Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

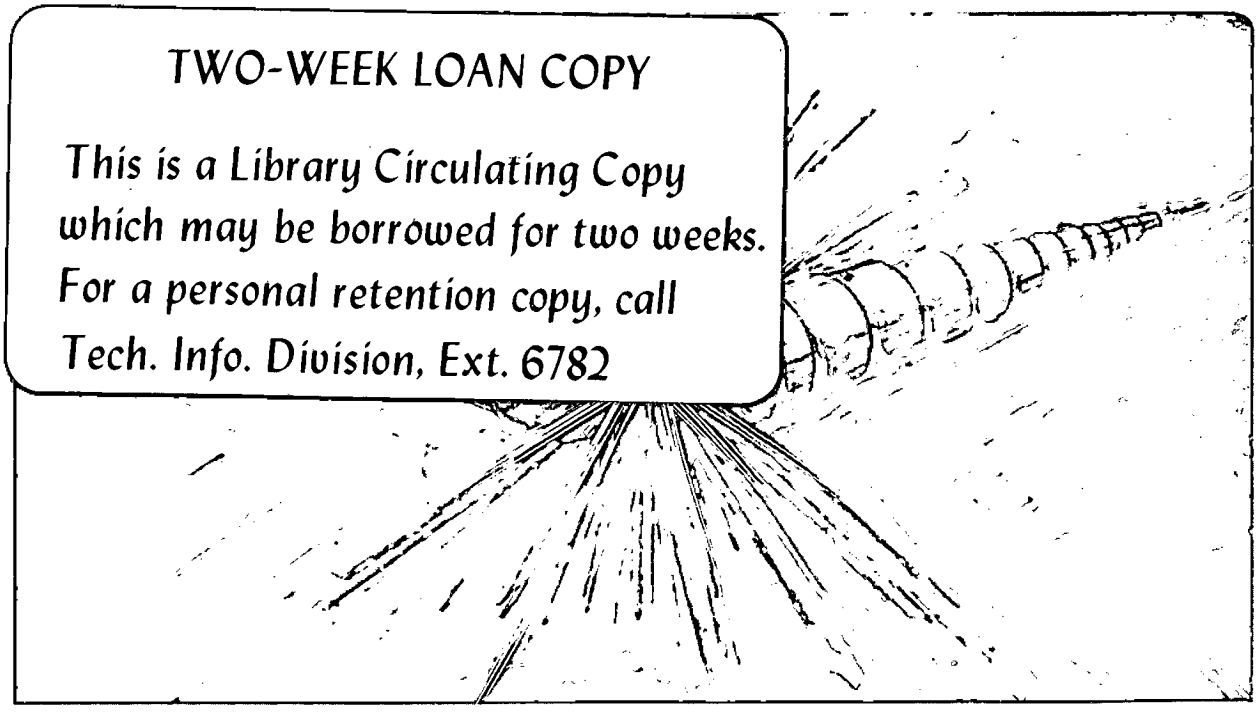
Accelerator & Fusion Research Division

Submitted to Physics of Fluids

A HAMILTONIAN FORMULATION OF GUIDING CENTER MOTION

Robert G. Littlejohn

June 1979



Prepared for the U. S. Department of Energy
under Contract W-7405-ENG-48

DISCLAIMER

To be submitted to Phys. Fluids.

LBL-9141

A HAMILTONIAN FORMULATION OF GUIDING CENTER MOTION*

Robert G. Littlejohn

Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

ABSTRACT

Nonrelativistic guiding center motion in the magnetic field $\underset{\sim}{B}=\underset{\sim}{B}(\underset{\sim}{x})$, with $\underset{\sim}{E}=0$, is studied using Hamiltonian methods. The drift equations are carried to second order in the perpendicular motion. The Hamiltonian methods which are used are described in detail in order to facilitate possible applications. Unusual mathematical techniques are called upon, especially the use of noncanonical coordinates in phase space. Lie transforms are used to carry out the perturbation expansion. Applications in kinetic theory, in the area of adiabatic invariants, and in other areas are anticipated.

*Work was supported by the Office of Fusion Energy of the U.S. Department of Energy under contract No. W-7405-ENG-48.

I. INTRODUCTION

In a recent paper[1] I have described the mathematical apparatus of a new approach to a Hamiltonian formulation of guiding center motion, and I have illustrated the method with the problem of nonrelativistic guiding center motion in the magnetic field $\underset{\sim}{B} = B(x,y)\hat{z}$. In this paper I will extend those results to the case of a nonrelativistic particle moving in a time-independent but otherwise arbitrary magnetic field $\underset{\sim}{B} = \underset{\sim}{B}(\underset{\sim}{x})$, with the electric field $\underset{\sim}{E} = 0$. Throughout this paper, except in Appendix A, a familiarity with the mathematical methods of Ref. 1 will be assumed.

The study of guiding center motion is essentially a problem in perturbation theory in classical mechanics. Although it has always been known that charged particle motion can be described in Hamiltonian terms, nevertheless most of the results that have been obtained in this area have been derived with non-Hamiltonian perturbation methods.[2-8] Therefore Hamiltonian methods have not found wide application in studies of guiding center motion in plasma physics, in spite of the great interest in the dynamics of plasmas in nonuniform magnetic fields. This is unfortunate, because Hamiltonian methods provide great computational advantages over non-Hamiltonian methods, as well as a formalism which is notable for its elegance and notational compactness.

The original Hamiltonian treatment of guiding center motion was given by Gardner,[9] who employed field line coordinates and mixed-variable generating functions in an algorithm to systematically remove the dependence of the Hamiltonian on gyrophase to all orders. Gardner's methods were elaborated upon by Taniuti,[10] Stern,[11] and others, who also used field line coordinates and mixed-variable generating functions.

Recently Mynick[12] has developed a theory of guiding center motion using Hamiltonian methods. Mynick has also used field line coordinates, but in contrast to the authors above he has used a combination of mixed-variable generating functions and Lie transforms. Mynick seems to be the first to have employed the great power of Lie transforms in guiding center work. By way of additional contrast, Mynick has used an ordering scheme which treats the parallel and perpendicular scale lengths with different ordering parameters. His results are perturbative, i.e. represented by power series, only in the parallel ordering parameter. The results are in closed form for the perpendicular ordering parameter.

In addition, some recent work by Meyer[13] has shown how the guiding center problem can be treated without using either mixed-variable generating functions or Lie transforms. Instead, Meyer has developed canonical transformations by appealing directly to the defining Poisson bracket relations. Furthermore, Meyer has avoided the use of field line coordinates. Meyer's work has many points in common with the theory presented in Ref. 1 and here, although the detailed nature of the connection remains to be established.

Two salient features of this work are the use of rectangular coordinates instead of field line coordinates in configuration space and the use of noncanonical coordinates in phase space. The latter especially calls into play certain unusual mathematical techniques, which are described in Ref. 1. In addition, the perturbation expansion which is used to eliminate the dependence of the Hamiltonian on gyrophase is effected by means of Lie transforms.

It is the primary purpose of this paper to provide the details of a Hamiltonian treatment of guiding center motion, rather than simply the resulting drift equations. The hope is that this paper will lay the groundwork for applications in kinetic theory and other areas. Therefore I will go into much more detail than would be necessary if only the drift equations were of interest.

Nevertheless, the most immediate and tangible results of this work are the drift equations, which are carried out to second order in the perpendicular motion of the guiding center. Using non-Hamiltonian methods, Northrop and Rome[8] have carried the drift equations to the same order under the same assumptions, viz. nonrelativistic motion in a static magnetic field. Therefore there is little that is new in the drift equations, although the form which is developed here for the second order guiding center position gives rise to equations of motion which are less complicated than those of Northrop and Rome. This may be seen most easily in Appendix A. Finally, I should note that a detailed comparison of these results with those of Northrop and Rome shows complete agreement.

Since there will perhaps be readers who will be interested only in the drift equations, and not in the Hamiltonian methods used to derive them, I have given in Appendix A a summary of the drift equations for a particularly convenient (but non-Hamiltonian) choice of guiding center variables, employing a notation which is as independent as possible of conventions established earlier in the paper. This Appendix should be especially useful for numerical or simulation work.

The organization of this paper is as follows. The basic purpose of Sec. 2 is to define the problem. In this section, we introduce three sets of phase space coordinates, which are called "physical particle variables." The last two sets especially have great physical immediacy, and their use has the important effect of banishing, once and for all, the magnetic vector potential $\underset{\sim}{A}$ from the formalism.

Sec. 3 contains a number of technical details of the algebra which must be used in a treatment of guiding center motion to second order. This algebra focuses on the system of unit vectors employed, and special attention is given to the perpendicular unit vectors. Most of this section would be unnecessary if the guiding center Hamiltonian were only carried to lowest order. Similarly, much of the algebraic details given in this section would be unavoidable in any treatment of guiding center motion to second order, whether it be Hamiltonian or not.

Secs. 4 and 5 are devoted to the Darboux transformation. Since there are a number of properties of the Darboux transformation which can be expressed in closed form, most notably the components of the Poisson tensor in the resulting coordinate system, these properties are derived and listed in Sec. 4. The Darboux transformation itself must be developed as power series in ε, and this development is carried out in Sec. 5.

In Sec. 6 we perform the averaging transformation, using Lie transforms, and obtain thereby the guiding center Hamiltonian as well as a set of guiding center variables. It turns out that the guiding center variables depend on the choice of perpendicular unit vectors which is made in the problem definition. In order to deal with this situation, we discuss

at length the degree of arbitrariness in the guiding center variables, and we prove that in any semicanonical coordinate system, such as seems to be necessary for a Hamiltonian treatment of any kind, a dependence on the choice of perpendicular unit vectors is unavoidable. In a noncanonical coordinate system, however, such a dependence can be eliminated, at least through second order. Indeed, the noncanonical guiding center variables used in Appendix A are free of such dependencies.

Finally, in Sec. 7 we discuss the results and suggest various extensions and applications.

2. PHYSICAL PARTICLE VARIABLES

In this section we will discuss three relatively simple coordinate systems in phase space. Of these, the first consists of a slight variation on the rectangular canonical coordinates $(\underset{\sim}{q},\underset{\sim}{p})$ which are usually used in a Hamiltonian formulation of the motion of a charged particle in a magnetic field. The other two coordinate systems are related in a simple manner to the instantaneous dynamical state of the particle and to the magnetic field at the particle position. Therefore the variables making up these coordinate systems will be called "physical particle variables," in contrast to guiding center variables, which will be introduced later. Of the three coordinate systems described in this section, only the first is a canonical system. In addition, we will establish certain notational conventions in this section.

The motion of a particle of charge e and mass m in a static magnetic field $\underset{\sim}{B}(\underset{\sim}{x})$ with $\underset{\sim}{E}=0$ may be described by the Hamiltonian

$$H(\underset{\sim}{q},\underset{\sim}{p}) = \frac{1}{2}\left[\underset{\sim}{p} - \frac{e}{mc}\underset{\sim}{A}(\underset{\sim}{q})\right]^2 \tag{2.1}$$

where $\underset{\sim}{A}$ is the magnetic vector potential satisfying $\underset{\sim}{B} = \nabla\times\underset{\sim}{A}$. This Hamiltonian differs slightly from the usual Hamiltonian for a charged particle. It does, however, give the correct equations of motion as long as the canonical coordinates $(\underset{\sim}{q},\underset{\sim}{p})$ are related to the particle's position $\underset{\sim}{x}$ and velocity $\underset{\sim}{v}$ by

$$\begin{aligned} \underset{\sim}{x} &= \underset{\sim}{q} \\ \underset{\sim}{v} &= \underset{\sim}{p} - \frac{e}{mc}\underset{\sim}{A}(\underset{\sim}{q}) \end{aligned} \tag{2.2}$$

To use the Hamiltonian (2.1) it should be remembered that the Hamiltonian has dimensions of energy/mass, i.e. (velocity)^2, and that the canonical momentum $\underset{\sim}{p}$ has dimensions of velocity.

With the ordering $(\underset{\sim}{q},\underset{\sim}{p})$, the Poisson tensor (which was called the σ-tensor in Ref. 1) has the following components:

$$\sigma^{ij} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \tag{2.3}$$

Here I represents the 3×3 identity matrix.

The guiding center approximation is introduced into the Hamiltonian (2.1) by replacing the charge e by e/ε, where ε is a formal expansion parameter. The result is

$$H(\underset{\sim}{q},\underset{\sim}{p}) = \frac{1}{2}\left[\underset{\sim}{p} - \frac{e}{\varepsilon mc}\underset{\sim}{A}(\underset{\sim}{q})\right]^2 \tag{2.4}$$

In addition, the transformation law (2.2) is modified as follows:

$$\begin{aligned} \underset{\sim}{x} &= \underset{\sim}{q} \\ \underset{\sim}{v} &= \underset{\sim}{p} - \frac{e}{\varepsilon mc}\underset{\sim}{A}(\underset{\sim}{q}) \end{aligned} \tag{2.5}$$

The Poisson tensor given in Eq. (2.3) does not change with the introduction of ε. Henceforth we will use Eqs. (2.4) and (2.5) instead of (2.1) and (2.2).

The parameter ε may be considered to be a variable, describing a family of systems, of which the one corresponding to $\varepsilon = 1$ is the physical system. The order of an expression is determined by its behaviour as $\varepsilon \to 0$, while the position $\underset{\sim}{x}$, velocity $\underset{\sim}{v}$, and fields $\underset{\sim}{A}$ and $\underset{\sim}{B}$ are held fixed. Thus the Hamiltonian is O(1), and the canonical

momentum $\underset{\sim}{p}$ is $O(\varepsilon^{-1})$.

The second coordinate system consists of the particle variables $\underset{\sim}{x}$ and $\underset{\sim}{v}$, which are related to $\underset{\sim}{q}$ and $\underset{\sim}{p}$ by Eq. (2.5). In this coordinate system, with the ordering $(\underset{\sim}{x},\underset{\sim}{v})$, the Poisson tensor has the form

$$\sigma^{ij} = \begin{pmatrix} 0 & I \\ -I & \frac{e}{\varepsilon mc} B \end{pmatrix} \tag{2.6}$$

where the symbol B represents the antisymmetric tensor which is dual to the magnetic field vector $\underset{\sim}{B}$:

$$B_{ij} = \varepsilon_{ijk} B_k \tag{2.7}$$

Here and in the remainder of this paper summation over repeated indices is to be understood. As for the Hamiltonian, it is especially simple in the $(\underset{\sim}{x},\underset{\sim}{v})$ coordinates:

$$H(\underset{\sim}{x},\underset{\sim}{v}) = \frac{1}{2} v^2 \tag{2.8}$$

An alternate form for the Poisson tensor, which is completely equivalent to Eq. (2.6), is sometimes useful. If we are given any two phase functions F and G, expressed in terms of $(\underset{\sim}{x},\underset{\sim}{v})$, then their Poisson bracket {F,G} is given by

$$\{F,G\} = \frac{\partial F}{\partial \underset{\sim}{x}} \cdot \frac{\partial G}{\partial \underset{\sim}{v}} - \frac{\partial F}{\partial \underset{\sim}{v}} \cdot \frac{\partial G}{\partial \underset{\sim}{x}} + \frac{1}{\varepsilon} \underset{\sim}{\Omega} \cdot \left(\frac{\partial F}{\partial \underset{\sim}{v}} \times \frac{\partial G}{\partial \underset{\sim}{v}} \right) \tag{2.9}$$

where the vector $\underset{\sim}{\Omega}$ is defined by

$$\underset{\sim}{\Omega} = \frac{e}{mc} \underset{\sim}{B} \tag{2.10}$$

The third coordinate system represents a kind of cylindrical coordinates in velocity space, with the local magnetic field vector indicating the direction of the cylinder axis. We write $B(\underset{\sim}{x}) = |\underset{\sim}{B}(\underset{\sim}{x})|$, and define the unit vector $\hat{b}(\underset{\sim}{x})$ by

$$\hat{b}(\underset{\sim}{x}) = \frac{\underset{\sim}{B}(\underset{\sim}{x})}{B(\underset{\sim}{x})} \tag{2.11}$$

It is convenient to assume that $B(\underset{\sim}{x})$ is bounded away from zero in the spatial region of interest. Not only does this guarantee that the vector $\hat{b}(\underset{\sim}{x})$ is continuous, but it is also a necessary condition for the validity of the guiding center approximation.

Two variables of the new coordinate system are defined in terms of the velocity $\underset{\sim}{v}$ and the vector $\hat{b}$. These are u and w, the instantaneous parallel and perpendicular velocities, respectively, and they are given by

$$u = \underset{\sim}{v}\cdot\hat{b}(\underset{\sim}{x}) \tag{2.12}$$

$$w = (v^2 - u^2)^{\frac{1}{2}} \tag{2.13}$$

Let us now introduce, in addition to $\hat{b}$, two more fields of unit vectors, which are called $\hat{\tau}_1(\underset{\sim}{x})$ and $\hat{\tau}_2(\underset{\sim}{x})$, and which are illustrated in Fig. 1. Taken together with $\hat{b}$, these form a right-handed set of unit vectors:

$$\hat{\tau}_1\cdot\hat{\tau}_1 = \hat{\tau}_2\cdot\hat{\tau}_2 = \hat{b}\cdot\hat{b} = 1 \tag{2.14}$$

$$\hat{\tau}_1\cdot\hat{\tau}_2 = \hat{\tau}_1\cdot\hat{b} = \hat{\tau}_2\cdot\hat{b} = 0 \tag{2.15}$$

$$\hat{b} = \hat{\tau}_1\times\hat{\tau}_2 \tag{2.16}$$

For the time being, we may assume that $\hat{\tau}_1$ and $\hat{\tau}_2$ are arbitrary, apart from the relations (2.14)-(2.16). Later we will consider the possibility

of a judicious choice for $\hat{\tau}_1$ and $\hat{\tau}_2$.

It is useful to define several more quantities, relative to the $\hat{\tau}_1$-$\hat{\tau}_2$ plane, i.e. the perpendicular plane. These quantities are shown in Fig. 2. First we define the perpendicular velocity vector $\underset{\sim}{v}_\perp$ by $\underset{\sim}{v}_\perp = \underset{\sim}{v} - u\hat{b}$. Next, the gyroradius vector $\underset{\sim}{r}$ is given by $\underset{\sim}{r} = \varepsilon(\hat{b}\times\underset{\sim}{v})/\Omega$, where Ω is the signed gyrofrequency:

$$\Omega = \frac{eB}{mc} = \hat{b}\cdot\underset{\sim}{\Omega} = \text{sign}(e)\,|\underset{\sim}{\Omega}| \tag{2.17}$$

It is convenient to introduce a velocity-dependent unit vector $\hat{c}$, which is in the direction of the perpendicular velocity vector $\underset{\sim}{v}_\perp$, so that $\underset{\sim}{v}_\perp = w\hat{c}$, or

$$\underset{\sim}{v} = u\hat{b} + w\hat{c} \tag{2.18}$$

In addition, we define another velocity-dependent unit vector $\hat{a}$, given by

$$\hat{a} = \hat{b}\times\hat{c} \tag{2.19}$$

The triad $(\hat{a},\hat{b},\hat{c})$ forms a right-handed set. The gyroradius vector $\underset{\sim}{r}$ is related to the unit vector $\hat{a}$ by $\underset{\sim}{r} = \varepsilon w\hat{a}/\Omega$. Finally, the gyrophase θ is defined as the angle, measured in a clockwise sense, between $\hat{\tau}_1$ and $\hat{a}$. Thus we have

$$\begin{aligned} \hat{a} &= \cos\theta\ \hat{\tau}_1 - \sin\theta\ \hat{\tau}_2 \\ \hat{c} &= -\sin\theta\ \hat{\tau}_1 - \cos\theta\ \hat{\tau}_2 \end{aligned} \tag{2.20}$$

Our third coordinate system in phase space consists of the six physical particle variables $(\underset{\sim}{x},u,\theta,w)$ as just defined. In these definitions

we have refrained from referring to circles or circular motion, because in general the motion is not exactly circular, and because in a theory which is to be systematic to any order we do not want to call upon concepts which are vague beyond lowest order. In this sense the terms "gyroradius" and "gyrofrequency" are imprecise. Therefore the definitions above may perhaps best be taken as closed-form, algebraic relations specifying a variable transformation $(\underset{\sim}{x},\underset{\sim}{v}) \rightarrow (\underset{\sim}{x},u,\theta,w)$.

Nevertheless, in the special case of a uniform magnetic field, these variables do have a physical meaning which is both simple and precise, because the perpendicular motion of the particle is circular. Fig. 3 shows the meaning of some of these variables in the case of a uniform magnetic field. In this case the guiding center positon $\underset{\sim}{X}$ is given by $\underset{\sim}{X} = \underset{\sim}{x} - \underset{\sim}{r}$ exactly, and it is the precise center of the circle of motion. Later we will discuss ways in which the definition of $\underset{\sim}{X}$ may be extended to the case of nonuniform fields in a manner which is systematic to all orders. For now, however, we simply use Fig. 3 for its suggestive value. For example, it may be seen that the unit vectors $\hat{a}$ and $\hat{c}$ rotate with the particle, in a clockwise direction (θ increasing) for a positive particle, and in a counterclockwise direction (θ decreasing) for a negative particle.

To complete the description of the $(\underset{\sim}{x},u,\theta,w)$ coordinate system, we need the Hamiltonian and the Poisson tensor. The former is easy to obtain:

$$H(\underset{\sim}{x},u,\theta,w) = \frac{1}{2}(w^2 + u^2) \tag{2.21}$$

As for the Poisson tensor, it may be obtained from Eq. (2.9) and the relation $\{z^i,z^j\} = \sigma^{ij}$, with $\underset{\sim}{z} = (\underset{\sim}{x},u,\theta,w)$ being taken as the six-dimensional coordinate vector. A little calculation gives the components of the Poisson tensor in the following form:

$$\{\underset{\sim}{x},\underset{\sim}{x}\} = 0 \tag{2.22a}$$

$$\{\underset{\sim}{x},u\} = \hat{b} \tag{2.22b}$$

$$\{\underset{\sim}{x},\theta\} = -\hat{a}/w \tag{2.22c}$$

$$\{\underset{\sim}{x},w\} = \hat{c} \tag{2.22d}$$

$$\{u,\theta\} = -\hat{a}\cdot\nabla\hat{b}\cdot\hat{c} - \hat{b}\cdot\nabla\hat{c}\cdot\hat{a} - \frac{u}{w}\,\hat{b}\cdot\nabla\hat{b}\cdot\hat{a} \tag{2.22e}$$

$$\{u,w\} = w\,\hat{c}\cdot\nabla\hat{b}\cdot\hat{c} + u\,\hat{b}\cdot\nabla\hat{b}\cdot\hat{c} \tag{2.22f}$$

$$\{\theta,w\} = \frac{\Omega}{\epsilon w} + \hat{c}\cdot\nabla\hat{c}\cdot\hat{a} + \frac{u}{w}\,\hat{b}\cdot(\nabla\times\hat{b}) \tag{2.22g}$$

Two notational conventions have been used in these equations and should be mentioned. First, for any pair of vectors $\underset{\sim}{Y}$ and $\underset{\sim}{Z}$, $\nabla\underset{\sim}{Y}\cdot\underset{\sim}{Z}$ means $(\nabla\underset{\sim}{Y})\cdot\underset{\sim}{Z}$ and not $\nabla(\underset{\sim}{Y}\cdot\underset{\sim}{Z})$. This convention will be followed throughout this paper. And second, the operator ∇ is to be taken at fixed (u,θ,w), and not at fixed $\underset{\sim}{y}$. This convention is followed whenever we are expressing any relation in the $(\underset{\sim}{x},u,\theta,w)$ coordinate system.

There are altogether 15 independent components of a general 6×6 antisymmetric matrix, which Eqs. (2.22) give for the component matrix σ^{ij}. Of these, ϵ appears in only one, as shown by Eq. (2.22g).

Eqs. (2.22) contain a number of different expressions involving unit vectors and their gradients. Expressions of this type occur more and more frequently as one proceeds with the guiding center problem, especially at higher orders. Therefore we turn now to a systematic study of the properties of these unit vectors.

3. PROPERTIES OF THE UNIT VECTORS

A number of simple but important properties of the unit vectors follow from the orthonormality conditions, Eqs. (2.14)-(2.16). We include in this list of properties the velocity dependent vectors $\hat{a}$ and $\hat{c}$, defined by Eq. (2.20), since in the remainder of the calculation these vectors are even more useful than $\hat{\tau}_1$ and $\hat{\tau}_2$. First, we express the identity tensor I and the vector operator $\hat{b}\times$ in terms of the unit vectors:

$$\hat{\tau}_1\hat{\tau}_1 + \hat{\tau}_2\hat{\tau}_2 + \hat{b}\hat{b} = \hat{a}\hat{a} + \hat{b}\hat{b} + \hat{c}\hat{c} = I \tag{3.1}$$

$$\hat{\tau}_2\hat{\tau}_1 - \hat{\tau}_1\hat{\tau}_2 = \hat{a}\hat{c} - \hat{c}\hat{a} = \hat{b}\times I \tag{3.2}$$

Next, we have the following relations involving the gradients of the unit vectors:

$$\begin{aligned} &\nabla\hat{\tau}_1\cdot\hat{\tau}_1 = \nabla\hat{\tau}_2\cdot\hat{\tau}_2 = \nabla\hat{a}\cdot\hat{a} = \nabla\hat{b}\cdot\hat{b} = \nabla\hat{c}\cdot\hat{c} = 0 \\ &\nabla\hat{\tau}_1\cdot\hat{b} = -\nabla\hat{b}\cdot\hat{\tau}_1, \quad \nabla\hat{\tau}_2\cdot\hat{b} = -\nabla\hat{b}\cdot\hat{\tau}_2, \quad \nabla\hat{\tau}_1\cdot\hat{\tau}_2 = -\nabla\hat{\tau}_2\cdot\hat{\tau}_1 \\ &\nabla\hat{a}\cdot\hat{b} = -\nabla\hat{b}\cdot\hat{a}, \quad \nabla\hat{c}\cdot\hat{b} = -\nabla\hat{b}\cdot\hat{c}, \quad \nabla\hat{a}\cdot\hat{c} = -\nabla\hat{c}\cdot\hat{a} \end{aligned} \tag{3.3}$$

Third, the normalization of $\hat{b}$ implies the following useful identities:

$$\hat{b}\times(\nabla\times\hat{b}) = -\hat{b}\cdot\nabla\hat{b} \tag{3.4}$$

$$\nabla\times\hat{b} = \hat{b}\times(\hat{b}\cdot\nabla\hat{b}) + \hat{b}[\hat{b}\cdot(\nabla\times\hat{b})] \tag{3.5}$$

$$\underset{\sim}{Y}\underset{\sim}{Z}:\nabla\nabla\hat{b}\cdot\hat{b} = -(\underset{\sim}{Y}\cdot\nabla\hat{b})\cdot(\underset{\sim}{Z}\cdot\nabla\hat{b}) \tag{3.6}$$

where $\underset{\sim}{Y}$ and $\underset{\sim}{Z}$ are any two vectors. In particular, Eq. (3.6) implies

$$\hat{b}\hat{b}:\nabla\nabla\hat{b}\cdot\hat{b} = -(\hat{b}\cdot\nabla\hat{b})^2 \tag{3.7}$$

In addition to the above, the vector $\hat{b}$ satisfies the following relation,

on account of the Maxwell equation $\nabla \cdot \underset{\sim}{B} = 0$:

$$\nabla \cdot \hat{b} = - \frac{\hat{b} \cdot \nabla B}{B} \tag{3.8}$$

The vector $\nabla \hat{\tau}_1 \cdot \hat{\tau}_2 = \nabla \hat{c} \cdot \hat{a}$ is of special importance, so we assign to it the symbol $\underset{\sim}{R}$:

$$\underset{\sim}{R} = \nabla \hat{\tau}_1 \cdot \hat{\tau}_2 = \nabla \hat{c} \cdot \hat{a} \tag{3.9}$$

The vector $\underset{\sim}{R}$ has the following geometrical interpretation. The vectors $\hat{\tau}_1$ and $\hat{\tau}_2$, which define the perpendicular plane, are a function of position $\underset{\sim}{x}$ and hence vary from point to point. This variation is partly due to the variation in the vector $\hat{b}$, to which $\hat{\tau}_1$ and $\hat{\tau}_2$ are orthogonal, and partly due to an arbitrariness in the definition of $\hat{\tau}_1$ and $\hat{\tau}_2$, which at this point in the work we are allowing for. Therefore if we examine the vectors $\hat{\tau}_1$ and $\hat{\tau}_2$ at some point P and at a neighboring point P', then these vectors and the perpendicular plane they define will be rotated at P' relative to their values at P. If the vectors $\hat{\tau}_1$ and $\hat{\tau}_2$ at P' are projected back onto the perpendicular plane at P, then they will be rotated by a certain angle $\Delta\psi$ relative to the vectors $\hat{\tau}_1$ and $\hat{\tau}_2$ at P, and the angle $\Delta\psi$ will, for small separations, be proportional to the distance between P and P'. Indeed, if we let $\Delta\underset{\sim}{x}$ be the displacement vector between P and P', then we have $\Delta\psi = \Delta\underset{\sim}{x} \cdot \underset{\sim}{R}$. In particular, the quantity $\hat{b} \cdot \underset{\sim}{R}$ represents the rate (in terms of radians per unit length) at which the vectors $\hat{\tau}_1$ and $\hat{\tau}_2$ "twist" as one moves along a magnetic field line.

These considerations are important when we consider the arbitrariness

in the definition of $\hat{\tau}_1$ and $\hat{\tau}_2$. Without as yet addressing the question of a possible judicious choice for $\hat{\tau}_1$ and $\hat{\tau}_2$, let us suppose that we have, in addition to $\hat{\tau}_1(\underset{\sim}{x})$ and $\hat{\tau}_2(\underset{\sim}{x})$, another pair of perpendicular unit vector fields $\hat{\tau}_1'(\underset{\sim}{x})$ and $\hat{\tau}_2'(\underset{\sim}{x})$. Both pairs are required to satisfy the relations in Eqs. (2.14)-(2.16), but beyond that their specification is arbitrary. Both pairs of unit vectors must lie in the perpendicular plane, so a relation of the following form must hold between them:

$$\begin{aligned} \hat{\tau}_1' &= \cos\phi\ \hat{\tau}_1 - \sin\phi\ \hat{\tau}_2 \\ \hat{\tau}_2' &= \sin\phi\ \hat{\tau}_1 + \cos\phi\ \hat{\tau}_2 \end{aligned} \tag{3.10}$$

where $\phi = \phi(\underset{\sim}{x})$ is in general dependent on position. We conclude that if $\hat{\tau}_1$ and $\hat{\tau}_2$ are given, then any other choice of perpendicular unit vectors is related to the given one by some rotation angle field $\phi(\underset{\sim}{x})$, and conversely.

Let us now consider how the various quantities defined in Sec. 2 change under the selection of a new set of perpendicular unit vectors, as shown by Eq. (3.10) and as specified by the field $\phi(\underset{\sim}{x})$. Following the notation above, we let primes represent the new quantities.

Clearly, the parallel and perpendicular velocities are invariant under such a change, i.e. $u' = u$, and $w' = w$. The gyrophase θ, on the other hand, changes by the amount ϕ, since θ' is the gyrophase relative to the $\hat{\tau}_1'$ direction:

$$\theta' = \theta - \phi(\underset{\sim}{x}) \tag{3.11}$$

Therefore of the coordinates $(\underset{\sim}{x},u,\theta,w)$, only θ depends on the choice of perpendicular unit vectors.

The unit vectors $\hat{a}$ and $\hat{c}$, which are defined in terms of $\hat{b}$ and the particle velocity $\underset{\sim}{v}$, are naturally invariant under the transformation indicated by Eq. (3.10). Nevertheless, the vector $\underset{\sim}{R}$, which can be expressed in terms of the gradients of $\hat{a}$ and $\hat{c}$ by $\underset{\sim}{R} = \nabla\hat{c}\cdot\hat{a}$, is not invariant:

$$\underset{\sim}{R}' = \nabla\hat{\tau}_1'\cdot\hat{\tau}_2' = \underset{\sim}{R} - \nabla\phi \tag{3.12}$$

In view of the geometrical interpretation of the vector $\underset{\sim}{R}$ which was given above, this result should not be surprising. Lest it seem paradoxical from a mathematical point of view, i.e. that $\hat{a}$ and $\hat{c}$ are invariant while $\underset{\sim}{R} = \nabla\hat{c}\cdot\hat{a}$ is not, we recall that the operator ∇ in the expression for $\underset{\sim}{R}$ is taken at fixed (u,θ,w), and that θ is not invariant. That is, the operator ∇, in this sense, is not invariant. It is interesting to observe that Eq. (3.12) is analogous to a gauge transformation for the magnetic vector potential $\underset{\sim}{A}$.

Let us now ask ourselves to what extent the vector $\underset{\sim}{R}$ can be brought into some simple form by an appropriate choice of perpendicular unit vectors. We might begin by asking if it is possible to choose $\hat{\tau}_1$ and $\hat{\tau}_2$ so that $\underset{\sim}{R} = 0$. The answer, as may be seen from Eq. (3.12), is no, because in general $\nabla\times\underset{\sim}{R} \neq 0$. Nevertheless, this line of reasoning raises an interesting point, namely that the curl of $\underset{\sim}{R}$ is invariant under a change of perpendicular unit vectors: $\nabla\times\underset{\sim}{R} = \nabla\times\underset{\sim}{R}'$. This in turn suggests that the vector $\nabla\times\underset{\sim}{R}$ can be expressed purely in terms of $\hat{b}$. Some algebra shows that this is indeed the case:

$$\nabla\times\underset{\sim}{R} = \frac{1}{2}\,\hat{b}\,[(b_{i,j}b_{j,i}) - (\nabla\cdot\hat{b})^2] + (\nabla\cdot\hat{b})(\hat{b}\cdot\nabla\hat{b}) - \hat{b}\cdot\nabla\hat{b}\cdot\nabla\hat{b} \tag{3.13}$$

Ordinary vector notation fails with the first term in $\hat{b}$, so index notation has been used, with commas representing differentiation. For example, $b_{i,j}$ means $\partial b_i/\partial x_j$. Eq. (3.13) is of use in computing the second order drifts.

Although we cannot make $\underset{\sim}{R} = 0$ by a choice of perpendicular unit vectors, it is possible to make one component of $\underset{\sim}{R}$ vanish by such a choice. Consider, for example, the component along $\hat{b}$. Suppose $\hat{b}\cdot\underset{\sim}{R} \neq 0$ with respect to some choice $\hat{\tau}_1$, $\hat{\tau}_2$ of perpendicular unit vectors. Then define $\phi(\underset{\sim}{x})$ by

$$\phi(\underset{\sim}{x}) = \int^{\underset{\sim}{x}} \underset{\sim}{R}(\underset{\sim}{x}')\cdot d\underset{\sim}{x}' \tag{3.14}$$

where the line integral is taken along a magnetic field line, and where the lower limit refers to some arbitrary initial value surface. Then $\hat{b}\cdot\nabla\phi = \hat{b}\cdot\underset{\sim}{R}$, and by Eq. (3.12) the change in unit vectors engendered by ϕ through Eq. (3.10) gives $\hat{b}\cdot\underset{\sim}{R}' = 0$.

This result can be strengthened. Let $\psi(\underset{\sim}{x})$ be any given scalar field. Then it is possible to choose a pair of perpendicular unit vectors such that $\hat{b}\cdot\underset{\sim}{R} = \psi$. To see this, let

$$\phi(\underset{\sim}{x}) = \int^{\underset{\sim}{x}} [\underset{\sim}{R}(\underset{\sim}{x}') - \psi(\underset{\sim}{x}')\hat{b}(\underset{\sim}{x}')]\cdot d\underset{\sim}{x}' \tag{3.15}$$

with the same integration conventions as in Eq. (3.14). Then $\hat{b}\cdot\nabla\phi = \hat{b}\cdot\underset{\sim}{R} - \psi$, and the conclusion follows. This result will be of use later.

The practical applications of guiding center theory fall into two broad classes, namely theoretical and computational. In computational

work it would not be desirable to choose perpendicular unit vectors according to the method of the last paragraph, because in order to determine $\hat{\tau}_1$ and $\hat{\tau}_2$ at a given point $\underset{\sim}{x}$ one would have to perform a numerical integration along field lines. For this kind of work it would be much better to have a local determination of perpendicular unit vectors. In theoretical studies, on the other hand, there is no harm in choosing perpendicular unit vectors in some nonlocal way, if it will simplify the resulting expressions. Later in this paper we will have opportunity to make some such choice.

It is possible to choose perpendicular unit vectors which depend only locally on the magnetic field direction $\hat{b}$. For example, one might let $\hat{\tau}_1$ and $\hat{\tau}_2$ be the principal normal and binormal unit vectors:

$$\hat{\tau}_1 = \frac{\hat{b}\cdot\nabla\hat{b}}{|\hat{b}\cdot\nabla\hat{b}|} \tag{3.16}$$

$$\hat{\tau}_2 = \hat{b}\times\hat{\tau}_1 \tag{3.17}$$

However, this choice has the disadvantage, from a theoretical standpoint, of producing discontinuities in $\hat{\tau}_1$ and $\hat{\tau}_2$ at an inflection point of a field line, and it is incapable of handling the case of straight field lines, which formed the subject of Ref. 1. In addition, it does not seem to cause any simplification in expressions which appear later in this work. Therefore we will make no further use of this possible choice for $\hat{\tau}_1$ and $\hat{\tau}_2$.

For most of the remainder of this paper, the gyrophase θ will appear only implicitly, through the unit vectors $\hat{a}$ and $\hat{c}$. As may be seen from

Eq. (2.20), these vectors are linear in $\sin\theta$ and $\cos\theta$, i.e. they are quantities purely of the first harmonic in θ. When these vectors are multiplied together, possibly in conjunction with contractions and spatial gradients, in general there will result terms of other multiples of the fundamental harmonic, i.e. a Fourier series in θ. The operation of projecting out the Fourier components of an expression is a familiar feature of perturbation theory for nearly periodic systems, and it is convenient at this point to elaborate upon the Fourier decomposition of various expressions which will be used later. The discussion will not be particularly deep or profound, since the highest harmonic we will encounter is the second, and relatively ad hoc techniques will suffice for our purposes. It is for the same reason that we do not introduce complex unit vectors.

Let us begin with quantities of the zeroth harmonic in θ. First we have the following two tensor operators, which are quadratic in $\hat{a}$ and $\hat{c}$, and which are of the zeroth harmonic:

$$\hat{a}\hat{a} + \hat{c}\hat{c} = I - \hat{b}\hat{b} \tag{3.18}$$

$$\hat{a}\hat{c} - \hat{c}\hat{a} = \hat{b}\times I \tag{3.19}$$

These were already mentioned in Eqs. (3.1) and (3.2). Next, the vectors $\hat{b}$, $\nabla\times\hat{b}$, and any other vector expressed purely in terms of $\hat{b}$, are, of course, of the zeroth harmonic. The vector $\underset{\sim}{R} = \nabla\hat{c}\cdot\hat{a}$ is also of the zeroth harmonic. Finally, we have the following scalars of the zeroth harmonic, which we abbreviate by giving them special symbols:

$$Z_0 = \hat{b}\cdot(\nabla\times\hat{b}) = \hat{c}\cdot\nabla\hat{b}\cdot\hat{a} - \hat{a}\cdot\nabla\hat{b}\cdot\hat{c} \tag{3.20}$$

$$Z_1 = \nabla\cdot\hat{b} \quad = \hat{a}\cdot\nabla\hat{b}\cdot\hat{a} + \hat{c}\cdot\nabla\hat{b}\cdot\hat{c} \tag{3.21}$$

$$Z_2 = \hat{b}\cdot\nabla\hat{c}\cdot\hat{a} \quad = \hat{b}\cdot\underset{\sim}{R} \tag{3.22}$$

The symbol Z is a mnemonic for "zeroth harmonic." Observe that Z_0 vanishes in a current free region of space, i.e. when $\nabla\times\underset{\sim}{B} = 0$, and that Z_2 can be made to take on any desired value by an appropriate choice of perpendicular unit vectors, as was noted above.

The principal vectors of the first harmonic are $\hat{a}$ and $\hat{c}$. In addition, we have the following scalars, in which the symbol F is a mnemonic for "first harmonic":

$$F_0 = \hat{b}\cdot\nabla\hat{b}\cdot\hat{a} \tag{3.23}$$

$$F_1 = \hat{b}\cdot\nabla\hat{b}\cdot\hat{c} \tag{3.24}$$

$$F_2 = \hat{a}\cdot\nabla\hat{c}\cdot\hat{a} = \hat{a}\cdot\underset{\sim}{R} \tag{3.25}$$

$$F_3 = \hat{c}\cdot\nabla\hat{c}\cdot\hat{a} = \hat{c}\cdot\underset{\sim}{R} \tag{3.26}$$

At the second harmonic, there are two tensor operators of importance, namely $\hat{a}\hat{a}$-$\hat{c}\hat{c}$ and $\hat{a}\hat{c}$+$\hat{c}\hat{a}$. From these we define the following scalars, in which the symbol S is a mnemonic for "second harmonic":

$$S_0 = \frac{1}{2}\,(\hat{a}\cdot\nabla\hat{b}\cdot\hat{c} + \hat{c}\cdot\nabla\hat{b}\cdot\hat{a}) \tag{3.27}$$

$$S_1 = \frac{1}{2}\,(\hat{a}\cdot\nabla\hat{b}\cdot\hat{a} - \hat{c}\cdot\nabla\hat{b}\cdot\hat{c}) \tag{3.28}$$

Using these definitions, let us rewrite the Poisson bracket relations in Eqs. (2.22e)-(2.22g) so as to show the Fourier decomposition of the terms:

$$\{u,\theta\} = \frac{1}{2} Z_0 - Z_2 - \frac{u}{w} F_0 - S_0 \tag{3.29a}$$

$$\{u,w\} = w(\frac{1}{2} Z_1 - S_1) + uF_1 \tag{3.29b}$$

$$\{\theta,w\} = \frac{\Omega}{\varepsilon w} + \frac{u}{w} Z_0 + F_3 \tag{3.29c}$$

We conclude this section by listing in Table I the derivatives and integrals with respect to θ of the various first and second harmonic quantities defined above. This table will be of use later.

4. THE DARBOUX TRANSFORMATION: FORMAL PROPERTIES

4.1. Preliminaries

In this section we will describe some of the formal properties of the Darboux transformation, which will take us from the $(\underset{\sim}{x},u,\theta,w)$ coordinate system in phase space to a new system, denoted by $(\underset{\sim}{X},U,\theta,J)$. This transformation is defined and justified on the basis of a straightforward and obvious extension of the methods of Ref. 1. We will postpone until Sec. 5 a derivation of explicit expressions for the Darboux transformation, and concentrate in this section on various closed-form results which can be obtained without those expressions. Most importantly, we will derive in this section the components of the Poisson tensor with respect to the $(\underset{\sim}{X},U,\theta,J)$ coordinate system.

4.2. Specification of the Darboux transformation

Following the pattern established in Ref. 1, we seek a set of five independent functions of $(\underset{\sim}{x},u,\theta,w)$, namely $\underset{\sim}{X}$, U, and J, which will satisfy the following Poisson bracket relations:

$$\{\theta,J\} = 1/\epsilon \tag{4.1}$$

$$\{\theta,\underset{\sim}{X}\} = 0 \tag{4.2}$$

$$\{\theta,U\} = 0 \tag{4.3}$$

$$\{J,\underset{\sim}{X}\} = 0 \tag{4.4}$$

$$\{J,U\} = 0 \tag{4.5}$$

The solution of these equations for the five unknown functions $(\underset{\sim}{X},U,J)$ will produce a "semicanonical" coordinate system in phase space,

namely $(\underset{\sim}{X},U,\theta,J)$, in which the variables θ and J are, one might say, "canonically decoupled" from the other four variables $(\underset{\sim}{X},U)$. The reason for choosing the symbols $\underset{\sim}{X}$ and U for four of the new coordinates will become apparent in a moment. As in Ref. 1, the Poisson bracket $\{\theta,J\}$ is required to take on the value $1/\varepsilon$ instead of 1 so that J will be $O(1)$ instead of $O(\varepsilon)$.

The transformation $(\underset{\sim}{x},u,\theta,w) \rightarrow (\underset{\sim}{X},U,\theta,J)$ will be called the Darboux transformation, because the solution to Eqs. (4.1)-(4.5) is obtained by applying the Darboux algorithm, as explained in detail in Ref. 1. In the new coordinates the unperturbed system, corresponding physically to rapid, circular gyrations, is separated from the perturbation, which corresponds to inhomogeneities in the magnetic field. The precise meaning of this statement will become clear in Sec. 5, when we obtain the Hamiltonian in the $(\underset{\sim}{X},U,\theta,J)$ coordinates. In addition, the semicanonical nature of the new coordinate system allows us to carry out an averaging transformation by means of Lie transforms, as will be shown in Sec. 6, the result of which is a Hamiltonian which is independent of θ.

To solve Eqs. (4.1)-(4.5) we introduce two differential operators, $d/d\lambda$ and $d/d\mu$, defined by

$$d/d\lambda = \{\ ,\theta\} \tag{4.6}$$

$$d/d\mu = \{\ ,J\} \tag{4.7}$$

The operator $d/d\mu$ is not determined until we have, at least in principle, a solution for J. Using these operators, Eqs. (4.1)-(4.5) can be written in the following form:

$$dJ/d\lambda = -1/\epsilon \tag{4.8}$$

$$d\underset{\sim}{x}/d\lambda = 0 \tag{4.9}$$

$$dU/d\lambda = 0 \tag{4.10}$$

$$d\underset{\sim}{x}/d\mu = 0 \tag{4.11}$$

$$dU/d\mu = 0 \tag{4.12}$$

Let us write $\underset{\sim}{z} = (\underset{\sim}{x},u,\theta,w)$, and consider the phase space curves $\underset{\sim}{z} = \underset{\sim}{z}(\lambda)$ which satisfy

$$d\underset{\sim}{z}/d\lambda = \{\underset{\sim}{z},\theta\} \tag{4.13}$$

These curves will be called the "θ-characteristics," because they are the characteristic curves of the partial differential operator contained in Eq. (4.6). Once the θ-characteristics have been determined, the solutions to Eqs. (4.8)-(4.10) follow immediately. Similarly, the curves $\underset{\sim}{z} = \underset{\sim}{z}(\mu)$ satisfying

$$d\underset{\sim}{z}/d\mu = \{\underset{\sim}{z},J\} \tag{4.14}$$

will be called the "J-characteristics," and they are used to solve Eqs. (4.11) and (4.12).

The defining equation for the θ-characteristics, Eq. (4.13), may be written out, using Eqs. (2.22) and (3.29). The result is

$$\frac{d\underset{\sim}{x}}{d\lambda} = -\frac{\hat{a}}{w} \tag{4.15}$$

$$\frac{du}{d\lambda} = \frac{1}{2} Z_0 - Z_2 - \frac{u}{w} F_0 - S_0 \tag{4.16}$$

$$\frac{dw}{d\lambda} = -\frac{\Omega}{\epsilon w} - \frac{u}{w} Z_0 - F_3 \tag{4.17}$$

For ϵ sufficiently small, the right hand side of Eq. (4.17) is dominated

by the term $-\Omega/\varepsilon w$, and w is seen to be a monotonic function of λ. Therefore Eq. (4.17) can be used to eliminate λ in favor of w, yielding the following set of differential equations, in which the equation for J, derived from Eq. (4.8), has been included:

$$\frac{d\underset{\sim}{X}}{dw} = \frac{\varepsilon \hat{\underset{\sim}{a}}}{D} \tag{4.18}$$

$$\frac{du}{dw} = \frac{\varepsilon}{D}\left[w\left(-\frac{1}{2}Z_0 + Z_2\right) + uF_0 + wS_0\right] \tag{4.19}$$

$$\frac{dJ}{dw} = \frac{w}{D} \tag{4.20}$$

Here the denominator D is given by

$$D = \Omega + \varepsilon(uZ_0 + wF_3) \tag{4.21}$$

Eqs. (4.18)-(4.20) are more useful than Eqs. (4.8) and (4.15)-(4.17) for a practical determination of the functions $(\underset{\sim}{X},U,J)$.

4.3. Geometrical interpretation of the coordinates $(\underset{\sim}{X},U,J)$

Let us give a geometrical interpretation to the θ-characteristics, and also to the functions $(\underset{\sim}{X},U,J)$ which are determined from them. We may assume for the sake of argument that we have a positive particle, so that $\Omega>0$. A similar sequence of deductions will go through for a negative particle. Let us also assume, as we did above, that ε is small enough that the term $-\Omega/\varepsilon w$ dominates the right hand side of Eq. (4.17). Then as λ increases, w decreases monotonically toward w=0. Therefore the θ-characteristics, which must lie on the surfaces θ = constant, converge inward toward the four-dimensional surface w=0. This surface is a singular surface, in the sense that it is a branch surface of the phase function θ, and a single point of this surface

is converged upon by a whole family of θ-characteristics. An entirely analogous behaviour for the θ-characteristics was observed in Ref. 1 and discussed there in greater detail.

Every point $\underset{\sim}{z} = (\underset{\sim}{x},u,\theta,w)$ of phase space (except those for which w=0) has a unique θ-characteristic passing through it, and that θ-characteristic, followed inward, reaches the surface w=0. Fig. 4 gives a schematic illustration of the θ-characteristics and certain quantities associated with them. When the surface w=0 has been reached, the $\underset{\sim}{x}$ and u coordinates take on certain values, which can be considered functions of the original point $\underset{\sim}{z}$. We will call these functions $\underset{\sim}{X}(\underset{\sim}{x},u,\theta,w)$ and $U(\underset{\sim}{x},u,\theta,w)$; they have the property that when w=0, $\underset{\sim}{X}=\underset{\sim}{x}$ and U=u. Effectively, the functions $\underset{\sim}{X}$ and U form a coordinate system on the surface w=0, which is being treated as an initial value surface for the θ-characteristics. The values of the functions $\underset{\sim}{X}$ and U elsewhere in phase space are found by propagating these functions along θ-characteristics, i.e. by assigning the same values of $\underset{\sim}{X}$ and U to any two points $\underset{\sim}{z}$ and $\underset{\sim}{z}'$ which lie on the same θ-characteristic. Clearly, the functions $\underset{\sim}{X}$ and U so constructed are constants of the θ-characteristics, and hence satisfy Eqs. (4.9) and (4.10).

As for the function $J(\underset{\sim}{x},u,\theta,w)$, we define it to be $-1/\varepsilon$ times the elapsed λ parameter between the point $\underset{\sim}{z} = (\underset{\sim}{x},u,\theta,w)$ and the w=0 point on the θ-characteristic passing through $\underset{\sim}{z}$. The resulting function satisfies Eq. (4.8), and it also satisfies the initial value condition J=0 when w=0.

4.4. Constants of the J-characteristics

According to the Darboux algorithm, the four functions $(\underset{\sim}{X},U)$ will

be constants of the J-characteristics everywhere in phase space, i.e. they will satisfy Eqs. (4.11) and (4.12), if they are constants of the J-characteristics on the initial value surface w=0. In order to analyze the J-characteristics on w=0, we need an expression for the function $J(\underset{\sim}{x},u,\theta,w)$ near w=0, so that Poisson brackets may be formed. That is, we need a solution to Eq. (4.20) as a power series in w. To lowest order in w, the result can be obtained by inspection; it is

$$J(\underset{\sim}{x},u,\theta,w) = \frac{w^2}{2D_0} + O(w^3) \tag{4.22}$$

where

$$D_0 = \Omega + \varepsilon u Z_0 \tag{4.23}$$

Now we may find the J-characteristics near w=0, using Eqs. (4.22) and (4.14). The resulting differential equations for the J-characteristics are

$$\frac{d\underset{\sim}{x}}{d\mu} = \frac{w\hat{c}}{D_0} + O(w^2) \tag{4.24}$$

$$\frac{du}{d\mu} = \frac{wuF_1}{D_0} + O(w^2) \tag{4.25}$$

$$\frac{dw}{d\mu} = O(w^2) \tag{4.26}$$

The right hand sides of all three of these equations go to zero as $w\to 0$, so that the J-characteristics on the surface w=0 consists of immobile points. Hence the functions $(\underset{\sim}{X},U)$, which take on the values $(\underset{\sim}{x},u)$ on w=0, are constants of the J-characteristics on w=0, and therefore also everywhere else in phase space. We conclude that the functions $(\underset{\sim}{X},U,J)$,

whose construction has been described but not yet explicitly demonstrated, satisfy Eqs. (4.8)-(4.12), and hence also Eqs. (4.1)-(4.5).

4.5. The Poisson tensor in the $(\underset{\sim}{X},U,\theta,J)$ coordinate system

Of the 15 independent components of the Poisson tensor in the $(\underset{\sim}{X},U,\theta,J)$ coordinate system, nine are given by Eqs. (4.1)-(4.5). The remaining six components, i.e. the Poisson brackets of the coordinates $(\underset{\sim}{X},U)$ among themselves, remain to be determined. The method we use for finding these Poisson brackets is exactly that used in Ref. 1; since the Poisson brackets of the variables $(\underset{\sim}{X},U)$ among themselves are constant along θ-characteristics, we can evaluate them on the initial value surface $w=0$. The results, expressed in terms of the variables $(\underset{\sim}{X},U)$, will then be valid everywhere in phase space. An interesting aspect of this procedure is that it gives results in closed form, i.e. not as a power series in ε.

In order to find the required Poisson brackets on the surface $w=0$, we need the functions $(\underset{\sim}{X},U)$ in a neighborhood of $w=0$, so that derivatives may be taken. Therefore, as we did above with the function J, we now solve Eqs. (4.18) and (4.19) as a power series in w. Again, to lowest order, the results can be written down practically by inspection:

$$\underset{\sim}{X}(\underset{\sim}{x},u,\theta,w) = \underset{\sim}{x} - \frac{\varepsilon w}{D_0}\,\hat{a} + O(w^2) \tag{4.27}$$

$$U(\underset{\sim}{x},u,\theta,w) = u - \frac{\varepsilon w u F_0}{D_0} + O(w^2) \tag{4.28}$$

Taking the Poisson brackets of these quantities among themselves and keeping track of the w-ordering gives, after some algebra,

$$\{X_i,X_j\} = \frac{\varepsilon}{D_0}(a_ic_j-a_jc_i) + O(w) \tag{4.29}$$

$$\{X_i,U\} = b_i + \frac{\varepsilon u}{D_0}(a_iF_1-c_iF_0) + O(w) \tag{4.30}$$

Then taking the limit $w\to 0$ and replacing $(\underset{\sim}{x},u)$ by $(\underset{\sim}{X},U)$ gives the following results, which are valid everywhere in phase space:

$$\{\underset{\sim}{X},\underset{\sim}{X}\} = \frac{\varepsilon}{\Omega+\varepsilon UZ_0}\hat{b}\times I \tag{4.31}$$

$$\{\underset{\sim}{X},U\} = \hat{b} + \frac{\varepsilon U}{\Omega+\varepsilon UZ_0}\hat{b}\times(\hat{b}\cdot\nabla\hat{b}) \tag{4.32}$$

In these equations all fields are evaluated at $\underset{\sim}{X}$, e.g. $\hat{b}$ means $\hat{b}(\underset{\sim}{X})$, and ∇ means $\partial/\partial\underset{\sim}{X}$. Eqs. (4.31) and (4.32), along with Eqs. (4.1)-(4.5), completely specify the Poisson tensor in the $(\underset{\sim}{X},U,\theta,J)$ coordinate system.

4.6. The Lagrange tensor

Because of the unfamiliarity of the manipulations used to derive Eqs. (4.31) and (4.32), it would be reassuring to check the self-consistency of the underlying theory. One way to do this is to compute the 4×4 component matrix of the Lagrange tensor (called the ω-tensor in Ref. 1) which corresponds to the 4×4 Poisson tensor given in Eqs. (4.31) and (4.32). According to the theory, the Lagrange tensor must be closed (see Eq. (2.12) of Ref. 1). Here we are dealing only with the reduced system of two degrees of freedom, described by the variables $(\underset{\sim}{X},U)$, because the overall Poisson tensor, including the variables (θ,J), has been brought into block diagonal form by Eqs. (4.1)-(4.5) (see Eq. (3.12) of Ref. 1).

Let us adopt the ordering $(\underset{\sim}{X},U) = (X,Y,Z,U)$ for the four phase space coordinates, and define, for the purposes of this demonstration, two vectors $\underset{\sim}{M}$ and $\underset{\sim}{N}$ by

$$\underset{\sim}{M} = \varepsilon\hat{b} \tag{4.33}$$

$$\underset{\sim}{N} = \underset{\sim}{\Omega} + \varepsilon U\nabla\times\hat{b} \tag{4.34}$$

The vector $\underset{\sim}{N}$ is closely related to the vector $\underset{\sim}{B}^*$ of Morozov and Solov'ev.[14] Using Eq. (3.5) it is then straightforward to show that

$$\{\underset{\sim}{X},\underset{\sim}{X}\} = \frac{\underset{\sim}{M}\times I}{\hat{b}\cdot\underset{\sim}{N}} \tag{4.35}$$

$$\{\underset{\sim}{X},U\} = \frac{\underset{\sim}{N}}{\hat{b}\cdot\underset{\sim}{N}} \tag{4.36}$$

and hence the Poisson tensor has the form

$$\sigma^{ij}_{(4)} = \frac{1}{\hat{b}\cdot\underset{\sim}{N}}\begin{pmatrix} 0 & -M_z & M_y & N_x \\ M_z & 0 & -M_x & N_y \\ -M_y & M_x & 0 & N_z \\ -N_x & -N_y & -N_z & 0 \end{pmatrix} \tag{4.37}$$

Here the subscript 4 has been appended to the symbol σ to indicate that we are dealing with the reduced 4×4 Poisson tensor in the variables $(\underset{\sim}{X},U)$.

On taking the negative of the inverse of $\sigma^{ij}_{(4)}$ we obtain the 4×4 Lagrange tensor $\omega_{(4)ij}$:

$$\omega_{(4)ij} = \frac{1}{\varepsilon} \begin{pmatrix} 0 & -N_z & N_y & M_x \\ N_z & 0 & -N_x & M_y \\ -N_y & N_x & 0 & M_z \\ -M_x & -M_y & -M_z & 0 \end{pmatrix} \tag{4.38}$$

Note that the expressions for the components of the Lagrange tensor are simpler than those of the Poisson tensor, in that they lack the denominator $\hat{b}\cdot\underset{\sim}{N}$. The tensor $\omega_{(4)}$ is closed, i.e. it satisfies

$$\frac{\partial \omega_{(4)ij}}{\partial z^k} + \frac{\partial \omega_{(4)jk}}{\partial z^i} + \frac{\partial \omega_{(4)ki}}{\partial z^j} = 0 \tag{4.39}$$

where $\underset{\sim}{z} = (\underset{\sim}{X},U)$, if the following relations hold:

$$\nabla\cdot\underset{\sim}{N} = 0 \tag{4.40}$$

$$\nabla\times\underset{\sim}{M} = \partial\underset{\sim}{N}/\partial U \tag{4.41}$$

It may be immediately verified that these two equations are valid, and hence that the Lagrange tensor $\omega_{(4)}$ is closed.

An important result may be obtained from the Lagrange tensor. Let us revert to the full six-dimensional coordinate set $\underset{\sim}{z} = (\underset{\sim}{X},U,\theta,J)$, and write ω_{ij} for the 6×6 Lagrange tensor. Then in accordance with Eqs. (4.1)-(4.5) we have

$$\omega_{ij} = \left(\begin{array}{cc:cc} \multicolumn{2}{c:}{\multirow{2}{*}{$\omega_{(4)}$}} & \multicolumn{2}{c}{\multirow{2}{*}{0}} \\ & & & \\ \hdashline \multicolumn{2}{c:}{\multirow{2}{*}{0}} & 0 & \varepsilon \\ \multicolumn{2}{c:}{} & -\varepsilon & 0 \end{array}\right) \tag{4.42}$$

Let us now put $\underset{\sim}{z}_c = (\underset{\sim}{q},\underset{\sim}{p})$ for the original canonical coordinates of Sec. 2. Since the quantities ω_{ij} are the Lagrange brackets of the coordinates $\underset{\sim}{z}$ among themselves, we have, using the notation of Ref. 1 for the matrix γ,

$$\omega_{ij} = \frac{\partial z_c^k}{\partial z^i} \gamma_{km} \frac{\partial z_c^m}{\partial z^j} \tag{4.43}$$

On taking the determinant of this relation we obtain

$$\det(\omega_{ij}) = \Delta^2 \tag{4.44}$$

where Δ is the Jacobian of the transformation $\underset{\sim}{z}_c = (\underset{\sim}{q},\underset{\sim}{p}) \rightarrow \underset{\sim}{z} = (\underset{\sim}{X},U,\theta,J)$:

$$\Delta = \det \left(\frac{\partial(\underset{\sim}{q},\underset{\sim}{p})}{\partial(\underset{\sim}{X},U,\theta,J)} \right) \tag{4.45}$$

From these relations and from Eq. (4.38) it is easy to find $|\Delta|$:

$$|\Delta| = |\Omega + \varepsilon U Z_0| \tag{4.46}$$

Therefore we have

$$d^3\underset{\sim}{q}\, d^3\underset{\sim}{p} = |\Omega + \varepsilon U Z_0|\, d^3\underset{\sim}{X}\, dU\, d\theta\, dJ \tag{4.47}$$

This relation is of obvious importance in any Vlasov kinetic treatment of a plasma which is expressed in the coordinates $(\underset{\sim}{X},U,\theta,J)$.

5. THE DARBOUX TRANSFORMATION: EXPLICIT EXPRESSIONS

In this section we will give explicit formulas for the Darboux transformation $(\underset{\sim}{x},u,\theta,w) \to (\underset{\sim}{X},U,\theta,J)$ and its inverse, expressed as power series in ε. To the order given the calculations are fairly simple and easily checked. In addition, we will give the Hamiltonian, also as a power series in ε, in the $(\underset{\sim}{X},U,\theta,J)$ coordinates.

5.1. Specification of the θ-characteristics

The Darboux transformation is found by solving Eqs. (4.18)-(4.20) for the θ-characteristics and for the evolution of the function J along them. To this end it is useful to imagine two points $\underset{\sim}{z}_i = (\underset{\sim}{x}_i,u_i,\theta_i,w_i)$ and $\underset{\sim}{z}_f = (\underset{\sim}{x}_f,u_f,\theta_f,w_f)$, the "initial" point and "final" point, which lie on the same θ-characteristic. In addition, we will call the values of the function J at the two points J_i and J_f. Since a θ-characteristic always lies on a contour surface of θ, we have $\theta_i = \theta_f$, and the subscripts on this variable can be dropped. As for the variables $\underset{\sim}{x}_f$, u_f, and J_f, we will find expressions, written as power series in ε, which give these quantities as functions of w_f, w_i, $\underset{\sim}{x}_i$, u_i, and J_i. Due to the form of the differential equations in Eqs. (4.18)-(4.20), w is regarded as the independent variable parametrizing the θ-characteristics, so both w_i and w_f appear in the expressions for $\underset{\sim}{x}_f$, u_f, and J_f. The quantities $\underset{\sim}{x}_i$, u_i and J_i are to be thought of as initial conditions for the functions $\underset{\sim}{x}_f$, u_f, and J_f; clearly, the determination of these functions completely specifies the θ-characteristics and the evolution of the quantity J along them.

The method we use for finding the functions $\underset{\sim}{x}_f$, u_f, and J_f has been called the method of parameter perturbations by Nayfeh.[15] The method is extremely simple; we put

$$\underset{\sim}{x}(w) = \underset{\sim}{x}_0 + \varepsilon\underset{\sim}{x}_1 + \varepsilon^2\underset{\sim}{x}_2 + O(\varepsilon^3) \tag{5.1}$$

$$u(w) = u_0 + \varepsilon u_1 + O(\varepsilon^2) \tag{5.2}$$

$$J(w) = J_0 + \varepsilon J_1 + O(\varepsilon^2) \tag{5.3}$$

in which the quantities $\underset{\sim}{x}_0$, $\underset{\sim}{x}_1$, etc., are to be regarded as functions of w. These expressions are substituted into Eqs. (4.18)-(4.20), all quantities are expanded out in powers of ε, and then collected order by order. For example, we have

$$\Omega(\underset{\sim}{x}) = \Omega(\underset{\sim}{x}_0) + \varepsilon\underset{\sim}{x}_1\cdot\nabla\Omega(\underset{\sim}{x}_0) + O(\varepsilon^2) \tag{5.4}$$

The result is a hierarchy of differential equations, which can be solved order by order.

The solution of the differential equations requires only trivial integrations. When the results are collected together, we obtain the following formulas, valid between any two points $\underset{\sim}{z}_i$ and $\underset{\sim}{z}_f$ on a θ-characteristic:

$$\underset{\sim}{x}_f = \underset{\sim}{x}_i + \frac{\varepsilon\hat{a}}{\Omega}(w_f-w_i) + \varepsilon^2\left\{\frac{1}{2\Omega}(w_f-w_i)^2\,\hat{a}\cdot\nabla\left(\frac{\hat{a}}{\Omega}\right)\right.$$
$$\left. - \frac{\hat{a}}{2\Omega^2}\left[(w_f^2-w_i^2)F_3 + 2u_i(w_f-w_i)Z_0\right]\right\} + O(\varepsilon^3) \tag{5.5}$$

$$u_f = u_i + \frac{\varepsilon}{2\Omega}\left[(w_f^2-w_i^2)(Z_2 - \tfrac{1}{2}Z_0 + S_0) + 2u_i(w_f-w_i)F_0\right] + O(\varepsilon^2) \tag{5.6}$$

$$J_f = J_i + \frac{1}{2\Omega}(w_f^2-w_i^2) + \varepsilon\left\{\frac{1}{6\Omega}(w_f-w_i)^2(2w_f+w_i)\hat{a}\cdot\nabla\left(\frac{1}{\Omega}\right)\right.$$
$$\left. - \frac{1}{6\Omega^2}\left[2(w_f^3-w_i^3)F_3 + 3u_i(w_f^2-w_i^2)Z_0\right]\right\} + O(\varepsilon^2) \tag{5.7}$$

In these formulas, all fields on the right hand side are evaluated at $\underset{\sim}{x}_i$.

5.2. The Darboux transformation and its inverse

Let us specialize the formulas above so as to obtain $\underset{\sim}{X}$, U, and J as functions of $(\underset{\sim}{x},u,\theta,w)$. To do this we identify $\underset{\sim}{z}_i$ with $\underset{\sim}{z}$ and $\underset{\sim}{z}_f$ with the w=0 point on the θ-characteristic passing through $\underset{\sim}{z}$. That is, we set $\underset{\sim}{x}_i=\underset{\sim}{x}$, $u_i=u$, $w_i=w$, and $J_i=J$, and also $\underset{\sim}{x}_f=\underset{\sim}{X}$, $u_f=U$, $w_f=0$, and $J_f=0$. These substitutions are in accordance with the definition and initial value properties of the functions $\underset{\sim}{X}$, U, and J, as described in Sec. 4, and they give the following:

$$\underset{\sim}{X}(\underset{\sim}{x},u,\theta,w) = \underset{\sim}{x} - \frac{\varepsilon w\hat{a}}{\Omega} + \varepsilon^2\left\{\frac{w^2}{2\Omega^2}\left[\Omega\hat{a}\cdot\nabla\left(\frac{\hat{a}}{\Omega}\right) + F_3\hat{a}\right] + \frac{uw}{\Omega^2}Z_0\hat{a}\right\} + O(\varepsilon^3) \tag{5.8}$$

$$U(\underset{\sim}{x},u,\theta,w) = u - \frac{\varepsilon}{2\Omega}\left[w^2\left(Z_2 - \frac{1}{2}Z_0 + S_0\right) + 2uwF_0\right] + O(\varepsilon^2) \tag{5.9}$$

$$J(\underset{\sim}{x},u,\theta,w) = \frac{w^2}{2\Omega} + \varepsilon\left[\frac{w^3}{6\Omega^3}(\hat{a}\cdot\nabla\Omega - 2\Omega F_3) - \frac{w^2u}{2\Omega^2}Z_0\right] + O(\varepsilon^2) \tag{5.10}$$

In these formulas the fields on the right hand side are evaluated at the particle positon $\underset{\sim}{x}$. Eqs. (5.8)-(5.10) form the Darboux transformation.

Note that through the $O(\varepsilon)$ term the quantity $\underset{\sim}{X}$ corresponds with the usual definition of the guiding center. Alternatively, we might say that $\underset{\sim}{X}$ coincides with the exact guiding center for a uniform magnetic field. It is on these grounds that we will call the variables $(\underset{\sim}{X},U,\theta,J)$ "guiding center variables," or, for reasons which will become apparent in the next section, "intermediate guiding center variables." The first

term of the expression for U needs no interpretation; it is the instantaneous parallel velocity. And the first term of the expression for J is, of course, proportional to the magnetic moment to lowest order. Note that J is negative for a negative particle. There is not much point in interpreting these formulas beyond these lowest order terms, because the higher order terms will change when we perform the averaging transformation, in Sec. 6.

Let us return to Eqs. (5.5)-(5.7) and swap the roles of $\underset{\sim}{z}_i$ and $\underset{\sim}{z}_f$. This will allow us to determine $\underset{\sim}{x}$, u, and J as functions of $\underset{\sim}{X}$, U, and w. That is, we set $\underset{\sim}{x}_i=\underset{\sim}{X}$, $u_i=U$, $w_i=0$, and $J_i=0$, and also $\underset{\sim}{x}_f=\underset{\sim}{x}$, $u_f=u$, $w_f=w$, and $J_f=J$. Doing so, we obtain

$$\underset{\sim}{x}(\underset{\sim}{X},U,\theta,w) = \underset{\sim}{X} + \frac{\varepsilon w\hat{a}}{\Omega} + \varepsilon^2\left\{\frac{w^2}{2\Omega^2}\left[\Omega\hat{a}\cdot\nabla\left(\frac{\hat{a}}{\Omega}\right) - F_3\hat{a}\right]\right.$$

$$\left. - \frac{Uw}{\Omega^2} Z_0\hat{a}\right\} + O(\varepsilon^3) \tag{5.11}$$

$$u(\underset{\sim}{X},U,\theta,w) = U + \frac{\varepsilon}{2\Omega}\left[w^2\left(Z_2 - \frac{1}{2} Z_0 + S_0\right) + 2UwF_0\right] + O(\varepsilon^2) \tag{5.12}$$

$$J(\underset{\sim}{X},U,\theta,w) = \frac{w^2}{2\Omega} + \varepsilon\left[-\frac{w^3}{3\Omega^3}(\hat{a}\cdot\nabla\Omega+\Omega F_3) - \frac{w^2U}{2\Omega^2} Z_0\right] + O(\varepsilon^2) \tag{5.13}$$

In these formulas the fields on the right hand side are evaluated at $\underset{\sim}{X}$, and ∇ means $\partial/\partial\underset{\sim}{X}$.

Eqs. (5.11)-(5.13) do not quite form the inverse of the Darboux transformation, because to have the inverse it is necessary to express the physical particle variables $(\underset{\sim}{x},u,\theta,w)$ in terms of the intermediate

guiding center variables $(\underset{\sim}{X},U,\theta,J)$. To do this, we first invert the series in Eq. (5.13) to find w as a function of $(\underset{\sim}{X},U,\theta,J)$. To the order given this series inversion is trivial, and it gives

$$w(\underset{\sim}{X},U,\theta,J) = (2\Omega J)^{1/2} + \varepsilon\left[\frac{(2\Omega J)}{3\Omega^2}(\hat{a}\cdot\nabla\Omega + \Omega F_3) + \frac{(2\Omega J)^{1/2}U}{2\Omega} Z_0\right] + O(\varepsilon^2) \tag{5.14}$$

This is then substituted into Eqs. (5.11) and (5.12), yielding

$$\underset{\sim}{x}(\underset{\sim}{X},U,\theta,J) = \underset{\sim}{X} + \frac{\varepsilon(2\Omega J)^{1/2}}{\Omega}\hat{a} + \varepsilon^2\left\{\frac{(2\Omega J)}{6\Omega^3}\left[3\Omega\hat{a}\cdot\nabla\hat{a} - \Omega F_3\hat{a} - (\hat{a}\cdot\nabla\Omega)\hat{a}\right] - \frac{(2\Omega J)^{1/2}U}{2\Omega^2} Z_0\hat{a}\right\} + O(\varepsilon^3) \tag{5.15}$$

$$u(\underset{\sim}{X},U,\theta,J) = U + \varepsilon\left[\frac{(2\Omega J)}{2\Omega}(Z_2 - \frac{1}{2} Z_0 + S_0) + \frac{(2\Omega J)^{1/2}U}{\Omega} F_0\right] + O(\varepsilon^2) \tag{5.16}$$

Again, all fields on the right hand side are evaluated at $\underset{\sim}{X}$. Eqs. (5.14)-(5.16) form the inverse of the Darboux transformation.

5.3. The Hamiltonian

It is now possible to find the Hamiltonian in the intermediate guiding center variables $(\underset{\sim}{X},U,\theta,J)$. It is obtained by simply substituting Eqs. (5.14) and (5.16) into (2.21), and this gives

$$H(\underset{\sim}{X},U,\theta,J) = \Omega J + \frac{1}{2}U^2 + \varepsilon\left[\frac{(2\Omega J)^{3/2}}{3\Omega^2}(\Omega F_3+\hat{a}\cdot\nabla\Omega)\right.$$

$$\left. + \frac{(2\Omega J)U}{2\Omega}(\frac{1}{2}Z_0+Z_2+S_0) + \frac{(2\Omega J)^{1/2}U^2}{\Omega}F_0\right] + O(\varepsilon^2) \qquad (5.17)$$

We should not expect this Hamiltonian to be independent of θ, and indeed, there are θ-dependent terms in the $O(\varepsilon)$ term of Eq. (5.17). The angle θ is a well defined function of the physical particle variables $(\underset{\sim}{x},\underset{\sim}{v})$, as indicated implicitly by Eqs. (2.18) and (2.20), and this particular functional form was chosen on the basis of two considerations. The first requirement was that θ should reduce to the exact gyrophase for a uniform magnetic field. (With sufficient care in the limiting process, this is equivalent to $\varepsilon\to 0$.) This requirement makes the unperturbed system "recurrent," in Kruskal's[16] terminology, and it causes the Hamiltonian to be independent of θ at lowest order. The second requirement was that θ should have a simple dependence on the physical particle variables. Neither of these requirements takes into consideration the higher order corrections in the guiding center expansion, and the result is a Hamiltonian which depends on θ beyond lowest order.

In spite of its θ-dependence, however, the Hamiltonian above may be used to obtain the well known, classic drifts, because the θ-dependence of H causes corrections only at an order in ε which is beyond these classic drifts. To see this, let us write H in the form

$$H(\underset{\sim}{X},U,\theta,J) = \Omega J + \frac{1}{2}U^2 + \varepsilon H_1(\underset{\sim}{X},U,\theta,J) + O(\varepsilon^2) \qquad (5.18)$$

and then use the Poisson bracket relations, given in Eqs. (4.1)-(4.5),

(4.31), and (4.32), to compute time derivatives. Let us carry the results to the highest order in ϵ which is compatible with an assumption of ignorance about the term ϵH_1. The Poisson bracket relations in Eqs. (4.31) and (4.32) are to be expanded in a power series in ϵ in this process.

The drifts themselves are found by computing $d\underset{\sim}{X}/dt$. Carried through $O(\epsilon)$, this is

$$\frac{d\underset{\sim}{X}}{dt} = \hat{b}U + \epsilon\left[\frac{1}{\Omega}\hat{b}\times(J\nabla\Omega) + \frac{U^2}{\Omega}\hat{b}\times(\hat{b}\cdot\nabla\hat{b}) + \hat{b}\frac{\partial H_1}{\partial U}\right] + O(\epsilon^2) \tag{5.19}$$

Evidently, the parallel motion of the guiding center can be found only through $O(1)$, because of the term in $\partial H_1/\partial U$. That is, we have

$$\left(\frac{d\underset{\sim}{X}}{dt}\right)_{\parallel} = \hat{b}U + O(\epsilon) \tag{5.20}$$

The perpendicular motion, on the other hand, can be found through $O(\epsilon)$:

$$\left(\frac{d\underset{\sim}{X}}{dt}\right)_{\perp} = \frac{\epsilon}{\Omega}\hat{b}\times(J\nabla\Omega + U^2\hat{b}\cdot\nabla\hat{b}) + O(\epsilon^2) \tag{5.21}$$

Mirroring effects are displayed by computing dU/dt:

$$\frac{dU}{dt} = -J\hat{b}\cdot\nabla\Omega + O(\epsilon) \tag{5.22}$$

Finally, we can compute the time derivatives of θ and J:

$$\frac{d\theta}{dt} = \frac{\Omega}{\epsilon} + O(1) \tag{5.23}$$

$$\frac{dJ}{dt} = -\frac{\partial H_1}{\partial\theta} + O(\epsilon) \tag{5.24}$$

In Eq. (5.24) we see that J has a time evolution at O(1). This evolution is, however, purely oscillatory at O(1), because the operator $\partial/\partial\theta$ projects out purely oscillatory terms in θ. Therefore J has a secular time evolution only at $O(\varepsilon)$. That J has a time evolution at all is, of course, a reflection of the fact that the Hamiltonian does depend on θ in terms beyond lowest order, and hence that J is a constant of the motion only to lowest order.

When the $O(\varepsilon)$ term in the Hamiltonian is made independent of θ by means of a near-identity coordinate transformation, all of the results expressed in Eqs. (5.20)-(5.24) become extended to one higher order. In particular, one obtains the second order perpendicular drifts. We now turn our attention to the averaging transformation, which will yield a Hamiltonian which is independent of θ.

6. THE GUIDING CENTER HAMILTONIAN

In this section we will develop a procedure for finding a near-identity transformation of the form $(\underset{\sim}{X},U,\theta,J) \to (\bar{\underset{\sim}{X}},\bar{U},\bar{\theta},\bar{J})$ such that the Hamiltonian in the new coordinates is independent of $\bar{\theta}$. The new variables will be called "averaged guiding center variables," and the new Hamiltonian K will be called the "guiding center Hamiltonian." The procedure involved is a variant of the Lie transform method, as detailed in Ref. 1. Using the guiding center Hamiltonian, we will be able to find, among other things, the second order perpendicular drifts.

6.1. The averaging transformation

According to the theory developed in Ref. 1, coordinate transformations associated with Hamiltonian flows preserve the functional form of the Poisson tensor, which in our case is given by Eqs. (4.1)-(4.5) and (4.31)-(4.32). These transformations were given the name "symplectic transformations" in Ref. 1, and they are, in a sense, canonical transformations expressed in noncanonical coordinates.

In order to develop an expression for a near-identity symplectic transformation, we consider a sequence of time-independent phase functions $g_1, g_2, \ldots$, which we will call the generators of the transformation. The generators are associated with a sequence $L_1, L_2, \ldots$ of "Lie operators," defined in terms of the Poisson bracket:

$$L_n = \varepsilon\{g_n, \ \} \tag{6.1}$$

The factor ε has been inserted into this definition in order to cancel the factor $1/\varepsilon$ in Eq. (4.1), so that the Lie operators L_n are O(1).

The Lie operators are in turn associated with a sequence T_1, T_2, ... of symplectic transformation operators, according to the rule

$$T_n = \exp(-\varepsilon^n L_n/n) \tag{6.2}$$

Finally, the T_n are multiplied together, giving an overall symplectic transformation T and its inverse T^{-1}:

$$T = \ldots T_3 T_2 T_1 \tag{6.3}$$

$$T^{-1} = T_1^{-1} T_2^{-1} T_3^{-1} \ldots \tag{6.4}$$

Under the action of the transformation T, the old variables $\underset{\sim}{z} = (\underset{\sim}{X}, U, \theta, J)$ go into new variables $\bar{\underset{\sim}{z}} = (\bar{\underset{\sim}{X}}, \bar{U}, \bar{\theta}, \bar{J})$ according to

$$\bar{\underset{\sim}{z}} = T\underset{\sim}{z} \tag{6.5}$$

$$\underset{\sim}{z} = T^{-1}\bar{\underset{\sim}{z}} \tag{6.6}$$

Likewise, the old Hamiltonian H is transformed into the new Hamiltonian K:

$$K = T^{-1}H \tag{6.7}$$

Our goal is to design the transformation T, i.e. to find the generators g_n, so that the new Hamiltonian K will be independent of $\bar{\theta}$. In addition, we demand that the transformation itself be free of secular terms.

To this end we expand the components of the Poisson tensor, which appear implicitly in Eq. (6.1), in power series in ε, and write

$$L_n = L_{n0} + \varepsilon L_{n1} + \varepsilon^2 L_{n2} + \ldots \tag{6.8}$$

where

$$L_{n0} = \frac{\partial g_n}{\partial \theta}\frac{\partial}{\partial J} - \frac{\partial g_n}{\partial J}\frac{\partial}{\partial \theta} \tag{6.9}$$

$$L_{n1} = \hat{b}\cdot\nabla g_n \frac{\partial}{\partial U} - \frac{\partial g_n}{\partial U}\hat{b}\cdot\nabla \tag{6.10}$$

and so forth. These are substituted into Eq. (6.2) and thence into Eqs. (6.3) and (6.4), giving

$$T = I - \varepsilon L_{10} + \frac{\varepsilon^2}{2}(-L_{20} + L_{10}^2 - 2L_{11}) + O(\varepsilon^3) \tag{6.11}$$

$$T^{-1} = I + \varepsilon L_{10} + \frac{\varepsilon^2}{2}(L_{20} + L_{10}^2 + 2L_{11}) + O(\varepsilon^3) \tag{6.12}$$

Finally, we write

$$H = \sum_{n=0}^{\infty} \varepsilon^n H_n \tag{6.13}$$

$$K = \sum_{n=0}^{\infty} \varepsilon^n K_n \tag{6.14}$$

and combine Eqs. (6.12)-(6.14) with (6.7) to get, to the lowest two orders,

$$K_0 = H_0 \tag{6.15}$$

$$L_{10}H_0 = K_1 - H_1 \tag{6.16}$$

For the purposes of this paper it will only be necessary to find the first generator, g_1, which is specified by Eq. (6.16).

6.2. The guiding center Hamiltonian

Let us apply the results above to the Hamiltonian in Eq. (5.17). First, from Eq. (6.15) we have

$$K_0(\bar{\underset{\sim}{X}},\bar{U},\bar{J}) = \Omega(\bar{X})\bar{J} + \frac{1}{2}\bar{U}^2 \tag{6.17}$$

Next, Eq. (6.16) is decomposed into its averaged and oscillatory parts in θ. The averaged part gives K_1:

$$K_1(\bar{\underset{\sim}{X}},\bar{U},\bar{J}) = \bar{J}\bar{U}(\frac{1}{2}Z_0 + Z_2) \tag{6.18}$$

The oscillatory part gives a differential equation for g_1:

$$\Omega\frac{\partial g_1}{\partial\theta} = -\frac{(2\Omega J)^{3/2}}{3\Omega^2}(\Omega F_3+\hat{a}\cdot\nabla\Omega) - \frac{(2\Omega J)U}{2\Omega}S_0$$

$$-\frac{(2\Omega J)^{1/2}U^2}{\Omega}F_0 \tag{6.19}$$

Using Table I, this is easily integrated, yielding

$$g_1 = \frac{(2\Omega J)^{3/2}}{3\Omega^3}(-\Omega F_2+\hat{c}\cdot\nabla\Omega) - \frac{(2\Omega J)U}{4\Omega^2}S_1 + \frac{(2\Omega J)^{1/2}U^2}{\Omega^2}F_1 \tag{6.20}$$

Here we may collect together the terms of K, writing out Z_0 and Z_2:

$$K(\bar{\underset{\sim}{X}},\bar{U},\bar{J}) = \Omega\bar{J} + \frac{1}{2}\bar{U}^2 + \varepsilon\bar{J}\bar{U}\left[\frac{1}{2}\hat{b}\cdot(\nabla\times\hat{b}) + \hat{b}\cdot\underset{\sim}{R}\right] + O(\varepsilon^2) \tag{6.21}$$

Of course, all fields on the right hand side are evaluated at the averaged guiding center position $\bar{\underset{\sim}{X}}$. K is the guiding center Hamiltonian.

6.3. The averaging transformation: explicit results

Using the result for g_1 and the transformation formulas in Eqs. (6.11) and (6.12), it is easy to write out explicit expressions for the averaging transformation $(\underset{\sim}{X},U,\theta,J) \to (\bar{\underset{\sim}{X}},\bar{U},\bar{\theta},\bar{J})$. We find the following:

$$\bar{\underset{\sim}{X}}(\underset{\sim}{X},U,\theta,J) = \underset{\sim}{X} + \frac{\varepsilon^2\hat{b}}{4\Omega^2}[-(2\Omega J)S_1 + 8(2\Omega J)^{1/2}UF_1] + O(\varepsilon^3) \tag{6.22}$$

$$\bar{U}(\underset{\sim}{X},U,\theta,J) = U + O(\varepsilon^2) \tag{6.23}$$

$$\bar{J}(\underset{\sim}{X},U,\theta,J) = J + \frac{\varepsilon}{6\Omega^3}[2(2\Omega J)^{3/2}(\Omega F_3+\hat{a}\cdot\nabla\Omega) + 3\Omega(2\Omega J)US_0 + 6\Omega(2\Omega J)^{1/2}U^2F_0] + O(\varepsilon^2) \tag{6.24}$$

$$\bar{\theta}(\underset{\sim}{X},U,\theta,J) = \theta + \frac{\varepsilon}{2\Omega^2}[2(2\Omega J)^{1/2}(-\Omega F_2+\hat{c}\cdot\nabla\Omega) - \Omega US_1 + 2\Omega(2\Omega J)^{-1/2}U^2F_1] + O(\varepsilon^2) \tag{6.25}$$

We need not write out the inverse of Eqs. (6.22)-(6.25), because to the order given it may be obtained simply by swapping $\underset{\sim}{z}$ and $\bar{\underset{\sim}{z}}$ and changing the sign of the correction terms.

Of perhaps greater importance than the above is the transformation connecting the averaged guiding center variables with the physical particle variables. This transformation is obtained by combining Eqs. (6.22)-(6.25) with Eqs. (5.8)-(5.10). The result is

$$\bar{\underset{\sim}{X}}(\underset{\sim}{x},u,\theta,w) = \underset{\sim}{x} - \frac{\varepsilon w\hat{a}}{\Omega} + \varepsilon^2\left\{\frac{w^2}{4\Omega^3}\left[-3\Omega S_1\hat{b} + 2(\hat{b}\cdot\nabla\Omega)\hat{b} + 2\Omega(\hat{b}\times\underset{\sim}{R})\right.\right.$$

$$\left.\left. - \nabla\Omega - (\hat{a}\hat{a}-\hat{c}\hat{c})\cdot\nabla\Omega\right] + \frac{wu}{\Omega^2}(Z_0\hat{a}+2F_1\hat{b})\right\} + O(\varepsilon^3) \qquad (6.26)$$

$$\bar{U}(\underset{\sim}{x},u,\theta,w) = u - \frac{\varepsilon}{2\Omega}[w^2(Z_2 - \tfrac{1}{2}Z_0+S_0) + 2uwF_0] + O(\varepsilon^2) \qquad (6.27)$$

$$\bar{\theta}(\underset{\sim}{x},u,\theta,w) = \theta + \frac{\varepsilon}{2\Omega^2}\left[2w(-\Omega F_2+\hat{c}\cdot\nabla\Omega) - \Omega uS_1\right.$$

$$\left. + 2\Omega\frac{u^2}{w}F_1\right] + O(\varepsilon^2) \qquad (6.28)$$

$$\bar{J}(\underset{\sim}{x},u,\theta,w) = \frac{w^2}{2\Omega} + \frac{\varepsilon}{2\Omega^3}\left[w^3\hat{a}\cdot\nabla\Omega + \Omega w^2u(S_0-Z_0)\right.$$

$$\left. +2\Omega wu^2F_0\right] + O(\varepsilon^2) \qquad (6.29)$$

In these expressions, all fields on the right hand side are evaluated at the physical particle position $\underset{\sim}{x}$.

For completeness, we give here the inverse of the transformation specified by Eqs. (6.26)-(6.29).

$$\underset{\sim}{x}(\bar{\underset{\sim}{X}},\bar{U},\bar{\theta},\bar{J}) = \bar{\underset{\sim}{X}} + \varepsilon\frac{(2\Omega\bar{J})^{1/2}}{\Omega}\hat{a} + \varepsilon^2\left\{\frac{(2\Omega\bar{J})}{4\Omega^3}\left[-\Omega S_1\hat{b} + 4(\hat{b}\cdot\nabla\Omega)\hat{b}\right.\right.$$

$$\left. - 2\Omega(\hat{b}\times\underset{\sim}{R}) - 3\nabla\Omega + (\hat{a}\hat{a}-\hat{c}\hat{c})\cdot\nabla\Omega\right] + \frac{(2\Omega\bar{J})^{1/2}\bar{U}}{4\Omega^2}\left[-8F_1\hat{b}\right.$$

$$\left.\left. - 2Z_0\hat{a} - \hat{b}\times(\hat{a}\cdot\nabla\hat{b}) - \hat{c}\cdot\nabla\hat{b}\right] - \frac{\bar{U}^2}{\Omega^2}\hat{b}\cdot\nabla\hat{b}\right\} + O(\varepsilon^3) \qquad (6.30)$$

$$u(\underset{\sim}{\bar{X}},\bar{U},\bar{\theta},\bar{J}) = \bar{U} + \frac{\epsilon}{2\Omega}\Big[(2\Omega\bar{J})(Z_2 - \frac{1}{2}Z_0 + S_0) + 2(2\Omega\bar{J})^{1/2}\bar{U}F_0\Big] + O(\epsilon^2) \tag{6.31}$$

$$\theta(\underset{\sim}{\bar{X}},\bar{U},\bar{\theta},\bar{J}) = \bar{\theta} + \frac{\epsilon}{2\Omega^2}\Big[2(2\Omega\bar{J})^{1/2}(\Omega F_2 - \hat{c}\cdot\nabla\Omega) + \Omega\bar{U}S_1 - 2\Omega(2\Omega\bar{J})^{-1/2}\bar{U}^2F_1\Big] + O(\epsilon^2) \tag{6.32}$$

$$w(\underset{\sim}{\bar{X}},\bar{U},\bar{\theta},\bar{J}) = (2\Omega\bar{J})^{1/2} + \frac{\epsilon}{2\Omega}\Big[(2\Omega\bar{J})^{1/2}\bar{U}(Z_0 - S_0) - 2\bar{U}^2F_0\Big] + O(\epsilon^2) \tag{6.33}$$

In these expressions, all fields on the right hand side are evaluated at the averaged guiding center position $\underset{\sim}{\bar{X}}$, and the vectors $\hat{a}$ and $\hat{c}$ are evaluated at $\bar{\theta}$.

Of all these relations, Eq. (6.29) is especially important. $\bar{J}$ is the adiabatic invariant associated with the gyration, and it is proportional to the magnetic moment, denoted here by μ:

$$\mu = \frac{e}{c}\bar{J} \tag{6.34}$$

According to this relation, μ is positive for particles of both signs of charge. The $O(\epsilon)$ term in Eq. (6.29) is in agreement with the old result derived originally by Kruskal.[4] The $O(\epsilon^2)$ term which would follow is to date unknown, except for magnetic fields of special symmetry.

6.4. Uniqueness of the averaged guiding center variables

It is important to ask to what extent the variables $(\bar{\underset{\sim}{X}},\bar{U},\bar{\theta},\bar{J})$, given by Eqs. (6.26)-(6.29), are unique, so that other guiding center variables of possible advantage may be selected in various applications. On the face of it, this is a formidable question, because the route from the physical particle variables in Sec. 2 to the averaged guiding center variables here is long, and it is punctuated with a number of reasonable but essentially arbitrary choices whose ultimate effect is not clear. To formulate an answer it is perhaps best to study the end product, especially in the light of Kruskal's[16] theory of "nice" variables, rather than to analyze in detail the method by which the end product was obtained.

Let us begin by listing, roughly in order of increasing specialization, some properties which the averaged guiding center variables satisfy. We may then examine the degrees of freedom which are introduced, step by step, as the listed properties, taken as restricting assumptions on the averaged guiding center variables, are relaxed.

First and foremost, the averaged guiding center variables are free of rapid oscillations to all orders, at least in the imagined and formal limit that the required power series are carried out to all orders. To state this property a little more precisely, we may say that the time derivative of the averaged guiding center variables is independent of the angle-like variable $\bar{\theta}$. This is the property of "niceness," and its exact definition involves the singling out of an angle-like variable whose time evolution, unlike that of the remaining variables, is non-zero at lowest order. (In the case at hand, the lowest order is $O(\epsilon^{-1})$.) In a noncanonical theory of guiding center motion, such as that developed by Northrop and Rome,[8] niceness is the only essential requirement. The overbar notation for our

variables here, as well as the word "averaged," are reminders that the variables are nice.

Second, the averaged guiding center variables form a semicanonical coordinate system in phase space. By this we mean that the set of six variables $(\underset{\sim}{\bar{X}},\bar{U},\bar{\theta},\bar{J})$ consists of two, namely $\bar{\theta}$ and $\bar{J}$, which are (apart from the factor $1/\varepsilon$) canonically conjugate, plus four more, namely $\underset{\sim}{\bar{X}}$ and $\bar{U}$, which have vanishing Poisson brackets with $\bar{\theta}$ and $\bar{J}$. Let us write $\underset{\sim}{\bar{\zeta}}$ for the four variables $\underset{\sim}{\bar{X}}$ and $\bar{U}$ collectively. Then the semicanonical requirement can be written as $\{\bar{\theta},\bar{J}\}=1/\varepsilon$ and $\{\underset{\sim}{\bar{\zeta}},\bar{\theta}\}=\{\underset{\sim}{\bar{\zeta}},\bar{J}\}=0$. With the given identification for $\underset{\sim}{\bar{\zeta}}$, this requirement is equivalent to the Poisson bracket relations in Eqs. (4.1)-(4.5).

Third and finally, the four variables $\underset{\sim}{\bar{\zeta}}$ satisfy Poisson bracket relations among themselves whose form is given by Eqs. (4.31)-(4.32). These relations, as well as those in Eqs. (4.1)-(4.5), were preserved under the symplectic averaging transformation.

Given all three of these requirements, the averaged guiding center variables are still not unique. Consider first the Poisson bracket relations. These relations are certainly preserved under any symplectic transformation. Conversely, if a transformation preserves the Poisson bracket relations, and if the transformation can be continuously connected with the identity transformation, then it is (questions of convergence aside) a symplectic transformation such as shown in Eq. (6.3).

Only a certain subclass of the symplectic transformations will preserve niceness, however. The members of this subclass are associated with generators g which are independent of θ. If we put $L = \varepsilon\{g, \}$ for such a generator and $T = \exp(-L)$, then it is easy to see that T takes any phase function which is independent of θ into another such function. (Here we are treating

factors of ϵ slightly differently than in Eqs. (6.1)-(6.2).) In particular, a $\bar{\theta}$-independent Hamiltonian goes into another such Hamiltonian, and hence niceness is preserved.

Such a transformation T takes the variables $\bar{z} = (\bar{\zeta},\bar{\theta},\bar{J})$ into a new set $T\bar{z} = \bar{z}' = (\bar{\zeta}',\bar{\theta}',\bar{J}')$ according to

$$\bar{\zeta}' = \exp(-L)\bar{\zeta} \tag{6.35}$$

$$\bar{\theta}' = \exp(-L)\bar{\theta} \tag{6.36}$$

$$\bar{J}' = \bar{J} \tag{6.37}$$

Since we are assuming that $\partial g/\partial\bar{\theta} = 0$, the action of T on the variable $\bar{J}$ can be written out explicitly. The action of T on the variables $\bar{\zeta}$ and $\bar{\theta}$, given by Eqs. (6.35)-(6.36), can be written as a power series in ϵ, assuming that g itself can be expanded in powers of ϵ. Explicitly, we have

$$\bar{X}' = \bar{X} + \epsilon\hat{b}\frac{\partial g}{\partial\bar{U}} + O(\epsilon^2 g) \tag{6.38}$$

$$\bar{U}' = \bar{U} - \epsilon\hat{b}\cdot\nabla g + O(\epsilon^2 g) \tag{6.39}$$

$$\bar{\theta}' = \bar{\theta} + \frac{\partial g}{\partial\bar{J}} + O(\epsilon g) \tag{6.40}$$

The transformation given by Eqs. (6.35)-(6.37) is the most general one which satisfies all three properties listed above, if we restrict consideration to transformations which can be continuously connected with the identity. It is interesting to observe that the degree of arbitrariness in the averaged guiding center variables, as indicated by this transformation, can also be achieved by modifying certain steps in the procedure used to derive the averaged guiding center variables $(\bar{X},\bar{U},\bar{\theta},\bar{J})$. For example, a suitable choice for g in Eqs. (6.38)-(6.40) will reproduce the effects of a redefinition of perpendicular unit vectors, as will be shown below.

In addition, if we had allowed for a constant of integration, depending on $(\underset{\sim}{X},U,J)$, on passing from Eq. (6.19) to Eq. (6.20), then the effect would be the same as the transformation above, with $g = \varepsilon c$ and c being the constant of integration.

Let us now suspend the third requirement, and ask for the general form of a transformation $(\underset{\sim}{\bar{\zeta}},\bar{\theta},\bar{J}) \rightarrow (\underset{\sim}{\bar{\zeta}}'',\bar{\theta}'',\bar{J}'')$, such that the double primed variables are nice, and such that the variables $\bar{\theta}''$ and $\bar{J}''$ are canonically decoupled from the variables $\underset{\sim}{\bar{\zeta}}''$ in the manner shown by Eqs. (4.1)-(4.5), but where the Poisson brackets of the variables $\underset{\sim}{\bar{\zeta}}''$ among themselves may take on whatever form they will. Certainly there is nothing sacred about the forms given in Eqs. (4.31)-(4.32). These forms came from our choice of coordinate system on the surface w=0, namely that which is naturally induced there by the $(\underset{\sim}{x},u)$ coordinate mesh. Although this choice was reasonable, it was not compelling.

We may answer this question first by noting that the symplectic transformation given by Eqs. (6.35)-(6.37), followed by a transformation of the form

$$\underset{\sim}{\bar{\zeta}}'' = \underset{\sim}{Z}(\underset{\sim}{\bar{\zeta}}') \tag{6.41}$$

$$\bar{\theta}'' = \bar{\theta}' \tag{6.42}$$

$$\bar{J}'' = \bar{J}' \tag{6.43}$$

where $\underset{\sim}{Z}$ is an arbitrary invertible transformation of four variables into four variables, will be a member of the class of transformations we seek. Because the second transformation mixes up the four variables $\underset{\sim}{\bar{\zeta}}'$ among themselves, but leaves $\bar{\theta}'$ and $\bar{J}'$ alone, Eqs. (4.31)-(4.32) will in general pass into a form with little resemblance to its antecedent, whereas the form of Eqs. (4.1)-(4.5) will remain invariant. An example of such a transformation $\underset{\sim}{Z}$ would be the transformation which leaves $\bar{U}$ unchanged but

which converts $\bar{\underset{\sim}{X}}$ into spherical (or toroidal) coordinates. Secondly and conversely, it is possible to argue that any transformation which preserves niceness as well as the form of Eqs. (4.1)-(4.5) is the composition of a transformation of the form of Eqs. (6.35)-(6.38) with one of the form of Eqs. (6.41)-(6.43).

When we abandon the second requirement, that the form of the Poisson brackets in Eqs. (4.1)-(4.5) hold, then we are left only with the requirement of niceness. It was argued in Ref. 1 that at least a semicanonical coordinate system is necessary in order to carry out Hamiltonian perturbation theory, although Hamiltonian mechanics itself can be made generally covariant. Therefore, for practical purposes, the relaxation of the second requirement amounts to an abandonment of Hamiltonian mechanics. Let us note, therefore, before taking leave of semicanonical coordinate systems, certain features which are common to all such systems.

Most outstandingly, the quantity $\bar{J}$ is common to all such systems, as shown by Eqs. (6.37) and (6.43). Kruskal has shown that $\bar{J}$ is an action integral associated with certain closed curves in phase space, called "rings." The detailed form of the action integral is equivalent to the Poisson bracket relation in Eq. (4.1). Rings are geometrical constructs which are based on the properties of nice variables and which are independent of coordinate system. It follows that $\bar{J}$ cannot change under a transformation of coordinates which preserves both niceness and the semicanonical Poisson bracket conditions. One might summarize this by saying that the adiabatic invariant associated with the gyration is unique.

Next, we note that the quantity $\bar{\theta}$ can change only by the addition of some function which is dependent on the other five nice variables,

as shown by Eq. (6.40). Geometrically, this amounts to a change in the origin of phase, different for each ring, which nevertheless leaves unchanged the relative reckoning of phase along any given ring.

Let us now relax all requirements except that of niceness. Let us write $\underset{\sim}{\bar{\xi}}$ for the five nice variables $(\underset{\sim}{\bar{X}},\bar{U},\bar{J})$ collectively, and ask for the most general variable transformation which preserves niceness. Kruskal has answered this question; it is

$$\bar{\theta}' = \bar{\theta} + f(\underset{\sim}{\bar{\xi}}) \tag{6.44}$$

$$\underset{\sim}{\bar{\xi}}' = \underset{\sim}{\Xi}(\underset{\sim}{\bar{\xi}}) \tag{6.45}$$

where f is an arbitrary function and where $\underset{\sim}{\Xi}$ is an arbitrary invertible transformation of five variables into five others. Note that Eq. (6.45) involves a much greater freedom of choice of variables than was allowed in the semicanonical coordinate systems. We will see later that this extra freedom makes non-Hamiltonian treatments of guiding center motion somewhat more convenient, for some purposes, than Hamiltonian treatments.

6.5. A judicious choice for perpendicular unit vectors

It may be seen in Eqs. (6.26)-(6.29) that all of the averaged guiding center variables except $\bar{J}$ depend on the choice of perpendicular unit vectors, as shown by their dependence on the vector $\underset{\sim}{R}$. In addition, the Hamiltonian K, shown in Eq. (6.21), depends on $\underset{\sim}{R}$ in the $O(\epsilon)$ term. That $\bar{J}$ does not must be a reflection of the fact that $\bar{J}$ can be defined in invariant terms, as was mentioned above. The deeper significance of this

observation is not clear, but it may be noted by way of providing a clue that $\bar{J}$ is the generator of displacements in $\bar{\theta}$.

In any case, those quantities which depend on $\underset{\sim}{R}$ through $Z_2 = \hat{b}\cdot\underset{\sim}{R}$ can be brought into a possibly simpler form by a judicious choice of perpendicular unit vectors, as shown in Sec. 3. As noted before, such a choice cannot be determined locally and hence is not useful for numerical work. But for theoretical or algebraic purposes, there is no harm in setting $\hat{b}\cdot\underset{\sim}{R}$ equal to any scalar field we like. In particular, if we take

$$\hat{b}\cdot\underset{\sim}{R} = -\frac{1}{2}\,\hat{b}\cdot(\nabla\times\hat{b}) \tag{6.46}$$

then the $O(\epsilon)$ term in the Hamiltonian K vanishes, and we have

$$K(\bar{\underset{\sim}{X}},\bar{U},\bar{J}) = \Omega\bar{J} + \frac{1}{2}\,\bar{U}^2 + O(\epsilon^2) \tag{6.47}$$

This choice of perpendicular unit vectors is equivalent to taking for the field $\phi(\underset{\sim}{x})$, appearing in Eqs. (3.11)-(3.12), the following:

$$\phi(\underset{\sim}{x}) = \int^{\underset{\sim}{x}} (\underset{\sim}{R} + \frac{1}{2}\,\nabla\times\hat{b})\cdot d\underset{\sim}{x}' \tag{6.48}$$

where the integrand is evaluated at $\underset{\sim}{x}'$ and the integral is taken along a field line. It is also equivalent to taking

$$g(\bar{\underset{\sim}{X}},\bar{U},\bar{J}) = -\bar{J}\,\phi(\bar{\underset{\sim}{X}}) \tag{6.49}$$

in Eqs. (6.35)-(6.40).

In addition to simplifying the Hamiltonian K, the assumed choice of

perpendicular unit vectors simplifies Eqs. (6.27) and (6.31), giving

$$\bar{U}(\underset{\sim}{x},u,\theta,w) = u - \frac{\varepsilon}{2\Omega}\left[w^2(S_0-Z_0) + 2uwF_0\right] + O(\varepsilon^2) \tag{6.50}$$

$$u(\underset{\sim}{\bar{X}},\bar{U},\bar{\theta},\bar{J}) = \bar{U} + \frac{\varepsilon}{2\Omega}\left[(2\Omega\bar{J})(S_0-Z_0) + 2(2\Omega\bar{J})^{1/2}\bar{U}F_0\right] + O(\varepsilon^2) \tag{6.51}$$

On taking the phase average, which agrees with the time average to lowest order, Eq. (6.51) gives an equation which provides an interpretation of the variable $\bar{U}$:

$$\bar{U} = \text{Avg.}(u) + \varepsilon\bar{J}\hat{b}\cdot(\nabla\times\hat{b}) + O(\varepsilon^2) \tag{6.52}$$

The variable $\bar{U}$ agrees with the variable $v_{\|}$ used by Northrop and Rome[8] through the order given. A different choice of perpendicular unit vectors could have been made which would cause $\bar{U}$ to be identical with Avg.(u), although it would also cause the Hamiltonian K to be more complicated. The effect of the $O(\varepsilon)$ term in Eq. (6.52) has been carefully discussed by Northrop and Rome.

No matter what choice is made for perpendicular unit vectors, however, it is impossible to rid the expression for $\underset{\sim}{\bar{X}}$, given in Eq. (6.26), of its dependence on $\underset{\sim}{R}$, which is through the term $\hat{b}\times\underset{\sim}{R}$. (The only exception is the case that $\nabla\times\underset{\sim}{R}$, given by Eq. (3.13), should vanish.) We shall return to this point later.

6.6. The equations of motion

Let us make the choice of perpendicular unit vectors implied by Eq. (6.46) and derive the equations of motion, which will give us,

among other things, the second order perpendicular drifts. The general case of an arbitrary choice of perpendicular unit vectors need not be given; the more complicated formulas which result in this case are easily worked out.

First let us compute $d\bar{\underset{\sim}{X}}/dt$. We have

$$\frac{d\bar{\underset{\sim}{X}}}{dt} = \{\bar{\underset{\sim}{X}},\bar{\underset{\sim}{X}}\}\cdot[\bar{J}\nabla\Omega + O(\varepsilon^2)] + \{\bar{\underset{\sim}{X}},\bar{U}\}[\bar{U} + O(\varepsilon^2)] \tag{6.53}$$

Taking the parallel and perpendicular components of this relative to $\hat{b}(\bar{\underset{\sim}{X}})$, we obtain

$$\left(\frac{d\bar{\underset{\sim}{X}}}{dt}\right)_{\parallel} = \hat{b}\bar{U} + O(\varepsilon^2) \tag{6.54}$$

$$\left(\frac{d\bar{\underset{\sim}{X}}}{dt}\right)_{\perp} = \frac{\varepsilon}{\Omega+\varepsilon\bar{U}\hat{b}\cdot(\nabla\times\hat{b})}\ \hat{b}\times(\bar{J}\nabla\Omega + \bar{U}^2\hat{b}\cdot\nabla\hat{b}) + O(\varepsilon^3) \tag{6.55}$$

Eq. (6.54) shows that $\bar{U}$ is actually the parallel velocity of the guiding center. Eq. (6.55) shows that the $O(\varepsilon^2)$ correction to the perpendicular drifts is proportional to the $O(\varepsilon)$ term, although this simplicity has been achieved at the price of making a special and not necessarily convenient choice for the definition of the guiding center position $\bar{\underset{\sim}{X}}$. Of course, Eq. (6.55) is easily expanded properly into a power series in ε.

Next, we may obtain the $O(\varepsilon)$ correction to the mirroring expression.

$$\frac{d\bar{U}}{dt} = -\bar{J}\ \frac{(\hat{b}\Omega + \varepsilon\bar{U}\nabla\times\hat{b})\cdot\nabla\Omega}{\Omega+\varepsilon\bar{U}\hat{b}\cdot(\nabla\times\hat{b})} + O(\varepsilon^2) \tag{6.56}$$

Eq. (3.5) has been used in writing this result in the form given.

Finally, we have the evolution of the gyrophase:

$$\frac{d\bar{\theta}}{dt} = \frac{\Omega(\bar{\underset{\sim}{X}})}{\varepsilon} + O(\varepsilon) \tag{6.57}$$

Of course we have $d\bar{J}/dt = 0$ to all orders.

6.7. Eliminating the dependence of $\bar{\underset{\sim}{X}}$ on $\underset{\sim}{R}$

One's intuition says that the guiding center position $\bar{\underset{\sim}{X}}$ should not depend upon the choice of perpendicular unit vectors, and hence that the appearance of the term $\hat{b}\times\underset{\sim}{R}$ in Eq. (6.26) represents a flaw or a shortcoming in the theory. Therefore we may ask if it is possible to choose a new set of averaged guiding center variables which are free of this term in the new quantities which correspond to $\bar{\underset{\sim}{X}}$. It is here that we call upon the discussion of subsection 6.4.

Any alternate definition for $\bar{\underset{\sim}{X}}$ must be nice, since niceness is the one inviolate requirement which averaged guiding center variables must satisfy. Northrop and Rome have used the expression "guiding point" for some arbitrary, nice definition of $\bar{\underset{\sim}{X}}$. There are many ways to define a variable $\bar{\underset{\sim}{X}}'$ which agrees with our $\bar{\underset{\sim}{X}}$ in any given number of leading terms of Eq. (6.26) and which is also nice, and at the $O(\varepsilon^2)$ term there is little physical reason for choosing one form over another. This may be seen from Eqs. (6.44)-(6.45), showing how a new set of nice variables can be created from an old set.

In particular, we may set

$$\bar{\underset{\sim}{X}}' = \bar{\underset{\sim}{X}} - \varepsilon^2 \frac{\bar{J}}{\Omega} \hat{b}\times\underset{\sim}{R} \tag{6.58}$$

and we have a variable $\bar{\underset{\sim}{X}}'$ which is both nice and independent of perpendicular unit vectors. Likewise, we can kill the term $Z_2 = \hat{b}\cdot\underset{\sim}{R}$ in Eq. (6.27) by putting

$$\bar{U}' = \bar{U} + \varepsilon\bar{J}[\hat{b}\cdot\underset{\sim}{R} + \frac{1}{2}\hat{b}\cdot(\nabla\times\hat{b})] \tag{6.59}$$

and we obtain a parallel velocity $\bar{U}'$ which is also independent of perpendicular unit vectors. This $\bar{U}'$ is identical to the $\bar{U}$ of Eq. (6.50) but obtained in a very different way.

Unfortunately, the variable $\bar{\underset{\sim}{X}}'$ shown in Eq. (6.58) cannot be used in a Hamiltonian theory, nor can any other nice alternatives which eliminate the dependence on the choice of perpendicular unit vectors. This can be seen by examining Eqs. (6.35)-(6.43) which give the most general coordinate transformation allowed in a semicanonical theory. Therefore it appears that the intuition referred to above is wrong, at least for Hamiltonian mechanics.

For certain applications, especially numerical ones, it is desirable to employ guiding center variables which are independent of the choice of perpendicular unit vectors. Furthermore, the Hamiltonian structure of the underlying theory may not be important in such work. Therefore we give, in Appendix A, a set of noncanonical variables and their equations of motion which would be useful for such purposes.

7. CONCLUSIONS

One shortcoming of this work is that it does not allow for time-dependent fields. Nevertheless, the mathematical techniques which were developed in Ref. 1 and applied to static magnetic fields in this paper can be extended in a straightforward manner to time-dependent electromagnetic fields. The results of this extension will be reported upon in future publications.

The Hamiltonian methods developed here seem to yield results with less labor than older methods, especially when carried beyond lowest order. Of course, there is a compensation in that there is more theory to be mastered, but this represents a kind of fixed overhead which does not increase as one proceeds to higher orders. For example, it seems feasible for one person working alone to extend the results of this paper to one higher order, although the amount of algebra is significant. I myself have carried out approximately half of this calculation, but I have not recorded it here because of its incompleteness and because it does not have much practical value. On the other hand, to the order given the equivalents of Eqs. (6.54)-(6.56) for general electromagnetic fields are unknown and may perhaps best be derived by these methods.

Even when carried to lowest order, however, the Hamiltonian methods presented here promise to be useful for the analysis of additional perturbations. For example, the effects of a small amplitude electromagnetic wave on single particle motion in a nonuniform background magnetic field have been studied by Grebogi, Kaufman, and Littlejohn.[17] In this analysis, the guiding center Hamiltonian in Eq. (6.47) is taken to

be the unperturbed system, to which perturbing terms representing the wave are added. The resulting Hamiltonian can then be treated by standard perturbation techniques.

Additional results in the realm of single particle motion can be obtained by iterating the Darboux algorithm. This will allow one to study the nearly periodic motion of the guiding center corresponding to the longitudinal bouncing and motion on the flux surfaces. The results obtained to date in this area[3,18] involve an averaging over the phase of the longitudinal bouncing motion, which is introduced after an averaging over the phase of gyration has been performed. If both phases are introduced before averaging, then the door is open to an analysis of resonances between gyration and bouncing. These resonances have an important effect on particle confinement in fusion devices of the mirror type, and a perturbation treatment should be especially useful in the so-called superadiabatic regime.

Self-consistent treatments of ensembles of particles are especially important in plasma physics. In the Vlasov approximation, Hamiltonian methods are well adapted to such treatments, and they have been applied in recent years to a number of different problems.[19-25] Nevertheless, for the case of nonuniform magnetic fields one has had to make do with non-Hamiltonian methods, such as are used with drift kinetic equations. Possibly the area of application of greatest value for the Hamiltonian methods of this paper will be in kinetic theory.

ACKNOWLEDGEMENTS

I would especially like to thank A. N. Kaufman for having created a research environment in which basic questions can be pursued for their own sake. I would also like to thank C. Grebogi for his careful reading of the manuscript.

APPENDIX A.

This appendix gives the guiding center equations of motion in a form which would be suitable for numerical integration or other purposes with a minimum of overhead of notational conventions. The formulas of this appendix are similar to those given by Northrop and Rome,[8] but they are somewhat simpler. For the numerical integration of systems of ordinary differential equations it is important for efficiency reasons that the "driving terms," represented below by the right hand sides of Eqs. (A.5)-(A.8), be as simple as possible. Therefore the definitions of the guiding center variables given below have been juggled so as to simplify the corresponding equations of evolution.

For those readers continuing from the main text, we note that the guiding center variables $(\underset{\sim}{X},U,J)$ used in this appendix are nice but noncanonical variables. The overbar notation has been dropped, and these variables are not to be confused with the intermediate guiding center variables of Sec. 4. The variables U and J are identical to $\bar{U}$ and $\bar{J}$ of Sec. 6, while $\underset{\sim}{X}$ is identical to $\bar{\underset{\sim}{X}}'$ of Eq. (6.58).

Let $\underset{\sim}{x}$ and $\underset{\sim}{v}$ be the particle's instantaneous position and velocity, let $\hat{b}$ be the unit vector in the direction of the magnetic field $\underset{\sim}{B}$, let $\Omega=eB/mc$ be the signed gyrofrequency, and let $u = \hat{b}\cdot\underset{\sim}{v}$ be the particle's instantaneous parallel velocity. Then the guiding center position $\underset{\sim}{X}$ may be defined as follows:

$$\underset{\sim}{X} = \underset{\sim}{x} - \frac{\varepsilon}{\Omega}\hat{b}\times\underset{\sim}{v} + \varepsilon^2\left\{\frac{1}{8\Omega^3}\left[4\underset{\sim}{v}_\perp(\underset{\sim}{v}_\perp\cdot\nabla\Omega) + 9v_\perp^2\hat{b}(\hat{b}\cdot\nabla\Omega) - 4v_\perp^2\nabla\Omega\right.\right.$$

$$\left.\left. + 6\Omega\hat{b}(\underset{\sim}{v}_\perp\cdot\nabla\hat{b}\cdot\underset{\sim}{v}_\perp)\right] + \frac{u}{\Omega^2}\left[\hat{b}\cdot(\nabla\times\hat{b})\hat{b}\times\underset{\sim}{v} + 2\hat{b}(\hat{b}\cdot\nabla\hat{b}\cdot\underset{\sim}{v}_\perp)\right]\right\} + O(\varepsilon^3) \qquad (A.1)$$

In this formula and those that follow, ε represents a mnemonic device for keeping track of the order of the terms. It should be set to unity in applications. Furthermore, the operator ∇ in expressions involving $\nabla\hat{b}$ is taken to operate only on the vector $\hat{b}$. One may think of $\nabla\hat{b}$ as a matrix M with components $M_{ij} = \partial b_j/\partial x_i$.

The parallel velocity of the guiding center U is defined as follows:

$$U = u - \frac{\varepsilon}{4\Omega}\left[2(\hat{b}\times\underset{\sim}{v})\cdot\nabla\hat{b}\cdot\underset{\sim}{v}_\perp - v_\perp^2\hat{b}\cdot(\nabla\times\hat{b}) + 4u\hat{b}\cdot\nabla\hat{b}\cdot(\hat{b}\times\underset{\sim}{v})\right] + O(\varepsilon^2) \qquad (A.2)$$

The quantity U is identical with the quantity $v_\parallel$ used by Northrop and Rome.

The adiabatic invariant of gyration J is related to the magnetic moment μ by $\mu = eJ/c$. It is given by

$$J = \frac{v_\perp^2}{2\Omega} + \frac{\varepsilon}{4\Omega^3}\left\{2v_\perp^2(\hat{b}\times\underset{\sim}{v})\cdot\nabla\Omega + \Omega u\left[2(\hat{b}\times\underset{\sim}{v})\cdot\nabla\hat{b}\cdot\underset{\sim}{v}_\perp - v_\perp^2\hat{b}\cdot(\nabla\times\hat{b})\right]\right.$$
$$\left. + 4\Omega u^2\hat{b}\cdot\nabla\hat{b}\cdot(\hat{b}\times\underset{\sim}{v})\right\} + O(\varepsilon^2) \qquad (A.3)$$

Note that J is negative for a negative particle.

Eqs. (A.1)-(A.3) are to be regarded as definitions of the guiding center variables in terms of the instantaneous particle variables $\underset{\sim}{x}$ and $\underset{\sim}{v}$. Therefore all fileds on the right hand sides, such as Ω, and $\hat{b}$, are evaluated at the instantaneous particle position $\underset{\sim}{x}$. In addition, note that these equations can be written in a number of different forms. The forms chosen are more or less arbitrary.

The kinetic energy $K = \frac{1}{2}mv^2$ of the particle can be expressed in terms of the guiding center variables. The relation is

$$K = m[\Omega(\underset{\sim}{X})J + \frac{1}{2}U^2] + O(\varepsilon^2) \tag{A.4}$$

The equations of motion of the guiding center variables are as follows. First, the parallel velocity:

$$\frac{dU}{dt} = -J\hat{b}\cdot\left[\nabla\Omega + \frac{\varepsilon U}{\Omega}(\hat{b}\cdot\nabla\hat{b})\times\nabla\Omega\right] + O(\varepsilon^2) \tag{A.5}$$

This is completely equivalent to Eq. (6.56). Next we have the parallel motion of the guiding center, by which we mean the component of $d\underset{\sim}{X}/dt$ which is in the direction $\hat{b}(\underset{\sim}{X})$. This is

$$\left(\frac{d\underset{\sim}{X}}{dt}\right)_{\parallel} = \hat{b}U + O(\varepsilon^2) \tag{A.6}$$

Finally, we have the perpendicular motion of the guiding center:

$$\begin{aligned}\left(\frac{d\underset{\sim}{X}}{dt}\right)_{\perp} = \frac{\varepsilon}{\Omega}\hat{b}\times(J\nabla\Omega + U^2\hat{b}\cdot\nabla\hat{b}) + \varepsilon^2\hat{b}\times\Bigg\{&-\frac{U^3}{\Omega^2}(\hat{b}\cdot\nabla\times\hat{b})\hat{b}\cdot\nabla\hat{b} \\ &+ \frac{JU}{\Omega}\Bigg[-\frac{\nabla\Omega}{\Omega}(\hat{b}\cdot\nabla\times\hat{b}) - \frac{1}{2}(\hat{b}\cdot\nabla\times\hat{b})\hat{b}\cdot\nabla\hat{b} + \frac{1}{2}\nabla(\hat{b}\cdot\nabla\times\hat{b}) \\ &- \hat{b}\times(\hat{b}\cdot\nabla\hat{b}\cdot\nabla\hat{b}) + (\nabla\cdot\hat{b})\hat{b}\times(\hat{b}\cdot\nabla\hat{b})\Bigg]\Bigg\} + O(\varepsilon^3)\end{aligned} \tag{A.7}$$

In this expression, the term in $\nabla(\hat{b}\cdot\nabla\times\hat{b})$ is not in a form which would be most convenient for numerical integration. When this term is expanded out, along with all the other terms multiplying JU above, there results

$$\frac{JU}{\Omega}\,\hat{b}\times\Big[\ \ldots\ \Big] = \frac{JU}{\Omega}\Big[\frac{\nabla\Omega\times(\nabla\times\hat{b})}{\Omega} - \frac{5}{2}\,(\nabla\cdot\hat{b})\hat{b}\cdot\nabla\hat{b} - \frac{1}{2}\,\hat{b}\hat{b}:\nabla\nabla\hat{b}$$

$$+\ \hat{b}\cdot\nabla\hat{b}\cdot\nabla\hat{b} - \frac{1}{2}\,\nabla(\nabla\cdot\hat{b}) + \frac{1}{2}\,\nabla^2\hat{b}\Big]_{\perp} \qquad \text{(A.8)}$$

where the symbol $\perp$ means to take the perpendicular projection relative to $\hat{b}(\underset{\sim}{X})$. Finally, we note that in Eqs. (A.5)-(A.8) all fields on the right hand sides are evaluated at the guiding center position $\underset{\sim}{X}$.

REFERENCES

1. R. G. Littlejohn, Lawrence Berkeley Laboratory Report No. LBL-8917, 1979. To be published in J. Math. Phys.
2. H. Alfvén, Cosmical Electrodynamics (Clarendon Press, Oxford, 1950).
3. T. G. Northrop, The Adiabatic Motion of Charged Particles (Interscience, New York, 1963).
4. M. D. Kruskal, "Elementary Orbit and Drift Theory," in Plasma Physics (International Atomic Energy Agency, Vienna, 1965).
5. A. Baños, J. Plasma Phys. 1, 305(1967).
6. G. E. Wilson, Phys. Fluids 12, 1673(1969).
7. P. Martín, "Second Order Drift Velocities," preprint, Dpto. Física, Universidad Simón Bolívar, Caracas, Venezuela, 1977.
8. T. G. Northrop and J. A. Rome, Phys. Fluids 21, 384(1978).
9. C. S. Gardner, Phys. Rev. 115, 791(1959).
10. T. Taniuti, New York University Report No. NYO-9755, 1961.
11. D. P. Stern, Goddard Space Flight Center Report No. X-641-71-56, 1977.
12. H. E. Mynick, Lawrence Berkeley Laboratory Report No. LBL-8366, 1979, and Ph.D. thesis, University of California, Berkeley, 1979.
13. H. Meyer (ORNL), private communication.
14. A. I. Morozov and L. S. Solov'ev, in Reviews of Plasma Physics Vol. 2 (M. A. Leontovich editor, Consultants Bureau, New York, 1966).
15. A. H. Nayfeh, Perturbation Methods (Wiley, New York, 1973).
16. M. D. Kruskal, J. Math. Phys. 3, 806(1962).
17. C. Grebogi, A. N. Kaufman, and R. G. Littlejohn, Lawrence Berkeley Laboratory Report No. LBL-8916, 1979.
18. T. G. Northrop, M. D. Kruskal, and C. S. Liu, Phys. Fluids 9, 1503(1966).

19. R. Dewar, Phys. Fluids 16, 1102(1973).

20. R. Dewar, J. Phys. A9, 2043(1976)

21. S. Johnston, Phys. Fluids 19, 93(1976).

22. J. R. Cary and A. N. Kaufman, Phys. Rev. Lett. 39, 402(1977).

23. S. Johnston and A. N. Kaufman, Phys. Rev. Lett. 40, 1266(1978).

24. S. Johnston, A. N. Kaufman, and G. L. Johnston, J. Plasma Phys. 20, 365(1978).

25. J. R. Cary, Ph.D. thesis, University of California, Berkeley, 1979. Available as Lawrence Berkeley Laboratory Report No. LBL-8185.

TABLE I. Derivatives and integrals of various quantities with respect to θ. The symbol X refers to any of the quantities in the first column.

X	$dX/d\theta$	$\int X d\theta$
$\hat{a}$	$\hat{c}$	$-\hat{c}$
$\hat{c}$	$-\hat{a}$	$\hat{a}$
F_0	F_1	$-F_1$
F_1	$-F_0$	F_0
F_2	F_3	$-F_3$
F_3	$-F_2$	F_2
S_0	$-2S_1$	$\frac{1}{2}S_1$
S_1	$2S_0$	$-\frac{1}{2}S_0$

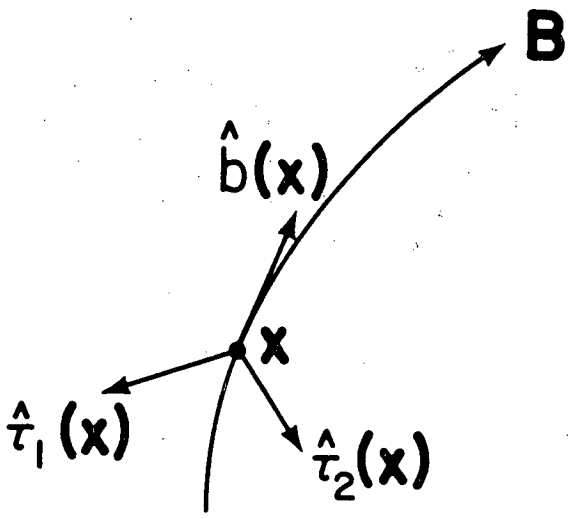
FIGURE CAPTIONS

Fig. 1. The three unit vectors $\hat{b}$, $\hat{\tau}_1$, and $\hat{\tau}_2$.

Fig. 2. The perpendicular plane. θ is the gyrophase to lowest order, and the unit vectors $\hat{a}$ and $\hat{c}$ rotate with the particle.

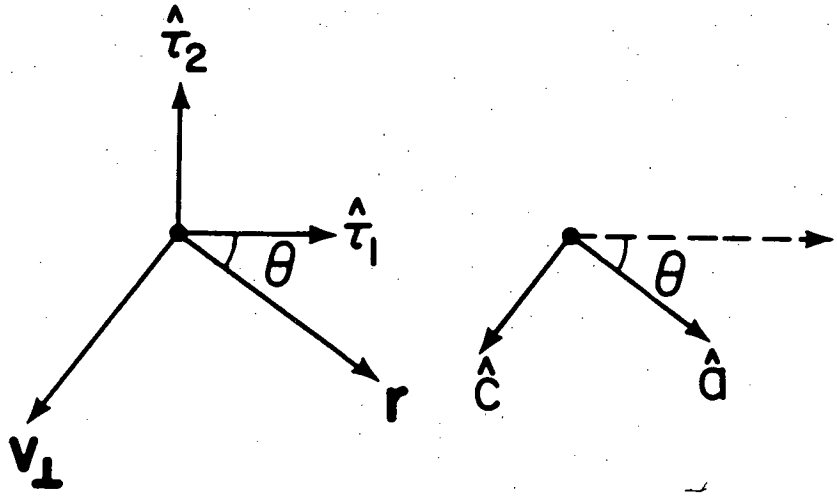
Fig. 3. Motion in a uniform magnetic field. $\underset{\sim}{X}$ is the guiding center position.

Fig. 4. A schmatic illustration of a θ-characteristic. The w=0 "plane" in the diagram actually represents a four-dimensional surface in phase space.



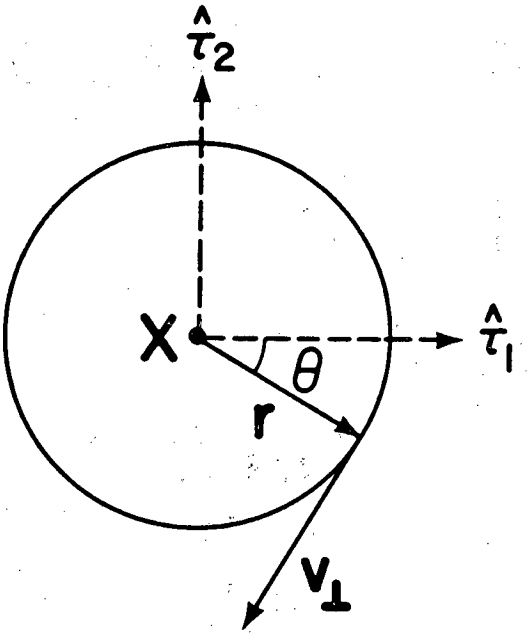
XBL 795-1522

Fig. 1.



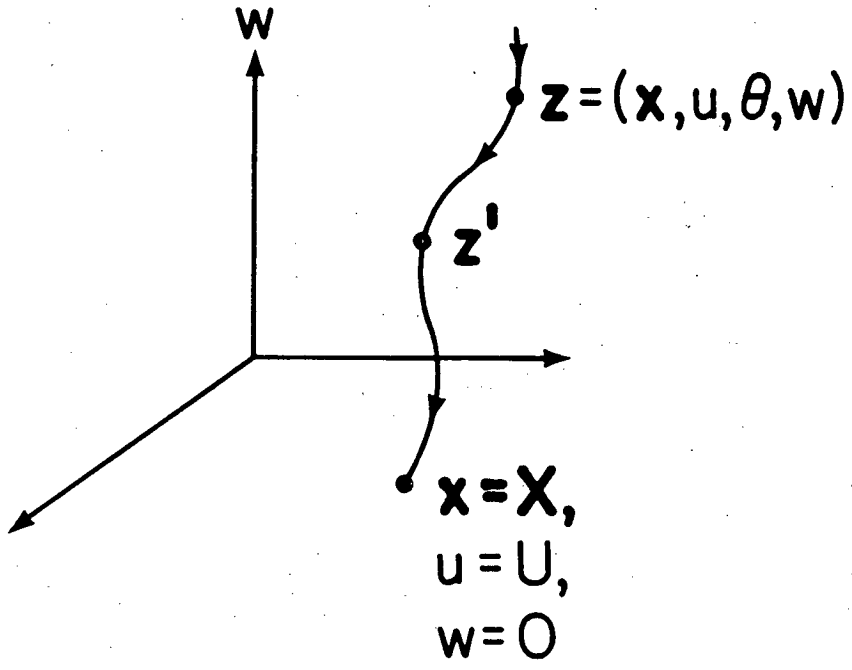
XBL 795-1521

Fig. 2.



XBL 795-1523

Fig. 3.



XBL 795-1524

Fig. 4.

TECHNICAL INFORMATION DEPARTMENT
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720