v} \cos(u\theta^I - v\theta^2) \right] \right\} \right\rangle \\ &= -\frac{1}{2} \left\langle \frac{\partial w_I}{\partial J_1} \frac{\partial H_I}{\partial \theta^I} + \frac{\partial w_I}{\partial J_2} \frac{\partial H_I}{\partial \theta^2} \right\rangle, \end{aligned} \quad (3-7.16)$$

$$\omega_1 \frac{\partial w_2}{\partial \theta^I} + \omega_2 \frac{\partial w_2}{\partial \theta^2} = \left\{ \frac{\partial w_I}{\partial J_1} \frac{\partial H_I}{\partial \theta^I} + \frac{\partial w_I}{\partial J_2} \frac{\partial H_I}{\partial \theta^2} \right\} \quad (3-7.17)$$

Therefore to order ϵ^2 secular terms come at the resonances $(r+u)\omega_1 - (s+v)\omega_2 = 0$ and $(r-u)\omega_1 - (s-v)\omega_2 = 0$, which cause the appearance of secondary resonant islands. Also if one gets closer one will find that there are similar island chains within islands. This self-similar structure is repeated on the smaller and smaller scale.

When the perturbation get even bigger, the islands of different resonances overlap and the good KAM surfaces between island chains disappear. The merger of the separated stochastic layers signifies the transition of the system from regional chaotic to globally stochastic.

To summarize the above discussion, when the perturbation contains only one Fourier term the motion is still integrable but the phase space topology is changed. When the perturbation contains more than one Fourier terms, motion near the resonance surfaces become stochastic, but KAM surfaces still exist when perturbation is small. The KAM surfaces are destroyed when the perturbation is large enough that neighboring islands overlap each other. Motion become globally chaotic.

3-8 A Numerical Example

Before ending this chapter, we present an example showing the effect of perturbations. The system resembles a magnetic field line Hamiltonian, which we discussed in Sec. 2-2.

The "field line" Hamiltonian is

$$\chi = \chi_0(\psi) + \chi_{I,1} \cos[2\pi(\varphi - \theta)] + \chi_{I,2} \cos[2\pi(\varphi - 2\theta)], \quad (3-8.1)$$

with $\chi_{I,1}$ and $\chi_{I,2}$ constants, and

$$\chi_0 = \iota_0 \psi + \frac{1}{2} \iota_I (\psi - \psi_0)^2. \quad (3-8.2)$$

The equivalent two dimensional time-independent Hamiltonian is that of Eq. (3-7.1), with

$$H_0 = \iota_0 I_\theta + \frac{1}{2} \iota_I (I_\theta - \psi_0)^2 + I_\varphi. \quad (3-8.3)$$

We integrated Hamilton's equations of motion numerically and plotted the surface of section of (ψ, θ) , or (I_θ, θ) . Fig. 3.2 shows the results. As we increase the perturbation the islands become bigger and the stochastic region become larger. At first, there are only the main resonant islands, then the secondary resonant islands appear, and finally, in Fig. 3.2(d), the two main island chains overlap with each other, and even islands within islands become visible.

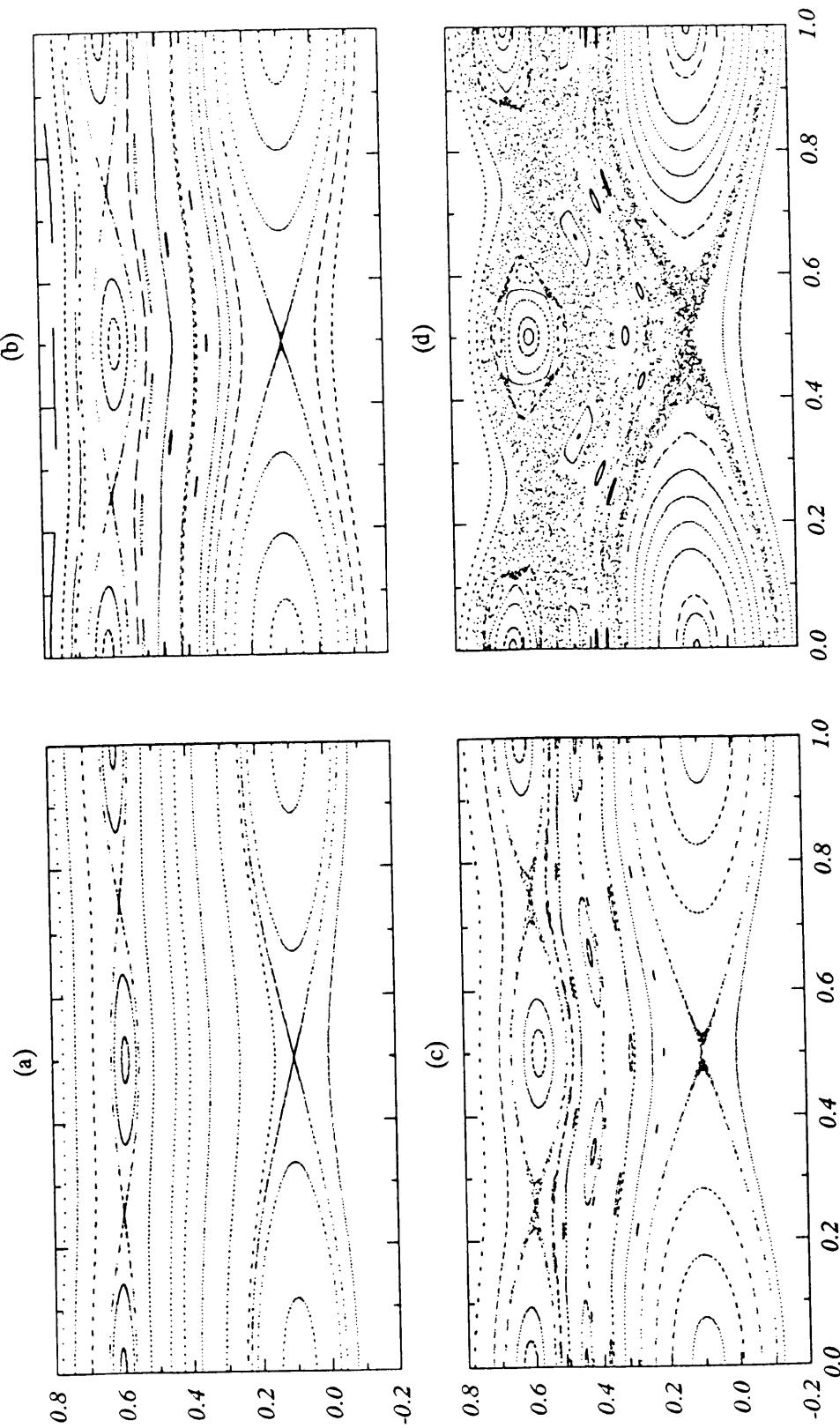


Figure 3.2 Surface of section (ψ, θ) for the field line Hamiltonian of Eq. (3-8.1). The perturbation parameters χ_{11} and χ_{22} increase from (a) to (d).

References

- 1 A. J. Lichtenberg and M. A. Lieberman, *Regular and Stochastic Motion* (Springer-Verlag, New York, 1983)
- 2 H. Goldstein, *Classical Mechanics* (Addison-Wesley Publishing Company, Inc., 1980)
- 3 V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer, New York, 1978)
- 4 R. G. Littlejohn, *J. Math. Phys.* **20**, 2445 (1979); **23**, 742 (1982)
- 5 J. R. Cray and R. G. Littlejohn, *Ann. Phys.* **151**, 1 (1983)
- 6 C. L. Siegal, *Ann. Math.* **42**, 806 (1941); *Math. Anal.* **128**, 144 (1954)
- 7 A. N. Kolmogorov, *Dokl. Akad. Nauk. USSR* **98**, 527 (1954);
J. Moser, *Nachr. Akad. Wiss. Göttingen. Math. Phys. K1*, 1 (1962);
V. I. Arnold, *Sov. math. Dokl.* **3**, 136 (1963)
- 8 V. I. Arnold, *Russian Math. Surveys* **18**, 85 (1964)
- 9 H. Poincaré, *Les Méthodes Nouvelles de la Mécanique Céleste* (Gauthier-Villars, Paris, 1892)
- 10 H. Von Zeipel, *Ark. Astron. Mat. Fys.* **11**, No.1 (1916)
- 11 G. Hori, *Astron. Soc. Japan* **18**, 287 (1966)
- 12 L. M. Garrido, *J. Math. Phys.* **10**, 1045 (1968)
- 13 A. Deprit, *Cel. Mech.* **1**, 12 (1969)
- 14 R. L. Dewar, *J. Phys. A* **9**, 2043 (1976)
- 15 R. A. Howland, *Cel. Mech.* **17**, 327 (1977)
- 16 C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill, Inc. 1978)

3. Hamiltonian Mechanics

17 G. D. Birkhoff, *Dynamical System* (American Mathematical Society, New York,

1927)

18 N. N. Bogoliubov and Y. A. Mitropolsky, *Asymptotic Methods in the Theory of
Nonlinear Oscillations* (Gordon and Breach, New York 1961)

19 M. D. Kruskal, J. Math. Phys. 3, 806 (1962)

CHAPTER 4

THE HIGHER ORDER DRIFT HAMILTONIAN IN A CURL FREE FIELD

In this chapter a research paper *the Exact and Drift Hamiltonian*, which is published in Physics Fluids B 4, page 2429, 1992, is reproduced.

4-1 Introduction

Particles in a fusion plasma have very long mean free paths compared to the size of the system. Adequate confinement for fusion implies that the trajectories of alpha particles (α 's) must remain in the plasma for at least a slowing down time, and thermal particles can only move a small fraction of the plasma radius during a collision time. A charged particle moves very rapidly on a nearly circular orbit about a magnetic field line. The center of this circle, which is called the guiding center, moves along the field line with a parallel velocity $v_{||} = \mathbf{v} \cdot \mathbf{B}/B$, and drifts slowly from one field line to another with a drift velocity $v_d \sim \epsilon v$, where $\epsilon \sim \rho/a$ is the ratio of the gyroradius to plasma size. The guiding center motion can be found to order ϵ by the standard drift equations first derived by Alfvén.¹ Since the Alfvén drift equations use only the lowest order of the Taylor expansion of the motion in ϵ , their validity is guaranteed only for a time that is short compared to $(\epsilon\Omega)^{-1}$, with $\Omega = qB/m$ the cyclotron frequency. However, the time scale t_d for a particle to move across the field lines a distance of order the plasma size, a , is $t_d \sim a/v_d$ and $t_d\Omega \sim 1/\epsilon^2$. In other words, $(\epsilon^2\Omega)^{-1}$ is the shortest time scale of interest in the calculation of particle

trajectories. An extreme situation is that of α particles slowing down in a tokamak reactor which has a rippled magnetic field due to the finite number of toroidal field coils. The gyroradius of the α 's is of order 10 centimeters and the characteristic spatial scale (the distance between toroidal field coils) is of order a few meters, so $\epsilon \sim \rho/a \approx 1/20$. The slowing down time t_s times the gyrofrequency can be as large as $t_s\Omega \sim 10^8$ and $t_s \sim (\epsilon^4 \Omega)^{-1}$. Alfvén's derivation gives little reason for confidence in using the drift equations on the time scale required. The next order drift terms in an expansion in ϵ have been calculated.²⁻⁵ In this chapter, we give a canonical Hamiltonian treatment of the guiding center problem in Boozer coordinates.

A major difference between this and earlier work on higher order drift Hamiltonian is the use of Boozer coordinates (ψ, θ, φ) . A general set of magnetic coordinates contains three coordinates (ψ, θ, φ) . One of the important properties of the magnetic coordinates is that the plasma pressure is a function of the toroidal flux coordinate ψ alone. The primary interest in particle trajectories in toroidal plasma equilibria is their deviation from the surfaces of constant pressure. The ψ -coordinate provides the simplest possible description of this deviation. The magnetic field lines also have a trivial form in magnetic coordinates, $\psi = \text{constant}$ and $\theta - \iota(\psi)\varphi = \text{constant}$ with θ a poloidal angle and φ a toroidal angle (Fig. 2.3), which is equivalent to the magnetic field having the contravariant representation

$$\mathbf{B} = \nabla\psi \times \nabla\theta + \iota(\psi)\nabla\varphi \times \nabla\psi.$$

Magnetic coordinates are arbitrary in one function of position, $\omega(x)$. If $(\psi, \theta_1, \varphi_1)$ are one set of magnetic coordinates then $(\psi, \theta_2, \varphi_2)$ are another set if $\varphi_2 = \varphi_1 + \omega(x)$ and $\theta_2 = \theta_1 + \iota(\psi)\omega(x)$. The original treatment of magnetic coordinates,⁶ Hamada coordinates, used the freedom of $\omega(x)$ to make the Jacobian of the magnetic coordinates constant on a ψ -surface. It was later demonstrated that the freedom of $\omega(x)$ could be used to simplify the covariant representation of the magnetic field⁷ to

$$\mathbf{B} = \mu_0[G(\psi)\nabla\varphi + I(\psi)\nabla\theta + \beta_*(\psi, \theta, \varphi)\nabla\psi].$$

This particular set of magnetic coordinates are called Boozer coordinates. The Boozer coordinates have a simple relation to the canonical coordinates of the Hamiltonian that corresponds to Alfvén drift equations. We call this Hamiltonian the standard drift Hamiltonian H_d .

A surprising feature of the standard drift Hamiltonian in Boozer coordinates is that the three dimensional properties of the magnetic field enter only through the field strength $B(\psi, \theta, \varphi)$.⁸ The exact Hamiltonian H_e , which is the Hamiltonian of the exact particle trajectories, on the other hand depends on the full metric tensor of magnetic coordinates, $g^{\psi\theta} = \nabla\psi \cdot \nabla\theta$, etc.; its dependence on the three spatial coordinates cannot be written in terms of the field strength alone. This opens the important possibility that the standard drift Hamiltonian can have an exact symmetry that is not shared by the exact Hamiltonian. In this chapter we show that the first order correction to the standard drift Hamiltonian depends on the metric tensor in a generic way. Previous work fails to show this feature since the only formulation of the drift equations that makes the existence of such a symmetry transparent is the formulation using Boozer coordinates.

The exact Hamiltonian depends on six variables. The Hamiltonian corresponding to the Alfvén drift equations, which we call the standard drift Hamiltonian, depends on only four variables. These four variables are basically the parallel velocity $v_{||}$ and the three components of the spatial position of the guiding center X . The standard drift Hamiltonian depends on only four variables instead of six, because the magnetic moment μ , which is determined by the perpendicular velocity v_\perp , $\mu = mv_\perp^2/2B$, is assumed to be conserved. The gyrophase θ_g , which is the coordinate conjugate to μ and rotates at the gyrofrequency, is irrelevant to the guiding center motion.

Let H_e be the Hamiltonian of six variables that describes the exact trajectory. We call the Hamiltonian that most accurately describes the guiding center motion using only four canonical variables the guiding center Hamiltonian H_g . The relationship between the exact and guiding center Hamiltonian is

$$H_\epsilon(\mu, \theta_g; P, Q) = H_g(\mu; P, Q) + H_\mu(\mu, \theta_g; P, Q),$$

with H_μ the difference between the exact and guiding center Hamiltonian. The vector $P = (p_1, p_2)$ and $Q = (q_1, q_2)$ are the four canonical variables of the guiding center Hamiltonian. If μ were exactly conserved, H_ϵ would be independent of θ_g and H_μ could be taken to be zero.

There are two important issues. First, how faithfully can the trajectories of the guiding center Hamiltonian H_g represent the trajectories of the exact Hamiltonian H_ϵ ? In other words, how long can the solutions of a four variable Hamiltonian (the guiding center Hamiltonian H_g) represent the solutions of a six variable Hamiltonian (the exact Hamiltonian H_ϵ)? Second, even if a sufficiently accurate guiding center Hamiltonian H_g exists, how accurate an approximation to this guiding center Hamiltonian is required? The standard drift Hamiltonian H_d contains only the lowest order terms in the Taylor expansion in ϵ . The relation between these two Hamiltonians, which have four canonical variables, is

$$H_g = H_d + \epsilon H_I + \epsilon^2 H_2 + \dots$$

The general principles of Hamiltonian mechanics imply that the important part of a perturbation, such as H_μ and H_I , is the part that resonates with the trajectories given by the lower order Hamiltonian, H_g or H_d .

While the direct answer to the above two question is difficult to obtain, the application of our results to the quasi-helically symmetric magnetic field greatly simplifies the comparison between the motion of the drift and exact trajectories, and helps to answer these questions.

Nührenberg and Zille⁹ have found numerical examples of three dimensional stellarator equilibria in which the field strength has the approximate form $B(\psi, \theta - N\varphi)$, with N an integer, while the metric tensor depends in a complicated way on all three spatial

coordinates. The idealization of these equilibria are said to have quasi-helical symmetry. The standard drift Hamiltonian of quasi-helical symmetric equilibria has an exactly conserved canonical momentum, so the neoclassical confinement properties of such equilibria are similar to those of a toroidally symmetric tokamak. Garren and Boozer¹⁰ later showed that although one can achieve exact quasi-helical symmetry on one ψ -surface, quasi-helical symmetry must be broken in the global expression for the field strength of toroidal equilibria by terms of order $1/A^3$ with A the aspect ratio of the torus. The metric tensor of such equilibria has terms that break helical symmetry that are of order $1/A$. Since quasi-helical symmetry is broken in very different orders, in $1/A$, in the exact and the standard drift Hamiltonian, quasi-helical equilibria provide a simple but important test of the long term accuracy of the drift equations. [Without violating any principles of Hamiltonian mechanics, one can even consider fully three dimension magnetic fields that have exact quasi-helical symmetry though such fields are inconsistent with plasma equilibria.]

We find that the higher order corrections in ε , the ratio of the gyroradius to plasma size, to the standard drift Hamiltonian depend in a generic way on the metric tensor. Such terms produce drifts that break the invariance of the standard drift Hamiltonian in quasi-helical equilibria. These drifts are of order $(\varepsilon^2/A)v$ while the breaking of the conserved momentum by the standard drift Hamiltonian is of order $(\varepsilon/A^3)v$ in self-consistent equilibria. The first effect of the breaking of the exact symmetry of the standard drift Hamiltonian is an oscillation in the conserved momentum p_λ of quasi-helical symmetry. The details of the breakdown in the conservation of p_λ in the drift and exact Hamiltonian will be the subject of chapter 6. Here we give the higher order correction to the standard drift Hamiltonian, which makes the drift Hamiltonian dependent on the metric as well as the field strength.

In Sec. 4-2, we give a brief discussion of the magnetic coordinates and the standard drift Hamiltonian. In Sec. 4-3, we give the exact Hamiltonian in the guiding center coordinates, in which four of the six canonical variables are the same as those of the drift Hamiltonian. In Sec. 4-4, we perform an expansion of the exact Hamiltonian, and transform the two canonical variables that correspond to the gyromotion into action-angle

variables. We give the first order correction to the standard drift Hamiltonian, which describes the drift motion to order ε^2 . The results are summarized in Sec. 4-5, so the reader who is interested only in the form of higher order drift Hamiltonian K_d can move directly to this section. We also calculate the second order correction, which is one order higher than the previous works and describes the drift motion to order ε^3 . Since our main interest is using Boozer coordinates to formulate the drift Hamiltonian, we give the second order correction in Appendices 4-A. In Appendix 4-D we give the drift Hamiltonian with first order correction in a pseudo-cartesian magnetic coordinates. In performing our calculations, we use the small dimensionless parameter ε to keep track of the ordering, but set $\varepsilon = 1$ at the end of the calculation.

4-2 The Magnetic Coordinates and the Drift Hamiltonian

In order to have a simple form of the Hamiltonian, we need to choose the coordinate system carefully. The trajectories of charged particles in magnetic fields are strongly tied to the magnetic field lines, so magnetic coordinates form a natural coordinate system. Magnetic coordinates^{6,7} and the standard drift Hamiltonian in magnetic coordinates^{8,11,12} have been studied previously. Here we briefly summarize the results of these studies.

A. The Magnetic Coordinates

Any magnetic field $\mathbf{B}(\mathbf{x})$ can be represented in canonical coordinates (ψ, θ, φ) as

$$\mathbf{B} = \nabla\psi \times \nabla\theta + \nabla\varphi \times \nabla\chi(\psi, \theta, \varphi), \quad (4-2.1)$$

with χ the field line Hamiltonian, i.e. the field lines are given by $d\psi/d\varphi = -\partial\chi/\partial\theta$ and $d\theta/d\varphi = \partial\chi/\partial\psi$. Magnetic surfaces exist if and only if there is a function $f(\mathbf{x})$ that satisfies

$$\mathbf{B} \cdot \nabla f(\mathbf{x}) = 0,$$

with $|\nabla f| \neq 0$ except along isolated curves. If magnetic surfaces exist, χ can be written as a function of ψ alone. The magnetic field then has the contravariant form⁴

$$\mathbf{B} = \nabla\psi \times \nabla\theta + \nabla\varphi \times \nabla\chi(\psi). \quad (4-2.2)$$

The physical interpretation of ψ , θ , φ and χ is given in Fig. 2.3. The toroidal flux inside a magnetic surface is ψ . The poloidal angle is θ and the toroidal angle is φ . The poloidal flux outside a magnetic surface is $-\chi$. The period of the poloidal angle θ and the period of the toroidal angle φ are unity instead of the conventional 2π . The vector potential \mathbf{A} associated with the field \mathbf{B} has the simple covariant form

$$\mathbf{A} = \psi\nabla\theta - \chi(\psi)\nabla\varphi. \quad (4-2.3)$$

This expression for the vector potential is a well-behaved, single-valued function of position throughout all space. A magnetic field associated with a plasma equilibrium, $\nabla p = \mathbf{j} \times \mathbf{B}$, also has a covariant representation⁷

$$\mathbf{B} = \mu_0[G(\psi)\nabla\varphi + I(\psi)\nabla\theta + \beta_*(\psi, \theta, \varphi)\nabla\psi], \quad (4-2.4)$$

where μ_0 is the permeability of free space, and the function β_* is closely related to the Pfirsch-Schlüter current. The physical interpretation of I and G is also given in Fig. 4.1. G is the poloidal current outside a magnetic surface, and I is the toroidal current enclosed by a magnetic surface.

In this chapter, we assume, for simplicity, that the magnetic field is a vacuum, or curl-free, field with good magnetic surfaces. In a curl-free field, the covariant form becomes

$$\mathbf{B} = \mu_0 G_0 \nabla\varphi, \quad (4-2.5)$$

where G_0 is the number of Amperes in the coils producing the field. The contravariant representation of the magnetic field and the vector potential, Eq.(4-2.2) and Eq.(4-2.3),

remain unchanged in the magnetic coordinates in which \mathbf{B} has a simple covariant form, Boozer coordinates.

B. The Standard Guiding Center Hamiltonian in the Magnetic Coordinates

The standard drift Hamiltonian can be derived from the drift Lagrangian. The Lagrangian of the exact trajectory of a charged particle in a magnetic field is

$$L = \frac{1}{2}mv^2 + q\mathbf{v} \cdot \mathbf{A}(\mathbf{x}), \quad (4-2.6)$$

with q the charge of the particle and $v^2 = v_\perp^2 + v_\parallel^2$ the square of the velocity. The drift Lagrangian given by Taylor¹¹ is

$$L_d = \frac{1}{2}mv_\parallel^2 + q\mathbf{v} \cdot \mathbf{A}(\mathbf{x}) - \mu B. \quad (4-2.7)$$

These two Lagrangians differ in the sign of $\mu B \approx mv_\perp^2/2$. The reason that μB is negative in the drift Lagrangian is that in the drift motion μB is a potential term instead of a kinetic energy term.

In the magnetic coordinates,

$$\mathbf{v} = \dot{\psi}\frac{\partial \mathbf{x}}{\partial \psi} + \dot{\theta}\frac{\partial \mathbf{x}}{\partial \theta} + \dot{\phi}\frac{\partial \mathbf{x}}{\partial \phi}, \text{ and } v_\parallel = \mathbf{v} \cdot \mathbf{B}/B.$$

By using the covariant form for the vector potential, Eq.(4-2.3), and the magnetic field, Eq.(4-2.5), as well as the orthogonality relations of general coordinates, $\nabla\xi^i \cdot (\partial\mathbf{x}/\partial\xi^j) = \delta_j^i$, one can trivially find the magnetic coordinate representation of the drift Lagrangian

$$L_d = \frac{1}{2}m\left(\frac{\mu_0 G_0}{B}\right)^2 \dot{\phi}^2 + q(\psi\dot{\theta} - \chi\dot{\phi}) - \mu B. \quad (4-2.8)$$

The canonical momenta are

$$p_\psi = 0; \quad p_\theta = q\psi; \quad p_\phi = m\left(\frac{\mu_0 G_0}{B}\right)^2 \dot{\phi} - q\chi. \quad (4-2.9)$$

The fact that p_ψ is zero implies that $\partial H_d / \partial p_\psi$ is ill defined and $\dot{p}_\psi = -\partial H_d / \partial \psi$ vanishes. If one writes out the Lagrangian equations of motion, one would find that there are one second order differential equation and two first order differential equations instead of three second order differential equations. The reason for this peculiarity is that although L_d is a function of six variables there are only four independent ones. The drift Hamiltonian has only four independent variables $\theta, \phi, p_\theta (= q\psi)$, and p_ϕ . The standard drift Hamiltonian ⁸ is

$$H_d(\theta, \phi; \psi, p_\phi) = \frac{I}{2m} \left(\frac{B(\psi, \theta, \phi)}{\mu_0 G_0} \right)^2 (p_\phi + q\chi(\psi))^2 + \mu B(\psi, \theta, \phi), \quad (4-2.10)$$

which depends on only the magnetic field strength.

Hamiltonian equations of motion are,

$$\dot{\psi} = -\frac{I}{q} \left[\frac{I}{m} \left(\frac{B}{\mu_0 G_0} \right)^2 (p_\phi + q\chi)^2 + \mu B \right] \frac{\partial B / \partial \theta}{B}; \quad (4-2.11a)$$

$$\dot{\theta} = \dot{\theta}_{||} + \dot{\theta}_{\perp}, \quad (4-2.11b)$$

with

$$\dot{\theta}_{||} = \iota \frac{I}{m} \left(\frac{B}{\mu_0 G_0} \right)^2 (p_\phi + q\chi),$$

and

$$\dot{\theta}_\perp = \frac{I}{q} \left[\frac{1}{m} \left(\frac{B}{\mu_0 G_0} \right)^2 (p_\phi + q\chi)^2 + \mu B \right] \frac{\partial B / \partial \psi}{B};$$

$$\dot{\phi} = \frac{I}{m} \left(\frac{B}{\mu_0 G_0} \right)^2 (p_\phi + q\chi); \quad (4-2.11c)$$

and

$$\dot{p}_\phi = - \left[\frac{1}{m} \left(\frac{B}{\mu_0 G_0} \right)^2 (p_\phi + q\chi)^2 + \mu B \right] \frac{\partial B / \partial \phi}{B}; \quad (4-2.11d)$$

with the rotational transform

$$\iota = d\chi(\psi) / d\psi. \quad (4-2.12)$$

The motion along the field satisfies

$$\dot{\psi} = 0 \text{ and } \dot{\theta} - \iota \dot{\phi} = 0.$$

Therefore $\dot{\theta}_\parallel$ and $\dot{\phi}$ describe the parallel motion, $\dot{\theta}_\perp$ and $\dot{\psi}$ describe the drift motion.
Since

$$p_\phi = q(\mu_0 G_0 \rho_\parallel - \chi),$$

with $\rho_\parallel = mv_\parallel/qB$, \dot{p}_ϕ cannot be interpreted simply as a parallel or drift motion. In some sense, p_ϕ gives the relation between the parallel velocity v_\parallel and the guiding center location. Notice that $\dot{\phi}$ and the parallel part of $\dot{\theta}$, $\dot{\theta}_\parallel$, are of order I . The drift part of $\dot{\theta}$, $\dot{\theta}_\perp$, $\dot{\psi}/\psi$ and \dot{p}_ϕ/p_ϕ are of order ε (because ψ and p_ϕ are large).

4-3 The Exact Hamiltonian in the Guiding Center Coordinates

The Hamiltonian of the exact trajectory of a charged particle in a magnetic field is

$$H = \frac{I}{2}mv^2. \quad (4-3.1)$$

The most often used canonical coordinates are the spatial position of the particle \mathbf{x} and conjugate momentum \mathbf{p} ,

$$\mathbf{p} = mv + q\mathbf{A}. \quad (4-3.2)$$

In an arbitrary set of spatial coordinates (ξ^1, ξ^2, ξ^3) , the canonical momenta are $p_i = \partial L / \partial \dot{\xi}^i = \mathbf{p} \cdot (\partial \mathbf{x} / \partial \xi^i)$. In the magnetic coordinates (ψ, θ, φ) , the vector potential is $\mathbf{A} = \psi \nabla \theta - \chi(\psi) \nabla \varphi$, Eq. (4-2.3). The velocity can also be written in the covariant form as

$$\mathbf{v} = \frac{v_{\parallel}}{B} \mathbf{B} + v_s \mathbf{e}_s + v_{\psi} \nabla \psi,$$

with $\mathbf{B} = \mu_0 G_0 \nabla \varphi$ and

$$\mathbf{e}_s = \nabla \theta - \iota(\psi) \nabla \varphi. \quad (4-3.3)$$

The vector \mathbf{e}_s lies roughly within the magnetic surfaces. The contravariant form for the magnetic field, Eq.(4-2.2), can be written as

$$\mathbf{B} = \nabla \psi \times \mathbf{e}_s, \quad (4-3.4)$$

which follows from Eq.(4-2.12) and Eq.(4-3.3).

Equation (4-3.4) implies that \mathbf{e}_s and $\nabla \psi$ are orthogonal to the \mathbf{B} -field. However, \mathbf{e}_s

and $\nabla\psi$ are not necessary orthogonal to each other. The canonical coordinates corresponding to the magnetic coordinates are:

$$\begin{aligned} \psi, \quad & p_\psi = mv_\psi; \\ \theta, \quad & p_\theta = mv_s + q\psi; \\ \varphi, \quad & p_\varphi = \frac{\mu_0 G_0}{B} mv_\parallel - mv_s - q\chi. \end{aligned} \tag{4-3.5}$$

The canonical coordinates $(\psi, p_\psi; \theta, p_\theta; \varphi, p_\varphi)$ are not analogous to those of the drift Hamiltonian. We therefore perform a canonical transformation so that the spatial variables are those of the guiding center (Ψ, Θ, φ) rather than those of the particle position (ψ, θ, φ) . Choosing the generating function

$$W(p_\psi, p_\theta, s, \Theta; p_\varphi, \varphi) = sp_\psi - p_\psi p_\theta/q - \Theta p_\theta - \varphi p_\varphi, \tag{4-3.6}$$

which is so called a generating function of the third type, the new canonical pairs are

$$\begin{aligned} s = mv_s/q = p_\theta/q - \psi, \quad & p = -mv_\psi = -p_\psi; \\ \Theta = \theta - mv_\psi/q = \theta - p_\psi/q, \quad & P_\Theta (= q\Psi) = q\psi + mv_s = p_\theta; \\ \varphi, \quad & p_\varphi. \end{aligned} \tag{4-3.7}$$

Equation (4-3.7) implies that the relationship between $\xi = (\psi, \theta, \varphi)$ and $\Xi = (\Psi, \Theta, \varphi)$ is

$$\psi = \Psi - s, \quad \theta = \Theta - p/q, \quad \text{and} \quad \varphi = \varphi. \tag{4-3.7'}$$

The location of the guiding center in the magnetic coordinates is (Ψ, Θ, φ) to the lowest

order in ε , the ratio of gyroradius to system size. To show this, we Taylor expand $X(\Psi, \Theta, \varphi)$ at (ψ, θ, φ) , and find

$$X(\Psi, \Theta, \varphi) = x(\psi, \theta, \varphi) + \varepsilon \frac{mv \times B}{qB^2} + o(\varepsilon^2),$$

which is the usual relation between the position of the particle x and the gyrocenter X .

The physical interpretation of the new canonical coordinates $(s, p; \Theta, P_\Theta; \varphi, p_\varphi)$ is as follows: s is the component of particle momentum in the e_s direction and lies roughly within the magnetic surfaces. The conjugate momentum to s is p , the component of the particle momentum perpendicular to the magnetic surface. The other two pairs of canonical variables Θ, P_Θ and φ, p_φ are the same as those of the drift Hamiltonian. The three spatial coordinates $P_\Theta (= q\Psi), \Theta, \varphi$ give the trajectory of the guiding center and p_φ , in some sense, gives the parallel velocity $v_{||}$.

The velocity of the particle can be written in the new coordinates $(s, p; \Theta, P_\Theta; \varphi, p_\varphi)$ as

$$\mathbf{v} = \frac{v_{||}}{B} \mathbf{B} + \frac{q}{m} s \mathbf{e}_s - \frac{l}{m} p \nabla \psi,$$

$$\text{with } v_{||} = \frac{B}{\mu_0 G_0} (p_\varphi + q\chi + qls).$$

The Hamiltonian of the exact trajectory in the coordinates $(s, p; \Theta, P_\Theta; \varphi, p_\varphi)$ is

$$H = \frac{l}{2m} \left(\frac{B}{\mu_0 G_0} \right)^2 (p_\varphi + q\chi + qls)^2 + \frac{l}{2m} (q^2 g^s s^2 - 2qg^c sp + g^\psi p^2), \quad (4-3.8)$$

where $g^s = \mathbf{e}_s \cdot \mathbf{e}_s$, $g^\psi = \nabla \psi \cdot \nabla \psi$ and $g^c = \nabla \psi \cdot \mathbf{e}_s$ are components of the metric tensor of the coordinates $(b, \mathbf{e}_s, \nabla \psi)$, with $b = B/B$. The other three metric components are $\mathbf{b} \cdot \mathbf{b} = 1$ and $\mathbf{b} \cdot \mathbf{e}_s = \mathbf{b} \cdot \nabla \psi = 0$. The metric tensor gives the shape of the magnetic surfaces. Using $B^2 = (\nabla \psi \times \mathbf{e}_s)^2$ which follows from Eq.(4-3.4), one trivially finds that the relation

between the strength of magnetic field and the components of the metric tensor is

$$B^2 = g^s g^\psi - (g^c)^2. \quad (4-3.9)$$

The exact Hamiltonian of Eq.(4-3.8) is defined in terms of the canonical coordinates $(s, p; \Theta, P_\Theta; \varphi, p_\varphi)$, but magnetic field and the metric are given in the coordinates $\xi = (\psi, \theta, \varphi)$. In other words, it is the magnetic field at the position of the particle ξ that affects the motion, not the field at the position of the guiding center $\Xi = (\Psi, \Theta, \varphi)$. With the relationship between ξ and Ξ given by Eq.(4-3.7'), the exact trajectories can be integrated using the exact Hamiltonian of Eq.(4-3.8).

4-4 Higher Order Drift Hamiltonian

To derive the higher order drift Hamiltonian, we require that the functions describing the magnetic field in Eq.(4-3.8) to be independent of s and p , which describe the gyromotion. The displacement of the particle position from the guiding center is small, s/Ψ and p/q are of order ε (see Eq.(4-3.7')), we can expand the magnetic field about the position Ξ . The expansion is complicated, however, due to the non-trivial nature of the parallel Hamiltonian

$$H^{\parallel} = \frac{I}{2m} \left(\frac{B}{\mu_0 G_0} \right)^2 (p_\varphi + q\chi + q\iota s)^2.$$

The quantities p_φ and $q\chi$ are both of order ε^{-1} , but their sum, $p_\varphi + q\chi$, is of order I .

We assume that the magnetic field is given in the following forms:

$$B = B(\psi, \theta, \varphi), \quad \chi = \chi(\psi), \quad \iota = \iota(\psi),$$

$$g^s = g^s(\psi, \theta, \varphi), \quad g^c = g^c(\psi, \theta, \varphi), \quad g^\psi = g^\psi(\psi, \theta, \varphi).$$

(4-4.1)

The Hamiltonian with the B -field in the coordinates $\Xi = (\Psi, \Theta, \varphi)$ is

$$H = H_0 + \varepsilon H_1 + \varepsilon^2 H_2 + o(\varepsilon^3), \quad (4-4.2a)$$

where

$$H_0 = \frac{I}{2m} \left(\frac{B}{\mu_0 G_0} \right)^2 (p_\varphi + q\chi)^2 + \frac{I}{2m} (q^2 g^s s^2 - 2qg^c sp + g^\psi p^2) \Big|_{\xi=\Xi} \quad (4-4.2b)$$

and

$$H_1 = -\frac{q}{2m} \left(\frac{B}{\mu_0 G_0} \right)^2 \iota_\Psi (p_\varphi + q\chi) s^2 - \frac{I}{m} \left(\frac{B}{\mu_0 G_0} \right)^2 (p_\varphi + q\chi)^2 \left(\frac{B_\Psi s + B_\Theta p/q}{B} \right) - \frac{I}{2m} [q^2 g_\Psi^s s^3 + q(g_\Theta^s - 2g_\Psi^c)s^2 p + (g_\Psi^\psi - 2g_\Theta^c)sp^2 + g_\Theta^\psi p^3/q] \Big|_{\xi=\Xi}, \quad (4-4.2c)$$

with ι , χ and their derivatives functions of Ψ alone, the field strength B , the components of the metric tensor and their derivatives given as functions of (Ψ, Θ, φ) ; and $f_{\Xi^i} = \partial f(\Xi)/\partial \Xi^i$. The second order Hamiltonian H_2 is given in Appendix 4-A. Now, all functions describing the magnetic field and their derivatives depend on the guiding center coordinates (Ψ, Θ, φ) only.

The slowly varying canonical coordinates are denoted by vector P and Q with $P = (P_\Theta = q\Psi, p_\varphi)$ and $Q = (\Theta, \varphi)$. The time scale of the rapid variables, p and s , in the exact Hamiltonian $H(p, s; P, Q)$ is of order ε^{-1} faster than the time scale of P and Q . This implies that the motion of p and s can be separated from P and Q , with p and s transformed into action-angle variables. We accomplish this transformation in two steps. First, we change $(p, s; P, Q)$ to $(\tilde{p}, \tilde{s}; \tilde{P}, \tilde{Q})$ canonically, so that \tilde{p} and \tilde{s} are orthogonal to each other in the lowest order Hamiltonian H_0 , leaving P and Q almost unchanged. Then, we carry

out another canonical transformation into gyrophase coordinates, i.e. from $(\tilde{p}, \tilde{s}; \tilde{P}, \tilde{Q})$ to $(J, \theta_g; \bar{P}, \bar{Q})$, where J and θ_g are action-angle variables in the lowest order of the Hamiltonian.

The generating function to change $(p, s; P, Q)$ to $(\tilde{p}, \tilde{s}; \tilde{P}, \tilde{Q})$ is

$$W' = \frac{I}{\varepsilon} \tilde{\mathbf{P}} \cdot \mathbf{Q} + \varepsilon \left\{ \tilde{p} \left[s - \frac{I}{2} \tilde{p} f(\tilde{P}, \tilde{Q}) \right] \right\}, \quad (4-4.3)$$

with $f(\tilde{P}, \tilde{Q}) = g^c(\tilde{P}, \tilde{Q}) / qg^s(\tilde{P}, \tilde{Q})$. One finds (see Appendix 4-B)

$$\varepsilon p = \varepsilon \tilde{p},$$

$$s(\tilde{p}, \tilde{s}; \tilde{P}, \tilde{Q}) = s_0(\tilde{p}, \tilde{s}; \tilde{P}, \tilde{Q}) + \varepsilon^2 s_2(\tilde{p}, \tilde{s}; \tilde{P}, \tilde{Q}) + o(\varepsilon^4);$$

$$\frac{I}{\varepsilon} \mathbf{P}(\tilde{p}, \tilde{s}; \tilde{P}, \tilde{Q}) = \frac{I}{\varepsilon} \tilde{\mathbf{P}} + \varepsilon \mathbf{P}_2(\tilde{p}, \tilde{s}; \tilde{P}, \tilde{Q}) + o(\varepsilon^3),$$

$$\mathbf{Q}(\tilde{p}, \tilde{s}; \tilde{P}, \tilde{Q}) = \tilde{\mathbf{Q}} + \varepsilon^2 \mathbf{Q}_2(\tilde{p}, \tilde{s}; \tilde{P}, \tilde{Q}) + o(\varepsilon^4),$$

(4-4.4)

where

$$s_0 = \tilde{s} + \tilde{p} f(\tilde{P}, \tilde{Q}),$$

$$s_2 = \frac{I}{2} \tilde{p}^3 \frac{\partial f(\tilde{P}, \tilde{Q})}{\partial \tilde{P}} \cdot \frac{\partial f(\tilde{P}, \tilde{Q})}{\partial \tilde{Q}},$$

$$\mathbf{P}_2 = -\frac{I}{2} \tilde{p}^2 \frac{\partial f(\tilde{P}, \tilde{Q})}{\partial \tilde{Q}},$$

$$\mathbf{Q}_2 = \frac{I}{2} \tilde{p}^2 \frac{\partial f(\tilde{P}, \tilde{Q})}{\partial \tilde{P}}.$$

(4-4.4')

The new Hamiltonian is

$$\begin{aligned}\tilde{H}(\tilde{p}, \tilde{s}; \tilde{\mathbf{P}}, \tilde{\mathbf{Q}}) &= H(p(\tilde{p}, \tilde{s}; \tilde{\mathbf{P}}, \tilde{\mathbf{Q}}), s(\tilde{p}, \tilde{s}; \tilde{\mathbf{P}}, \tilde{\mathbf{Q}}); \mathbf{P}(\tilde{p}, \tilde{s}; \tilde{\mathbf{P}}, \tilde{\mathbf{Q}}), Q(\tilde{p}, \tilde{s}; \tilde{\mathbf{P}}, \tilde{\mathbf{Q}})) \\ &= \tilde{H}_0(\tilde{p}, \tilde{s}; \tilde{\mathbf{P}}, \tilde{\mathbf{Q}}) + \varepsilon \tilde{H}_1(\tilde{p}, \tilde{s}; \tilde{\mathbf{P}}, \tilde{\mathbf{Q}}) + \varepsilon^2 \tilde{H}_2(\tilde{p}, \tilde{s}; \tilde{\mathbf{P}}, \tilde{\mathbf{Q}}) + o(\varepsilon^3),\end{aligned}\quad (4-4.5a)$$

where

$$\tilde{H}_0 = \frac{l}{2m} \left(\frac{B(\tilde{\mathbf{P}}, \tilde{\mathbf{Q}})}{\mu_0 G_0} \right)^2 \left[\tilde{p}_\varphi + q\chi(\tilde{\mathbf{P}}_\theta/q) \right]^2 + \frac{l}{2} \left[C(\tilde{\mathbf{P}}, \tilde{\mathbf{Q}}) \tilde{s}^2 + D(\tilde{\mathbf{P}}, \tilde{\mathbf{Q}}) \tilde{p}^2 \right], \quad (4-4.5b)$$

$$\tilde{H}_1 = H_1(\tilde{p}, s_0; \tilde{\mathbf{P}}, \tilde{\mathbf{Q}}) + \frac{l}{2m} \left(\frac{B(\tilde{\mathbf{P}}, \tilde{\mathbf{Q}})}{\mu_0 G_0} \right)^2 \left(\mathbf{P}_2 \cdot \frac{\partial}{\partial \tilde{\mathbf{P}}} \right) \left[\tilde{p}_\varphi + q\chi(\tilde{\mathbf{P}}_\theta/q) \right]^2, \quad (4-4.5c)$$

with

$$C(\tilde{\mathbf{P}}, \tilde{\mathbf{Q}}) = q^2 g^s(\tilde{\mathbf{P}}, \tilde{\mathbf{Q}})/m,$$

$$D(\tilde{\mathbf{P}}, \tilde{\mathbf{Q}}) = B^2(\tilde{\mathbf{P}}, \tilde{\mathbf{Q}})/mg^s(\tilde{\mathbf{P}}, \tilde{\mathbf{Q}}).$$

The second order term of Eq.(4-4.5a), \tilde{H}_2 , is given in Appendix 4-A. One needs to keep in mind that $p_\varphi + q\chi$ is of order l although p_φ and $q\chi$ are of order ε^{-1} . This is the reason that the second term in Eq.(4-4.5c) survives.

The perpendicular part of \tilde{H}_0 of Eq.(4-4.5b) is,

$$\tilde{H}_0^\perp = \frac{l}{2} \left[C(\tilde{\mathbf{P}}, \tilde{\mathbf{Q}}) \tilde{s}^2 + D(\tilde{\mathbf{P}}, \tilde{\mathbf{Q}}) \tilde{p}^2 \right].$$

The functions C and D change slowly because they depend on only the slow variables $\tilde{\mathbf{P}}$

and \tilde{Q} . The canonical variables \tilde{p} and \tilde{s} oscillate at the fast gyrofrequency. Thus, the perpendicular Hamiltonian \tilde{H}_0^\perp behaves like a slowly varying harmonic oscillator. We use the generating function

$$W'' = \frac{I}{\epsilon} \bar{P} \cdot \tilde{Q} + \epsilon W_I(J, \tilde{s}; \bar{P}, \tilde{Q}) \quad (4-4.6a)$$

with

$$W_I = \int \sqrt{2JR(\bar{P}, \tilde{Q}) - R^2(\bar{P}, \tilde{Q})\tilde{s}^2} d\tilde{s}, \quad (4-4.6b)$$

to change \tilde{p} and \tilde{s} into the action-angle variables J and θ_g , with \bar{P} and \tilde{Q} only changed slightly. The function $R(\bar{P}, \tilde{Q})$ is

$$R(\bar{P}, \tilde{Q}) = \sqrt{C(\bar{P}, \tilde{Q})/D(\bar{P}, \tilde{Q})} = qg^s(\bar{P}, \tilde{Q})/B(\bar{P}, \tilde{Q}), \quad (4-4.6c)$$

and is of order ϵ . We apply a procedure similar to that of transformation from $(p, s; P, Q)$ to $(\tilde{p}, \tilde{s}; \bar{P}, \tilde{Q})$ to find the relationship between $(\tilde{p}, \tilde{s}; \bar{P}, \tilde{Q})$ and $(J, \theta_g; \bar{P}, \tilde{Q})$. One then finds without difficulty that

$$\epsilon \tilde{p}(J, \theta_g; \bar{P}, \tilde{Q}) = \epsilon \tilde{p}_0(J, \theta_g; \bar{P}, \tilde{Q}) + \epsilon^3 \tilde{p}_2(J, \theta_g; \bar{P}, \tilde{Q}) + o(\epsilon^5),$$

$$\tilde{s}(J, \theta_g; \bar{P}, \tilde{Q}) = \tilde{s}_0(J, \theta_g; \bar{P}, \tilde{Q}) + \epsilon^2 \tilde{s}_2(J, \theta_g; \bar{P}, \tilde{Q}) + o(\epsilon^4),$$

$$\frac{I}{\epsilon} \tilde{P}(J, \theta_g; \bar{P}, \tilde{Q}) = \frac{I}{\epsilon} \bar{P} + \epsilon \tilde{P}_2(J, \theta_g; \bar{P}, \tilde{Q}) + o(\epsilon^3),$$

$$\tilde{Q}(J, \theta_g; \bar{P}, \tilde{Q}) = \tilde{Q} + \epsilon^2 \tilde{Q}_2(J, \theta_g; \bar{P}, \tilde{Q}) + o(\epsilon^4),$$

(4-4.7)

where

$$\begin{aligned}
\tilde{p}_0 &= \cos \theta_g \sqrt{2JR(\bar{\mathbf{P}}, \bar{\mathbf{Q}})}, \\
\tilde{s}_0 &= \sin \theta_g \sqrt{2J/R(\bar{\mathbf{P}}, \bar{\mathbf{Q}})}, \\
\tilde{p}_2 &= -\frac{\partial}{\partial \bar{\mathbf{P}}} W_I(J, \tilde{s}_0; \bar{\mathbf{P}}, \bar{\mathbf{Q}}) \cdot \frac{\partial}{\partial \bar{\mathbf{Q}}} \tilde{p}_0(J, \theta_g; \bar{\mathbf{P}}, \bar{\mathbf{Q}}), \\
\tilde{s}_2 &= -\frac{\partial}{\partial \bar{\mathbf{P}}} W_I(J, \tilde{s}_0; \bar{\mathbf{P}}, \bar{\mathbf{Q}}) \cdot \frac{\partial}{\partial \bar{\mathbf{Q}}} \tilde{s}_0(J, \theta_g; \bar{\mathbf{P}}, \bar{\mathbf{Q}}), \\
\tilde{\mathbf{P}}_2 &= \frac{\partial}{\partial \bar{\mathbf{Q}}} W_I(J, \tilde{s}_0; \bar{\mathbf{P}}, \bar{\mathbf{Q}}), \\
\tilde{\mathbf{Q}}_2 &= -\frac{\partial}{\partial \bar{\mathbf{P}}} W_I(J, \tilde{s}_0; \bar{\mathbf{P}}, \bar{\mathbf{Q}}).
\end{aligned} \tag{4-4.7'}$$

The new Hamiltonian in the coordinates $(J, \theta_g; \bar{\mathbf{P}}, \bar{\mathbf{Q}})$ is

$$\bar{H}(J, \theta_g; \bar{\mathbf{P}}, \bar{\mathbf{Q}}) = \bar{H}_0(J; \bar{\mathbf{P}}, \bar{\mathbf{Q}}) + \varepsilon \bar{H}_I(J, \theta_g; \bar{\mathbf{P}}, \bar{\mathbf{Q}}) + \varepsilon^2 \bar{H}_2(J, \theta_g; \bar{\mathbf{P}}, \bar{\mathbf{Q}}) + o(\varepsilon^3), \tag{4-4.8a}$$

where

$$\bar{H}_0 = \frac{1}{2m} \left(\frac{B(\bar{\mathbf{P}}, \bar{\mathbf{Q}})}{\mu_0 G_0} \right)^2 \left[\bar{p}_\varphi + q\chi(\bar{\mathbf{P}}_\vartheta/q) \right]^2 + \frac{q}{m} JB(\bar{\mathbf{P}}, \bar{\mathbf{Q}}) \tag{4-4.8b}$$

and

$$\bar{H}_I = \tilde{H}_I(\tilde{p}_0, \tilde{s}_0; \bar{\mathbf{P}}, \bar{\mathbf{Q}}) + \frac{1}{2m} \left(\frac{B(\bar{\mathbf{P}}, \bar{\mathbf{Q}})}{\mu_0 G_0} \right)^2 \left(\tilde{\mathbf{P}}_2 \cdot \frac{\partial}{\partial \bar{\mathbf{P}}} \right) \left[\bar{p}_\varphi + q\chi(\bar{\mathbf{P}}_\vartheta/q) \right]^2. \tag{4-4.8c}$$

The second order term \bar{H}_2 is given in Appendix 4-A.

The lowest order Hamiltonian \bar{H}_0 is independent of the gyrophase θ_g and J is equivalent to μ , which is the magnetic moment in Eq.(4-2.10), $\mu = (q/m)J$. The lowest order Hamiltonian in the ϵ expansion gives the standard drift Hamiltonian. We derive the drift Hamiltonian to order ϵ^2 by eliminating the θ_g dependence in \bar{H}_1 and \bar{H}_2 (see Appendix 4-C). The new drift Hamiltonian, to order ϵ , is

$$K_d(J; \bar{P}, \bar{Q}) = H_d + \epsilon V, \quad (4-4.9a)$$

with H_d the standard drift Hamiltonian, and

$$V = \langle \bar{H}_1 \rangle_{\theta_g} = [\bar{p}_{\varphi} + q\chi(\bar{P}_{\Theta}/q)]\Delta(\bar{P}_{\Theta}/q, \bar{\Theta}, \bar{\varphi}) \quad (4-4.9b)$$

and

$$\Delta = -J \frac{B}{2m(\mu_0 G_0)^2} \left[\iota \Psi g^{\Psi} + \left(\iota \frac{\partial}{\partial \Theta} + \frac{\partial}{\partial \varphi} \right) g^c - \frac{g^c}{g^s} \left(\iota \frac{\partial}{\partial \Theta} + \frac{\partial}{\partial \varphi} \right) g^s \right]. \quad (4-4.9c)$$

The correction to the motion of the standard drift Hamiltonian due to the first order term of Eq.(4-4.9a), V , is

$$\dot{\bar{\Theta}}_I^{\parallel} = \iota \Delta, \quad (4-4.10a)$$

$$\dot{\bar{\varphi}}_I = \Delta, \quad (4-4.10b)$$

and

$$\dot{\bar{\Theta}}_I^{\perp} = [\bar{p}_{\varphi} + q\chi(\bar{P}_{\Theta}/q)] \frac{\partial \Delta}{\partial \bar{P}_{\Theta}}, \quad (4-4.10c)$$

$$\dot{\bar{\Psi}}_I = -\frac{I}{q} [\bar{p}_{\varphi} + q\chi(\bar{P}_{\Theta}/q)] \frac{\partial \Delta}{\partial \bar{\Theta}}, \quad (4-4.10d)$$

$$\dot{\bar{p}}_{\varphi I} = -[\bar{p}_\varphi + q\chi(\bar{P}_\theta/q)] \frac{\partial \Delta}{\partial \bar{\varphi}}. \quad (4-4.10e)$$

The reader should notice that $\dot{\bar{\Theta}}_I^\parallel$ and $\dot{\bar{\varphi}}_I$ are of order ε , and $\dot{\bar{\Theta}}_I^\perp$, $\dot{\bar{\Psi}}_I/\bar{\Psi}$ and $\dot{\bar{p}}_{\varphi I}/\bar{p}_\varphi$ are of order ε^2 .

The relations between (P, Q) and $(J, \theta_s; \bar{P}, \bar{Q})$ are obtained by using Equations (4-4.4), (4-4.4'), (4-4.7), (4-4.7'), (3-6.14) and (3-6.15),

$$\begin{aligned} \varepsilon^{-1} P &= \varepsilon^{-1} \bar{P} + \varepsilon \bar{P}_2(J, \theta_s; \bar{P}, \bar{Q}) + o(\varepsilon^3), \\ Q &= \bar{Q} + \varepsilon^2 \bar{Q}_2(J, \theta_s; \bar{P}, \bar{Q}) + o(\varepsilon^4). \end{aligned} \quad (4-4.11)$$

After averaging over the gyrophase,

$$\begin{aligned} \langle \bar{P}_2 \rangle &= -\frac{l}{2} J G_I \frac{\partial G_2}{\partial \bar{Q}}, \\ \langle \bar{Q}_2 \rangle &= \frac{l}{2} J G_I \frac{\partial G_2}{\partial \bar{P}}, \end{aligned} \quad (4-4.11')$$

with $G_I = g^s/B$ and $G_2 = g^c/g^s$. The difference between the guiding center given the exact Hamiltonian, (P, Q) , and the one given by the drift Hamiltonian, (\bar{P}, \bar{Q}) , is of order ε^2 .

4-5 Conclusions

For a vacuum field with good magnetic surfaces, we give the exact Hamiltonian of a particle with charge q in Boozer coordinates (Fig. 2.3). In these coordinates, the magnetic

field has both the contravariant form

$$\mathbf{B} = \nabla\psi \times \nabla\theta + \nabla\varphi \times \nabla\chi(\psi),$$

and the covariant form

$$\mathbf{B} = \mu_0 G_0 \nabla\varphi.$$

For a given magnetic field $\mathbf{B}(x)$, the magnetic coordinates (ψ, θ, φ) can be determined numerically,^{7,13} i.e. the magnetic field can be written as

$$B = B(\psi, \theta, \varphi), \quad \chi = \chi(\psi), \quad \iota = d\chi(\psi)/d\psi,$$

$$g^s = g^s(\psi, \theta, \varphi), \quad g^c = g^c(\psi, \theta, \varphi), \quad g^\Psi = g^\Psi(\psi, \theta, \varphi),$$

with the metric components, $g^s = \mathbf{e}_s \cdot \mathbf{e}_s$, $g^\Psi = \nabla\psi \cdot \nabla\psi$ and $g^c = \nabla\psi \cdot \mathbf{e}_s$. The vector $\mathbf{e}_s = \nabla\theta - \iota(\psi)\nabla\varphi$ lies roughly within the magnetic surfaces.

We find the canonical coordinates (s, Θ, φ) with the canonical momenta $(p_s = p, P_\Theta = q\Psi, p_\varphi)$ for the exact Hamiltonian

$$H_e(p, s; P, Q) = \frac{1}{2m} \left(\frac{B(\xi)}{\mu_0 G_0} \right)^2 (p_\varphi + q\chi(\Psi - s) + q\iota(\Psi - s)s)^2 + \frac{1}{2m} [q^2 g^s(\xi) s^2 - 2qg^c(\xi)sp + g^\Psi(\xi)p^2],$$

with $\xi = (\psi, \theta, \varphi) = \Xi - \delta$, $\Xi = (\Psi, \Theta, \varphi)$ and $\delta = (s, p/q, 0)$. In the magnetic coordinates, the position of the particle is $\xi = (\psi, \theta, \varphi)$ and the position of the guiding center is $\Xi = (\Psi, \Theta, \varphi)$. The canonical coordinates p and s are the variables describing the fast gyromotion. The other four canonical coordinates, $P = (P_\Theta = q\Psi, p_\varphi)$ and $Q = (\Theta, \varphi)$, are the variables describing the guiding center motion. Integrating the exact Hamiltonian H_e in the coordinates $(p, s; P, Q)$, (P, Q) gives the motion of the guiding

center.

By separating the fast gyromotion (p,s) from slow motion (P,Q) and Taylor expanding the magnetic field about the location of the guiding center Ξ , we transform (p,s) into action-angle variables (J,θ_g) , keeping (P,Q) almost unchanged. We show that the lowest order Hamiltonian is the standard drift Hamiltonian and the higher order Hamiltonians depend on the gyrophase.

$$H_\epsilon(J, \theta_g; \bar{P}, \bar{Q}) = H_d(J; \bar{P}, \bar{Q}) + H'(J, \theta_g; \bar{P}, \bar{Q}),$$

with H' the sum of the higher order terms and

$$H_d = \frac{I}{2m} \left(\frac{B(\bar{\Xi})}{\mu_0 G_0} \right)^2 \left[\bar{p}_\phi + q\chi(\bar{\Psi}) \right]^2 + \frac{q}{m} J B(\bar{\Xi}),$$

with $\bar{\Xi} = (\bar{\Psi}, \bar{\Theta}, \bar{\varphi})$ and $\bar{P}_\Theta = q\bar{\Psi}$. The difference between (P,Q) and (\bar{P},\bar{Q}) is of order ϵ^2 . The canonical momentum J is equivalent to the magnetic moment μ in Eq.(4-2.10), $\mu = qJ/m$. The first and second order corrections to the standard drift Hamiltonian are derived by using the Lie transformation. The new drift Hamiltonian with the first order correction in $\epsilon \approx \rho/a$, the gyroradius to system size, is

$$K_d(J; \bar{P}, \bar{Q}) = H_d + [\bar{p}_\phi + q\chi(\bar{\Psi})] \Delta(\bar{\Xi}),$$

with

$$\Delta = -J \frac{B(\bar{\Xi})}{2m(\mu_0 G_0)^2} \left[\frac{d\iota(\bar{\Psi})}{d\bar{\Psi}} g^*(\bar{\Xi}) + \left(\iota \frac{\partial}{\partial \bar{\Theta}} + \frac{\partial}{\partial \bar{\varphi}} \right) g^c(\bar{\Xi}) - \frac{g^c(\bar{\Xi})}{g^*(\bar{\Xi})} \left(\iota \frac{\partial}{\partial \bar{\Theta}} + \frac{\partial}{\partial \bar{\varphi}} \right) g^*(\bar{\Xi}) \right].$$

The difference between (P,Q) and (\bar{P},\bar{Q}) can be ignored since it is of order ϵ^2 . To integrate the drift trajectories using this new drift Hamiltonian, one needs to give the magnetic field in the magnetic coordinates (ψ, θ, φ) , then set $(\psi, \theta, \varphi) = (\bar{\Psi}, \bar{\Theta}, \bar{\varphi})$.

One of the biggest differences between the exact Hamiltonian and the standard drift Hamiltonian is that the exact Hamiltonian depends on the metric tensor elements while the standard drift Hamiltonian does not. The first order correction is of great interest because it depends on the metric tensor elements in a generic way. The generic dependence follows from the fact that the elements of the metric tensor must be periodic functions of \bar{Q} , $(\bar{\Theta}, \bar{\varphi})$. Whenever the symmetry of the exact Hamiltonian H_e is broken by the metric, the symmetry of the drift Hamiltonian K_d is also broken.

Appendix 4-A: Second Order Hamiltonians

The second order Hamiltonian in Eq.(4-4.2a) is

$$\begin{aligned}
H_2 = & \frac{q^2}{8m} \left(\frac{B}{\mu_0 G_0} \right)^2 \iota_{\Psi}^2 s^4 + \frac{q}{3m} \left(\frac{B}{\mu_0 G_0} \right)^2 \iota_{\Psi\Psi} (p_{\Psi} + q\chi) s^3 \\
& + \frac{q}{m} \left(\frac{B}{\mu_0 G_0} \right)^2 \iota_{\Psi} (p_{\Psi} + q\chi) \left(\frac{B_{\Psi} s + B_{\Theta} p/q}{B} \right) s^2 \\
& + \frac{1}{2m} \left(\frac{B}{\mu_0 G_0} \right)^2 (p_{\Psi} + q\chi)^2 \left[\left(\frac{B_{\Psi} s + B_{\Theta} p/q}{B} \right)^2 + \frac{1}{B} \left(s \frac{\partial}{\partial \Psi} + \frac{p}{q} \frac{\partial}{\partial \Theta} \right)^2 B \right] \\
& + \frac{1}{4mq^2} \left[q^4 g_{\Psi\Psi}^s s^4 + 2q^3 \left(g_{\Psi\Theta}^s - g_{\Psi\Psi}^c \right) s^3 p + q^2 \left(g_{\Theta\Theta}^s - 4g_{\Psi\Theta}^c + g_{\Psi\Psi}^{\Psi} \right) s^2 p^2 \right. \\
& \quad \left. + 2q \left(g_{\Psi\Theta}^{\Psi} - g_{\Theta\Theta}^c \right) s p^3 + g_{\Theta\Theta}^{\Psi} p^4 \right] \Big|_{\xi=\Xi}, \tag{4-A.1}
\end{aligned}$$

with $f_{\Xi^i \Xi^j} = \partial^2 f / \partial \Xi^i \partial \Xi^j$.

The second order Hamiltonian in Eq.(4-4.5a) is

$$\begin{aligned}
\bar{H}_2 = & H_2(\bar{p}, s_0; \bar{P}, \bar{Q}) + \frac{1}{m} \left(\frac{B(\bar{P}, \bar{Q})}{\mu_0 G_0} \right)^2 [\bar{p}_{\Psi} + q\chi(\bar{P}_{\Theta}/q)]^2 \left[\bar{P}_2 \cdot \frac{\partial B(\bar{P}, \bar{Q}) / \partial \bar{P}}{B(\bar{P}, \bar{Q})} \right] \\
& + \left(\bar{P}_2 \cdot \frac{\partial}{\partial \bar{P}} \right) H_0^{\perp}(\bar{p}, s_0; \bar{P}, \bar{Q}) + \left(s_2 \frac{\partial}{\partial s_0} + Q_2 \cdot \frac{\partial}{\partial \bar{Q}} \right) H_0(\bar{p}, s_0; \bar{P}, \bar{Q}) \\
& - \frac{q}{2m} \left(\frac{B(\bar{P}, \bar{Q})}{\mu_0 G_0} \right)^2 \iota_{\Psi}(\bar{P}_{\Theta}/q) s_0^2 \left(\bar{P}_2 \cdot \frac{\partial}{\partial \bar{P}} \right) (\bar{p}_{\Psi} + q\chi(\bar{P}_{\Theta}/q)) \\
& - \frac{1}{m} \left(\frac{B(\bar{P}, \bar{Q})}{\mu_0 G_0} \right)^2 \left(\frac{B_{\Psi}(\bar{P}, \bar{Q}) s_0 + B_{\Theta}(\bar{P}, \bar{Q}) \bar{p}/q}{B(\bar{P}, \bar{Q})} \right) \left(\bar{P}_2 \cdot \frac{\partial}{\partial \bar{P}} \right) (\bar{p}_{\Psi} + q\chi(\bar{P}_{\Theta}/q))^2, \tag{4-A.2}
\end{aligned}$$

with $H_0^{\perp} = \frac{1}{2m} [q^2 g^s(\bar{P}, \bar{Q}) s_0^2 - 2q g^c(\bar{P}, \bar{Q}) s_0 \bar{p} + g^{\Psi}(\bar{P}, \bar{Q}) \bar{p}^2]$.

The second order Hamiltonian in Eq.(4-4.8a) is

$$\begin{aligned}
 \bar{H}_2 &= \tilde{H}_2(\tilde{p}_0, \tilde{s}_0; \bar{\mathbf{P}}, \bar{\mathbf{Q}}) + \frac{I}{m} \left(\frac{B(\bar{\mathbf{P}}, \bar{\mathbf{Q}})}{\mu_0 G_0} \right)^2 \left[\bar{p}_\varphi + q\chi(\bar{P}_\Theta/q) \right]^2 \left[\tilde{\mathbf{P}}_2 \cdot \frac{\partial B(\bar{\mathbf{P}}, \bar{\mathbf{Q}})/\partial \bar{\mathbf{P}}}{B(\bar{\mathbf{P}}, \bar{\mathbf{Q}})} \right] \\
 &\quad + \left(\tilde{\mathbf{P}}_2 \cdot \frac{\partial}{\partial \bar{\mathbf{P}}} \right) \tilde{H}_0^\perp(\tilde{p}, \tilde{s}_0; \bar{\mathbf{P}}, \bar{\mathbf{Q}}) + \left(\tilde{p}_2 \frac{\partial}{\partial \tilde{p}_0} + \tilde{s}_2 \frac{\partial}{\partial \tilde{s}_0} + \tilde{Q}_2 \cdot \frac{\partial}{\partial \bar{\mathbf{Q}}} \right) \tilde{H}_0(\tilde{p}_0, \tilde{s}_0; \bar{\mathbf{P}}, \bar{\mathbf{Q}}) \\
 &\quad - \frac{q}{2m} \left(\frac{B(\bar{\mathbf{P}}, \bar{\mathbf{Q}})}{\mu_0 G_0} \right)^2 \iota_\Psi(\bar{P}_\Theta/q) \tilde{s}_0^2 \left(\tilde{\mathbf{P}}_2 \cdot \frac{\partial}{\partial \bar{\mathbf{P}}} \right) (\bar{p}_\varphi + q\chi(\bar{P}_\Theta/q)) \\
 &\quad - \frac{I}{m} \left(\frac{B(\bar{\mathbf{P}}, \bar{\mathbf{Q}})}{\mu_0 G_0} \right)^2 \left(\frac{B_\Psi(\bar{\mathbf{P}}, \bar{\mathbf{Q}}) \tilde{s}_0 + B_\Theta(\bar{\mathbf{P}}, \bar{\mathbf{Q}}) \tilde{p}_0 / q}{B(\bar{\mathbf{P}}, \bar{\mathbf{Q}})} \right) \left(\tilde{\mathbf{P}}_2 \cdot \frac{\partial}{\partial \bar{\mathbf{P}}} \right) (\bar{p}_\varphi + q\chi(\bar{P}_\Theta/q))^2 \\
 &\quad + \frac{I}{m} \left(\frac{B(\bar{\mathbf{P}}, \bar{\mathbf{Q}})}{\mu_0 G_0} \right)^2 \left[\tilde{\mathbf{P}}_2 \cdot \frac{\partial}{\partial \bar{\mathbf{P}}} (\bar{p}_\varphi + q\chi(\bar{P}_\Theta/q)) \right] \left[P_2(\tilde{p}_0, \tilde{s}_0; \bar{\mathbf{P}}, \bar{\mathbf{Q}}) \cdot \frac{\partial}{\partial \bar{\mathbf{P}}} (\bar{p}_\varphi + q\chi(\bar{P}_\Theta/q)) \right].
 \end{aligned} \tag{4-A.3}$$

Appendix 4-B: Derivation of Equation (4-4.4)

The generating function, which changes $(p, s; \mathbf{P}, Q)$ to $(\tilde{p}, \tilde{s}; \tilde{\mathbf{P}}, \tilde{Q})$, is a function of the old coordinates and the new momenta,

$$W^I(\tilde{p}, s; \tilde{\mathbf{P}}, Q) = \frac{l}{\epsilon} \tilde{\mathbf{P}} \cdot Q + \epsilon \left\{ \tilde{p} \left[s - \frac{l}{2} \tilde{p} f(\tilde{\mathbf{P}}, Q) \right] \right\}, \quad (4-B.1)$$

with $f(\tilde{\mathbf{P}}, Q) = g^c(\tilde{\mathbf{P}}, Q) / qg^s(\tilde{\mathbf{P}}, Q)$, which is of order l/ϵ . Therefore

$$\epsilon p = \frac{\partial W^I}{\partial s} = \epsilon \tilde{p},$$

$$\tilde{s} = \frac{\partial W^I}{\partial \tilde{p}} = s - \tilde{p} f(\tilde{\mathbf{P}}, Q),$$

$$\frac{l}{\epsilon} \mathbf{P} = \frac{\partial W^I}{\partial Q} = \frac{l}{\epsilon} \tilde{\mathbf{P}} - \epsilon \frac{l}{2} \tilde{p}^2 \frac{\partial f(\tilde{\mathbf{P}}, Q)}{\partial Q},$$

$$\tilde{Q} = \frac{\partial W^I}{\partial \tilde{\mathbf{P}}} = Q - \epsilon^2 \frac{l}{2} \tilde{p}^2 \frac{\partial f(\tilde{\mathbf{P}}, Q)}{\partial \tilde{\mathbf{P}}}.$$

(4-B.2)

We need the old variables to be expressed as functions of the new variables. To invert Eq.(4-B.2), we assume

$$s(\tilde{p}, \tilde{s}; \tilde{\mathbf{P}}, \tilde{Q}) = s_0(\tilde{p}, \tilde{s}; \tilde{\mathbf{P}}, \tilde{Q}) + \epsilon s_1(\tilde{p}, \tilde{s}; \tilde{\mathbf{P}}, \tilde{Q}) + \epsilon^2 s_2(\tilde{p}, \tilde{s}; \tilde{\mathbf{P}}, \tilde{Q}) + o(\epsilon^3),$$

$$\frac{l}{\epsilon} \mathbf{P}(\tilde{p}, \tilde{s}; \tilde{\mathbf{P}}, \tilde{Q}) = \frac{l}{\epsilon} \tilde{\mathbf{P}} + P_l(\tilde{p}, \tilde{s}; \tilde{\mathbf{P}}, \tilde{Q}) + \epsilon P_2(\tilde{p}, \tilde{s}; \tilde{\mathbf{P}}, \tilde{Q}) + o(\epsilon^2),$$

$$Q(\tilde{p}, \tilde{s}; \tilde{\mathbf{P}}, \tilde{Q}) = \tilde{Q} + \epsilon Q_1(\tilde{p}, \tilde{s}; \tilde{\mathbf{P}}, \tilde{Q}) + \epsilon^2 Q_2(\tilde{p}, \tilde{s}; \tilde{\mathbf{P}}, \tilde{Q}) + o(\epsilon^3).$$

(4-B.3)

We expand Eq.(4-B.2) in Taylor series,

$$Q = \tilde{Q} + \varepsilon^2 \frac{1}{2} \tilde{p}^2 \frac{\partial f(\tilde{P}, \tilde{Q})}{\partial \tilde{P}} + o(\varepsilon^4).$$

Obviously Q_1 and the term of $o(\varepsilon^3)$ are zero. Thus

$$\frac{l}{\varepsilon} P = \frac{l}{\varepsilon} \tilde{P} - \varepsilon \frac{1}{2} \tilde{p}^2 \frac{\partial f(\tilde{P}, \tilde{Q})}{\partial \tilde{Q}} + o(\varepsilon^3),$$

$$s = \tilde{s} + \tilde{p}f(\tilde{P}, \tilde{Q}) + \varepsilon^2 \tilde{p}Q_2 \cdot \frac{\partial f(\tilde{P}, \tilde{Q})}{\partial \tilde{Q}} + o(\varepsilon^4).$$

Compare with Eq.(4-B.3), we have

$$s_0 = \tilde{s} + \tilde{p}f(\tilde{P}, \tilde{Q}),$$

$$s_2 = \frac{1}{2} \tilde{p}^3 \frac{\partial f(\tilde{P}, \tilde{Q})}{\partial \tilde{P}} \cdot \frac{\partial f(\tilde{P}, \tilde{Q})}{\partial \tilde{Q}},$$

$$P_2 = -\frac{1}{2} \tilde{p}^2 \frac{\partial f(\tilde{P}, \tilde{Q})}{\partial \tilde{Q}},$$

$$Q_2 = \frac{1}{2} \tilde{p}^2 \frac{\partial f(\tilde{P}, \tilde{Q})}{\partial \tilde{P}}.$$

(4-B.4)

and $P_1 = s_0 = 0$.

Appendix 4-C: Proof of the First Order Correction to the Standard Drift Hamiltonian

The first order correction to the standard drift Hamiltonian is

$$V = \langle \bar{H}_I \rangle, \quad (4-C.1)$$

with

$$\bar{H}_I = \bar{H}_I(\bar{p}_0, \bar{s}_0; \bar{P}, \bar{Q}) + \frac{I}{2m} \left(\frac{B(\bar{P}, \bar{Q})}{\mu_0 G_0} \right)^2 \left(\tilde{P}_2 \cdot \frac{\partial}{\partial \bar{P}} \right) [\bar{p}_\varphi + q\chi(\bar{P}_\theta/q)]^2. \quad (4-C.2)$$

In the second term only \tilde{P}_2 dependent on the gyrophase.

$$\tilde{P}_2 = \frac{\partial}{\partial \bar{Q}} W_I(J, \bar{s}_0; \bar{P}, \bar{Q}), \quad (4-C.3)$$

where

$$W_I = \int \sqrt{2JR(\bar{P}, \bar{Q}) - R^2(\bar{P}, \bar{Q})\bar{s}^2} d\bar{s},$$

$$\bar{s}_0 = \sin \theta_g \sqrt{2J/R(\bar{P}, \bar{Q})}.$$

We have

$$\tilde{P}_2 = \bar{s}_0 \sqrt{JR/2} \sqrt{1 - \bar{s}_0^2} \frac{R}{2J} \frac{\partial R/\partial \bar{Q}}{R} = J \frac{\partial R/\partial \bar{Q}}{R} \sin \theta_g \cos \theta_g, \quad (4-C.4)$$

$$\langle \tilde{P}_2 \rangle = 0.$$

Thus

$$\langle \bar{H}_I \rangle = \langle \tilde{H}_I(\tilde{p}_0, \tilde{s}_0; \bar{\mathbf{P}}, \bar{\mathbf{Q}}) \rangle. \quad (4-C.5)$$

From Eq.(4-4.5c),

$$\begin{aligned} \tilde{H}_I(\tilde{p}_0, \tilde{s}_0; \bar{\mathbf{P}}, \bar{\mathbf{Q}}) &= H_I(\tilde{p}_0, s_0(\tilde{p}_0, \tilde{s}_0; \bar{\mathbf{P}}, \bar{\mathbf{Q}}); \bar{\mathbf{P}}, \bar{\mathbf{Q}}) \\ &\quad + \frac{1}{2m} \left(\frac{B(\bar{\mathbf{P}}, \bar{\mathbf{Q}})}{\mu_0 G_0} \right)^2 \left(P_2(\tilde{p}_0; \bar{\mathbf{P}}, \bar{\mathbf{Q}}) \cdot \frac{\partial}{\partial \bar{\mathbf{P}}} \right) [\bar{p}_\phi + q\chi(\bar{P}_\theta/q)]^2. \end{aligned} \quad (4-C.6)$$

In H_I (see Eq.(4-4.2c)), the only term being non-zero after average over θ_g is

$$-\frac{q}{2m} \left(\frac{B(\bar{\mathbf{P}}, \bar{\mathbf{Q}})}{\mu_0 G_0} \right)^2 \iota_\Psi(\bar{P}_\theta/q)(\bar{p}_\phi + q\chi(\bar{P}_\theta/q)) s_0^2(\tilde{p}_0, \tilde{s}_0; \bar{\mathbf{P}}, \bar{\mathbf{Q}}).$$

And

$$\langle s_0^2(\tilde{p}_0, \tilde{s}_0; \bar{\mathbf{P}}, \bar{\mathbf{Q}}) \rangle = \langle (\tilde{s}_0 + \tilde{p}_0 f(\bar{\mathbf{P}}, \bar{\mathbf{Q}}))^2 \rangle = J \left[\frac{1}{R} + R f^2(\bar{\mathbf{P}}, \bar{\mathbf{Q}}) \right].$$

Since $R(\bar{\mathbf{P}}, \bar{\mathbf{Q}}) = qg^s(\bar{\mathbf{P}}, \bar{\mathbf{Q}})/B(\bar{\mathbf{P}}, \bar{\mathbf{Q}})$,

$$f(\bar{\mathbf{P}}, \bar{\mathbf{Q}}) = g^c(\bar{\mathbf{P}}, \bar{\mathbf{Q}})/qg^s(\bar{\mathbf{P}}, \bar{\mathbf{Q}}) = \frac{1}{R} \frac{g^c}{B}.$$

Hence

$$\langle s_0^2 \rangle = J \left(\frac{1}{R} + R f^2 \right) = \frac{J}{R} \left(1 + \frac{g^c}{B^2} \right) = \frac{J}{qg^s/B} \left(\frac{B^2 + g^c}{B^2} \right).$$

By using Eq.(4-3.9), we have

$$\langle s_0^2 \rangle = J \frac{g^s}{qB}.$$

Thus

$$\langle H_I \rangle = -J \frac{B(\bar{P}, \bar{Q})}{2m(\mu_0 G_0)^2} (\bar{p}_\varphi + q\chi(\bar{P}_\Theta/q)) \iota_\psi(\bar{P}_\Theta/q) g^\psi(\bar{P}, \bar{Q}). \quad (4-C.7)$$

In the second term of Eq.(4-C.6), only $P_2(\tilde{p}_0; \bar{P}, \bar{Q})$ dependent on θ_g . With Eq.(4-4.4') and Eq.(4-4.7'),

$$\langle P_2(\tilde{p}_0; \bar{P}, \bar{Q}) \rangle = -\frac{1}{2} \frac{\partial f(\bar{P}, \bar{Q})}{\partial \bar{Q}} \langle \tilde{p}_0^2 \rangle = -\frac{1}{2} JR(\bar{P}, \bar{Q}) \frac{\partial f(\bar{P}, \bar{Q})}{\partial \bar{Q}}.$$

The second term becomes

$$\begin{aligned} & -\frac{J}{2m} \left(\frac{B}{\mu_0 G_0} \right)^2 (\bar{p}_\varphi + q\chi) R \left(\iota \frac{\partial f}{\partial \Theta} + \frac{\partial f}{\partial \varphi} \right) \\ &= -J \frac{B}{2m(\mu_0 G_0)^2} (\bar{p}_\varphi + q\chi) \left[\left(\iota \frac{\partial}{\partial \Theta} + \frac{\partial}{\partial \varphi} \right) g^c - \frac{g^c}{g^s} \left(\iota \frac{\partial}{\partial \Theta} + \frac{\partial}{\partial \varphi} \right) g^s \right]. \end{aligned} \quad (4-C.8)$$

Therefore, we have

$$\langle \tilde{H}_I \rangle = -J \frac{B}{2m(\mu_0 G_0)^2} (\bar{p}_\varphi + q\chi) \left[\iota_\psi g^\psi + \left(\iota \frac{\partial}{\partial \Theta} + \frac{\partial}{\partial \varphi} \right) g^c - \frac{g^c}{g^s} \left(\iota \frac{\partial}{\partial \Theta} + \frac{\partial}{\partial \varphi} \right) g^s \right]. \quad (4-C.9)$$

Appendix 4-D: The Hamiltonian Near the Magnetic Axis

In the region that near the magnetic axis, certain metric tensor elements become singular in the magnetic coordinates (ψ, θ, φ) . We therefore introduce a pseudo-cartesian magnetic coordinates (ζ, γ, φ) which is closely related to Boozer coordinates (ψ, θ, φ) . These two coordinates are related by a canonical transformation. The generating function of this transformation is

$$F(\theta, \gamma) = -\frac{1}{2} \gamma^2 \cot(2\pi\theta). \quad (4-D.1)$$

The relations between (ζ, γ, φ) and (ψ, θ, φ) are $\psi = -\partial F / \partial \theta = \pi \gamma^2 / \sin^2(2\pi\theta)$ and $\zeta = -\partial F / \partial \gamma = \gamma \cot(2\pi\theta)$, or more explicitly

$$\begin{cases} \zeta = \sqrt{\psi/\pi} \cos(2\pi\theta) \\ \gamma = \sqrt{\psi/\pi} \sin(2\pi\theta) \end{cases}, \quad (4-D.2a)$$

and

$$\begin{cases} \psi = \pi(\zeta^2 + \gamma^2) \\ \theta = \frac{1}{2\pi} \tan^{-1}\left(\frac{\gamma}{\zeta}\right) \end{cases}. \quad (4-D.2b)$$

The contravariant representation of the magnetic field is

$$\mathbf{B} = \nabla \zeta \times \nabla \gamma + \nabla \varphi \times \nabla \chi. \quad (4-D.3)$$

The vector potential has the covariant form

$$\mathbf{A} = \zeta \nabla \gamma - \chi \nabla \varphi. \quad (4-D.4)$$

The covariant representation of the field is

$$\mathbf{B} = \mu_0 \left(G \nabla \varphi + \tilde{I} \nabla \gamma + \tilde{\beta}_* \cdot \nabla \zeta \right), \quad (4\text{-D.5})$$

with $\tilde{I} = \frac{\zeta}{2\pi(\zeta^2 + \gamma^2)} I + 2\pi\gamma\beta_*$ and $\tilde{\beta}_* = \frac{-\gamma}{2\pi(\zeta^2 + \gamma^2)} I + 2\pi\zeta\beta_*$. In a curl-free field, the covariant form remain the same as Eq.(4-2.5), $\mathbf{B} = \mu_0 G_0 \nabla \varphi$.

Now we introduce two vectors,

$$\mathbf{e}_\zeta = \nabla \zeta + 2\pi i \gamma \nabla \varphi, \quad (4\text{-D.6a})$$

and

$$\mathbf{e}_\gamma = \nabla \gamma - 2\pi i \zeta \nabla \varphi. \quad (4\text{-D.6b})$$

Observe that $\mathbf{B} = \mathbf{e}_\zeta \times \mathbf{e}_\gamma$. The velocity has the covariant form

$$\mathbf{v} = v_{||} \frac{\mathbf{B}}{B} + v_\zeta \mathbf{e}_\zeta + v_\gamma \mathbf{e}_\gamma. \quad (4\text{-D.7})$$

The canonical coordinates corresponding to the pseudo-cartesian magnetic coordinates are:

$$\zeta, \quad p_\zeta = mv_\zeta;$$

$$\gamma, \quad p_\gamma = mv_\gamma + q\zeta;$$

$$\varphi, \quad p_\varphi = \frac{\mu_0 G_0}{B} mv_{||} - 2\pi i (\zeta mv_\gamma - \gamma mv_\zeta) - q\chi. \quad (4\text{-D.8})$$

After a canonical transformation using the generating function

$$W = sp_\zeta - p_\zeta p_\gamma/q - \Gamma p_\gamma - \varphi p_\varphi,$$

the new canonical coordinates in the guiding center coordinates are

$$\begin{aligned} s &= p_\gamma/q - \zeta, & p &= -p_\zeta; \\ \Gamma &= \gamma - p_\zeta/q, & P_\Gamma (= qZ) &= p_\gamma; \\ \varphi, & & p_\varphi. & \end{aligned} \tag{4-D.9}$$

The relation between the position of the particle $\xi = (\zeta, \gamma, \varphi)$ and the position of the guiding center $\Xi = (Z, \Gamma, \varphi)$ is $\xi = \Xi - \delta$, with $\delta = (s, p/q, 0)$. It can be trivially shown by Taylor expansion that

$$X(Z, \Gamma, \varphi) = x(\zeta, \gamma, \varphi) + \varepsilon \frac{mv \times B}{qB^2} + o(\varepsilon^2).$$

The Hamiltonian of the exact trajectory in the pseudo-cartesian magnetic coordinates is

$$H_e = \frac{1}{2m} \left(\frac{B}{\mu_0 G_0} \right)^2 [p_\varphi + q\chi + 2\pi i(\zeta qs + \gamma p)]^2 + \frac{1}{2m} (q^2 G^s s^2 - 2qG^c sp + G^p p^2), \tag{4-D.10}$$

where $G^s = \mathbf{e}_\gamma \cdot \mathbf{e}_\gamma$, $G^p = \mathbf{e}_\zeta \cdot \mathbf{e}_\zeta$ and $G^c = \mathbf{e}_\zeta \cdot \mathbf{e}_\gamma$. The relation between the field strength and the metrics is similar to Eq.(4-3.9),

$$B^2 = G^s G^p - (G^c)^2.$$

After expansion and canonical transformation similar to those we did in Sec. 4-4, we obtain the drift Hamiltonian with first order correction in the pseudo-cartesian magnetic coordinates,

$$K_d = H_d + \varepsilon H_I, \quad (4\text{-D.11a})$$

with

$$H_d = \frac{l}{2m} \left(\frac{B(\Xi)}{\mu_0 G_0} \right)^2 [p_\varphi + q\chi(\Psi)]^2 + \frac{q}{m} J B(\Xi) \quad (4\text{-D.11b})$$

and

$$H_I = [p_\varphi + q\chi(\Psi)]\Delta, \quad (4\text{-D.11c})$$

where

$$\begin{aligned} \Delta = & -J \frac{B(\Xi)}{2m(\mu_0 G_0)^2} \left\{ 2\pi \left[G^p(\Xi) \left(l + 2\pi l_\Psi Z^2 \right) + G^s(\Xi) \left(l + 2\pi l_\Psi \Gamma^2 \right) + 4\pi l_\Psi Z \Gamma G^c \right] \right. \\ & \left. + \left(2\pi l Z \frac{\partial}{\partial \Gamma} + \frac{\partial}{\partial \varphi} \right) G^c(\Xi) - \frac{G^c(\Xi)}{G^s(\Xi)} \left(2\pi l Z \frac{\partial}{\partial \Gamma} + \frac{\partial}{\partial \varphi} \right) G^s(\Xi) \right\}. \end{aligned} \quad (4\text{-D.11d})$$

with $\Xi = (Z, \Gamma, \varphi)$, $P_\Gamma = qZ$ and $\Psi = \pi(Z^2 + \Gamma^2)$.

The Hamiltonian we give in this appendix is not only valid in the region near the magnetic axis but in all space. However, the matrices in the pseudo-cartesian magnetic coordinates are much more complicate than those in Boozer coordinates. One can generally avoid using the Hamiltonian in this pseudo-cartesian magnetic coordinates since the particle trajectories near the magnetic axis is very interested in plasma confinement.

References

- 1 H. Alfvén, Ark. Mat. Astron. Fys. **27A**, 1 (1940)
- 2 C. S. Gardner, Phys. Rev. **115**, 791 (1959); Phys. Fluids **9**, 1997 (1966)
- 3 T. G. Northrop and J. A. Rome, Phys. Fluids **21**, 384 (1978)
- 4 H. E. Mynick, Phys. Rev. Lett. **43**, 1019 (1979); Phys. Fluids **23**, 1888 (1980)
- 5 R. G. Littlejohn, J. Math. Phys. **20**, 2445 (1979); Phys. Fluids **24**, 1730 (1981)
- 6 S. Hamada, Nucl. Fusion **2**, 23 (1962); J. M. Greene and J. L. Johnson, phys. Fluids **5**, 510 (1962)
- 7 A. H. Boozer, Phys. Fluids **24**, 1999 (1981); Phys. Fluids **25**, 520 (1982)
- 8 A. H. Boozer, Phys. Fluids **27**, 2441 (1984)
- 9 J. Nührenberg and R. Zille, Phys. Lett. **A129**, 113 (1988)
- 10 D. A. Garren and A. H. Boozer, Phys. Fluids **B3**, 2822 (1991)
- 11 J. B. Taylor, Phys. Fluids **7**, 767 (1964)
- 12 A. H. Boozer, Phys. Fluids **23**, 904 (1980); Phys. Fluids **25**, 575 (1982); Phys. Fluids **26**, 496 (1983); Phys. Fluids **B3**, 875 (1991)
- 13 G. Kuo-Petravic, A. H. Boozer, J. A. Rome and R. H. Fowler, J. Comp. Phys. **51**, 261 (1983)
- 14 A. J. Lichtenberg and M. A. Lieberman, *Regular and Stochastic Motion* (Springer-Verlag, New York Inc., 1983)
- 15 A. Deprit, Cel. Mech. **1**, 12 (1969)

CHAPTER 5

THE HIGHER ORDER DRIFT HAMILTONIAN IN A FULL FIELD

In this chapter the formulation developed in chapter 4 is extended into a time-independent full electromagnetic field with plasma equilibrium. A noncanonical Hamiltonian method, which differs from that in chapter 4, is used.

5-1 The Exact Hamiltonian in Boozer Coordinates

In Boozer coordinates an magnetic field with plasma equilibrium has both simple contravariant representation

$$\mathbf{B} = \nabla\psi \times \nabla\theta + \nabla\varphi \times \nabla\chi(\psi), \quad (5-1.1)$$

and simple covariant representation¹

$$\mathbf{B} = \mu_0[G(\psi)\nabla\varphi + I(\psi)\nabla\theta + \beta_*(\psi, \theta, \varphi)\nabla\psi]. \quad (5-1.2)$$

The vector potential corresponding to Eq. (5-1.1) is

$$\mathbf{A} = \psi \nabla \theta - \chi \nabla \varphi. \quad (5-1.3)$$

Generally β_* is much smaller than $G + U$; $\beta_*/(G + U)$ is of order 1%. This important feature will be used later when we define our guiding center position. In the remaining part of this chapter we occasionally refer to the Boozer coordinates (ψ, θ, φ) as ξ or ξ^i , with $i = 1, 2, 3$.

The Hamiltonian of the exact particle trajectories, of mass m and charge e (a different notation q was used in chapter 4), in a static electromagnetic field is

$$H = \frac{1}{2} m v^2 + e \phi, \quad (5-1.4)$$

where ϕ is electric potential. In ordinary spatial coordinates \mathbf{x} , the canonical momenta are

$$\mathbf{p}_x = m \mathbf{v} + e \mathbf{A}.$$

In Boozer coordinates (ψ, θ, φ) , the velocity has a covariant representation

$$\mathbf{v} = \frac{v_{||}}{B} \mathbf{B} + v_s \mathbf{e}_s + v_\psi \nabla \psi, \quad (5-1.5)$$

with \mathbf{B} in the covariant representation of Eq. (5-1.2). The vector \mathbf{e}_s is defined as

$$\mathbf{e}_s = \nabla \theta - \iota(\psi) \nabla \varphi, \quad (5-1.6)$$

which lies roughly within the magnetic surfaces. The vectors \mathbf{e}_s and $\nabla\psi$ are both perpendicular to field lines, but are not necessarily perpendicular to each other. The contravariant representation of the magnetic field can also be written as

$$\mathbf{B} = \nabla\psi \times \mathbf{e}_s. \quad (5-1.7)$$

In Boozer coordinates ξ , the canonical momenta are $p_i = \mathbf{p}_x \cdot (\partial\mathbf{x}/\partial\xi^i)$. By using the covariant representation of the velocity and the vector potential, Eq. (5-1.5) and Eq. (5-1.3), and the dual relation $(\partial\mathbf{x}/\partial\xi^i) \cdot \nabla\xi^j = \delta_i^j$, the canonical momenta can be trivially found. They are

$$\begin{aligned} p_\psi &= \mu_0 \frac{\beta_*}{B} V - u, \\ p_\theta &= \mu_0 \frac{I}{B} V + s + e\psi \\ p_\varphi &= \mu_0 \frac{G}{B} V - ls - e\chi, \end{aligned} \quad (5-1.8)$$

where

$$V = mv_{||},$$

$$u = -mv_\psi,$$

$$s = mv_s.$$

(5-1.8')

The notations we use here are different from those of chapter 4. In chapter 4, $p = -mv_\psi$ and $s = mv_s/e$, p and s were canonical to each other in the case of vacuum field. Here u and s are not canonical any more. This noncanonical property hampers the canonical treatment of the higher order drift Hamiltonian and this is why the noncanonical Hamiltonian method is employed in deriving the higher order drift Hamiltonian.

The Hamiltonian in terms of s , u , V , ψ , θ and φ is

$$H = \frac{1}{2m}V^2 + \frac{1}{2m}(g^s s^2 - 2g^c s u + g^\psi u^2) + e\phi, \quad (5-1.9)$$

where the components of the metric tensor are $g^s = \mathbf{e}_s \cdot \mathbf{e}_s$, $g^c = \mathbf{e}_s \cdot \nabla \psi$ and $g^\psi = \nabla \psi \cdot \nabla \psi$. These metric components and the electric potential ϕ are functions of Boozer coordinates (ψ, θ, φ) . By using Eq. (5-1.7) one can show that

$$B^2 = g^s g^\psi - (g^c)^2. \quad (5-1.10)$$

Eq. (5-1.8) can easily be inverted so that s , u , V are functions of the canonical variables $\mathbf{p} = \mathbf{p}_\xi$ and $\mathbf{q} = \boldsymbol{\xi}$. The canonical Hamiltonian can thus be obtained.

If the Hamiltonian has no obvious symmetry in the canonical coordinates $(\mathbf{p}_\xi, \boldsymbol{\xi})$, there is no disadvantage to having the Hamiltonian in the noncanonical coordinates $(s, u, V, \psi, \theta, \varphi)$. Therefore we also calculate the Poisson tensor here. The action forms, according to the definition of Eq. (3-3.3), $\rho_i = p_l \frac{\partial q^l}{\partial z^i}$, are

$$\rho_\psi = p_\psi,$$

$$\rho_\theta = p_\theta,$$

$$\rho_\phi = p_\phi$$

(5-1.11)

and $\rho_s = \rho_u = \rho_v = 0$. p_ψ , p_θ and p_ϕ are defined in Eq. (5-1.8). The Lagrange tensor is easily derived from the action forms using Eq. (3-3.5),

$$\omega_{ij} = \frac{\partial p_l}{\partial z^i} \frac{\partial q^l}{\partial z^j} - \frac{\partial q^l}{\partial z^i} \frac{\partial p_l}{\partial z^j} = \frac{\partial \rho_j}{\partial z^i} - \frac{\partial \rho_i}{\partial z^j}.$$

The Lagrange brackets are

$$\omega_{s\theta} = -\omega_{\theta s} = l,$$

$$\omega_{s\phi} = -\omega_{\phi s} = -l,$$

$$\omega_{u\psi} = -\omega_{\psi u} = -l,$$

$$\omega_{v\psi} = -\omega_{\psi v} = \mu_0 \frac{\beta_*}{B},$$

$$\omega_{v\theta} = -\omega_{\theta v} = \mu_0 \frac{I}{B},$$

$$\omega_{v\varphi} = -\omega_{\varphi v} = \mu_0 \frac{G}{B},$$

$$\omega_{v\theta} = -\omega_{\theta v} = e + \mu_0 \left[\left(\frac{I}{B} \right)_\varphi - \left(\frac{\beta_*}{B} \right)_\theta \right] V,$$

$$\omega_{v\varphi} = -\omega_{\varphi v} = -eI - \iota_v s + \mu_0 \left[\left(\frac{G}{B} \right)_\varphi - \left(\frac{\beta_*}{B} \right)_\varphi \right] V,$$

$$\omega_{\theta\varphi} = -\omega_{\varphi\theta} = \mu_0 \left[\left(\frac{G}{B} \right)_\theta - \left(\frac{I}{B} \right)_\varphi \right] V,$$
(5-1.12)

and others are zeros. In the above equations and the remainder of this chapter $()_{\xi^i}$ is $\partial()/\partial\xi^i$ and f_{ξ^i} is $\partial f/\partial\xi^i$ with f any function. Inverting the Lagrange tensor we obtain the Poisson tensor. The Poisson brackets are

$$J^{vu} = -J^{uv} = e + \frac{\iota_v I}{G+U} s + \frac{\mu_0}{G+U} \left[G \left(\frac{I}{B} \right)_\varphi - I \left(\frac{G}{B} \right)_\varphi - G \left(\frac{\beta_*}{B} \right)_\theta + I \left(\frac{\beta_*}{B} \right)_\varphi + \beta_* \left(\frac{G}{B} \right)_\theta - \beta_* \left(\frac{I}{B} \right)_\varphi \right] V,$$

$$J^{v\theta} = -J^{\theta v} = \frac{B}{G+U} \left[\left(\frac{G}{B} \right)_\theta - \left(\frac{I}{B} \right)_\varphi \right] V,$$

$$J^{v\theta} = -J^{\theta v} = -\frac{G}{G+U},$$

$$J^{v\varphi} = -J^{\varphi v} = \frac{I}{G+U},$$

$$J^{uV} = -J^{Vu} = \frac{\iota_\psi B}{\mu_0(G+u)} s - \frac{B}{G+u} \left[\left(\frac{G}{B} \right)_\psi + \iota \left(\frac{I}{B} \right)_\psi - \iota \left(\frac{\beta_*}{B} \right)_\theta - \left(\frac{\beta_*}{B} \right)_\varphi \right] V,$$

$$J^{u\psi} = -J^{\psi u} = I,$$

$$J^{u\theta} = -J^{\theta u} = -\iota \frac{\beta_*}{G+u},$$

$$J^{u\varphi} = -J^{\varphi u} = -\frac{\beta_*}{G+u},$$

$$J^{v\theta} = -J^{\theta v} = -\iota \frac{B}{\mu_0(G+u)},$$

$$J^{v\varphi} = -J^{\varphi v} = -\frac{B}{\mu_0(G+u)},$$

(5-1.13)

and others are zero. The equation of motion in the noncanonical coordinates $(s, u, V, \psi, \theta, \varphi)$ can be trivially derived by using Eq. (3-3.1).

5-2 Higher Order Drift Hamiltonian

In this section we first give an outline of how to obtain the drift Hamiltonian with the first order correction. Secondly, we give the details of the coordinate transformation from the particle position to the guiding center position using Darboux's theorem.² Then, we calculate the higher order drift Hamiltonian. Finally, we calculate the Poisson brackets in

the new coordinates.

To derive the higher order drift Hamiltonian, we start with Eq. (5-1.8) and Eq. (5-1.9). The two variables s and u are the momenta in perpendicular directions. We now transform these two coordinates into gyrophase θ_s and another variable L , which is equivalent to the instantaneous magnetic moment. The transformation is as follows,

$$\begin{aligned} s &= \sqrt{2mL} \left[\sqrt{\frac{B}{g^s}} \sin(\theta_s + \alpha) + \frac{g^c}{g^s} \sqrt{\frac{g^s}{B}} \cos(\theta_s + \alpha) \right], \\ u &= \sqrt{2mL} \sqrt{\frac{g^s}{B}} \cos(\theta_s + \alpha). \end{aligned} \tag{5-2.1}$$

In Eq. (5-2.1) the initial phase $\alpha(\xi)$ is an arbitrarily well behaved function. The freedom of choosing $\alpha(\xi)$ is the freedom of choosing a gauge. It is worthwhile to point out that the transformation of Eq. (5-2.1) is defined globally instead of locally. The relation between the canonical variables (p, q) and the new variables $(\theta_s, L, V, \psi, \theta, \varphi)$ is $q = \xi$ and

$$\begin{aligned} p_\psi &= \mu_0 \frac{\beta_*}{B} V - u(\theta_s, L, \psi, \theta, \varphi), \\ p_\theta &= \mu_0 \frac{I}{B} V + s(\theta_s, L, \psi, \theta, \varphi) + e\psi, \\ p_\varphi &= \mu_0 \frac{G}{B} V - us(\theta_s, L, \psi, \theta, \varphi) - e\chi, \end{aligned} \tag{5-2.2}$$

where $s(\theta_g, L, \psi, \theta, \varphi)$ and $u(\theta_g, L, \psi, \theta, \varphi)$ are given by Eq. (5-2.1). The Hamiltonian in these particle position coordinates $(\theta_g, L, V, \psi, \theta, \varphi)$ is

$$H = \frac{I}{2m}V^2 + LB(\xi) + e\phi(\xi). \quad (5-2.3)$$

Although the Hamiltonian is independent of the gyrophase θ_g , the quantity L is not a constant of motion due to cross term of the Poisson brackets J^{ij} .

In the next step we perform a coordinate transformation using Darboux's theorem to decouple the motion of θ_g and L from other variables. Here is the outline of the strategy. We transform $(\theta_g, L, V, \psi, \theta, \varphi)$ to $(\theta_g, \mu, \bar{V}, \Psi, \Theta, \Phi)$ such that the new variables satisfy the following equations,

$$\{\theta_g, \mu\} = \frac{e}{m}, \quad (5-2.4a)$$

$$\{\theta_g, \bar{V}\} = \{\theta_g, \Xi^i\} = 0, \quad (5-2.4b)$$

$$\{\mu, \bar{V}\} = \{\mu, \Xi^i\} = 0, \quad (5-2.4c)$$

where Ξ^i are Ψ, Θ and Φ as $i = 1, 2, 3$. We will show later that (Ψ, Θ, Φ) gives the position of the guiding center. By solving the above equations, we get

$$L = \mu + \varepsilon L_I(\theta_g, \mu, \bar{V}, \Xi) + o(\varepsilon^2),$$

$$V = \bar{V} + \varepsilon V_I(\theta_g, \mu, \bar{V}, \Xi) + o(\varepsilon^2),$$

$$\xi = \Xi + \varepsilon \xi_1(\theta_g, \mu, \bar{V}, \Xi) + \varepsilon^2 \xi_2(\theta_g, \mu, \bar{V}, \Xi) + o(\varepsilon^3). \quad (5-2.5)$$

The Hamiltonian in coordinates $(\theta_g, \mu, \bar{V}, \Psi, \Theta, \Phi)$ is

$$H = H_0 + \varepsilon H_I + o(\varepsilon^2), \quad (5-2.6)$$

with

$$\begin{aligned} H_0 &= \frac{l}{2m} \bar{V}^2 + \mu B(\Xi) + e\phi(\Xi), \\ H_I &= \frac{l}{m} V_I \bar{V} + L_I B(\Xi) + \mu \left(\xi_I^i \frac{\partial}{\partial \Xi^i} \right) B(\Xi) \\ &\quad + e \left[\left(\xi_I^i \frac{\partial}{\partial \Xi^i} \right) + \frac{l}{2} \left(\xi_I^i \frac{\partial}{\partial \Xi^i} \right)^2 + \left(\xi_2^i \frac{\partial}{\partial \Xi^i} \right) \right] \phi(\Xi). \end{aligned} \quad (5-2.6')$$

The Hamiltonian depends on the gyrophase θ_g , and the Poisson brackets between θ_g , μ and the other four variables are decoupled. A simple Lie transformation can eliminate the θ_g dependence in the Hamiltonian. The result of the Lie transformation is the drift Hamiltonian with the first order correction,

$$K_d(\mu; \bar{V}, \Psi, \Theta, \Phi) = H_d(\mu; \bar{V}, \Psi, \Theta, \Phi) + \varepsilon K_I(\mu; \bar{V}, \Psi, \Theta, \Phi), \quad (5-2.7)$$

with H_d the standard drift Hamiltonian and K_I the first order correction

$$H_d = \frac{I}{2m} \bar{V}^2 + \mu B(\Xi) + e\phi(\Xi),$$

$$K_I = \langle H_I \rangle, \quad (5-2.7')$$

where $\langle \rangle$ refers to average over gyrophase θ_s . Thus the first order correction to the standard drift Hamiltonian is

$$\begin{aligned} K_I = & \frac{I}{m} \langle V_I \rangle \bar{V} + \langle L_I \rangle B(\Xi) + \mu \left(\langle \xi_I^i \rangle \frac{\partial}{\partial \Xi^i} \right) B(\Xi) \\ & + e \left[\left(\langle \xi_I^i \rangle \frac{\partial}{\partial \Xi^i} \right) + \frac{I}{2} \left\langle \left(\xi_I^i \frac{\partial}{\partial \Xi^i} \right)^2 \right\rangle + \left(\langle \xi_2^i \rangle \frac{\partial}{\partial \Xi^i} \right) \right] \phi(\Xi). \end{aligned} \quad (5-2.8)$$

A. Darboux Transformation

In this part we first derive the coordinate transformation which satisfies Eq. (5-2.4a) and Eq. (5-2.4b). We will show later (in part C) that the new coordinates also satisfy Eq. (5-2.4c).

We now begin to perform the Darboux transformation. The characteristic curve of Eq. (5-2.4a) and Eq. (5-2.4b) satisfies³

$$\frac{dz^i}{d\tau} = \{z^i, \theta_s\}, \quad (5-2.9)$$

with $z^i(\theta_s, \mu, \bar{V}, \Psi, \Theta, \Phi)$ being each of the old coordinates $(\theta_s, L, V, \psi, \theta, \varphi)$ as $i = 1, \dots, 6$. The parameter τ is a time like variable and the gyrophase θ_s is the Hamiltonian of the

characteristic curve. Eq. (5-2.4a) and Eq. (5-2.4b) can also be written as

$$\frac{d\mu}{d\tau} = -\frac{e}{m}, \quad (5-2.10a)$$

$$\frac{d\bar{V}}{d\tau} = \frac{d\Xi^i}{d\tau} = 0. \quad (5-2.10b)$$

We eliminate the parameter τ from Eq. (5-2.9) and Eq. (5-2.10a), get

$$\frac{dz^i}{d\mu} = \frac{m}{e} \{ \theta_i, z^i \}. \quad (5-2.11)$$

By integrating Eq. (5-2.11) we can obtain the transformation from the variables $(\theta_i, L, V, \psi, \theta, \varphi)$ to the new variables $(\theta_i, \mu, \bar{V}, \Psi, \Theta, \Phi)$ which satisfy Eq.(5-2.4).

The Poisson brackets in Eq. (5-2.11) can easily be obtained by first calculating the Lagrange tensor from the action forms, then inverting the Lagrange tensor. The detailed procedure is given in Appendix 5-A. Here are the Poisson brackets needed,

$$\{ \theta_i, L \} = J_o^{\theta_i L} + \varepsilon J_i^{\theta_i L} + o(\varepsilon^2),$$

$$\{ \theta_i, V \} = \varepsilon J_i^{\theta_i V} + o(\varepsilon^2),$$

$$\{ \theta_i, \psi \} = \varepsilon J_i^{\theta_i \psi},$$

$$\{ \theta_i, \theta \} = \varepsilon J_i^{\theta_i \theta} + \varepsilon^2 J_2^{\theta_i \theta},$$

$$\{\theta_\epsilon, \varphi\} = \epsilon J_1^{\theta_\epsilon \varphi} + \epsilon^2 J_2^{\theta_\epsilon \varphi}, \quad (5-2.12)$$

with

$$J_0^{\theta_\epsilon L} = \frac{e}{m},$$

$$\begin{aligned} J_1^{\theta_\epsilon L}(\theta_\epsilon, L, \psi, \theta, \varphi) &= \frac{l}{m} \left(s_\psi + \frac{\iota_\psi I}{G+U} s + \frac{G}{G+U} u_\theta - \frac{I}{G+U} u_\varphi \right) \\ &\quad + \frac{V}{m} \frac{\mu_0}{G+U} \left[G \left(\frac{I}{B} \right)_\psi - I \left(\frac{G}{B} \right)_\psi \right], \end{aligned}$$

$$\begin{aligned} J_1^{\theta_\epsilon V}(\theta_\epsilon, L, \psi, \theta, \varphi) &= \frac{l}{m} \frac{B}{\mu_0(G+U)} \left[\iota(s_L u_\theta - u_L s_\theta) + (s_L u_\varphi - u_L s_\varphi) - \iota_\psi s s_L \right] \\ &\quad + \frac{V}{m} \frac{B}{G+U} \left\{ s_L \left[\left(\frac{G}{B} \right)_\psi + \iota \left(\frac{I}{B} \right)_\psi \right] + u_L \left[\left(\frac{G}{B} \right)_\theta - \left(\frac{I}{B} \right)_\varphi \right] \right\}, \end{aligned}$$

$$J_1^{\theta_\epsilon \psi}(\theta_\epsilon, L, \psi, \theta, \varphi) = -\frac{l}{m} s_L,$$

$$J_1^{\theta_\epsilon \theta}(\theta_\epsilon, L, \psi, \theta, \varphi) = -\frac{l}{m} \frac{G}{G+U} u_L,$$

$$J_2^{\theta_\epsilon \theta}(\theta_\epsilon, L, \psi, \theta, \varphi) = \frac{l}{m} \frac{\iota \beta_*}{G+U} s_L,$$

$$J_1^{\theta_\epsilon \varphi}(\theta_\epsilon, L, \psi, \theta, \varphi) = \frac{l}{m} \frac{I}{G+U} u_L,$$

$$J_2^{\theta_\epsilon \varphi}(\theta_\epsilon, L, \psi, \theta, \varphi) = \frac{l}{m} \frac{\beta_*}{G+U} s_L.$$

(5-2.13)

In ordering the Poisson brackets we used the assumption that β_* is much smaller than $G + \mu$, $\beta_*/(G + \mu)$ is of order ϵ . This assumption is true for toroidal fusion devices. We show in Appendix B that we can define the guiding center with equal accuracy by neglecting terms with β_* in order ϵ . However, by using this freedom of choosing the guiding center we will obtain a much simpler drift Hamiltonian.

To solve Eq. (5-2.11) perturbatively we expand $z^i(\theta_g, \mu, \bar{V}, \Psi, \Theta, \Phi)$ as $z^i = z_0^i + \epsilon z_1^i + \epsilon^2 z_2^i + \dots$. From Eq. (5-2.8) we know that we only need to find the transformation to order ϵ for L and V and to order ϵ^2 for ξ . We obtain the equations of transformation order by order,

$$\frac{dL_0}{d\mu} = 1, \text{ and } \frac{dV_0}{d\mu} = \frac{d\xi_0^i}{d\mu} = 0;$$

$$\frac{dz_I^i}{d\mu} = \frac{m}{e} J_I^{\theta_g z^i}(\theta_g, L_0, \xi_0);$$

$$\frac{d\xi_2^i}{d\mu} = \frac{m}{e} \left[J_2^{\theta_g \xi^i}(\theta_g, L_0, \xi_0) + L_I \frac{\partial}{\partial L_0} J_I^{\theta_g \xi^i}(\theta_g, L_0, \xi_0) + \xi_I^j \frac{\partial}{\partial \xi_0^j} J_I^{\theta_g \xi^i}(\theta_g, L_0, \xi_0) \right].$$
(5-2.14)

We choose the initial condition for Eq. (5-2.14) be that the particle position is at the guiding center when gyroradius is zero, which means

$$(L, V, \psi, \theta, \phi)|_{\mu=0} = (0, \bar{V}, \Psi, \Theta, \Phi). \quad (5-2.14')$$

After integrating Eq. (5-2.14), we obtain

$$L = \mu + \varepsilon L_I(\theta_g, \mu, \bar{V}, \Xi) + o(\varepsilon^2),$$

$$V = \bar{V} + \varepsilon V_I(\theta_g, \mu, \bar{V}, \Xi) + o(\varepsilon^2),$$

$$\xi = \Xi + \varepsilon \xi_I(\theta_g, \mu, \bar{V}, \Xi) + \varepsilon^2 \xi_2(\theta_g, \mu, \bar{V}, \Xi) + o(\varepsilon^3),$$

(5-2.15)

with

$$L_I = \frac{2}{3e} \mu \left(s_\varphi + \frac{\iota_\varphi I}{G+U} s + \frac{G}{G+U} u_\theta - \frac{I}{G+U} u_\phi \right) + \frac{I}{e} \mu \bar{V} \frac{\mu_o}{G+U} \left[G \left(\frac{I}{B} \right)_\varphi - I \left(\frac{G}{B} \right)_\varphi \right],$$

$$\begin{aligned} V_I &= \frac{I}{2e} \frac{B}{\mu_o(G+U)} \left[\iota(su_\theta - us_\theta) + (su_\phi - us_\phi) - \iota_\varphi s^2 \right] \\ &\quad + \frac{I}{e} \bar{V} \frac{B}{G+U} \left\{ s \left[\left(\frac{G}{B} \right)_\varphi + \iota \left(\frac{I}{B} \right)_\varphi \right] + u \left[\left(\frac{G}{B} \right)_\theta - \left(\frac{I}{B} \right)_\phi \right] \right\}, \end{aligned}$$

$$\psi_I = -\frac{I}{e} s,$$

$$\theta_I = -\frac{I}{e} \frac{G}{G+U} u,$$

$$\varphi_I = \frac{I}{e} \frac{I}{G+U} u;$$

$$\langle \psi_2 \rangle = \frac{I}{6e^2} \left[2 \langle s_\varphi^2 \rangle + \frac{\iota_\varphi I}{G+U} \langle s^2 \rangle + \frac{G}{G+U} (3 \langle s_\theta u \rangle + \langle su_\theta \rangle) - \frac{I}{G+U} (3 \langle s_\phi u \rangle + \langle su_\phi \rangle) \right],$$

$$\begin{aligned}
\langle \theta_2 \rangle &= \frac{I}{6e^2} \frac{G}{G+I} \left[(\langle s_\psi u \rangle + 3\langle su_\psi \rangle) + \frac{\iota_\psi I}{G+I} \langle su \rangle + 2 \frac{G}{G+I} \langle u_\Theta^2 \rangle - 2 \frac{I}{G+I} \langle u_\Phi^2 \rangle \right] \\
&\quad + \frac{I}{2e^2} \left(\frac{G}{G+I} \right)_\psi \langle su \rangle, \\
\langle \varphi_2 \rangle &= -\frac{I}{6e^2} \frac{I}{G+I} \left[(\langle s_\psi u \rangle + 3\langle su_\psi \rangle) + \frac{\iota_\psi I}{G+I} \langle su \rangle + 2 \frac{G}{G+I} \langle u_\Theta^2 \rangle - 2 \frac{I}{G+I} \langle u_\Phi^2 \rangle \right] \\
&\quad - \frac{I}{2e^2} \left(\frac{I}{G+I} \right)_\psi \langle su \rangle.
\end{aligned} \tag{5-2.16}$$

In Eq. (5-2.16) we only give the average of the second order ψ_2 , θ_2 and φ_2 instead of their exact expression because only the averages are needed. All the functions, such as G , I , s etc., in Eq. (5-2.16) are functions of the new variables, e.g. $G(\Psi) = G(\psi)|_{\psi=\Psi}$.

Now we define the spatial variables (Ψ, Θ, Φ) be the guiding center of the trajectory. Our definition of the guiding center given by $X(\Psi, \Theta, \Phi)$ is different from the conventional definition

$$X_c(\Psi_c, \Theta_c, \Phi_c) = x(\psi, \theta, \varphi) + \frac{mv \times B}{eB^2},$$

with X_c the conventional guiding center and x the particle position. We show in Appendix 5-B that the guiding center given by our definition differs the one given by the usual definition only by $(\beta_e/G + I)s/e$, which is of order ε^2 .

B. The Higher Order Drift Hamiltonian

The first order correction to the standard drift Hamiltonian is given in Eq. (5-2.8). We now proceed to calculate the averages in Eq. (5-2.8). The functions s and u are given in Eq. (5-2.1), with the functions α, B, g^s and g^c be functions of (Ψ, Θ, Φ) . The averages are

$$\langle \xi_i \rangle = 0 \text{ for } i = 1, 2, 3,$$

$$\langle L_I \rangle = \frac{I}{e} \mu \bar{V} \frac{\mu_0}{G + U} \left[G \left(\frac{I}{B} \right)_\Psi - I \left(\frac{G}{B} \right)_\Psi \right],$$

$$\langle V_I \rangle = -\frac{m}{2e} \frac{\mu B}{\mu_0 (G + U)} \left\{ \frac{g^s}{B} \left[\iota \left(\frac{g^c}{g^s} \right)_\Theta + \left(\frac{g^c}{g^s} \right)_\Phi \right] + \iota_\Psi \frac{g^s}{B} + 2(\iota \alpha_\Theta + \alpha_\Phi) \right\},$$

$$\left\langle \left(\xi_i \frac{\partial}{\partial \Xi^i} \right)^2 \right\rangle = \frac{m}{e^2} \frac{\mu}{B} \left[\nabla \Psi \frac{\partial}{\partial \Psi} + (\nabla \Theta - \iota \nabla \Phi) \left(\frac{G}{G + U} \frac{\partial}{\partial \Theta} - \frac{I}{G + U} \frac{\partial}{\partial \Phi} \right) \right]^2,$$

$$\begin{aligned} \langle \psi_2 \rangle &= \frac{m}{6e^2} \mu \left\{ 2 \left(\frac{g^s}{B} \right)_\Psi + \frac{\iota_\Psi I}{G + U} \frac{g^s}{B} + \frac{G}{G + U} \left[\frac{g^s}{g^c} \left(\frac{g^c}{B} \right)_\Theta + \left(\frac{g^c}{B} \right)_\Phi \right] \right. \\ &\quad \left. - \frac{I}{G + U} \left[\frac{g^s}{g^c} \left(\frac{g^c}{B} \right)_\Phi + \left(\frac{g^c}{B} \right)_\Theta \right] + 2 \frac{G}{G + U} \alpha_\Theta - 2 \frac{I}{G + U} \alpha_\Phi \right\}, \end{aligned}$$

$$\begin{aligned} \langle \theta_2 \rangle &= \frac{m}{6e^2} \mu \frac{G}{G + U} \left\{ \frac{g^c}{g^s} \left(\frac{g^s}{B} \right)_\Psi + \left(\frac{g^c}{B} \right)_\Psi + \frac{\iota_\Psi I}{G + U} \frac{g^c}{B} + 2 \frac{G}{G + U} \left(\frac{g^s}{B} \right)_\Theta \right. \\ &\quad \left. - 2 \frac{I}{G + U} \left(\frac{g^s}{B} \right)_\Phi - 2 \alpha_\Psi \right\} + \frac{m}{2e^2} \mu \left(\frac{G}{G + U} \right)_\Psi \frac{g^c}{B}, \end{aligned}$$

$$\begin{aligned} \langle \phi_2 \rangle &= -\frac{m}{6e^2} \mu \frac{I}{G + U} \left\{ \frac{g^c}{g^s} \left(\frac{g^s}{B} \right)_\Psi + \left(\frac{g^c}{B} \right)_\Psi + \frac{\iota_\Psi I}{G + U} \frac{g^c}{B} + 2 \frac{G}{G + U} \left(\frac{g^s}{B} \right)_\Theta \right. \\ &\quad \left. - 2 \frac{I}{G + U} \left(\frac{g^s}{B} \right)_\Phi - 2 \alpha_\Psi \right\} - \frac{m}{2e^2} \mu \left(\frac{I}{G + U} \right)_\Psi \frac{g^c}{B}. \end{aligned}$$

(5-2.17)

In Eq. (5-2.17) the function $\alpha(\Psi, \Theta, \Phi)$ is a function free to be chosen. This freedom may be used to simplify the Hamiltonian, but at this stage it is not clear what is the best choice for $\alpha(\Psi, \Theta, \Phi)$.

The drift Hamiltonian with the first order correction is

$$K_d(\mu; \bar{V}, \Psi, \Theta, \Phi) = H_d(\mu; \bar{V}, \Psi, \Theta, \Phi) + \varepsilon K_1(\mu; \bar{V}, \Psi, \Theta, \Phi), \quad (5-2.18)$$

with H_d the standard drift Hamiltonian and K_1 the first order correction

$$\begin{aligned} H_d &= \frac{l}{2m} \bar{V}^2 + \mu B(\Xi) + e\phi(\Xi), \\ K_1 &= \frac{l}{m} \langle V_l \rangle \bar{V} + \langle L_l \rangle B(\Xi) + e \left[\frac{l}{2} \left\langle \left(\xi_l \frac{\partial}{\partial \Xi^l} \right)^2 \right\rangle + \left(\langle \xi_l \rangle \frac{\partial}{\partial \Xi^l} \right) \right] \phi(\Xi). \end{aligned} \quad (5-2.19)$$

C. The Poisson Brackets in the Guiding Center Coordinates $(\theta_s, \mu, \bar{V}, \Psi, \Theta, \Phi)$

The Poisson brackets of the new coordinates are constant along the characteristic curve of Eq. (5-2.4a) and Eq. (5-2.4b). To prove this we have,

$$\frac{d}{d\tau} \{Z^i, Z^j\} = \{\{Z^i, Z^j\}, \theta_s\} = -\{\{Z^j, \theta_s\}, Z^i\} - \{\{\theta_s, Z^i\}, Z^j\} = 0,$$

$$\frac{d}{d\tau} \{ \mu, Z^i \} = \{ \{ \mu, Z^i \}, \theta_s \} = - \{ \{ Z^i, \theta_s \}, \mu \} - \{ \{ \theta_s, \mu \}, Z^i \} = 0,$$

with Z^i or Z^j being any of the $(\bar{V}, \Psi, \Theta, \Phi)$. Thus we only need to calculate $\{\mu, Z^i\}$ and $\{Z^i, Z^j\}$ at a certain point on the characteristic curve of Eq. (5-2.4a) and Eq. (5-2.4b). We choose the point to be on the initial surface $(\theta_s, 0, \bar{V}, \Psi, \Theta, \Phi)$.

On the initial surface with Eqs. (5-2.2) and (5-2.14'), the action forms are

$$\rho_\Psi = \mu_0 \frac{\beta.(\Psi, \Theta, \Phi)}{B(\Psi, \Theta, \Phi)} \bar{V},$$

$$\rho_\Theta = \mu_0 \frac{I(\Psi)}{B(\Psi, \Theta, \Phi)} \bar{V} + e\Psi,$$

$$\rho_\Phi = \mu_0 \frac{G(\Psi)}{B(\Psi, \Theta, \Phi)} \bar{V} - e\chi(\Psi),$$

$$\rho_{\theta_s} = \rho_{\bar{V}} = 0.$$

(5-2.20)

The action form ρ_μ is undetermined on the initial surface, but it is not needed. This will become clear later. The Lagrange brackets on the initial surface are

$$\omega_{v\Psi} = -\omega_{\Psi v} = \mu_0 \frac{\beta.}{B},$$

$$\omega_{v\Theta} = -\omega_{\Theta v} = \mu_0 \frac{I}{B},$$

$$\omega_{\bar{V}\phi} = -\omega_{\phi\bar{V}} = \mu_0 \frac{G}{B},$$

$$\omega_{\psi\theta} = -\omega_{\theta\psi} = e + \mu_0 \left[\left(\frac{I}{B} \right)_{\psi} - \left(\frac{\beta_*}{B} \right)_{\theta} \right] \bar{V},$$

$$\omega_{\psi\phi} = -\omega_{\phi\psi} = -e\iota + \mu_0 \left[\left(\frac{G}{B} \right)_{\psi} - \left(\frac{\beta_*}{B} \right)_{\phi} \right] \bar{V},$$

$$\omega_{\theta\phi} = -\omega_{\phi\theta} = \mu_0 \left[\left(\frac{G}{B} \right)_{\theta} - \left(\frac{I}{B} \right)_{\phi} \right] \bar{V},$$

$$\omega_{\theta_t z^i} = -\omega_{z^i \theta_t} = 0.$$

(5-2.21)

Other Lagrange brackets are unknown because ρ_μ is unknown. Assuming the unknown Lagrange brackets to be arbitrary functions and inverting the Lagrange tensor, we obtain

$$\{\bar{V}, \Psi\} = -\{\Psi, \bar{V}\} = -\frac{I}{e} \frac{\bar{V}}{D G + I} \left[\left(\frac{G}{B} \right)_{\theta} - \left(\frac{I}{B} \right)_{\phi} \right],$$

$$\{\bar{V}, \Theta\} = -\{\Theta, \bar{V}\} = -\frac{\iota}{D} \frac{\mu_0 B}{G + I} + \frac{I}{e} \frac{\bar{V}}{D G + I} \left[\left(\frac{G}{B} \right)_{\psi} - \left(\frac{\beta_*}{B} \right)_{\phi} \right],$$

$$\{\bar{V}, \Phi\} = -\{\Phi, \bar{V}\} = -\frac{I}{D} \frac{\mu_0 B}{G + I} - \frac{I}{e} \frac{\bar{V}}{D G + I} \left[\left(\frac{I}{B} \right)_{\psi} - \left(\frac{\beta_*}{B} \right)_{\theta} \right],$$

$$\{\Psi, \Theta\} = -\{\Theta, \Psi\} = -\frac{I}{e} \frac{I}{D} \frac{G}{G + I},$$

$$\{\Psi, \Phi\} = -\{\Phi, \Psi\} = \frac{I}{e} \frac{1}{D} \frac{I}{G+U},$$

$$\{\Theta, \Phi\} = -\{\Phi, \Theta\} = -\frac{I}{e} \frac{1}{D} \frac{\beta_*}{G+U},$$

$$\{\mu, Z^i\} = 0,$$

(5-2.22)

where

$$D = I + \frac{I}{e} \frac{\mu_0}{G+U} \left\{ \left[G \left(\frac{I}{B} \right)_\psi - I \left(\frac{G}{B} \right)_\psi \right] - \left[G \left(\frac{\beta_*}{B} \right)_\theta - I \left(\frac{\beta_*}{B} \right)_\phi \right] + \beta_* \left[\left(\frac{G}{B} \right)_\theta - \left(\frac{I}{B} \right)_\phi \right] \right\} \bar{V}. \quad (5-2.22')$$

The last equation in Eq. (5-2.22) shows that the new coordinates satisfy Eq. (5-2.4c). Unlike the coordinate transformation, the Poisson brackets are exact.

5-3 Conclusions

We derived the Hamiltonian of the exact trajectories and the higher order Hamiltonian of the drift trajectories of a charged particle in a static electromagnetic field by using Boozer coordinates (ψ, θ, φ) , in which the magnetic field has both the contravariant representation

$$\mathbf{B} = \nabla \psi \times \nabla \theta + \nabla \varphi \times \nabla \chi(\psi)$$

and the covariant representation

$$\mathbf{B} = \mu_0 [G(\psi) \nabla \varphi + I(\psi) \nabla \theta + \beta_*(\psi, \theta, \varphi) \nabla \psi].$$

The results are summarized here. We also further simplify the drift Hamiltonian for tokamaks and stellarators.

The exact Hamiltonian in Boozer coordinates (ψ, θ, φ) is

$$H = \frac{1}{2m} V^2 + \frac{1}{2m} [g^s(\psi, \theta, \varphi) s^2 - 2g^c(\psi, \theta, \varphi) su + g^\psi(\psi, \theta, \varphi) u^2] + e\phi(\psi, \theta, \varphi),$$

where the components of the metric tensor are $g^s = \mathbf{e}_s \cdot \mathbf{e}_s$, $g^c = \mathbf{e}_s \cdot \nabla \psi$ and $g^\psi = \nabla \psi \cdot \nabla \psi$ with $\mathbf{e}_s = \nabla \theta - \iota \nabla \varphi$. Three of the six independent variables are the particle position in Boozer coordinates (ψ, θ, φ) . The other three can be chosen as s , u and V . The two variables s and u are the perpendicular momenta, and the variable V is the parallel momentum. The six variables $(s, u, V, \psi, \theta, \varphi)$ are noncanonical. The Poisson brackets needed for the equations of motion are given in Eq. (5-1.13). The rotational transform ι , the poloidal current G and the toroidal current I are functions of ψ only. The function β_* and the field strength B depend on (ψ, θ, φ) . The notation $()_{\xi^i}$ refers to $\partial()/\partial \xi^i$ and ι_ψ refers to $d\iota/d\psi$. One can also choose the canonical momenta p_ψ , p_θ and p_φ as the independent variables instead of s , u and V . The relation between s , u , V and p_ψ , p_θ , p_φ can be obtained from Eq. (5-1.8), which is

$$s = \frac{G(\psi)(p_\theta - e\psi) - I(\psi)(p_\varphi + e\chi(\psi))}{G(\psi) + \iota(\psi)I(\psi)},$$

$$u = p_\psi - \frac{\beta_*(\psi, \theta, \varphi)}{G(\psi) + \iota(\psi)I(\psi)} [\iota(\psi)(p_\theta - e\psi) + (p_\varphi + e\chi(\psi))],$$

$$V = \frac{B(\psi, \theta, \varphi)}{\mu_0[G(\psi) + \iota(\psi)I(\psi)]} [\iota(\psi)(p_\theta - e\psi) + (p_\varphi + e\chi(\psi))].$$

To find the guiding center from the exact trajectories, one can first integrate the equations of motion of the exact Hamiltonian, and then do a simple transformation as follows,

$$\Psi = \psi + \frac{l}{e}s,$$

$$\Theta = \theta + \frac{l}{e} \frac{G(\psi)}{G(\psi) + \iota(\psi)I(\psi)} u,$$

$$\Phi = \varphi - \frac{l}{e} \frac{I(\psi)}{G(\psi) + \iota(\psi)I(\psi)} u.$$

The (Ψ, Θ, Φ) give the guiding center.

The drift Hamiltonian with first order correction is

$$K_d(\mu; \bar{V}, \Psi, \Theta, \Phi) = H_d(\mu; \bar{V}, \Psi, \Theta, \Phi) + \varepsilon K_1(\mu; \bar{V}, \Psi, \Theta, \Phi),$$

with H_d the standard drift Hamiltonian and K_1 the first order correction

$$H_d = \frac{l}{2m} \bar{V}^2 + \mu B(\Xi) + e\phi(\Xi),$$

$$K_I = \frac{I}{m} \langle V_I \rangle \bar{V} + \langle L_I \rangle B(\Xi) + e \left[\frac{1}{2} \left\langle \left(\xi_I^i \frac{\partial}{\partial \Xi^i} \right)^2 \right\rangle + \left(\langle \xi_2^i \rangle \frac{\partial}{\partial \Xi^i} \right) \right] \phi(\Xi),$$

with

$$\langle L_I \rangle = \frac{I}{e} \mu \bar{V} \frac{\mu_0}{G + I} \left[G \left(\frac{I}{B} \right)_\Psi - I \left(\frac{G}{B} \right)_\Psi \right],$$

$$\langle V_I \rangle = -\frac{m}{2e} \frac{\mu B}{\mu_0 (G + I)} \left\{ \frac{g^s}{B} \left[\iota \left(\frac{g^c}{g^s} \right)_\Theta + \left(\frac{g^c}{g^s} \right)_\Phi \right] + \iota_\Psi \frac{g^s}{B} + 2(\iota \alpha_\Theta + \alpha_\Phi) \right\},$$

$$\left\langle \left(\xi_I^i \frac{\partial}{\partial \Xi^i} \right)^2 \right\rangle = \frac{m}{e^2} \frac{\mu}{B} \left[\nabla \Psi \frac{\partial}{\partial \Psi} + (\nabla \Theta - \iota \nabla \Phi) \left(\frac{G}{G + I} \frac{\partial}{\partial \Theta} - \frac{I}{G + I} \frac{\partial}{\partial \Phi} \right) \right]^2,$$

$$\begin{aligned} \langle \Psi_2 \rangle &= \frac{m}{6e^2} \mu \left\{ 2 \left(\frac{g^s}{B} \right)_\Psi + \frac{\iota_\Psi I}{G + I} \frac{g^s}{B} + \frac{G}{G + I} \left[\frac{g^s}{g^c} \left(\frac{g^c}{B} \right)_\Theta + \left(\frac{g^c}{B} \right)_\Phi \right] \right. \\ &\quad \left. - \frac{I}{G + I} \left[\frac{g^s}{g^c} \left(\frac{g^c}{B} \right)_\Phi + \left(\frac{g^c}{B} \right)_\Theta \right] + 2 \frac{G}{G + I} \alpha_\Theta - 2 \frac{I}{G + I} \alpha_\Phi \right\}, \end{aligned}$$

$$\begin{aligned} \langle \Theta_2 \rangle &= \frac{m}{6e^2} \mu \frac{G}{G + I} \left\{ \frac{g^c}{g^s} \left(\frac{g^s}{B} \right)_\Psi + \left(\frac{g^c}{B} \right)_\Psi + \frac{\iota_\Psi I}{G + I} \frac{g^c}{B} + 2 \frac{G}{G + I} \left(\frac{g^s}{B} \right)_\Theta \right. \\ &\quad \left. - 2 \frac{I}{G + I} \left(\frac{g^s}{B} \right)_\Phi - 2 \alpha_\Psi \right\} + \frac{m}{2e^2} \mu \left(\frac{G}{G + I} \right)_\Psi \frac{g^c}{B}, \end{aligned}$$

$$\begin{aligned} \langle \varphi_2 \rangle &= -\frac{m}{6e^2} \mu \frac{I}{G + I} \left\{ \frac{g^c}{g^s} \left(\frac{g^s}{B} \right)_\Psi + \left(\frac{g^c}{B} \right)_\Psi + \frac{\iota_\Psi I}{G + I} \frac{g^c}{B} + 2 \frac{G}{G + I} \left(\frac{g^s}{B} \right)_\Theta \right. \\ &\quad \left. - 2 \frac{I}{G + I} \left(\frac{g^s}{B} \right)_\Phi - 2 \alpha_\Psi \right\} - \frac{m}{2e^2} \mu \left(\frac{I}{G + I} \right)_\Psi \frac{g^c}{B}. \end{aligned}$$

In the drift Hamiltonian, the variable \bar{V} is the parallel momentum. The other three variables $\Xi = (\Psi, \Theta, \Phi)$ is the guiding center position. The drift Hamiltonian does not depend on the gyrophase. The magnetic moment μ is a constant of motion. The four independent variables $(\bar{V}, \Psi, \Theta, \Phi)$ are noncanonical. The Poisson brackets needed for the equations of motion are given in Eq. (5-2.22) and Eq. (5-2.22'). All the functions are functions of (Ψ, Θ, Φ) , i.e. $f(\Psi, \Theta, \Phi) = f(\psi, \theta, \phi)|_{(\psi, \theta, \phi) = (\Psi, \Theta, \Phi)}$ with f being any of the functions. The function $\alpha(\Psi, \Theta, \Phi)$ is an arbitrarily well behaved function free to be chosen.

All the functions $f(\Psi, \Theta, \Phi)$, such as B , g^s and etc., are periodic in Θ and Φ . For a given magnetic field $B(x)$, they can be determined numerically in the forms of Fourier expansion,^{1, 4}

$$f(\Psi, \Theta, \Phi) = \sum_{mn} f_{mn}(\Psi) \exp[2\pi i(m\Theta - n\Phi)].$$

The derivative of these function can be calculated very easily. Because the field strength B and metric tensor g^s , g^v , g^c are some Fourier series, any choice of the function $\alpha(\Psi, \Theta, \Phi)$ can not simplify $\langle V_i \rangle$ and $\langle \psi_2 \rangle$ at the same time. We thus choose α be zero.

Comparing with previous chapter, the main conclusions are the same. The standard drift Hamiltonian only depends on the field strength (a scalar) while the exact Hamiltonian depends on the vector property of the magnetic field. The higher order correction to the standard drift Hamiltonian also depends on the vector property of the field. When the toroidal current I and the function β_* in Eq. (5-1.3) are zero, i.e. vacuum field, the higher order drift Hamiltonian we derived here is identical with that of chapter 4. The main difference is that we use four noncanonical variable, the parallel velocity \bar{V} and the position

of the guiding center (Ψ, Θ, Φ) . This arises purely because of Ψ and Θ are not canonical in the case of full field. If one insists on using canonical coordinates, exotic coordinates have to be used instead of guiding center position.

The drift Hamiltonian certainly looks messy at this stage. However we can simplify this Hamiltonian significantly in the cases of tokamak and stellarator. The electric potential is almost constant within a magnetic surface, $\phi_\psi \gg \phi_\theta, \phi_\phi$. In a tokamak or stellarator the poloidal current G is much larger than the toroidal current I , I/G is of a few percent, and so is G_ψ/I_ψ . We thus obtain for tokamak and stellarator,

$$\frac{I}{m} \langle V_I \rangle \bar{V} = -\frac{I}{2e} \frac{\mu \bar{V}}{G} \left\{ g^s \left[l \left(\frac{g^c}{g^s} \right)_\theta + \left(\frac{g^c}{g^s} \right)_\phi \right] + l_\psi g^\psi \right\},$$

$$\langle L_I \rangle B = \frac{I}{e} \mu \bar{V} \mu_0 I_\psi,$$

$$\frac{e}{2} \left\langle \left(\xi_I \frac{\partial}{\partial \Xi^i} \right)^2 \right\rangle \phi = \frac{I}{2e} m \mu \frac{g^\psi}{B} \phi_{\psi\psi},$$

$$e \langle \xi_2 \rangle \frac{\partial}{\partial \Xi^i} \phi = \frac{I}{6e} m \mu \left[2 \left(\frac{g^\psi}{B} \right)_\psi + \left(\frac{g^s}{g^c} + 1 \right) \left(\frac{g^c}{B} \right)_\theta \right] \phi_\psi.$$

We use the exact Poisson brackets in order to keep the conservation of the Hamiltonian.

It is noteworthy that the standard drift Hamiltonian only contains the $E \times B$ drift, the $B \times \nabla B$ drift and the curvature drift. The first order correction to the standard Hamiltonian brings in the so called finite-Larmor-radius effect and the polarization drift. The standard

drift Hamiltonian depends on the field strength only. The first order correction to the standard drift Hamiltonian depends on the metric tensor like the exact Hamiltonian. Because the metric tensor elements are periodic in Θ and Φ , whenever the symmetry of the exact Hamiltonian is broken, the symmetry of the drift Hamiltonian is broken also.

Appendix 5-A: The Poisson Brackets for Darboux Transformation in Sec. 5-2

To calculate the Poisson brackets $\{\theta_i, z^i\}$ in Eq. (5-2.11) we first find the actions

$$\rho_\psi = p_\psi,$$

$$\rho_\theta = p_\theta,$$

$$\rho_\varphi = p_\varphi,$$

(5-A.1)

and $\rho_{\theta_i} = \rho_L = \rho_V = 0$. The canonical momenta p_ψ , p_θ and p_φ given in Eq. (5-2.2). The Lagrange brackets are

$$\omega_{\theta_i \psi} = -\omega_{\psi \theta_i} = -\mu_{\theta_i},$$

$$\omega_{\theta_i \theta} = -\omega_{\theta \theta_i} = s_{\theta_i},$$

$$\omega_{\theta_i \varphi} = -\omega_{\varphi \theta_i} = -ls_{\theta_i},$$

$$\omega_{L \psi} = -\omega_{\psi L} = -\mu_L,$$

$$\omega_{L \theta} = -\omega_{\theta L} = s_L,$$

$$\omega_{L\phi} = -\omega_{\phi L} = -\omega_L,$$

$$\omega_{V\psi} = -\omega_{\psi V} = \mu_0 \frac{\beta_*}{B},$$

$$\omega_{V\theta} = -\omega_{\theta V} = \mu_0 \frac{I}{B},$$

$$\omega_{V\varphi} = -\omega_{\varphi V} = \mu_0 \frac{G}{B},$$

$$\omega_{\psi\theta} = -\omega_{\theta\psi} = e + s_\psi + u_\theta + \mu_0 \left[\left(\frac{I}{B} \right)_\psi - \left(\frac{\beta_*}{B} \right)_\theta \right] V,$$

$$\omega_{\psi\varphi} = -\omega_{\varphi\psi} = -e\iota - \iota s_\psi + u_\varphi - \iota_\psi s + \mu_0 \left[\left(\frac{G}{B} \right)_\psi - \left(\frac{\beta_*}{B} \right)_\varphi \right] V,$$

$$\omega_{\theta\varphi} = -\omega_{\varphi\theta} = -\iota s_\theta - s_\varphi + \mu_0 \left[\left(\frac{G}{B} \right)_\theta - \left(\frac{I}{B} \right)_\varphi \right] V,$$

(5-A.2)

and others are zeros. The required Poisson brackets are

$$\{\theta_i, L\} = J_0^{\theta_i L} + \varepsilon J_1^{\theta_i L} + \varepsilon^2 J_2^{\theta_i L},$$

$$\{\theta_i, V\} = \varepsilon J_1^{\theta_i V} + \varepsilon^2 J_2^{\theta_i V},$$

$$\{\theta_i, \psi\} = \varepsilon J_1^{\theta_i \psi},$$

$$\begin{aligned}\{\theta_t, \theta\} &= \varepsilon J_1^{\theta_t \theta} + \varepsilon^2 J_2^{\theta_t \theta}, \\ \{\theta_t, \varphi\} &= \varepsilon J_1^{\theta_t \varphi} + \varepsilon^2 J_2^{\theta_t \varphi},\end{aligned}\tag{5-A.3}$$

where

$$\begin{aligned}J_0^{\theta_t L} &= \frac{e}{m}, \\ J_1^{\theta_t L} &= \frac{l}{m} \left(s_\psi + \frac{i_\psi I}{G+U} s + \frac{G}{G+U} u_\theta - \frac{I}{G+U} u_\varphi \right) + \frac{V}{m} \frac{\mu_o}{G+U} \left[G \left(\frac{I}{B} \right)_\psi - I \left(\frac{G}{B} \right)_\psi \right], \\ J_2^{\theta_t L} &= -\frac{l}{m} \frac{\beta_*}{G+U} (i s_\theta + s_\varphi) + \frac{V}{m} \frac{\mu_o}{G+U} \left\{ \beta_* \left[\left(\frac{G}{B} \right)_\theta - \left(\frac{I}{B} \right)_\varphi \right] - \left[G \left(\frac{\beta_*}{B} \right)_\theta - I \left(\frac{\beta_*}{B} \right)_\varphi \right] \right\}; \\ J_1^{\theta_t V} &= \frac{l}{m} \frac{B}{\mu_o (G+U)} \left[s_L (i u_\theta + u_\varphi) - u_L (i s_\theta + s_\varphi) - i_\psi s s_L \right] \\ &\quad + \frac{V}{m} \frac{B}{G+U} \left\{ s_L \left[\left(\frac{G}{B} \right)_\psi + i \left(\frac{I}{B} \right)_\psi \right] + u_L \left[\left(\frac{G}{B} \right)_\theta - \left(\frac{I}{B} \right)_\varphi \right] \right\}, \\ J_2^{\theta_t V} &= -\frac{V}{m} \frac{B}{G+U} s_L \left[i \left(\frac{\beta_*}{B} \right)_\theta + \left(\frac{\beta_*}{B} \right)_\varphi \right]; \\ J_1^{\theta_t \psi} &= -\frac{l}{m} s_L; \\ J_1^{\theta_t \theta} &= -\frac{l}{m} \frac{G}{G+U} u_L,\end{aligned}$$

$$J_1^{\theta, \varphi} = \frac{l}{m} \frac{I}{G + u} u_L,$$

$$J_2^{\theta, \varphi} = \frac{l}{m} \frac{\beta_*}{G + u} s_L.$$
(5-A.4)

In ordering the Poisson brackets we used the assumption that β_* is much smaller than $G + u$, $\beta_*/(G + u)$ is of order ε .

Appendix 5-B: Proof of Our Definition for the Guiding Center

In this appendix we prove the validity of the guiding center definition given in Sec. 5-2. In Sec. 5-2 we defined the guiding center to be $X(\Psi, \Theta, \Phi)$. The relation between the guiding center $X(\Psi, \Theta, \Phi)$ and the location of the particle $x(\psi, \theta, \varphi)$ is,

$$\begin{aligned}\psi &= \Psi - \frac{I}{e}s, \\ \theta &= \Theta - \frac{I}{e} \frac{G}{G+U} u, \\ \varphi &= \Phi + \frac{I}{e} \frac{I}{G+U} u.\end{aligned}\tag{5-B.1}$$

Let the conventional definition of the guiding center be $X_c(\Psi_c, \Theta_c, \Phi_c)$. The relation between the position of the particle $x(\psi, \theta, \varphi)$ and the guiding center $X_c(\Psi_c, \Theta_c, \Phi_c)$ is

$$x(\psi, \theta, \varphi) = X_c(\Psi_c, \Theta_c, \Phi_c) - \varepsilon \frac{mv \times B}{eB^2} + o(\varepsilon^2).\tag{5-B.2}$$

By using Eq. (5-1.3), Eq. (5-1.5) and Eq. (5-1.8'), one find

$$\frac{mv \times B}{eB^2} = \frac{\mu_0}{eB^2} [(G+U)s\nabla\theta \times \nabla\varphi + (Gu - i\beta_s s)\nabla\varphi \times \nabla\psi - (Iu + \beta_s s)\nabla\psi \times \nabla\theta].$$

The dual relations is $2\partial x/\partial\xi^i = J\varepsilon_{ijk}\nabla\xi^j \times \nabla\xi^k$, with the Jacobian J ,

$$J = \frac{I}{(\nabla\psi \times \nabla\theta) \cdot \nabla\varphi} = \frac{\mu_0(G + uI)}{B^2},$$

which follows from Eq. (5-1.3) and Eq. (5-1.7). Therefore we obtain

$$\frac{mv \times B}{eB^2} = \frac{I}{e} \left(s \frac{\partial x}{\partial \psi} + \frac{Gu - \iota\beta_* s}{G + uI} \frac{\partial x}{\partial \theta} - \frac{Iu + \beta_* s}{G + uI} \frac{\partial x}{\partial \varphi} \right). \quad (5-B.3)$$

Eq. (5-B.2) and Eq. (5-B.3) imply

$$\psi = \Psi_c - \frac{I}{e}s,$$

$$\theta = \Theta_c - \frac{I}{e} \left(\frac{G}{G + uI} u - \iota \frac{\beta_*}{G + uI} s \right),$$

$$\varphi = \Phi_c + \frac{I}{e} \left(\frac{I}{G + uI} u + \frac{\beta_*}{G + uI} s \right).$$

Since β_* is much smaller than $(G + uI)$ as we mentioned early, the guiding center defined by $X(\Psi, \Theta, \Phi)$ remains close (with a small difference of $[\beta_*/(G + uI)]s/e$) to the conventional $X_c(\Psi_c, \Theta_c, \Phi_c)$.

References

- 1 A. H. Boozer, Phys. Fluids **24**, 1999 (1981); Phys. Fluids **25**, 520 (1982)
- 2 R. G. Littlejohn, J. Math. Phys. **20**, 2445 (1979); Phys. Fluids **24**, 1730 (1981)
- 3 A. Broman, *Introduction to Partial Differential Equations* (Dover Publications, Inc., New York, 1989)
- 4 G. Kuo-Petraovic, A. H. Boozer, J. A. Rome and R. H. Fowler, J. Comp. Phys. **51**, 261 (1983)

CHAPTER 6

A NUMERICAL STUDY OF THE DRIFT APPROXIMATION

In this chapter a numerical study of the guiding center approximation is presented. The trajectories of the standard drift Hamiltonian H_d , the high order drift Hamiltonian K_d and the exact Hamiltonian H_e are compared for a quasihelical symmetric magnetic field. The results show that the differences of the phase space structures between the standard drift Hamiltonian and the exact Hamiltonian can be predicted by the high order drift Hamiltonian.

6-1 The Hamiltonians

The magnetic coordinates we use in this chapter are Boozer coordinates (ψ, θ, φ) . In Boozer coordinates a magnetic field \mathbf{B} has both simple covariant and contravariant forms,

$$\mathbf{B} = \mu_0 [G(\psi) \nabla \varphi + I(\psi) \nabla \theta + \beta_*(\psi, \theta, \varphi) \nabla \psi], \quad (6-1.1)$$

$$\mathbf{B} = \nabla \psi \times \nabla \theta + \nabla \varphi \times \nabla \chi(\psi). \quad (6-1.2)$$

Here ψ is the toroidal magnetic flux inside a magnetic surface, and θ and φ are the poloidal and toroidal angles, the periods of which are one instead of 2π . The poloidal flux outside a magnetic surface is $-\chi$. The function G is the poloidal current outside a magnetic surface and the function I is the toroidal current inside a magnetic surface. The function β_* is closely related to the Pfirsch-Schlüter current. The constant μ_0 is the

permeability of free space. In a vacuum field, $I = \beta_* = 0$ and Eq. (6-1.1) becomes

$$\mathbf{B} = \mu_0 G_0 \nabla \varphi. \quad (6-1.3)$$

The poloidal current G_0 is a constant, which is the total current in the poloidal coils in amperes. In this chapter we assume the magnetic field is a vacuum field for simplicity.

The exact Hamiltonian of a charged particle in a magnetic field is

$$H_e = \frac{I}{2} mv^2. \quad (6-1.4)$$

The Hamiltonian has a canonical representation in the magnetic coordinates. The six canonical coordinates are $(s, p; \theta, p_\theta; \varphi, p_\varphi)$. The physical interpretation of these variables is as follows: the general coordinate s and its conjugate momentum p are the components of the particle momentum perpendicular to the magnetic field line. The other two pairs of canonical variables θ, p_θ and φ, p_φ are the same as those of the drift Hamiltonian. The three spatial coordinates $p_\theta (= e\psi)$, θ , φ give the trajectory of the guiding center and p_φ , in some way, gives the parallel velocity $v_{||}$. The exact Hamiltonian has the form

$$H_e = \frac{I}{2m} \left(\frac{B(\xi)}{\mu_0 G_0} \right)^2 [p_\varphi + q\chi(\psi - s) + e\iota(\psi - s)s]^2 + \frac{I}{2m} [e^2 g^s(\xi)s^2 - 2eg^c(\xi)sp + g^\psi(\xi)p^2] \quad (6-1.5)$$

In Eq. (6-1.5), $\xi = (\psi - s, \theta - p/e, \varphi)$ is the position of the particle. The function ι is the rotational transform, $\iota = \chi'$. The metric components of the magnetic coordinates are defined as

$$g^s = \mathbf{e}_s \cdot \mathbf{e}_s, \quad g^\psi = \nabla \psi \cdot \nabla \psi \quad \text{and} \quad g^c = \nabla \psi \cdot \mathbf{e}_s. \quad (6-1.6)$$

with the vector $\mathbf{e}_s = \nabla \theta - \iota \nabla \varphi$ lying roughly within the magnetic surfaces. There is a

simple relation between the strength of magnetic field and the components of the metric tensor

$$B^2 = g^s g^\psi - (g^c)^2. \quad (6-1.7)$$

The metric tensor gives the geometry of the magnetic surfaces. In the exact Hamiltonian of Eq. (6-1.5), the magnetic field, which is in terms of the poloidal flux, the rotational transform and the metric, depends on the location of the particle ξ . The physical explanation is that the particle's motion is affected by the magnetic field at the position of the particle, not the field at the position of the guiding center. The first term of Eq. (6-1.5) is the parallel kinetic energy and the second term is the perpendicular kinetic energy. The perpendicular term is dominated by the gyromotion.

The exact trajectories of the charged particles in a nonuniform magnetic field are extremely complicated. However the time scale of the gyromotion is much faster than the time scale of the drift motion across the field line in most of the cases of interest. This means that there exists an adiabatic invariant, the magnetic moment. The magnetic moment is the action to the lowest order of the p - s gyromotion,

$$\mu \equiv \frac{I}{2\pi} \oint \frac{e}{m} p ds. \quad (6-1.8)$$

It is not difficult to show that the Hamiltonian for the guiding center, to the lowest order, is

$$H_d(\mu, \theta, p_\theta; \varphi, p_\varphi) = \langle H_e \rangle = \frac{I}{2m} \left(\frac{B(\psi, \theta, \varphi)}{\mu_0 G_0} \right)^2 [p_\varphi + e\chi(\psi)]^2 + \mu B(\psi, \theta, \varphi). \quad (6-1.9)$$

The general coordinates are θ and φ , and the conjugate momenta are $p_\theta = e\psi$ and $p_\varphi = e(\mu_0 G_0 \rho_{||} - \chi)$ with the parallel gyroradius $\rho_{||} = mv_{||}/eB$. The Hamiltonian H_d gives the equations of motion to the same order of Alfvén's drift equations of Eq. (1-17)

$$\mathbf{v}_d = \frac{\mathbf{E} \times \mathbf{B}}{B^2} + \frac{v_\perp^2}{2\omega} \frac{\mathbf{B} \times \nabla B}{B^2} + \frac{mv_\parallel^2}{e} \frac{\mathbf{B} \times \boldsymbol{\kappa}}{B^2}.$$

We call this Hamiltonian the standard drift Hamiltonian. The magnetic moment μ is a constant in the drift approximation.

In Chapter 4 we derived the first order correction to the standard drift Hamiltonian ¹

$$H_I = \epsilon_{\parallel} \gamma(\Xi) \mu B(\Xi), \quad (6-1.10)$$

with the geometry factor γ

$$\gamma(\Xi) = -\frac{I}{2B(\Xi)} \left[g^\psi(\Xi) \frac{d\iota(\psi)}{d\psi} + \frac{\mu_0 G_0}{B^2(\Xi)} g^s(\Xi) \mathbf{B}(\Xi) \cdot \nabla \left(\frac{g^c(\Xi)}{g^s(\Xi)} \right) \right]. \quad (6-1.11)$$

and the parallel gyroradius to system size ϵ_{\parallel}

$$\epsilon_{\parallel} = \rho_{\parallel} \frac{B}{\mu_0 G_0} = \frac{B(\Xi) [p_\phi + e\chi(\psi)]}{e(\mu_0 G_0)^2} \quad (6-1.12)$$

and $\Xi = (\psi, \theta, \varphi)$, the position of the guiding center. The high order drift Hamiltonian is

$$K_d = H_d + H_I. \quad (6-1.13)$$

The exact Hamiltonian depends on the vector property of the magnetic field, i.e. the field strength B and the components of the metric tensor of the magnetic coordinates g^s , g^ψ and g^c . The standard drift Hamiltonian, however, depends on the strength of the magnetic field only. Therefore it is possible that the standard drift Hamiltonian, H_d , can have a symmetry which the exact Hamiltonian, H_e , does not have. The first order correction to the standard drift Hamiltonian of Eq. (6-1.10) depends on the metric tensor in

a generic way.

A very good example is the quasihelically symmetric stellarator. The magnetic field strength of such a stellarator has the form of $B(\psi, \theta - N\varphi)$ to the order of A^{-2} , with the integer N the number of the field period and A the aspect ratio of the torus. The metric tensor, or the shape of the magnetic surfaces, does not have such a symmetry as the field strength. A typical magnetic surface of quasihelically symmetric fields is shown in Fig. 6.1.

With a trivial canonical transformation, via the generating function

$$S = (\alpha + N\varphi_h) p_\theta + \varphi_h p_\varphi, \quad (6-1.14)$$

the canonical coordinates of the Hamiltonians of Eqs. (6-1.5, 9, 13) can be transformed into the helical coordinates,

$$\alpha = \theta - N\varphi,$$

$$p_\alpha = p_\theta = e\psi,$$

$$\varphi_h = \varphi,$$

$$p_h = p_\varphi + Np_\theta.$$

(6-1.15)

The Hamiltonians of Eqs. (6-1.5, 9, 13) have the same form except that p_φ changes to p_h , χ changes to $\chi_h = \chi - N\psi$ and ι changes to $\iota_h = \iota - N$, i.e.

$$\begin{aligned} H_e(s, p; \alpha, p_\alpha; \varphi_h, p_h) &= \frac{1}{2m} \left(\frac{B}{\mu_0 G_0} \right)^2 (p_h + q\chi_h + e\iota_h s)^2 \\ &\quad + \frac{1}{2m} (e^2 g^s s^2 - 2eg^c sp + g^\psi p^2), \end{aligned} \quad (6-1.16)$$

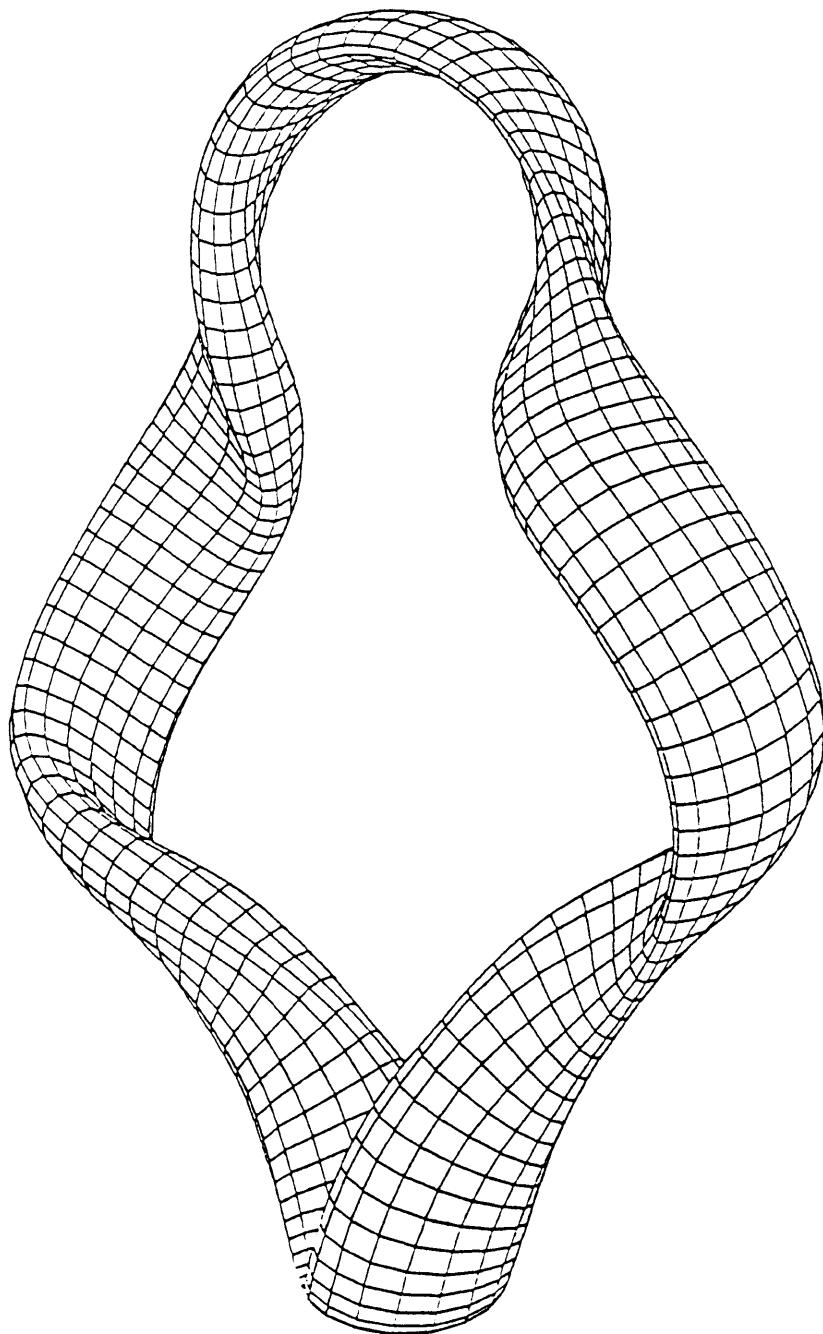


Figure 6.1 The magnetic flux surface of a quasi-helically symmetric stellarator of period $N = 4$ and aspect ratio $A = 8$ (provided by D. Anderson from the Torsatron/Stellarator Laboratory of the University of Wisconsin, Madison)

$$H_d(\mu; p_h; \alpha, p_\alpha) = \frac{I}{2m} \left(\frac{B}{\mu_0 G_0} \right)^2 (p_h + e\chi_h)^2 + \mu B, \quad (6-1.17)$$

$$K_d(\mu; \alpha, p_\alpha; \varphi_h, p_h) = H_d + \varepsilon_{\parallel} \gamma \mu B, \quad (6-1.18)$$

with $\gamma = -\frac{I}{2B} \left[g^\psi \dot{\iota}_h' + \frac{\mu_0 G_0}{B^2} g^s \mathbf{B} \cdot \nabla \left(\frac{g^c}{g^s} \right) \right]$ and $\varepsilon_{\parallel} = \frac{B}{e(\mu_0 G_0)^2} (p_h + e\chi_h)$.

For simplicity, we consider a magnetic field that has an exact quasihelical symmetry. The magnetic field strength only depends on two coordinates (ψ, α) and the metric depends on all three coordinates $(\psi, \alpha, \varphi_h)$. This means that the helical momentum p_h is a constant of motion for the standard drift Hamiltonian H_d , but not for the exact Hamiltonian H_e , nor for the high order drift Hamiltonian K_d .

6-2 The Numerical Results

A. The Trajectories of H_d

We start with the standard drift Hamiltonian H_d . Since the magnetic moment μ and the helical momentum p_h are constants of motion, the standard drift Hamiltonian has only one degree of freedom and is integrable.

Integrating the equations of motion one can find that there are two types of orbit: passing and trapped. In the $\psi \sim \theta$ plane a passing trajectory is like a circle and a trapped trajectory is like a banana (see Fig. 6.2). To understand these we first ignore the drift motion and just consider the parallel motion along the field line. The magnetic field lines spire around the torus. The field strength, which acts as a potential, is strong on the inside of the torus and weak on the outside. If a particle has a high enough energy it can pass the potential barrier and become a passing particle. If not, it becomes a trapped particle. Now we add the drift motion. There are two parts in drift motion, $\dot{\psi}$ and $\dot{\alpha}_\perp$, following from

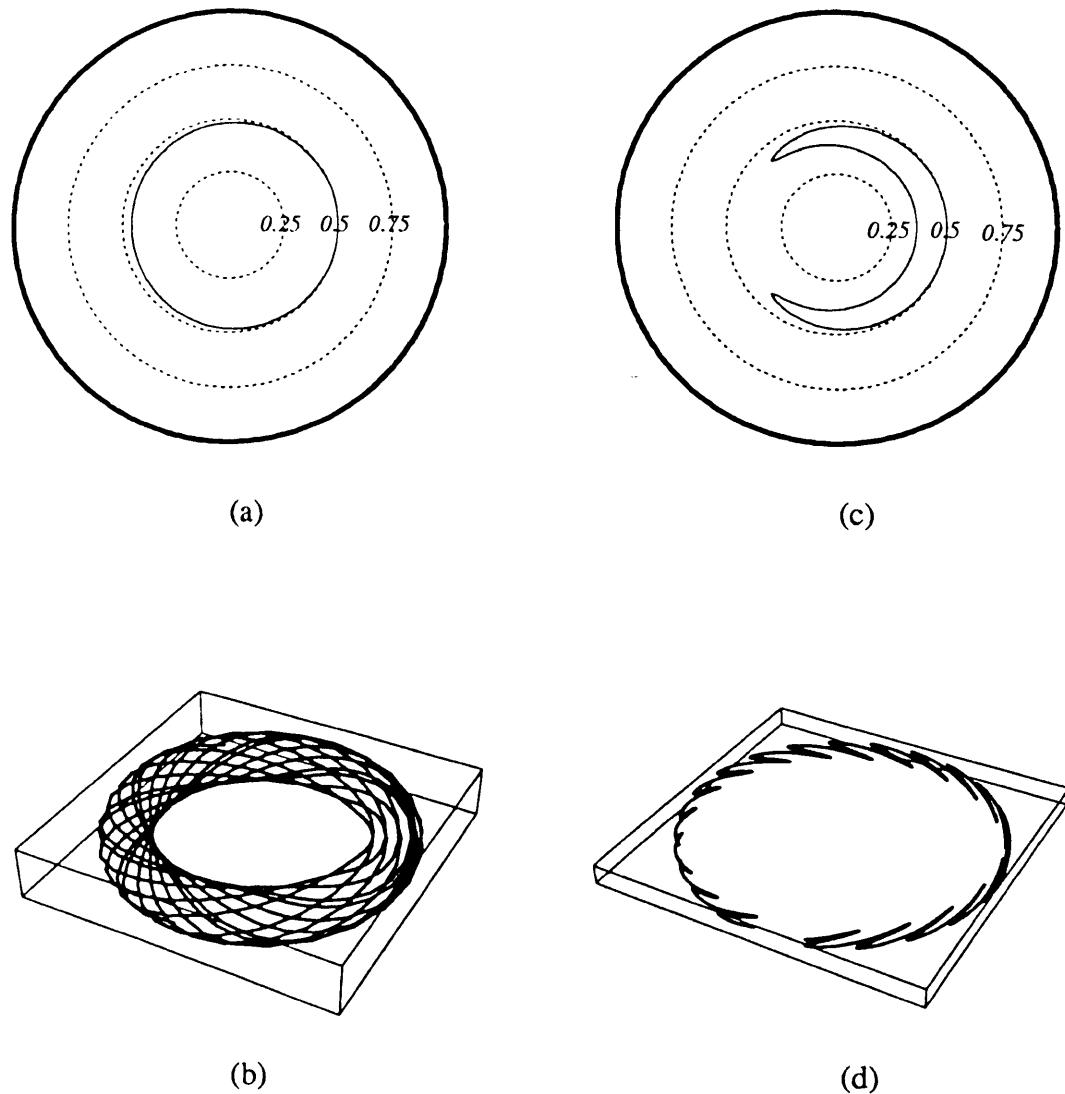


Figure 6.2 Orbits of the standard drift Hamiltonian in a quasihelical symmetric magnetic field. (a) A passing trajectory view in $\psi \sim \theta$ plane. (b) The 3-D view of the passing trajectory. (c) A trapped trajectory view in $\psi \sim \theta$ plane. (d) The 3-D view of the trapped trajectory.

Eq. (4-2.11),

$$\dot{\psi} = -\frac{I}{e} \left[\frac{1}{m} \left(\frac{B}{\mu_0 G_0} \right)^2 (p_h + e\chi_h)^2 + \mu B \right] \frac{\partial B / \partial \alpha}{B},$$

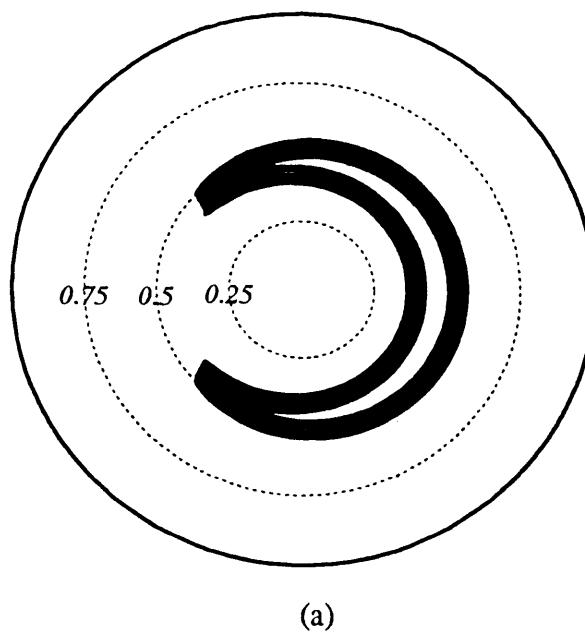
$$\dot{\alpha}_\perp = \frac{I}{e} \left[\frac{1}{m} \left(\frac{B}{\mu_0 G_0} \right)^2 (p_h + e\chi_h)^2 + \mu B \right] \frac{\partial B / \partial \psi}{B}.$$

The $\dot{\psi}$ drift takes the orbits across the flux surfaces so that the trapped trajectory becomes a banana instead of an arc and the center of the circle formed by a passing trajectory is shifted from the magnetic axis. The $\dot{\alpha}_\perp$ drift makes the whole orbit precess around the torus (in the φ_h direction).

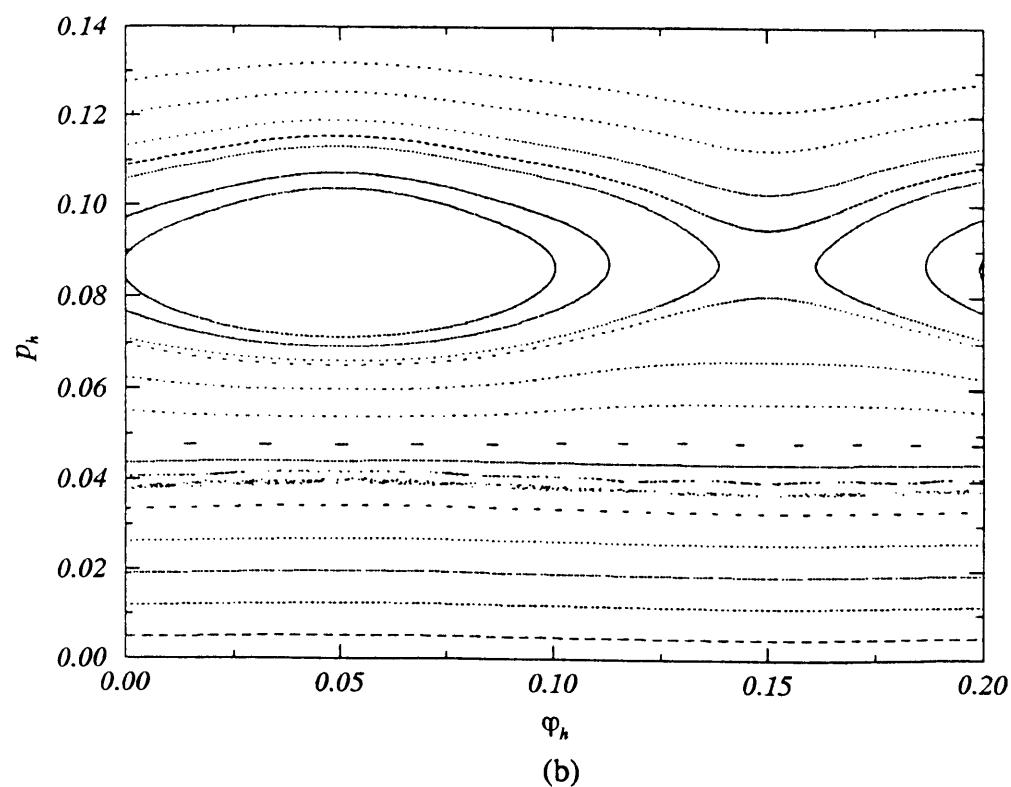
It is worthwhile to point out that because of the conservation of the helical momentum the trajectories always have fixed location with respect to the magnetic surfaces.

B. The Trajectories of K_d

Now we turn to the drift Hamiltonian with first order correction of Eq. (6-1.18), $K_d = H_d + \epsilon_{\parallel} \gamma \mu B$. The first order correction is usually much smaller than H_d . Thus we can treat the correction as a perturbation. The primary effect of this perturbation is to make the trapped or passing trajectories oscillate in and out about their home flux surfaces (see Fig. 6.3(a)). Those trajectories near the trapped and passing boundary become unpredictable. A trapped trajectory can change to a passing trajectory and become trapped again after a while, and *vise versa*. This is because that the φ_h symmetry is broken and p_h is no longer conserved. On the (p_h, φ_h) surface of section (SOS) most of the p_h invariant surfaces are distorted but still exist (see Fig. 6.3(b)). Those near the passing and trapped boundary become stochastic. There are a few small island chains in the barely trapped region, and they are usually not visible. There is one big island in the deep trapped region. When the magnetic moment μ is fixed, the location of this big island changes as the energy (or the pitch angle) is changed.



(a)



(b)

Figure 6.3 Trajectory of the high order drift Hamiltonian. (a) A trapped trajectory view in $\psi \sim \theta$ plane.
 (b) The (p_h, φ_h) surface of section.

The standard textbook example in the canonical perturbation method is to start with the action-angle variables of the integral part, H_d , then look for the resonant terms in the perturbation, $H_I = \epsilon_{\parallel} \gamma \mu B$. In the standard drift Hamiltonian, H_d , the canonical coordinates p_h and φ_h are in action-angle form, but $p_\alpha (= e\psi)$ and α are not. In principle (ψ, α) can be transformed to an action-angle form, but the transformation can only be done numerically instead of analytically. Thus H_I cannot be written in the action-angle coordinates of H_d . However this should not hamper our understanding of the phase space structure. The islands must come from the resonant surfaces of H_d .

In order to find the resonant surfaces of H_d we look for the closed orbits of H_d , because these closed orbits form the resonant surfaces. We integrate the trajectories of H_d for one period of (ψ, φ) motion (one transit for a passing orbit or one bounce for a trapped orbit) and record the change in φ_h , $\Delta\varphi_h$. If $\Delta\varphi_h = l/n$ with n an integer number, then the orbit closes onto itself after n periods. If $\Delta\varphi_h = 0$, then the orbit does not precess around the torus. We launch the particles at the bottom of the magnetic well with different helical momenta p_h for fixed magnetic moment μ and energy H_d , and we make $p_h \sim \Delta\varphi_h$ plots (see Fig. 6.4). There are three kinds: all passing, all trapped, part passing and part trapping. In most areas the $p_h \sim \Delta\varphi_h$ curves are very flat. For passing trajectories $\Delta\varphi_h$ is at about l/l_h , and indicates that the passing trajectories almost follow the magnetic field lines. For trapped trajectories $\Delta\varphi_h$ is small compared to unity because the bounce motion of trapped particles is much faster than the drift motion (the precession around the torus). The $p_h \sim \Delta\varphi_h$ curve has a sharp slope in the barely trapped region. $\Delta\varphi_h$ has a larger value for a barely trapped trajectory because the particle spends more time at the top of the potential barrier and therefore has a longer bounce period. The big island in the (p_h, φ_h) SOS plot comes from those "sitting" banana orbits (i.e. those with $\Delta\varphi_h = 0$). The small island chains come from those barely trapped orbits with $\Delta\varphi_h = l/n$.

C. The Trajectories of H_e

When the guiding center of the exact Hamiltonian, H_e of Eq. (6-1.16), is plotted in the

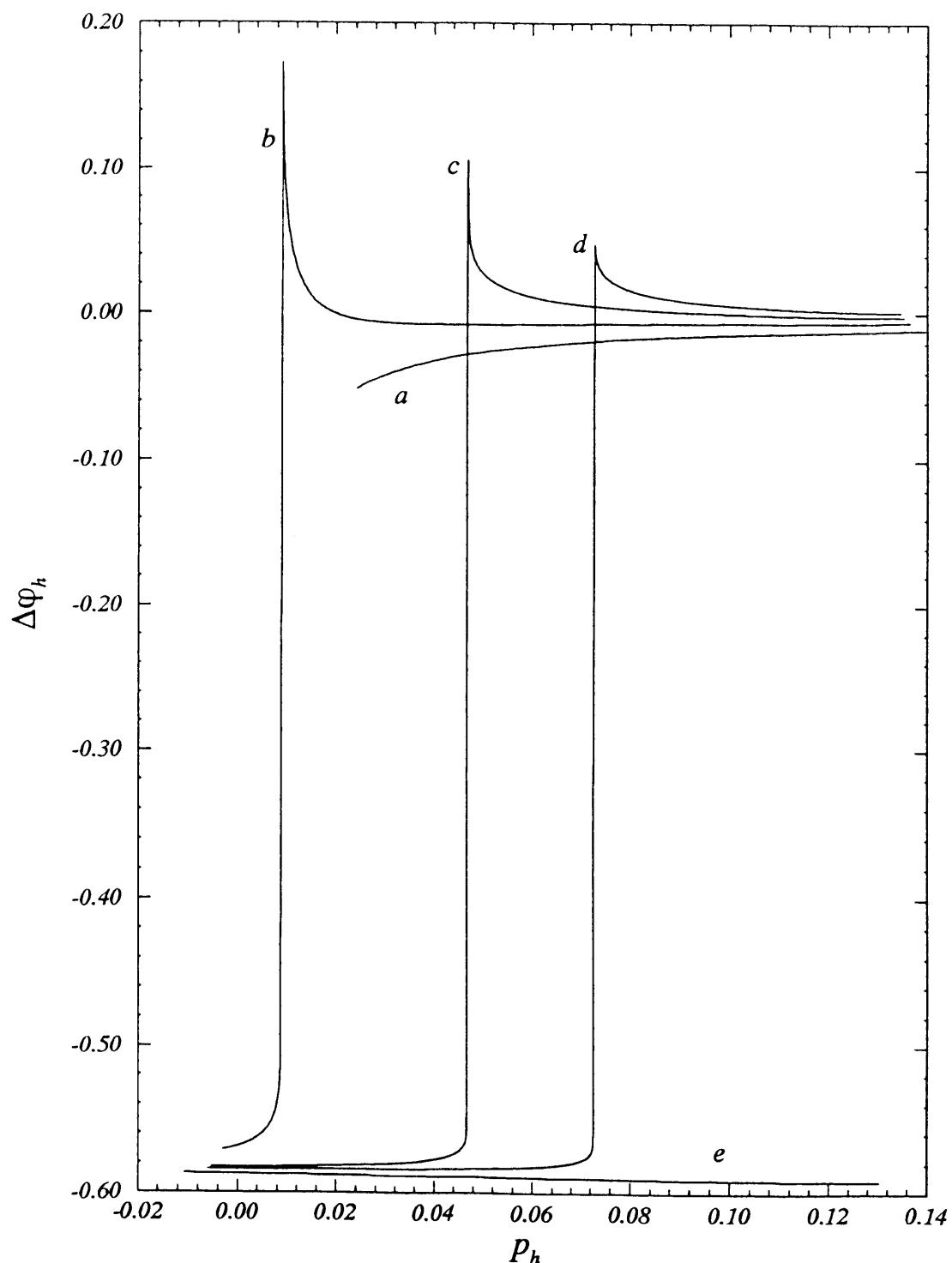


Figure 6.4 The $p_h - \Delta\varphi_h$ plots. The value of $\Delta\varphi_h$ is the change in φ_h after one repeat of the motion in $\psi \sim \theta$ plane for a given p_h . *a*: all trajectories are trapped. *b*, *c* and *d*: part of the trajectories are trapped and part of the ones are passing. *e*: all trajectories are passing.

$\psi \sim \theta$ plane, the orbits of the guiding center oscillate in and out like those of the higher order drift Hamiltonian (see Fig. 6.6(a)). The phase space structure of the exact Hamiltonian can not be compared with the one of the drift Hamiltonian directly, because the exact Hamiltonian H_e depends on six coordinates instead of four. In these six coordinates s and p are the ones describing the fast gyromotion and are of little interest. Furthermore if we calculate

$$\mu_{sp} = \frac{1}{2\pi} \oint p \, ds$$

as we integrate along the trajectories, μ_{sp} is almost a constant (see Fig. 6.5). For a typical shot we follow a particle for a time of $10^6 \omega^{-1}$ with ω the gyrofrequency. We fit $t \sim \mu_{sp}$ with a straight line. The slope of this line is of order 10^{-16} and the standard deviation over the average value of μ_{sp} , $\sigma_{sd}/\bar{\mu}_{sp}$, is of order 10^{-3} . The constancy of μ_{sp} implies that the two fast coordinates (s, p) are almost separated from the other four coordinates $(\alpha, p_\alpha; \varphi_h, p_h)$, which describe the motion of the guiding center. Thus we can compare the sub-phase-space $(\alpha, p_\alpha; \varphi_h, p_h)$ of the exact Hamiltonian H_e with the phase space of the higher order drift Hamiltonian K_d and ignore the (s, p) part. Figure 6.6 (b) shows a (p_h, φ_h) SOS of the sub-phase-space $(\alpha, p_\alpha; \varphi_h, p_h)$ of the exact Hamiltonian. Those "invariant" surfaces are a little fuzzy. We believe this is caused by the oscillation of μ_{sp} rather than its diffusion since μ_{sp} hardly diffuses for the time we integrate along the trajectories. The basic structure of this SOS is the same as the one of the high order drift Hamiltonian of Fig. 6.3 (b)

6-3 Conclusions

We numerically study the effect of the first order correction to the standard drift Hamiltonian for a vacuum field with exact quasihelical symmetry. We find that the phase space structure of the standard drift Hamiltonian is changed due to the first order correction. Because of the "singular" property at the passing and trapped boundaries and

6. A Numerical Study of the Drift Approximation

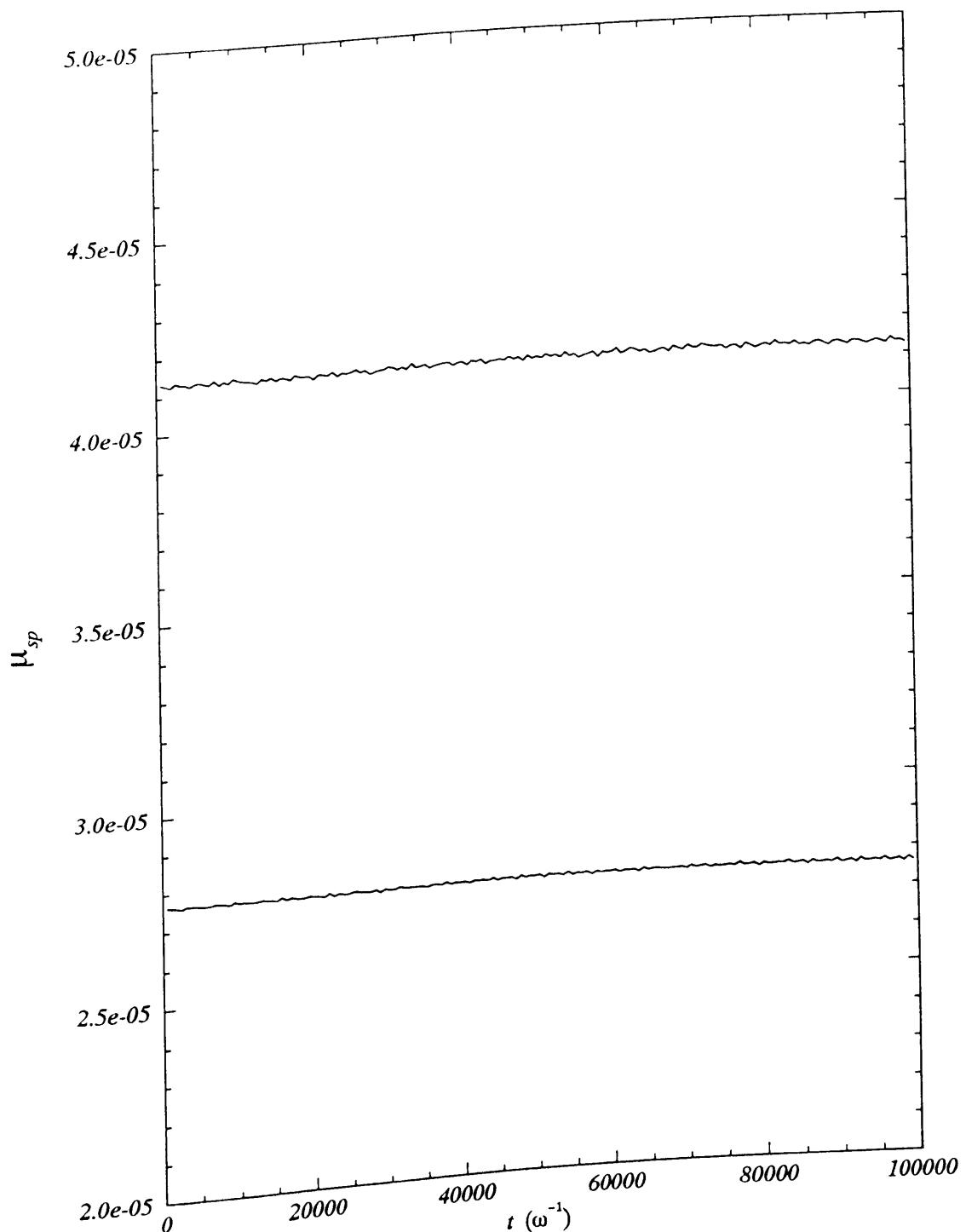
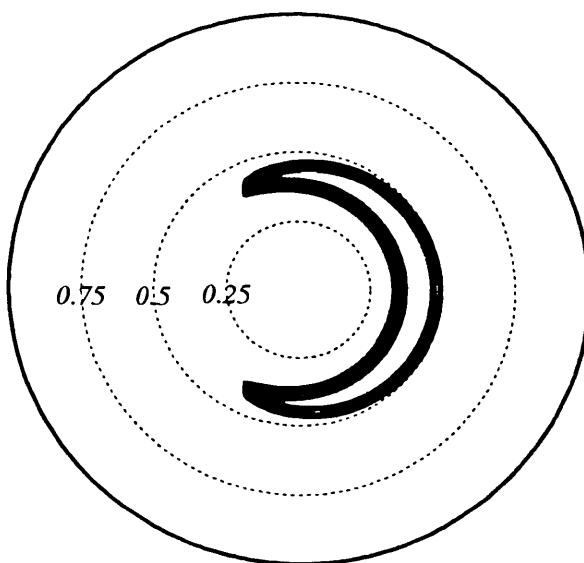
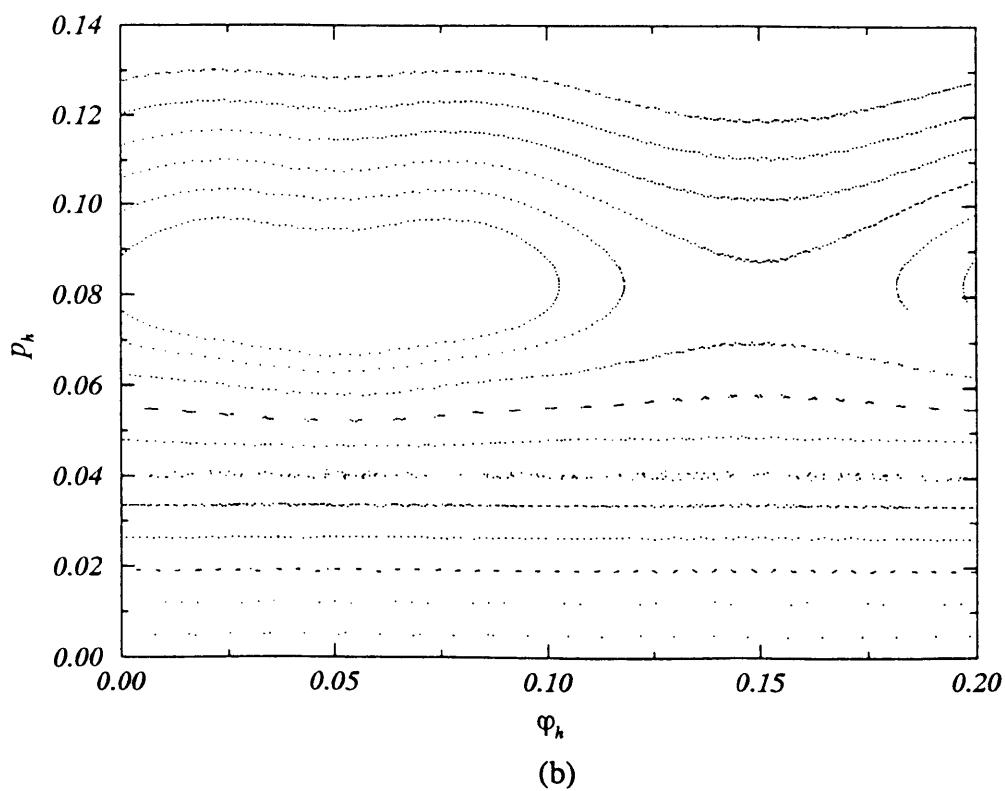


Figure 6.5 The quantity μ_{sp} changes with time for two different trajectories.



(a)



(b)

Figure 6.6 Trajectory of the exact Hamiltonian. (a) A trapped trajectory of guiding center view in $\psi \sim \theta$ plane. (b) The (p_h, φ_h) surface of section for the sub-phase-space of $(\alpha, \psi; \varphi_h, p_h)$.

the dense distribution of the low mode number resonant surfaces in the barely trapped region, the region near the passing trapped boundary becomes chaotic. The most important effect is the resonance of the "sitting" banana orbits. The oscillation of the "sitting" banana orbits may cause the particle prompt lost. Furthermore, if the gyroradius big enough, the majority of the trapped trajectories will become stochastic and the particle confinement is lost.

References

- ¹ Q. Yao and A. H. Boozer, Phys. Fluids B **4**, 2429 (1992)
- ² H. Alfvén, Ark. Mat Astron. Fys. **27A**, 1 (1940)
- ³ J. Nührenburg and R. Zille, Phys. Lett. A **129**, 113 (1988)
- ⁴ D. A. Garren and A. H. Boozer, Phys. Fluids B **3**, 2822 (1991)

CHAPTER 7 SUMMARY

We investigate the guiding center approximation for the electromagnetic field of a toroidal configuration. The first order correction to the standard drift Hamiltonian is derived in the magnetic coordinates (ψ, θ, φ) , in which the magnetic field has both simple covariant and contravariant representations

$$\mathbf{B}(\psi, \theta, \varphi) = \mu_0 [G(\psi)\nabla\varphi + I(\psi)\nabla\theta + \beta_*(\psi, \theta, \varphi)\nabla\psi],$$

$$\mathbf{B}(\psi, \theta, \varphi) = \nabla\psi \times \nabla\theta + \iota(\psi)\nabla\varphi \times \nabla\psi.$$

The results of the numerical study on the effect of the first order correction are also presented for a special case of a quasihelically symmetric field.

We start with the exact Hamiltonian of a charged particle

$$H_e = \frac{1}{2}mv_{||}^2 + \frac{1}{2}mv_{\perp}^2 + e\phi.$$

The first term is the parallel kinetic energy, the second term is the perpendicular kinetic energy and the last term is the electric potential energy. When this Hamiltonian is represented in the magnetic coordinates (ψ, θ, φ) , the perpendicular kinetic energy becomes

$$\frac{1}{2}mv_{\perp}^2 = \frac{1}{2m}(g^s s^2 - 2g^c s u + g^{\psi} u^2),$$

where s and u are the two variables describing the fast gyromotion, and the three functions, g^s , g^c and g^ψ , are the metric components of the magnetic coordinates. The two fast variables are generally not canonical to each other except for the special case of a vacuum magnetic field. There are three important steps in deriving the high order drift Hamiltonian: 1) Taylor expanding the exact Hamiltonian at the guiding center with a small parameter ε , which is the ratio of gyroradius to system size; 2) Transferring the two fast coordinates, s and u , into action-angle variables, μ and θ_g (magnetic momentum and gyrophase); 3) Eliminating the gyrophase dependence. We show

$$H_e(\mu, \theta_g; Z) = H_d(\mu; Z) + H_I(\mu; Z) + H_\mu(\mu, \theta_g; Z).$$

with H_d the standard drift Hamiltonian, H_I the first order correction to the standard drift Hamiltonian, and H_μ the sum of all the higher order terms. Compared with H_e , H_d is of order 1, H_I is of order ε and the leading order term in H_μ is of order ε^2 . The procedure we use in step three can be applied to derive the higher order drift Hamiltonian. However, the result becomes more complicated. Besides, the most important correction terms to H_d are the lowest order ones that resonant with H_d , which is in H_I .

Since the magnetic coordinates (ψ, θ, φ) are used, a major difference between the standard drift Hamiltonian and the exact Hamiltonian becomes obvious. The exact Hamiltonian depends on the vector properties of magnetic fields, i.e. both the field strength and the geometry of the flux surfaces. The standard drift Hamiltonian, however, depends only on the field strength. Because of this difference the standard drift Hamiltonian can have extra symmetries which the exact Hamiltonian does not have. We show that the first order correction, H_I , depends on the geometry of the flux surfaces.

The most interesting example is the quasihelically symmetric stellarator, for which the field strength has a symmetry and the flux surfaces do not. In the quasihelically symmetric field, trajectories of the standard drift Hamiltonian are perfectly confined owing to the conservation of the helical momentum. Trajectories of the exact Hamiltonian may not be confined for the lack of constraints. We numerically study the trajectories of both the drift Hamiltonian with the first order correction and the exact Hamiltonian for a quasihelically

symmetric field. We find that in most of the phase space the invariant KAM surfaces of the standard drift Hamiltonian are perturbed but still exist. In the region near the “sitting” banana orbits (the banana orbits do not precess around the torus) a big island appears. In physical space this means those “sitting” banana orbits oscillate in and out about their home flux surfaces. When the big island is close enough to the wall, the oscillation can cause prompt particle loss. The results from the high order drift Hamiltonian agree with those of the exact Hamiltonian.

The work we present in this dissertation is a beginning of the study of the guiding center approximation. Further investigation of this subject is still needed for both theoretical and experimental interest.

The quasihelically symmetric stellarator is a new generation of the stellarator fusion devices. Because of its helical symmetry in the field strength, it has the advantage of having a similar particle confinement property to axis-symmetric devices such as the tokamak, but it does not require the toroidal plasma current. We believe it is important to include the first order correction to the standard drift Hamiltonian in the particle simulation codes for such a stellarator. The first order correction may affect the particle confinement property due to the flux surfaces having no helical symmetry. It is worthwhile to point out that for the axis-symmetric devices like tokamaks the first order correction may not be as important as for quasihelically symmetric stellarators, because the symmetry of the flux surfaces is broken at the same order as the field strength. However, the ripples in the field strength, which are caused by the finite number of toroidal field coils, may have the same effect as the first order correction on quasihelical symmetric stellarator.

Another important aspect that we have not looked into is the effect of gyromotion. It is not known if the adiabatic invariant, the magnetic moment, is bounded or not during the time of interest. Even if it is bounded, its oscillation may still cause the discrepancy between the trajectories of the exact Hamiltonian and the drift Hamiltonian.

BIBLIOGRAPHY

- Alfvén, H. (1940). *Ark. Mat. Astron. Fys.* **27A**, 1.
- Arnold, V. I. (1963). *Sov. math. Dokl.* **3**, 136.
- Arnold, V. I. (1964). *Russian Math. Surveys* **18**, 85.
- Arnold, V. I. (1978). *Mathematical Methods of Classical Mechanics* (Springer, New York).
- Bender, C. M. and S. A. Orszag (1978). *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill, Inc.).
- Birkhoff, G. D. (1927). *Dynamical System* (American Mathematical Society, New York).
- Bogoliubov, N. N. and Y. A. Mitropolsky (1961). *Asymptotic Methods in the Theory of Nonlinear Oscillations* (Gordon and Breach, New York).
- Boozer, A. H. (1980). *Phys. Fluids* **23**, 904.
- Boozer, A. H. (1981). *Phys. Fluids* **24**, 1999.
- Boozer, A. H. (1982). *Phys. Fluids* **25**, 520.
- Boozer, A. H. (1982). *Phys. Fluids* **25**, 575.
- Boozer, A. H. (1983). *Phys. Fluids* **26**, 496.
- Boozer, A. H. (1983). *Phys. Fluids* **26**, 1288.
- Boozer, A. H. (1984). *Phys. Fluids* **27**, 2441.
- Boozer, A. H. (1991). *Phys. Fluids* **B3**, 875.
- Boozer, A. H. (1991). Lecture notes on *Plasma Physics*.
- Broman, A. (1989). *Introduction to Partial Differential Equations* (Dover Publications, Inc., New York)
- Chen, F. F. (1974). *Introduction to Plasma Physics* (Plenum Press, New York).
- Cray, J. R. and R. G. Littlejohn (1983). *Ann. Phys.* **151**, 1.
- Deprit, A. (1969). *Cel. Mech.* **1**, 12.
- Dewar, R. L. (1976). *J. Phys.* **A9**, 2043.
- Gardner, C. S. (1959). *Phys. Rev.* **115**, 791.
- Gardner, C. S. (1966). *Phys. Fluids* **9**, 1997.
- Garrido, L. M. (1968). *J. Math. Phys.* **10**, 1045.
- Garren, D. A. and A. H. Boozer (1991). *Phys. Fluids B* **3**, 2822.

- Goldstein, H. (1980). *Classical Mechanics* (Addison-Wesley Publishing Company, Inc.).
- Greene, J. M. and J. L. Johnson (1962). Phys. Fluids **5**, 510.
- Hamada, S. (1962). Nucl. Fusion **2**, 23.
- Hori, G. (1966). Astron. Soc. Japan **18**, 287.
- Howland, R. A. (1977). Cel. Mech. **17**, 327.
- Kolmogorov, A. N. (1954). Dokl. Akad. Nauk. USSR **98**, 527.
- Kruskal, M. D. and R. M. Kulsrud (1958). Phys. Fluids **1**, 265.
- Kruskal, M. D. (1962). J. Math. Phys. **3**, 806.
- Kuo-Petraovic, G., A. H. Boozer, J. A. Rome and R. H. Fowler (1983). J. Comp. Phys. **51**, 261.
- Lichtenberg, A. J. and M. A. Lieberman (1983). *Regular and Stochastic Motion* (Springer-Verlag, New York).
- Littlejohn, R. G. (1979). J. Math. Phys. **20**, 2445.
- Littlejohn, R. G. (1981). Phys. Fluids **24**, 1730.
- Littlejohn, R. G. (1982). J. Math. Phys. **23**, 742.
- Moser, J. (1962). Nachr. Akad. Wiss. Göttingen. Math. Phys. **K1**, 1.
- Mynick, H. E. (1979). Phys. Rev. Lett. **43**, 1019.
- Mynick, H. E. (1980). Phys. Fluids **23**, 1888.
- Northrop, T. G. and J. A. Rome (1978). Phys. Fluids **21**, 384.
- Nührenburg, J. and R. Zille (1988). Phys. Lett. A **129**, 113.
- Poincaré, H. (1892). *Les Méthodes Nouvelles de la Mécanique Céleste* (Gauthier-Villars, Paris).
- Siegel, C. L. Ann. (1941). Math. **42**, 806.
- Siegel, C. L. Ann. Math. (1954). Anal. **128**, 144.
- Taylor, J. B. (1964). Phys. Fluids **7**, 767.
- Von Zeipel, H. (1916). Ark. Astron. Mat. Fys. **11**, No.1.
- Yao, Q. and A. H. Boozer (1992). Phys. Fluids B **4**, 2429.

VITA

Qun Yao

The author was born in Shanghai, China, March 15, 1962 and married to Jing Gao on February 8, 1988. He attended East China Normal University, graduated in June of 1985 with the Bachelor of Science degree in physics. He had his Master of Science in chemistry in 1988 from the same university. In 1989 the author enter the College of William and Mary in Virginia to pursue his graduate career in physics. He was a teaching assistant from June 1989 to May 1991. He has been a research assistant since June 1991. The author was a Visiting Reseascher at the Torsatron/Stellarator Laboratory in the University of Wisconsin, Madison from June to August in 1991. He has been a Visiting Scholar at MIT Plasma Fusion Center since June 1992. The author is a member of the Chinese Chemistry Society and a member of the American Physics Society.