Lawrence Berkeley National Laboratory
Recent Work

Title

A GUIDING CENTER HAMILTONLAN: A NEW APPROACH

Permalink

https://escholarship.org/uc/item/64h9v22b

Author

Littlejohn, R.G.

Publication Date

1979-04-01

A GUIDING CENTER HAMILTONIAN: A NEW APPROACH

Robert G. Littlejohn

April 1979

DISCLAIMER

A GUIDING CENTER HAMILTONIAN: A NEW APPROACH*

Robert G. Littlejohn

Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

ABSTRACT

A Hamiltonian treatment of the guiding center problem is
given which employs noncanonical coordinates in phase space.
Separation of the unperturbed system from the perturbation is
achieved by using a coordinate transformation suggested by a
theorem of Darboux. As a model to illustrate the method, motion
in the magnetic field $\underset{\sim}{B} = B(x,y)\hat{z}$ is studied. Lie transforms
are used to carry out the perturbation expansion.

1. INTRODUCTION

In this paper I will report on a new approach to a Hamiltonian formu-
lation of the guiding center problem, an approach which leads to a remark-
ably deep insight into the formal structure of classical Hamiltonian mechan-
ics. This insight is not new, in the sense that the natural mathematical
apparatus for an abstract description of Hamiltonian mechanics is that of
differential geometry, and differential geometry has been exhaustively stud-
ied by mathematicians. Nevertheless, even those mathematicians who have
explicitly concerned themselves with Hamiltonian mechanics have tended to
use a language and a notation which is difficult for most physicists. Among
these we might mention Abraham[1], Vinogradov and Kupershmidt,[2] and Arnold.[3]

As a result, very little of the abstract point of view of Hamiltonian

mechanics has found its way into the physics literature, and certainly

not into the more familiar textbooks.[4-6] In addition, for most appli-

cations of interest in physics, even quite theoretical ones, a descrip-

tion of Hamiltonian mechanics which focuses on the differential geometry

of phase space may be deemed to be unncecessarily academic and impractical.

The guiding center problem appears to be an exception, however, since for

this problem one is virtually compelled to employ noncanonical coordinates

in phase space.

The term "the guiding center problem" refers to a certain perturbative

expansion of the solution to the equations of motion of a charged particle

in a given electromagnetic field. The perturbation expansion is based on

an approximation, the "guiding center approximation," which may be roughly

described by saying that electromagnetic effects dominate over inertial

effects. This problem is of great interest and importance in plasma phys-

ics and astrophysics, and over the years various means have been devised for

effecting this perturbative development.[7-14] All of these methods involve

an enormous amount of algebraic manipulations, which has hindered studies

into higher order effects. For example, there still remains some contro-

versy over certain second order terms. This work has arisen out of an

attempt to find a better way to solve this problem.

If the differential equations of motion for the guiding center problem

are written down without regard to their Hamiltonian origin, then it is

straightforward but laborious to subject these equations to a systematic

perturbative treatment, yielding the guiding center expansion. The required

perturbation methods, which are designed for systems of ordinary differential

equations with nearly periodic solutions, were largely developed by Krylov

and Bogoliubov,[15] Bogoliubov and Mitropolski,[16] and Kruskal.[17]   The

work of Kruskal is especially significant, because he showed how the

perturbative solutions relate to action integrals and adiabatic invari-

ants in the case that the system of ordinary differential equations can

be derived from a Hamiltonian.

Similar perturbative methods exist for Hamiltonian systems. These

methods are older than their non-Hamiltonian counterparts, having been

developed originally by Poincaré,[18] and they are the standard methods

found in textbooks.[5,6,19-21] If a system can be analyzed with Hamilton-

ian perturbation methods, then it is much better to do so than to use

non-Hamiltonian methods. The reason is that the equations of motion in

Hamiltonian mechanics are derivable from a scalar function, namely the

Hamiltonian, so that one can deal with a scalar instead of a vector.

Similar considerations apply to coordinate transformations, which in

Hamiltonian mechanics are specified by a scalar, namely the generating

function of the canonical transformation. This advantage becomes greatly

enhanced as one proceeds to higher and higher orders.

Unfortunately, the Hamiltonian for the guiding center problem, which

will be discussed in detail in Section 4 below, cannot be easily analyzed

by the standard methods of Poincaré. The reason is that the relation

between the canonical momentum p and the physical variables x and v des-

cribing the motion of the particle involves the use of the magnetic vector

potential A. That is, the introduction of the vector potential is the

price one must pay in order to use Hamiltonian mechanics. This in itself

would not be so bad, except that in the guiding center approximation the

transformation yielding p from x and v mixes up the ordering scheme, so

that there is no clear separation between the unperturbed system and the

perturbation. This difficulty is not inherent to the problem, but only to a Hamiltonian description of the problem in terms of the usual set of canonically conjugate q's and p's.

In this paper we take an approach to the guiding center problem which preserves the best features of the perturbation method of Poincaré, and yet avoids the use of the vector potential. These goals are accomplished by employing noncanonical coordinate systems in phase space. This step leads one to think more in terms of a geometrical picture of phase space dynamics, and less in terms of coordinate representations with respect to to canonically conjugate (q,p) pairs. One result is a heightened appreciation for the role of differential geometry in the formalism of Hamiltonian mechanics.

Sections 2 and 3 of this paper are included for the sake of establishing certain notation conventions and for the sake of completeness. Section 2 develops some of the essentials of a covariant formulation of Hamiltonian mechanics. This presentation is intentionally and necessarily incomplete, due to lack of space; for example, certain propositions are stated without proof. Unfortunately, there does not seem to be anything in print which covers this subject except in the abstract language of modern mathematics.

In addition, in Section 2 we prove a certain theorem, Theorem 1, which is not at all profound, but which seems heretofore not to have been articulated in quite the same manner, and which is crucial to our perturbation development in Section 5. In Section 3 we discuss in detail a theorem of Darboux, pertaining to the existence of canonical coordinates, which is central in our choice of coordinates in phase space.

In Section 4 we set up the Hamiltonian for the motion of a charged particle in the guiding center approximation. The case studied is that of a nonrelativistic particle in a static magnetic field with a high degree of symmetry, namely $B(x) = B(x,y)\hat{z}$. Although this is a very special case, it serves to illustrate the novel mathematical techniques described in this paper. The application of the same techniques to more realistic problems is straightforward and will be reported upon in forthcoming publications. In Section 4 we use a procedure suggested by the proof of Darboux's theorem to construct a certain "semicanonical" coordinate system in phase space, preparing the Hamiltonian for a standard perturbation analysis, along the lines of the method of Poincaré.

In Section 5 we carry out the perturbation expansion to second order in the guiding center approximation. The expansion is based on the perturbation method of Poincaré, but it differs in two significant ways. One way is that canonical transformations are expressed in terms of their Lie generators, instead of the more conventional mixed-variable generating functions. That is, we use a variant of the so-called Lie transform method, which has been pioneered by Hori,[22] Deprit,[23] Dewar,[24] and others. The second way is that a system of phase space coordinates is used which is noncanonical.

Finally, in Section 6 we discuss various technical aspects of the method and possible extensions and generalizations.

2. A COVARIANT FORMULATION OF HAMILTONIAN MECHANICS

In this section we outline some of the essential features of Hamiltonian mechanics in the context of an arbitrary coordinate system in phase space. To do this it is necessary to call upon the formalism of differential geometry. A relatively accessible source for a more thorough coverage of this subject is the recent textbook by Arnold.[3]

We will denote a coordinate system on phase space by the symbol z or z^i, representing 2N coordinates. N is the number of degrees of freedom of the Hamiltonian system. When these coordinates are some choice of the usual q's and p's, we will call them <u>canonical</u> coordinates, and refer to a canonical coordinate system. In this section, when we refer to canonical coordinates we will decompose the 2N coordinates z^i into q's and p's as follows:

$$z^i = (q_1,\ldots,q_N,p_1,\ldots,p_N) \tag{2.1}$$

Canonical coordinates are to be regarded as a special case, and unless we state the contrary, the coordinates z^i are not to be interpreted as necessarily representing a canonical coordinate system.

A convenient place to begin a covariant formulation of Hamiltonian mechanics is with the Lagrange Brackets. If z represents a set of canonical coordinates, and if $\bar{z}$ represents a set of 2N independent functions of z, then $\bar{z}$ may be interpreted as a possibly noncanonical coordinate system in phase space. The Lagrange Bracket of the quantity $\bar{z}^i$ with the quantity $\bar{z}^j$ will be denoted by the symbol $\bar{\omega}_{ij}$, which, according to the definition, is given by

$$\bar{\omega}_{ij} = \sum_k^N \left(\frac{\partial q_k}{\partial \bar{z}^i} \frac{\partial p_k}{\partial \bar{z}^j} - \frac{\partial p_k}{\partial \bar{z}^i} \frac{\partial q_k}{\partial \bar{z}^j} \right) \tag{2.2}$$

It is convenient to introduce a certain constant, antisymmetric, orthogonal 2N×2N matrix γ, which is represented here by its partition into four N×N matrices:

$$\gamma = \begin{pmatrix} 0 & | & I \\ \text{------} & + & \text{------} \\ -I & | & 0 \end{pmatrix} \tag{2.3}$$

In terms of the matrix γ, the Lagrange Brackets $\bar{\omega}_{ij}$ can be written as follows:

$$\bar{\omega}_{ij} = \frac{\partial z^k}{\partial \bar{z}^i} \gamma_{kl} \frac{\partial z^l}{\partial \bar{z}^j} \tag{2.4}$$

Here and throughout this section summation over repeated indices is understood.

The Poisson Bracket of two phase functions f and g will be denoted by $\{f,g\}$. The Poisson Brackets of the coordinates $\bar{z}$ among themselves are of special importance, and we denote these quantities by $\bar{\sigma}^{ij}$. According to the definition of the Poisson Bracket, we have

$$\bar{\sigma}^{ij} = \{\bar{z}^i, \bar{z}^j\} = \sum_{k}^{N} \left(\frac{\partial \bar{z}^i}{\partial q_k} \frac{\partial \bar{z}^j}{\partial p_k} - \frac{\partial \bar{z}^i}{\partial p_k} \frac{\partial \bar{z}^j}{\partial q_k} \right) \tag{2.5}$$

This can also be written in terms of the matrix γ, as follows:

$$\bar{\sigma}^{ij} = \frac{\partial \bar{z}^i}{\partial z^k} \gamma_{kl} \frac{\partial \bar{z}^j}{\partial z^l} \tag{2.6}$$

In Eqs. (2.4) and (2.6) there may be recognized the transformation laws for the components of second rank tensors of the covariant and contravariant types, respectively. According to this interpretation, $\bar{\omega}_{ij}$ and $\bar{\sigma}^{ij}$ are the components of two tensors with respect to the coordinate system $\bar{z}$. When the coordinate system z is arbitrary, i.e. not

necessarily canonical, or when no distinction need by made between two coordinate systems, we will drop the overbars and write simply ω_{ij} or σ^{ij} for the components of the two tensors with respect to the coordinate system $\underset{\sim}{z}$.

The following connections between the ω tensor and the σ tensor are important. By the well-known properties of the Lagrange Brackets and Poisson Brackets, we have, in any coordinate system,

$$\omega_{ij}\sigma^{kj} = \delta^k_i \tag{2.7}$$

In addition, it is easy to see that $\omega_{ij} = \sigma^{ij} = \gamma_{ij}$ if and only if the coordinate system $\underset{\sim}{z}$ is canonical.

The ω and σ tensors can be viewed in the abstract, apart from their component representations. For the σ tensor, the relation between the two points of view is given by

$$\sigma = \sigma^{ij} \frac{\partial}{\partial z^i} \otimes \frac{\partial}{\partial z^j} = \sum_k \frac{\partial}{\partial q_k} \wedge \frac{\partial}{\partial p_k} \tag{2.8}$$

Thus, for example, the Poisson Bracket of two phase functions f and g can be regarded as the value of the σ tensor on the differentials of the two functions:

$$\{f, g\} = \sigma(df, dg) = \frac{\partial f}{\partial z^i} \sigma^{ij} \frac{\partial g}{\partial z^j} \tag{2.9}$$

Likewise, the ω tensor can be regarded as a 2-form:

$$\omega = \frac{1}{2} \, \omega_{ij} \, dz^i \wedge dz^j = \sum_k^N dq_k \wedge dp_k \tag{2.10}$$

The 2-form ω is <u>nondegenerate</u>, meaning

$$\det(\omega_{ij}) \neq 0 \tag{2.11}$$

It is also <u>closed</u>, meaning $d\omega = 0$, or

$$\frac{\partial \omega_{ij}}{\partial z^k} + \frac{\partial \omega_{jk}}{\partial z^i} + \frac{\partial \omega_{ki}}{\partial z^j} = 0 \tag{2.12}$$

A manifold, such as Hamiltonian phase space, which is endowed with a closed, nondegenerate 2-form is said to be a <u>symplectic</u> manifold.

The fact that ω is closed is especially important. It implies and is implied by the Jacobi identity:

$$\{f,\{g,h\}\} + \{g,\{h,f\} + \{h,\{f,g\} = 0 \tag{2.13}$$

We do not allow the 2-form ω to depend on time, since to do so causes the Poincaré invariants to depend on time. That is, we demand

$$\frac{\partial \omega_{ij}}{\partial t} = 0 \tag{2.14}$$

From a practical point of view, this means that most time-dependent transformations $z = z(q,p,t)$, taking us from a canonical coordinate system to an arbitrary system, must be excluded. Time-dependent canonical trans-

formations are an exception, since $\omega_{ij} = \gamma_{ij} =$ constant in any canonical system. A dynamical system described by a time-dependent Hamiltonian H may be treated by the well-known procedure of taking t and $-H$ as canonically conjugate coordinates in an extended phase space of N+1 degrees of freedom. In this paper there will be no need to consider either time-dependent coordinate transformations or time-dependent Hamiltonians.

An important example of a noncanonical coordinate system in phase space is afforded by the dynamical system consisting of a nonrelativistic particle of mass m and charge e moving in a given, static magnetic field $\underset{\sim}{B}(\underset{\sim}{x})$. The usual canonical coordinates $(\underset{\sim}{q},\underset{\sim}{p})$ for the phase space of this system are given in terms of the particle's position $\underset{\sim}{x}$ and velocity $\underset{\sim}{v}$ by

$$\underset{\sim}{q} = \underset{\sim}{x}$$
$$\underset{\sim}{p} = m\underset{\sim}{v} + \frac{e}{c}\underset{\sim}{A}(\underset{\sim}{x})$$

(2.15)

where $\underset{\sim}{A}(\underset{\sim}{x})$ is a vector potential corresponding to the magnetic field $\underset{\sim}{B}(\underset{\sim}{x})$. The coordinates $(\underset{\sim}{x},\underset{\sim}{v})$ parametrize phase space equally as well as $(\underset{\sim}{q},\underset{\sim}{p})$, but they are noncanonical. Using Eq. (2.5), the components of the σ tensor with respect to this coordinate system are easily obtained:

$$\{x_i,x_j\} = 0$$
$$\{x_i,v_j\} = -\{v_i,x_j\} = \frac{1}{m}\delta_{ij}$$
$$\{v_i,v_j\} = \frac{e}{m^2 c} B_{ij}$$

(2.16)

where

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

(2.17)

The components of the σ tensor can be written in matrix form, with the ordering $\underset{\sim}{z} = (\underset{\sim}{x}, \underset{\sim}{v})$:

$$\sigma^{ij} = \frac{1}{m} \left(\begin{array}{c|c} 0 & I \\ \hline -I & \dfrac{e}{mc}B \end{array} \right) \qquad (2.18)$$

Here the symbol B represents the magnetic field tensor, defined in Eq. (2.17). The components of the 2-form ω in the same coordinate system are given by

$$\omega_{ij} = m \left(\begin{array}{c|c} -\dfrac{e}{mc}B & I \\ \hline -I & 0 \end{array} \right) \qquad (2.19)$$

Observe that the fact that ω is closed implies the Maxwell equation $\nabla \cdot \underset{\sim}{B} = 0$.

Let us now turn our attention to Hamilton's equations of motion and their consequences. These equations are easily cast into a generally covariant form by using the Poisson Bracket and Eq. (2.9). The result is

$$\frac{dz^i}{dt} = \{z^i, H\} = \sigma^{ij} \frac{\partial H}{\partial z^j} \qquad (2.20)$$

One may say that the Hamiltonian transforms as a scalar under arbitrary time-independent coordinate transformations.

As an example of Hamilton's equations in a noncanonical coordinate system, consider the $(\underset{\sim}{x}, \underset{\sim}{v})$ coordinate used in Eqs. (2.15)-(2.19). The Hamiltonian in the $(\underset{\sim}{q}, \underset{\sim}{p})$ coordinates is

$$H(\underset{\sim}{q},\underset{\sim}{p}) = \frac{1}{2m}(\underset{\sim}{p} - \frac{e}{c}\underset{\sim}{A}(\underset{\sim}{q}))^2 \qquad (2.21)$$

In the $(\underset{\sim}{x},\underset{\sim}{v})$ system this becomes, using Eq. (2.15),

$$H(\underset{\sim}{x},\underset{\sim}{v}) = \frac{1}{2}mv^2 \qquad (2.22)$$

Then the equations of motion are

$$\frac{d}{dt}\begin{pmatrix} \underset{\sim}{x} \\ \underset{\sim}{v} \end{pmatrix} = \frac{1}{m}\left(\begin{array}{c|c} 0 & I \\ \hline -I & \frac{e}{mc}B \end{array}\right)\begin{pmatrix} \frac{\partial H}{\partial \underset{\sim}{x}} \\ \frac{\partial H}{\partial \underset{\sim}{v}} \end{pmatrix} \qquad (2.23)$$

$$= \begin{pmatrix} \underset{\sim}{v} \\ \frac{e}{mc}\underset{\sim}{v} \times \underset{\sim}{B} \end{pmatrix}$$

These are, of course, the Newton-Lorentz equations. The "nonphysical" magnetic vector potential $\underset{\sim}{A}$ disappears from the formalism when the $(\underset{\sim}{x},\underset{\sim}{v})$ coordinates are used.

Let us now return to Hamilton's equations of motion and replace the parameter t, describing the trajectories in phase space, with the nondescript parameter λ. This is done because in two applications in this paper, one in the proof of Darboux's theorem and one in the perturbation analysis of Sec. 5, the trajectories which arise from Hamilton's equations have nothing to do with the time evolution of a dynamical system. This replacement also avoids some inessential confusion over our disallowal of time-dependent coordinate transformations.

Let $\underset{\sim}{S}(\underset{\sim}{z}_0,\lambda)$ be the solution to Hamilton's equations which satisfies $\underset{\sim}{z} = \underset{\sim}{z}_0$ at $\lambda = 0$. That is, $\underset{\sim}{S}(\underset{\sim}{z}_0,\lambda)$ satisfies

$$\frac{\partial S^i}{\partial \lambda} = \sigma^{ij} \frac{\partial H}{\partial z^j} \qquad (2.24)$$

where the right-hand side is evaluated at $\underset{\sim}{z} = \underset{\sim}{S}(\underset{\sim}{z}_0,\lambda)$, and also $\underset{\sim}{S}(\underset{\sim}{z}_0,0) = \underset{\sim}{z}_0$ for all $\underset{\sim}{z}_0$. We assume the equivalent of a time-independent system, meaning that Hamilton's equations are autonomous, so that

$$\underset{\sim}{S}(\underset{\sim}{S}(\underset{\sim}{z}_0,\lambda_1),\lambda_2) = \underset{\sim}{S}(\underset{\sim}{z}_0,\lambda_1 + \lambda_2) \qquad (2.25)$$

for all $\underset{\sim}{z}_0$, λ_1,λ_2. This is an elementary result from the theory of ordinary differential equations,[25] and it gives rise to an interpretation of the solution S as a representation of a one-parameter group of diffeomorphisms of phase space onto itself. In view of their origin from Hamilton's equations, these diffeomorphisms are called symplectic diffeomorphisms, and the group is called a Hamiltonian flow.

Symplectic diffeomorphisms can be regarded as mappings of phase space onto itself in a manner independent of coordinate representation, or, in conjunction with a given coordinate system $\underset{\sim}{z}$, they can be regarded as mappings of $\mathbb{R}^{2N}$ onto itself. Of course, the underlying Hamiltonian H and symplectic 2-form ω are implicit. The latter point of view is more useful to us here, because it encourages us to think of symplectic diffeomorphisms as λ-dependent coordinate transformations. That is, we associate a coordinate transformation $\underset{\sim}{z} \to \underset{\sim}{\bar{z}}$ with $\underset{\sim}{\bar{z}} = \underset{\sim}{S}(\underset{\sim}{z},\lambda)$; we will call such a coordinate transformation a symplectic transformation.

For the purposes of perturbation theory it is useful to associate a symplectic transformation with a linear operator, which we denote by $T(\lambda)$. This operator acts on the vector space of phase functions and maps it into

itself, according to the rule

$$(T(\lambda)f)(\underset{\sim}{z}) = f(\underset{\sim}{S}(\underset{\sim}{z},\lambda)) \qquad \qquad (2.26)$$

for any phase function f. That is, $Tf = f \circ \underset{\sim}{S}$. The set $\{T(\lambda) | \lambda \in \mathbb{R}\}$ forms a linear representation of the Hamiltonian flow, and the group multiplication law, corresponding to Eq. (2.25), is

$$T(\lambda_1)T(\lambda_2) = T(\lambda_1 + \lambda_2) \qquad \qquad (2.27)$$

A suitable basis for the Lie algebra of the T representation of the Hamiltonian flow is the operator L, defined by

$$Lf = \{H,f\} \qquad \qquad (2.28)$$

for any phase function f. With these definitions, Hamilton's equations can be written

$$\frac{d}{d\lambda} T(\lambda) = -LT(\lambda) \qquad \qquad (2.29)$$

with solution

$$T(\lambda) = \exp(-\lambda L) \qquad \qquad (2.30)$$

It is well-known that the solutions to Hamilton's equations of motion in the usual (q,p) language give rise to canonical transformations. With respect to an arbitrary coordinate system in phase space, symplectic trans-

formations are the proper generalization of canonical transformations, or at least the regular canonical transformations.[6] Moreover, these transformations play a privileged role among all possible transformations, in spite of the covariant formalism being pursued here, because the 2-form ω is invariant under Hamiltonian flows. This invariance can be stated in several different but equivalent ways. One way is to say that symplectic diffeomorphisms with respect to a canonical coordinate system yield canonical transformations. Another way is to state the invariance of the first Poincaré invariant, which is the integral of ω over some surface in phase space.

For our purposes we choose a third way. We consider some coordinate system $\underset{\sim}{z}$, with respect to which ω has components $\omega_{ij}(\underset{\sim}{z})$, which are to be regarded as definite functions of $\underset{\sim}{z}$. Under an arbitrary change of coordinates $\underset{\sim}{z} \to \underset{\sim}{\bar{z}}$ the components of ω go into $\bar{\omega}_{ij}(\underset{\sim}{\bar{z}})$, which we consider to be functions of the new coordinates $\underset{\sim}{\bar{z}}$, according to the usual rule for covariant tensors:

$$\bar{\omega}_{ij}(\underset{\sim}{\bar{z}}) = \frac{\partial z^k}{\partial \bar{z}^i} \frac{\partial z^\ell}{\partial \bar{z}^j} \omega_{k\ell}(\underset{\sim}{z}) \tag{2.31}$$

However, if the transformation $\underset{\sim}{z} \to \underset{\sim}{\bar{z}}$ is a symplectic transformation, then the invariance of ω means $\bar{\omega}_{ij}(\underset{\sim}{\bar{z}}) = \omega_{ij}(\underset{\sim}{\bar{z}})$. Thus we have the following theorem:

Theorem 1. The functional form of the components of the 2-form ω (and hence also of the σ tensor) is invariant under symplectic transformations.

We will make use of this theorem in Section 5.

3. DARBOUX'S THEOREM

An axiomatic approach to Hamiltonian mechanics begins with the 2-form ω, assumed to be closed and nondegenerate, and then develops the consequences of these assumptions, such as the Jacobi identity. The approach taken in most textbooks on classical mechanics, on the other hand, is to prove theorems such as the Jacobi identity by employing a canonical coordinate system. The axiomatic approch is equivalent to the textbook approach only if it can be shown that a canonical coordinate system actually exists, i.e. a coordinate system such that $\omega_{ij} = \gamma_{ij}$. That one (and hence a whole class) does exist is a consequence of Darboux's theorem, which we shall prove in this section.

For the purposes of Darboux's theorem, it is convenient to decompose a set $\underset{\sim}{z}$ of canonical coordinates into q's and p's in the following order:

$$\underset{\sim}{z} = (q_1, p_1, \ldots, q_N, p_N) \tag{3.1}$$

Corresponding to this ordering, the matrix γ has the form

$$\gamma_{ij} = \begin{pmatrix} \begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array} & & & 0 \\ & \begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array} & & \\ & & \ddots & \\ 0 & & & \begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array} \end{pmatrix} \tag{3.2}$$

This ordering differs from that used in Section 2.

We shall denote phase space by ϕ, representing a 2N dimensional manifold. The construction of canonical coordintes given in the proof of Darboux's theorem generally holds only locally, i.e. in some finite

neighborhood of a given point. We shall, in this section, ignore all questions of the region of applicability of the construction, and speak as if it were valid for all of ϕ. With this understanding, we may state the theorem.

Theorem 2 (Darboux's Theorem). Let there be given a closed, non-degenerate 2-form ω on ϕ and a coordinate system z with respect to which ω has components ω_{ij}. Then there exists a coordinate transformation $z \to \bar{z}$ such that the components $\bar{\omega}_{ij}$ of ω with respect to the new coordinates have the form $\bar{\omega}_{ij} = \gamma_{ij}$. Furthermore, any one of the new coordinates $\bar{z}^i$, considered as a function of the old coordinates z, can be chosen at will.

We remark that if the original coordinate system z is canonical itself, then the constructive proof of Darboux's theorem gives a method of determining a canonical transformation $z \to \bar{z}$ in which one of the new coordinates $\bar{z}^i(z)$ takes on a specified form. It is in this context that Darboux's theorem will be used in Section 4.

Darboux's theorem is proved by induction, using the following lemma:

Lemma. Let there be given the hypotheses of Darboux's theorem. Then there exists a coordinate transformation $z \to \bar{z}$ such that the components $\bar{\omega}_{ij}$ of ω with respect to the new coordinates $\bar{z}$ have the form

$$\bar{\omega}_{ij} = \begin{pmatrix} \Omega_{ij} & 0 \\ \hline 0 & \begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix} \end{pmatrix} \tag{3.3}$$

Furthermore, any one of the new coordinates $\bar{z}^i(z)$ can be chosen at will.

To show how this lemma implies Darboux's theorem, we develop some simple corollaries of the lemma. To do this, it is convenient to label

the new coordinates $\bar{z}$ as follows:

$$\bar{z} = (Z, q, p) \tag{3.4}$$

where the new coordinates Z, corresponding to the Ω_{ij} block in Eq. (3.3), represent 2N-2 functions $Z^i(z)$. First of all, we note that the (2N-2)×(2N-2) matrix Ω_{ij} is antisymmetric. Next, since ω is non-degenerate, we have $\det(\bar{\omega}_{ij}) \neq 0$, and hence also $\det(\Omega_{ij}) \neq 0$. Then, since ω is closed, we have

$$\frac{\partial \bar{\omega}_{ij}}{\partial \bar{z}^k} + \frac{\partial \bar{\omega}_{jk}}{\partial \bar{z}^i} + \frac{\partial \bar{\omega}_{ki}}{\partial \bar{z}^j} = 0 \tag{3.5}$$

If the index k in this equation is set to 2N-1 or 2N, corresponding to the new coordinates q or p, and if neither i nor j takes on these values, then two terms vanish according to Eq. (3.3), since we have $\bar{\omega}_{jk} = \bar{\omega}_{ki} = 0$, and the remaining term gives

$$\frac{\partial \Omega_{ij}}{\partial q} = \frac{\partial \Omega_{ij}}{\partial p} = 0 \tag{3.6}$$

Hence the quantities Ω_{ij} depend only on the new coordinates Z. When none of the indices i,j,k takes on the value 2N-1 or 2N, Eq. (3.5) becomes

$$\frac{\partial \Omega_{ij}}{\partial z^k} + \frac{\partial \Omega_{jk}}{\partial z^i} + \frac{\partial \Omega_{ki}}{\partial z^j} = 0 \tag{3.7}$$

In Eqs. (3.6) and (3.7), the indices i,j,k run over the numbers 1,...,2N-2, corresponding to the coordinates Z.

The result of these corollaries is that the quantities Ω_{ij} are the components with respect to the coordinate system $\underset{\sim}{Z}$ of a certain closed, nondegenerate 2-form Ω on some manifold Φ of dimensionality 2N-2. The manifold Φ can be identified with a submanifold of ϕ, as will be shown later. Hence on Φ the 2-form Ω satisfies the hypotheses of Darboux's theorem, and by the lemma there exists a coordinate transformation $\underset{\sim}{Z} \to \underset{\sim}{\overline{Z}}$, taking the components Ω'_{ij} into $\overline{\Omega}_{ij}$, such that one more pair of q,p coordinates is constructed, and such that one more step toward the form of Eq. (3.2) has been taken. After N applications of the lemma, Darboux's theorem is proved.

The proof of the lemma is constructive. We will call the program for the construction of the coordinates $\underset{\sim}{\overline{z}} = (\underset{\sim}{Z},q,p)$ the <u>Darboux</u> <u>algorithm.</u>

By hypothesis, ω is nondegenerate, so $\det(\omega_{ij}) \neq 0$. Therefore we can define a tensor σ with components σ^{ij} according to Eq. (2.7), and from this, a Poisson Bracket according to Eq. (2.9). When we perform a coordinate transformation $\underset{\sim}{z} \to \underset{\sim}{\overline{z}}$, the components $\overline{\sigma}^{ij}$ of the σ tensor with respect to the new coordinates $\underset{\sim}{\overline{z}}$ are the Poisson Brackets of the new coordinates among themselves. With the definition $\underset{\sim}{\overline{z}} = (\underset{\sim}{Z},q,p)$, we demand the following form for these Poisson Brackets:

$$\{q,p\} = 1 \tag{3.8}$$

$$\{Z^i,q\} = 0 \tag{3.9}$$

$$\{Z^i,p\} = 0 \tag{3.10}$$

$$\{Z^i,Z^j\} = \Sigma^{ij} \tag{3.11}$$

The precise form of the quantities Σ^{ij} is immaterial, although they will automatically be the components of a (2N-2)×(2N-2), antisymmetric, invertible matrix, since the form of $\overline{\sigma}^{ij}$ is given by

$$\bar{\sigma}^{ij} = \begin{pmatrix} \Sigma^{ij} & 0 \\ \hline 0 & \begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix} \end{pmatrix} \tag{3.12}$$

Clearly, Eqs. (3.8)-(3.11) are equivalent to Eq. (3.12) which in turn is equivalent to Eq. (3.3).

First we solve Eq. (3.8). We pick some function $q(z)$ on ϕ for one of the new coordinates; the other 2N-1 functions, $p(z)$ and $z^i(z)$, will then be constrained by Eqs. (3.8)-(3.10). In terms of the given function $q(z)$, Eq. (3.8) is a first-order, linear inhomogeneous partial differential equation for the unknown function $p(z)$. Such an equation always has a solution,[26] which may be found by integrating along the characteristics of the partial differential operator.

In this case the characteristics are the curves $z = z(\lambda)$ which are the solutions to the following set of ordinary differential equations:

$$\frac{dz^i}{d\lambda} = \{z^i, q\} \tag{3.13}$$

These characteristics are the trajectories which result upon treating $q(z)$ as a Hamiltonian. Therefore we will call them "q-characteristics." The parameter λ, which is suggestive of time, is a real number parametrizing the trajectories. It is natural to treat the operator $d/d\lambda$ as a field of tangent vectors, and to write

$$\frac{d}{d\lambda} = \sum_{ij} \sigma^{ij} \frac{\partial q}{\partial z^j} \frac{\partial}{\partial z^i} \tag{3.14}$$

A picture of the solution $p(z)$ to Eq. (3.8) is useful; see Fig. 1. In this figure, Q represents a contour surface of constant q, i.e. a 2N-1 dimensional manifold. Because q is constant along any q-characteristic, every q-characteristic lies in some such contour surface, such as the q-characteristic C_q in the figure. To find $p(z)$, we choose a 2N-1 dimensional manifold P_0, cutting all the Q surfaces. P_0 is arbitrary, excecept that it must be nowhere tangent to any Q surface, since that would result in dq ∧ dp = 0 and preclude the use of q and p as new coordinates. The surface P_0 is to be taken as an initial value surface for $p(z)$; for example, it is convenient to take $p(z) = 0$ for $z \in P_0$. For $z \notin P_0$, $p(z)$ is defined as the negative of the elapsed λ parameter, relative to P_0, of the q-characteristic passing through z. From Eq. (3.14) it then follows that

$$\frac{dp}{d\lambda} = \{p,q\} = -1 \qquad (3.15)$$

and Eq. (3.8) is satisfied.

Next we want to solve Eq. (3.9) for 2N-2 functions $Z^i(z)$ which are independent of each other and also of q and p. Considering q as given and p and Z as unknowns, Eq. (3.9) is the same partial differential equation as Eq. (3.8), except that it is homogeneous. Such an equation possesses 2N-1 independent solutions, so we seem to have one more solution than we need. Actually, we do not, because q itself satisfies the differential equation, i.e. $\{q,q\} = 0$, and the remaining 2N-2 solutions are left for the Z^i.

To construct the solutions $Z^i(z)$ to Eq. (3.9), observe that these functions must be constant along q-characteristics:

$$\frac{dz^i}{d\lambda} = 0 \qquad\qquad (3.16)$$

The z^i may be found by constructing a coordinate system on the surface P_0, in which q is one of the coordinates and the other 2N-2 coordinates are z^i. This defines $z^i(z)$ for $z \in P_0$. For $z \notin P_0$, the values $z^i(z)$ are propagated along q-characteristics so that $z^i(z) = z^i(z')$ whenever z and z' are on the same q-characteristic. The result clearly satisfies Eq. (3.16), and hence also Eq. (3.9).

The functions $z^i(z)$ so constructed are not unique, since any invertible transformation of the form $\bar{Z} = \bar{Z}(Z,q)$, taking Z into $\bar{Z}$, gives a new set of solutions. Such a transformation can be regarded as a coordinate transformation on P_0.

When we turn to Eq. (3.10), we see that the z^i must satisfy further constraints. The latitude we have in the choice of the z^i, as mentioned in the last paragraph, is useful here, because by a proper choice of the coordinate system (Z,q) on P_0 it is possible to satisfy Eqs. (3.9) and (3.10) simultaneously.

The characteristics of Eq. (3.10) are found by treating p(z) as a Hamiltonian, and we will call them the "p-characteristics." They are the solutions $z = z(\mu)$ of the ordinary differential equations

$$\frac{dz^i}{d\mu} = \{z^i, p\} \qquad\qquad (3.17)$$

As before, we may define a tangent vector field $d/d\mu$ by

$$\frac{d}{d\mu} = \sum_{ij} \sigma^{ij} \frac{\partial p}{\partial z^j} \frac{\partial}{\partial z^i} = \{ \ , p\} \qquad\qquad (3.18)$$

The functions $Z^i(z)$ are to be simultaneous constants of the q-characteristics and the p-characteristics. An arbitrary pair of Hamiltonian flows does not in general possess simultaneous constants, since the diffeomorphisms belonging to the two flows do not in general commute. It may be shown, however, that two Hamiltonian flows commute if and only if the Poisson Bracket of the two Hamiltonians is a constant. In the case at hand, the q-flow and the p-flow commute, since $\{q,p\} = 1$.

To construct the $Z^i(z)$, we first select some contour surface Q_0 of $q(z)$, and form the 2N-2 dimensional manifold Φ which is the intersection of this surface with P_0, as shown in Fig. 2. The manifold Φ is the same one mentioned earlier, on which the 2-form Ω is defined. Within Φ we construct a coordinate system by arbitrarily choosing 2N-2 independent functions $Z^i(z)$. Thus the $Z^i(z)$ are defined for $z \in \Phi$. The values $Z^i(z)$ are then propagated along the p-characteristics passing through Φ. These characteristics lie entirely in one contour surface of p, namely P_0. Therefore the $Z^i(z)$ are now defined for $z \in P_0$, and they are constants of the p-characteristics on this surface. The definition of the Z^i is then extended to all of ϕ by propagating along q-characteristics, as shown in Fig. 2. Thus, finally, the $Z^i(z)$ are defined on all of phase space, and they are constants of the q-characteristics everywhere in ϕ.

The last step is to show that the $Z^i(z)$ are constants of the p-characteristics, not just on P_0, but everywhere in ϕ. To do this, consider the quantities $\{Z^i,p\}$, which are known to vanish on the surface P_0. To find their values elsewhere, we compute their derivatives along the q-characteristics, using Eqs. (3.14), (3.8) and (3.9):

$$\frac{d}{d\lambda}\{z^i,p\} = \{\{z^i,p\},q\}$$

$$= \{\{q,p\},z^i\} + \{\{z^i,q\},p\} = 0$$

(3.19)

Hence the $\{z^i,p\}$ vanish everywhere in ϕ, and Eq. (3.10) is satisfied

The Jacobi identity has entered at this point, and it is here that the

fact that ω is closed, which implies the Jacobi identity, has been

called upon.

This completes our proof of Darboux's theorem. Although it may be

regarded as primarily of theoretical interest, we will make a practical

application of it in the next section.

4. APPLICATION OF DARBOUX'S THEOREM TO THE GUIDING CENTER PROBLEM

4.1. Preliminaries

Eqs. (2.15) and (2.21) describe the motion of a nonrelativistic

charged particle in a static magnetic field. For the purposes of this

section and the next, we want to modify these equations in three steps.

The first step is to introduce a dimensionless perturbation parameter

ϵ by replacing the charge e by e/ϵ. Then when the solutions to the equa-

tions of motion are developed in powers of ϵ, the result is the "guiding

center approximation." Although the true solution is found in the end

by setting $\epsilon = 1$, it us useful to consider ϵ to be a variable, describing

a family of systems. In particular, we shall speak of the order of an

expression in terms of its behavior as $\epsilon \to 0$, it being understood that

the particle variables $\underset{\sim}{x}$ and $\underset{\sim}{v}$ and the fields $\underset{\sim}{A}$ and $\underset{\sim}{B}$ are to be held

constant in this limiting process. For example, the gyroradius $mv_\perp c/eB$

is of order one, meaning $O(\epsilon)$, and the gyrofrequency eB/mc is of order -1,

meaning $O(\epsilon^{-1})$. The physical meaning of the limit $\epsilon \to 0$ is that the

particle motion is dominated by a nearly circular, rapid gyration of small gyroradius, which samples only small variations in the magnetic field during a single gyroperiod. The physical meaning of this limit is discussed in greater detail by Northrop,[8] and some of the delicate mathematical aspects of the limit are discussed by Kruskal.[9,27]

The second step is to suppress the constants e, m and c for the sake of notational convenience. These constants are easily restored by a dimensional analysis. The resulting Hamiltonian is

$$H(\underset{\sim}{q},\underset{\sim}{p}) = \frac{1}{2}(\underset{\sim}{p} - \frac{1}{\varepsilon}\underset{\sim}{A}(\underset{\sim}{q}))^2 \tag{4.1}$$

and the relation between the particle variables $(\underset{\sim}{x},\underset{\sim}{v})$ and the canonical coordinates $(\underset{\sim}{q},\underset{\sim}{p})$ is

$$\begin{aligned}
\underset{\sim}{x} &= \underset{\sim}{q} \\
\underset{\sim}{v} &= \underset{\sim}{p} - \frac{1}{\varepsilon}\underset{\sim}{A}(\underset{\sim}{q})
\end{aligned} \tag{4.2}$$

The third step is to restrict consideration to magnetic fields of the form $\underset{\sim}{B}(\underset{\sim}{x}) = B(x,y)\hat{z}$, and furthermore to consider only particle motion in the x-y plane. The problem thereby becomes two-dimensional, and we write $\underset{\sim}{x} = (x,y)$, $\underset{\sim}{v} = (v_x,v_y)$, etc. The magnetic field can be treated as a scalar in the two-dimensional problem; we assume $B > 0$ in the region of space under consideration.

4.2. Two coordinate transformations

In this section we will subject the Hamiltonian in Eq. (4.1) to a sequence of coordinate transformations. The first is given by Eq. (4.2); it was discussed in Sec. 2 in greater detail. Under the coordinate trans-

formation $(\underset{\sim}{q},\underset{\sim}{p}) \to (\underset{\sim}{x},\underset{\sim}{v})$, the Hamiltonian becomes

$$H(\underset{\sim}{x},\underset{\sim}{v}) = \frac{1}{2}v^2 \qquad (4.3)$$

The components σ^{ij} of the σ tensor in this coordinate system can be conveniently represented by giving the formula for the Poisson Bracket of two phase functions f and g:

$$\{f,g\} = \frac{\partial f}{\partial \underset{\sim}{x}} \cdot \frac{\partial g}{\partial \underset{\sim}{v}} - \frac{\partial f}{\partial \underset{\sim}{v}} \cdot \frac{\partial g}{\partial \underset{\sim}{x}} + \frac{1}{\varepsilon} \underset{\sim}{B} \cdot \left(\frac{\partial f}{\partial \underset{\sim}{v}} \times \frac{\partial g}{\partial \underset{\sim}{v}} \right) \qquad (4.4)$$

This is easily seen to be equivalent to Eq. (2.16). Note that ε appears explicitly in the Poisson Bracket.

The second coordinate transformation is motivated by the form of the solution for a uniform magnetic field, which corresponds to the limit $\varepsilon \to 0$. A picture of the particle motion for the case that B(x,y) is uniform is shown in Fig. 3, assuming a positively charged particle. The following definitions, relating to the second coordinate transformation, are valid for an arbitrary field B(x,y), but their physical interpretation is most simple in the uniform case.

First we define a unit vector $\hat{\underset{\sim}{b}}$ along the magnetic field B. According to previous conventions, we have $\hat{\underset{\sim}{b}} = \hat{\underset{\sim}{z}}$. Next we define a unit vector $\hat{\underset{\sim}{c}}$ in the direction of the particle's velocity:

$$\underset{\sim}{v} = v\hat{\underset{\sim}{c}} \qquad (4.5)$$

Finally, we define a unit vector $\hat{\underset{\sim}{a}}$ by $\hat{\underset{\sim}{a}} = \hat{\underset{\sim}{b}} \times \hat{\underset{\sim}{c}}$. Thus the triad $(\hat{\underset{\sim}{a}},\hat{\underset{\sim}{b}},\hat{\underset{\sim}{c}})$ forms a right-handed set. Note that for a uniform magnetic field $\hat{\underset{\sim}{a}}$ is

in the direction of the gyroradius vector $\underset{\sim}{r}$, which is the displacement between the guiding center position $\underset{\sim}{X}$ and the position of the particle $\underset{\sim}{x}$:

$$\underset{\sim}{x} = \underset{\sim}{X} + \underset{\sim}{r} \tag{4.6}$$

In the units chosen, we have, for a uniform magnetic field,

$$\underset{\sim}{r} = \frac{\varepsilon v}{B} \, \underset{\sim}{\hat{a}} \tag{4.7}$$

Fig. 3 also shows the gyrophase θ, which we define as the angle between $\underset{\sim}{\hat{a}}$ and the x-axis, measured in a clockwise sense. Using this angle, we may state the relations between the triad $(\underset{\sim}{\hat{a}}, \underset{\sim}{\hat{b}}, \underset{\sim}{\hat{c}})$ and $(\underset{\sim}{\hat{x}}, \underset{\sim}{\hat{y}}, \underset{\sim}{\hat{z}})$:

$$\begin{aligned}
\underset{\sim}{\hat{c}} &= -\sin\theta \, \underset{\sim}{\hat{x}} - \cos\theta \, \underset{\sim}{\hat{y}} \\
\underset{\sim}{\hat{a}} &= \cos\theta \, \underset{\sim}{\hat{x}} - \sin\theta \, \underset{\sim}{\hat{y}} \\
\underset{\sim}{\hat{b}} &= \underset{\sim}{\hat{z}}
\end{aligned} \tag{4.8}$$

In the uniform field limit, θ evolves linearly in time with frequency B/ε.

We now make the coordinate transformation $(x, y, v_x, v_y) \rightarrow (x, y, \theta, v)$. The Hamiltonian keeps the form of Eq. (4.3), but the Poisson Bracket changes, as indicated here by the components of the σ tensor:

$$\begin{aligned}
\{x_i, x_j\} &= 0 \\
\{\underset{\sim}{x}, v\} &= \underset{\sim}{\hat{c}} \\
\{\underset{\sim}{x}, \theta\} &= -\underset{\sim}{\hat{a}}/v \\
\{\theta, v\} &= \frac{B}{\varepsilon v}
\end{aligned} \tag{4.9}$$

4.3. The Darboux algorithm

The third coordinate transformation is not trivial, and requires some motivation. Consider a Hamiltonian $H(q,p)$. A typical strategy in Hamiltonian perturbation theory is to find a canonical transformation $(q,p) \to (\bar{q},\bar{p})$ such that the new Hamiltonian K is independent of one or more (perhaps all) of the new generalized coordinates $\bar{q}$. To be specific, suppose it is made independent of one new coordinate, say $\bar{q}_1$. Then none of the equations of motion for the other $2N-1$ phase coordinates depends on $\bar{q}_1$, i.e. the $\bar{q}_1$ time evolution is decoupled from the evolution of all the other phase coordinates. In addition, the conjugate momentum $\bar{p}_1$ is a constant of the motion.

It may be seen from Eq. (2.20) that such a strategy does not work so easily in the case of a noncanonical coordinate system. The Hamiltonian may be independent of one of the coordinates z^i, but it does not follow in general that some other coordinate will be a constant of the motion or that the given coordinate will decouple from the others. The reason is that consideration must be given to the components of the σ tensor, which in general depend on z. Consider, for example, the Hamiltonian in Eq. (4.3) and the σ tensor given by Eq. (4.9). These give the following equations of motion:

$$\frac{dx}{dt} = v\hat{c}$$

$$\frac{dv}{dt} = 0 \qquad\qquad (4.10)$$

$$\frac{d\theta}{dt} = \frac{B(x)}{\varepsilon}$$

Thus, although $\partial H / \partial \theta = 0$, θ is not decoupled from the other variables.

It may not be necessary, however, to have a canonical coordinate system in order for the usual strategy of Hamiltonian perturbation theory to work. Consider, for example, the components of the σ tensor shown in Eq. (3.12), with respect to the coordinate system $(Z_1,..,Z_{N-2},q,p)$. Such a coordinate system could be considered "semicanonical," because of the relations in Eqs. (3.8)-(3.10). If $\partial H/\partial q = 0$ in a coordinate system of this type, then p is a constant of the motion, and q is decoupled from the other coordinates. There is no need for the other 2N-2 coordinates $\underset{\sim}{Z}$ to fall into canonically conjugate pairs, and in fact it may be desirable that they not do so.

These considerations suggest that we transform from the coordinates (x,y,θ,v) to a new, semicanonical set (X,Y,θ,J), in which θ remains unchanged and J is canonically conjugate to θ, i.e. $\{\theta,J\} = 1$. The other two coordinates X and Y are to have vanishing Poisson Brackets with both θ and J, but beyond that their form remains to be determined. As it turns out, these two quantities are related to a kind of generalized guiding center position.

Evidently, the coordinate transformation we desire is the result of one application of the Darboux algorithm to the coordinate set (x,y,θ,v), with θ chosen as the new generalized coordinate q, with J corresponding to p, and with $\underset{\sim}{X} = (X,Y)$ corresponding to the (2N-2)-vector $\underset{\sim}{Z}$. Actually, it is desirable to modify the form of Eqs. (3.8)-(3.10) slightly, and ask for solutions J, $\underset{\sim}{X}$ to the set

$$\{\theta,J\} = 1/\varepsilon \qquad\qquad (4.11)$$

$$\{\underset{\sim}{X},\theta\} = 0 \qquad\qquad (4.12)$$

$$\{\underset{\sim}{X},J\} = 0 \qquad\qquad (4.13)$$

The form of Eq. (4.11), which is in contrast to $\{\theta, J\} = 1$, is chosen so that the solution J will be of order zero, i.e. $O(1)$, instead of $O(\epsilon)$.

To solve these equations we will need the θ-characteristics, i.e. the trajectories which result from treating θ as a Hamiltonian. We put $d/d\lambda = \{\ , \theta\}$ and use Eq. (4.9) to get the following differential equations for the θ-characteristics:

$$\frac{d\underset{\sim}{x}}{d\lambda} = -\frac{\hat{\underset{\sim}{a}}}{v} \qquad (4.14)$$

$$\frac{dv}{d\lambda} = -\frac{B(\underset{\sim}{x})}{\epsilon v} \qquad (4.15)$$

Likewise, Eqs. (4.11) and (4.12) can be written in terms of the parameter λ:

$$\frac{dJ}{d\lambda} = -\frac{1}{\epsilon} \qquad (4.16)$$

$$\frac{d\underset{\sim}{X}}{d\lambda} = 0 \qquad (4.17)$$

To get a picture of the θ-characteristics we may examine Eq. (4.15). Since we are assuming $B > 0$, Eq. (4.15) shows that as the parameter λ increases the θ-characteristics move monotonically inward on the surfaces $\theta = $ constant toward the two-dimensional surface $v = 0$, which we shall call V_0. The projections of some of these characteristics onto the v_x-v_y plane are shown in Fig. 4. It is clear that V_0 is a singular surface for the differential equations in Eqs. (4.14) and (4.15), since a single point on this surface is converged upon by a whole family of θ-characteristics,

each one corresponding to a different value of θ. That this is so is not surprising, since θ has a branch point at $v = 0$. The singular nature of the θ-characteristics on this surface will cause us to make certain slight alterations in the Darboux algorithm, as it was presented in Sec. 3.

4.4. Obtaining J

To proceed, it is useful to eliminate the parameter λ from Eqs. (4.14)-(4.16) in favor of v. Since v depends monotonically on λ, this change of independent variable is permissible, and it gives

$$\frac{d\underset{\sim}{x}}{dv} = \frac{\varepsilon}{B} \hat{\underset{\sim}{a}} \tag{4.18}$$

$$\frac{dJ}{dv} = \frac{v}{B} \tag{4.19}$$

Although these equations depend upon the unspecified function $B(\underset{\sim}{x})$ and cannot, therefore, be integrated in closed form, nevertheless a perturbative solution in powers of ε is easily obtained. Since every θ-characteristic meets the surface V_0, the simplest initial condition to assume for the function J is $J = 0$ when $v = 0$. Then integrating Eq. (4.19) by parts and using Eq. (4.18) in an iterative manner yields the formal solution

$$J(\underset{\sim}{x},\theta,v) = \sum_{n=0}^{\infty} \frac{(-\varepsilon)^n v^{n+2}}{(n+2)!} L^n \frac{1}{B(\underset{\sim}{x})} \tag{4.20}$$

where L is the Lie operator defined by

$$L = \frac{1}{B(\underset{\sim}{x})} \hat{\underset{\sim}{a}} \cdot \frac{\partial}{\partial \underset{\sim}{x}} \tag{4.21}$$

The function J is our solution to Eq. (4.11). Note that to lowest order it is the magnetic moment of gyration:

$$J = \frac{v^2}{2B} + O(\varepsilon) \qquad (4.22)$$

The surface V_0 corresponds, in the sense that it is the initial value surface for J, to the surface P_0 in Fig. 2 and in the discussion of the Darboux algorithm in Section 3. Nevertheless, it fails to correspond to P_0 in that it is two-dimensional instead of three-dimensional. This failure is a result of the singularity of θ on $v = 0$, and it causes V_0 to correspond, in a somewhat different sense, to the surface Φ in Fig. 2. These considerations are a warning to be careful in following the Darboux algorithm.

4.5. Obtaining $\underset{\sim}{X}$

We proceed with the construction of a simultaneous solution to Eqs. (4.12) and (4.13) as follows. First we determine the J-characteristics on V_0. We let μ be the real parameter associated with these characteristics, i.e. we put $d/d\mu = \{\ ,J\}$. In an arbitrary region of phase space the equations defining the J-characteristics are complicated, due to the complicated form of Eq. (4.20). But when $v = 0$, they simplify greatly, yielding

$$\frac{d\underset{\sim}{x}}{d\mu} = 0 \qquad (4.23)$$

$$\frac{dv}{d\mu} = 0 \qquad (4.24)$$

Eq. (4.24) is no surprise, because the J-characteristics must remain in a J contour surface, which is $v = 0$ by construction. As for Eq. (4.23), it tells us that the J-characteristics on V_0 are not curves at all, but rather immobile points.

Next we select a coordinate system on V_0, which is to correspond to the coordinates $\underset{\sim}{Z}$ on Φ as described in Section 3, and hence also to the quantities $\underset{\sim}{X}$ in Eqs. (4.12) and (4.13). The simplest and most obvious coordinate system is the rectangular system $\underset{\sim}{x}$ supplied by the original problem. It is for this reason that we use the symbol $\underset{\sim}{X}$ here instead of $\underset{\sim}{Z}$. Therefore we define, for points on V_0,

$$X(\underset{\sim}{x}, v = 0, \theta) = \underset{\sim}{x} \qquad (4.25)$$

The quantities $\underset{\sim}{X}$ are now propagated along J-characteristics in order to satisfy

$$\{\underset{\sim}{X}, J\} = 0 \qquad (4.26)$$

on V_0. But since the J-characteristics are just points, there is nothing to this step, and Eq. (4.26) is automatically satisfied on V_0.

The quantities $\underset{\sim}{X}$ are next propagated along θ-characteristics to extend their definition to all of phase space. The two-dimensional surface V_0 reaches all of four-dimensional phase space by following θ-characteristics because a whole family of θ-characteristics meets any given point of V_0. The result is that the value of the function $\underset{\sim}{X}$ at any given phase point $\underset{\sim}{z} = (x, \theta, v)$ is found by following the θ-characteristic passing through $\underset{\sim}{z}$ until it reaches $v = 0$. This is shown schematically in Fig. 5. The coordinate θ has been suppressed in the figure in order to make a drawing possible. By this definition, we have

$$\{X,\theta\} = 0 \qquad\qquad (4.27)$$

everywhere in phase space.

Exactly as was done in Section 3, we can prove that $d/d\lambda\{X,J\} = 0$, so that Eq. (4.26) is satisfied, not just on V_0, but everywhere in phase space. It is not at all easy to verify Eq. (4.26) directly, using the solution for J given in Eq. (4.20) and that for X given below.

At this point we find an explicit expression for the function $X(x,\theta,v)$. This is obtained from Eq. (4.18), by means of an iterated integration by parts, exactly as Eq. (4.20) was obtained. Eq. (4.25) serves as initial conditions. The result is

$$X(x,\theta,v) = \exp(-\varepsilon vL)x \qquad\qquad (4.28)$$

where the Lie operator L is defined in Eq. (4.21). It is interesting to note that when this series is carried through $O(\varepsilon)$, the result is the guiding center position:

$$X = x - \frac{\varepsilon v}{B}\hat{a} + O(\varepsilon^2) \qquad\qquad (4.29)$$

This may be compared to Eqs. (4.6) and (4.7) for the case of the uniform magnetic field.

Our ability to express the solution X in terms of a simple Lie series is probably fortuitous. For example, the analogous situation does not obtain for the guiding center problem in three dimensions. Nevertheless, some of the many properties of these series[28] will be of use to us here.

For example, Eq. (4.28) may be inverted to solve for $\underset{\sim}{x}$:

$$\underset{\sim}{x}(X,\theta,v) = \exp(+\epsilon vL)\underset{\sim}{X} \qquad (4.30)$$

In this equation the Lie operator L is given by

$$L = \frac{1}{B(\underset{\sim}{X})} \; \hat{\underset{\sim}{a}} \cdot \frac{\partial}{\partial \underset{\sim}{X}} \qquad (4.31)$$

which is to be contrasted with Eq. (4.21). Lie operators are best regarded as operators which take functions into other functions, so that the independent variables in question are dummies. Therefore in what follows we shall usually not explicitly indicate the independent variables in the Lie operator itself, it being understood that they are the same as those of the operand. Eqs. (4.28) and (4.30) are examples of this convention.

4.6. Obtaining the σ tensor

We now have an explicit form for the variable transformation $(x,\theta,v) \rightarrow (X,\theta,J)$, given by Eqs. (4.20) and (4.28). In order to make use of the new coordinate system, we need in addition the components of the σ tensor with respect to the new coordinates. Of the six independent components of the 4×4 antisymmetric component matrix σ^{ij}, five were determined by the construction of the new coordinates, as shown in Eqs. (4.11)-(4.13). The remaining component corresponds to the one independent component of the 2×2 matrix Σ^{ij}, which is shown in Eq. (3.12). This remaining component is the Poisson Bracket {X,Y}, which according to Eq. (3.6) can depend only on X, i.e. not on θ or J.

Consider the Poisson Bracket $\{X,Y\}$ at an arbitrary phase point $z = (X,\theta,J)$. It is easily established that this Poisson Bracket is constant along both θ- and J-characteristics, i.e. that

$$\frac{d}{d\lambda}\{X,Y\} = \frac{d}{d\mu}\{X,Y\} = 0 \tag{4.32}$$

Effectively, this is an application of Poisson's theorem: the Poisson Bracket of any two constants of a Hamiltonian flow is another such constant. Therefore $\{X,Y\}$ can be evaluated at any point on the θ-characteristic which passes through $z = (X,\theta,J)$, and the result will be the same as at z itself. Clearly, the most convenient point to make such an evaluation is on V_0.

In order to find $\{X,Y\}$ on V_0 it is necessary to compute $\{X,Y\}$ in the neighborhood of V_0 and then to let $v \to 0$. In this regard, it may be seen that Eq. (4.28) can be considered a power series in v as well as in ε. Writing this series out, and using Eq. (4.8), we have

$$X = x - \frac{\varepsilon v}{B}\cos\theta + O(v^2)$$
$$Y = y + \frac{\varepsilon v}{B}\sin\theta + O(v^2) \tag{4.33}$$

Then a direct computation of the Poisson Bracket, using Eq. (4.9), gives

$$\{X,Y\} = -\frac{\varepsilon}{B(x)} + O(v) \tag{4.34}$$

But when we let $v \to 0$, x becomes identical with X, and we obtain

$$\{X,Y\} = -\frac{\varepsilon}{B(X)} \tag{4.35}$$

By the arguments above, this is valid at any point $(X,\theta,\underset{\sim}{J})$ of phase

space. As predicted, $\{X,Y\}$ depends only on $\underset{\sim}{X}$.

Altogether, in the coordinate system (X,Y,θ,J) the components

of the σ matrix are

$$
\sigma^{ij} = \begin{pmatrix} 0 & -\dfrac{\varepsilon}{B(\underset{\sim}{X})} & 0 & 0 \\[2ex] +\dfrac{\varepsilon}{B(\underset{\sim}{X})} & 0 & 0 & 0 \\[2ex] 0 & 0 & 0 & \dfrac{1}{\varepsilon} \\[2ex] 0 & 0 & -\dfrac{1}{\varepsilon} & 0 \end{pmatrix} \tag{4.36}
$$

That is, we can write the Poisson Bracket of two functions f and g in

terms of the coordinates (X,θ,J) as follows:

$$
\{f,g\} = \frac{\varepsilon}{B(\underset{\sim}{X})}\left(\frac{\partial f}{\partial Y}\frac{\partial g}{\partial X} - \frac{\partial f}{\partial X}\frac{\partial g}{\partial Y}\right) + \frac{1}{\varepsilon}\left(\frac{\partial f}{\partial \theta}\frac{\partial g}{\partial J} - \frac{\partial f}{\partial J}\frac{\partial g}{\partial \theta}\right) \tag{4.37}
$$

4.7. Iterating the Darboux algorithm

At this point it is interesting to consider what would happen if

another iteration of the Darboux algorithm were carried out, representing

a coordinate change $(X,Y,\theta,J) \to (Q,P,\theta,J)$, which would bring the σ tensor

into the form $\sigma^{ij} = \gamma_{ij}/\varepsilon$. Except for the factor $1/\varepsilon$, which is a minor

consideration, we would then have constructed, by means of a number of

noncanonical intermediaries, an overall canonical transformation

$(q_x, q_y, p_x, p_y) \to (Q,P,\theta,J)$. According to the theory in Section 3, the new

coordinates Q and P would be functions of $\underset{\sim}{X}$ alone, and they would satisfy

$\{Q,P\} = 1/\varepsilon$.

The functions Q and P of $\underset{\sim}{X}$ which are produced by a second iteration of the Darboux algorithm cannot be constructed perturbatively, as were $\underset{\sim}{X}$ and J. Nevertheless, these functions are related in a simple manner to the well-known Euler Potentials,[29] which are usually denoted by α and β :

$$Q(\underset{\sim}{X}) = \beta(\underset{\sim}{X})/\epsilon$$
$$P(\underset{\sim}{X}) = \alpha(\underset{\sim}{X})/\epsilon$$

(4.38)

The functions α and β satisfy

$$\nabla\alpha\times\nabla\beta = \underset{\sim}{B}$$

(4.39)

which in our two-dimensional field configuration becomes

$$B(X,Y) = \frac{\partial\alpha}{\partial X}\frac{\partial\beta}{\partial Y} - \frac{\partial\alpha}{\partial Y}\frac{\partial\beta}{\partial X}$$

(4.40)

From this and Eq. (4.37) it is easy to show that $\{Q,P\} = 1/\epsilon$. Incidentally, we see that Darboux's theorem implies the existence of Euler Potentials, at least for the two-dimensional field configuration considered here.

In the remainder of this paper we choose to use the coordinates $\underset{\sim}{X}$ instead of the Euler Potentials α and β , i.e. we choose to remain with the semicanonical coordinate system (X,Y,θ,J). This is done for several reasons. In the first place, what we gain by using canonical coordinates is the ability to use standard textbook formulas for Hamiltonian mechanics, while what we lose is that we must deal with Euler Potentials, which are nonphysical in the same sense that the vector potential $\underset{\sim}{A}$ is nonphysical. On the other hand, Eq. (4.37) shows that the Poisson Bracket in the $(X,\theta,\underset{\sim}{J})$

coordinate system is not excessively complicated in comparison to the usual formula for a canonical coordinate system. In the second place, when the guiding center problem is generalized to three-dimensional fields and is analyzed along the lines presented here, there results a set of four noncanonical variables, corresponding to the two variables (X,Y) given here. These four variables cannot be transformed into two canonically conjugate pairs except by using functions which are much less familiar than the Euler Potentials. That is, the two-dimensional problem is a special case, in that the second application of the Darboux algorithm is solvable in terms of well-known functions. To treat the general case, it seems better to stay with noncanonical or semicanonical coordinate systems, and this we shall do also in the special two-dimensional case.

4.8. The Hamiltonian

Let us now consider the inverse of the transformation $(x,\theta,v) \to (X,\theta,J)$, which we will need in order to express the Hamiltonian in terms of the new coordinates. To begin with, we have in Eq. (4.20) the quantity J expressed as a function of (x,θ,v). Using Eq. (4.30), J may be expressed as a function of (X,θ,v). In the process of eliminating x in favor of X, there results a double infinite series involving the operator L. This can be collapsed back into a single series, yielding finally

$$J(X,\theta,v) = \sum_{n=0}^{\infty} \frac{\varepsilon^n v^{n+2}}{n!(n+2)} L^n\left(\frac{1}{B(X)}\right) \qquad (4.41)$$

Next, we invert this series to obtain v as a function of (X,θ,J). Carried out through second order, this gives

$$v(\underset{\sim}{X},\theta,J) = (2BJ)^{1/2} + \varepsilon\frac{(2BJ)}{3B^2}(\hat{\underset{\sim}{a}}\cdot\nabla B)$$

$$+ \varepsilon^2 \frac{(2BJ)^{3/2}}{72\ B^4}[9B(\hat{\underset{\sim}{a}}\hat{\underset{\sim}{a}}:\ \nabla\nabla B) - 7(\hat{\underset{\sim}{a}}\cdot\nabla B)^2]$$

(4.42)

This can then be substituted into Eq. (4.30) to obtain $\underset{\sim}{x}$ as a function of $(\underset{\sim}{X},\theta,J)$:

$$\underset{\sim}{x}(\underset{\sim}{X},\theta,J) = \underset{\sim}{X} + \varepsilon\ \frac{(2BJ)^{1/2}}{B}\ \underset{\sim}{a} - \varepsilon^2\ \frac{(2BJ)}{6B^3}(\hat{\underset{\sim}{a}}\cdot\nabla B)\hat{\underset{\sim}{a}}$$

$$+ \varepsilon^3\ \frac{(2BJ)^{3/2}}{72\ B^5}[-3B(\hat{\underset{\sim}{a}}\hat{\underset{\sim}{a}}:\ \nabla\nabla B) + 5(\hat{\underset{\sim}{a}}\cdot\nabla B)^2]\hat{\underset{\sim}{a}}$$

(4.43)

In Eqs. (4.42) and (4.43), B means $B(\underset{\sim}{X})$ and ∇ means $\partial/\partial\underset{\sim}{X}$. These two formulas give the desired inverse transformation, $(\underset{\sim}{X},\theta,J) \to (\underset{\sim}{x},\theta,v)$.

Finally, we can use Eqs. (4.3) and (4.42) to find the Hamiltonian in the $(\underset{\sim}{X},\theta,J)$ coordinate system. The result is

$$H(\underset{\sim}{X},\theta,J) = BJ + \varepsilon\ \frac{(2BJ)^{3/2}}{3B^2}(\hat{\underset{\sim}{a}}\cdot\nabla B)$$

$$+ \varepsilon^2\ \frac{(2BJ)^2}{24B^4}[3B(\hat{\underset{\sim}{a}}\hat{\underset{\sim}{a}}:\ \nabla\nabla B) - (\hat{\underset{\sim}{a}}\cdot\nabla B)^2] + O(\varepsilon^3)$$

(4.44)

In the next section we will follow the usual strategy of Hamiltonian perturbation theory in order to find a transformation which will make H independent of θ. The result will be a Hamiltonian for the guiding center motion.

5. THE GUIDING CENTER HAMILTONIAN

In this section the Hamiltonian in Eq. (4.44) is subjected to a near-identity coordinate transformation of the form $(X,\theta,J) \rightarrow (\bar{X},\bar{\theta},\bar{J})$ such that three criteria are fulfilled. First, the new Hamiltonian is to be independent of $\bar{\theta}$. Second, the transformation is to be free of secular terms. And third, the new coordinates are to be semi-canonical in the same sense that the old ones are, so that $\bar{J}$ will be a constant of the motion (the generalized magnetic moment) and so that the time evolution of $\bar{\theta}$ will decouple from that of the other phase coordinates. The first two criteria are standard in Hamiltonian perturbation theory for nearly periodic systems; the third is a novel element, arising from our use of noncanical coordinates in phase space.

We are not looking for canonical transformations, in the usual sense, because our coordinate system is noncanonical. However, on the strength of Theorem 1, we do want to use symplectic transformations, since these will cause the third criterion to be fulfilled. Although these coordinate transformations are very much like canonical transformations, being in a sense canonical transformations expressed in noncanonical coordinates, it is nevertheless awkward to express them in terms of the usual mixed variable generating functions. Instead, we express these symplectic transformations in terms of a set of Lie generators, following the theory outlined in Section 2. That is, we will use a variant of the Lie transform method.[21-24]

Consider a sequence $w_1, w_2, \ldots$ of time-independent phase functions, and associated operators $L_1, L_2, \ldots$ which are defined on analogy to Eq. (2.28):

$$L_n f = \varepsilon\{w_n, f\} \tag{5.1}$$

for any phase function f. The factor ε has been introduced into this definition because the Poisson Bracket given in Eq. (4.37) has a term which is $O(\varepsilon^{-1})$.

Next, each of these functions is used to generate a symplectic transformation, according to the formula

$$T_n = \exp\left(-\frac{\varepsilon^n L_n}{n}\right) \tag{5.2}$$

The factor $1/n$ is included in order to make the resulting formulas follow as closely as possible the conventions of Cary.[30] Finally, a symplectic transformation T is constructed by multiplying together the T_n:

$$T = \ldots T_3 T_2 T_1 \tag{5.3}$$

$$T^{-1} = T_1^{-1} T_2^{-1} T_3^{-1} \ldots \tag{5.4}$$

These operators are expanded as power series in ε by multiplying together the exponential series associated with Eq. (5.2). To obtain the correct ordering in powers of ε it is necessary to take account of the fact that the operators L_n consist of a $O(1)$ part and an $O(\varepsilon^2)$ part, according to Eq. (4.37). Therefore we define two more series of operators, as follows:

$$M_n f = \frac{\partial w_n}{\partial \theta} \frac{\partial f}{\partial J} - \frac{\partial w_n}{\partial J} \frac{\partial f}{\partial \theta} \tag{5.5}$$

and

$$N_{n+2} = \frac{1}{B}\left(\frac{\partial w_n}{\partial Y}\frac{\partial f}{\partial X} - \frac{\partial w_n}{\partial X}\frac{\partial f}{\partial Y}\right) \tag{5.6}$$

so that

$$L_n = M_n + \epsilon^2 N_{n+2} \tag{5.7}$$

When the operators T and T^{-1} are expressed in terms of the M and N operators, the results are, through third order in ϵ,

$$\begin{aligned}
T &= I - \epsilon M_1 + \frac{1}{2}\epsilon^2(-M_2 + M_1^2) + \frac{1}{6}\epsilon^3(-2M_3 - 6N_3 \\
&\quad - M_1^3 + 3M_2 M_1) + O(\epsilon^4)
\end{aligned} \tag{5.8}$$

$$\begin{aligned}
T^{-1} &= I + \epsilon M_1 + \frac{1}{2}\epsilon^2(M_2 + M_1^2) + \frac{1}{6}\epsilon^3(2M_3 + 6N_3 \\
&\quad + M_1^3 + 3M_1 M_2) + O(\epsilon^4)
\end{aligned} \tag{5.9}$$

In terms of the coordinates $\underset{\sim}{z} = (X,\theta,J)$ and $\underset{\sim}{\bar{z}} = (\bar{X},\bar{\theta},\bar{J})$, we may say, somewhat loosely,

$$\underset{\sim}{\bar{z}} = T\underset{\sim}{z} \tag{5.10}$$

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

As was noted in Section 4, the independent variables of the Lie operators M_n and N_n which appear in the expansion of T are the same as those of the operand.

When the symplectic transformation T is applied to the Hamiltonian H, there results a new Hamiltonian K, according to

$$K = T^{-1}H \qquad (5.12)$$

In this equation we expand both K and H in powers of ε:

$$H = \sum_{n=0}^{\infty} \varepsilon^n H_n \qquad (5.13)$$

$$K = \sum_{n=0}^{\infty} \varepsilon^n K_n \qquad (5.14)$$

Then using Eq. (5.9) and collecting terms gives a hierarchy of equations, which through second order can be expressed as follows:

$$0 = K_0 - H_0 \qquad (5.15)$$

$$M_1 H_0 = K_1 - H_1 \qquad (5.16)$$

$$M_2 H_0 = 2(K_2 - H_2) - M_1(H_1 + K_1) \qquad (5.17)$$

These equations are written in this form because they are to be regarded as partial differential equations for the w_n, which specify the transformation T. To see this, note that

$$M_n H_0 = B \frac{\partial w_n}{\partial \theta} \qquad (5.18)$$

The perturbation expansion is carried out by selecting the w_n, order by order, so that K is independent of θ, and so that the w_n contain only purely oscillatory terms in θ. The resulting w_n are

$$w_1 = \frac{(2BJ)^{3/2}}{3B^3}(\hat{c} \cdot \nabla B) \qquad (5.19)$$

$$w_2 = \frac{(2BJ)^2}{24B^5}\hat{a}\hat{c}: (3B\nabla\nabla B - \nabla B\nabla B) \tag{5.20}$$

The new Hamiltonian K, which we may justifiably call the guiding center Hamiltonian, is given by

$$K(\bar{\underset{\sim}{X}},\bar{J}) = B\bar{J} + \varepsilon^2\frac{\bar{J}^2}{4B^2}[B\nabla^2 B - 3(\nabla B)^2] + O(\varepsilon^3) \tag{5.21}$$

where B means $B(\bar{\underset{\sim}{X}})$ and where ∇ means $\partial/\partial\bar{\underset{\sim}{X}}$.

The equations of motion resulting from K are immediate; the effect of the ε ordering of the Poisson Bracket should be noted.

$$\frac{d\bar{\underset{\sim}{X}}}{dt} = \frac{\varepsilon\hat{b}}{B}\times\left\{\bar{J}\nabla B + \varepsilon^2\frac{\bar{J}^2}{4}\nabla\left[\frac{\nabla^2 B}{B} - \frac{3(\nabla B)^2}{B^2}\right]\right\} \tag{5.22}$$

$$+ O(\varepsilon^5)$$

$$\frac{d\bar{\theta}}{dt} = \frac{B}{\varepsilon} + \varepsilon\frac{\bar{J}}{2B^2}[B\nabla^2 B - 3(\nabla B)^2] + O(\varepsilon^3) \tag{5.23}$$

$$\frac{d\bar{J}}{dt} = 0 \tag{5.24}$$

The first term of Eq. (5.22) is the so-called "grad B drift."

Finally, the relation (5.10) can be written out, connecting $\underset{\sim}{z}$ and $\bar{\underset{\sim}{z}}$. This gives

$$\bar{\underset{\sim}{X}} = \underset{\sim}{X} + \frac{\varepsilon^3}{B}\hat{b}\times\nabla\left[\frac{(2BJ)^{3/2}}{3B^3}(\hat{\underset{\sim}{c}}\cdot\nabla B)\right] + O(\varepsilon^4) \tag{5.25}$$

$$\bar{\theta} = \theta + \varepsilon \, \frac{(2BJ)^{1/2}}{B^2} (\hat{\underset{\sim}{c}} \cdot \nabla B)$$

(5.26)

$$+ \, \varepsilon^2 \frac{(2BJ)}{12B^4} \, \hat{\underset{\sim}{a}}\hat{\underset{\sim}{c}} : \, (3B\nabla\nabla B - 5\nabla B\nabla B) + 0(\varepsilon^3)$$

$$\bar{J} = J + \varepsilon \frac{(2BJ)^{3/2}}{3B^3} (\hat{\underset{\sim}{a}} \cdot \nabla B) + \varepsilon^2 \, \frac{(2BJ)^2}{48B^5} [\, (7\hat{\underset{\sim}{a}}\hat{\underset{\sim}{a}} + 9\hat{\underset{\sim}{c}}\hat{\underset{\sim}{c}}) : \nabla B\nabla B$$

(5.27)

$$+ \, 3B(\hat{\underset{\sim}{a}}\hat{\underset{\sim}{a}} - \hat{\underset{\sim}{c}}\hat{\underset{\sim}{c}}) : \nabla\nabla B] + 0(\varepsilon^3)$$

In all cases these formulas have been carried out to the highest order which is consistent with the knowledge of only w_1 and w_2.

By combining Eqs. (5.25)-(5.27) with (4.20) and (4.28) the variables $(\bar{\underset{\sim}{X}}, \bar{\theta}, \bar{J})$ can be expressed in terms of $(\underset{\sim}{x}, \underset{\sim}{v})$. We remark that although the convergence of the series in Eqs. (5.25)-(5.27) is questionable, the convergence of the series in Eqs. (4.20) and (4.28) is easy to establish for sufficiently small values of ε and for $1/B$ a real analytic function of $\underset{\sim}{x}$. The practical utility of perturbation series may not be lost even if the series are divergent.

6. DISCUSSION AND CONCLUSIONS

The use of the transformation given in Eqs. (4.20) and (4.28), which we may call the Darboux transformation, is the most unusual element in the approach taken in this paper to a perturbation problem. There is nothing new, however, in the function which this transformation serves. The Darboux transformation fulfills the purpose of isolating the unperturbed system from the perturbation, and it is exactly the difficulty of achieving this separation that has made previous Hamiltonian treatments of guiding center motion so nonstandard in appearance and awkward in execution. In addition, the Darboux transformation yields a set of variables

which are natural to the unperturbed system, since to lowest order X and J are constants of the motion and θ evolves linearly in time. The importance of these two goals--the isolation of the unperturbed system and the choice of an appropriate set of coordinates for the unperturbed system--has been made very clear, on the basis of an invariant, geometrical picture of phase space orbits, in a seminal paper by Kruskal[17] on nearly periodic systems. These goals are common to both Hamiltonian and non-Hamiltonian systems, and the Darboux transformation forms a kind of bridge between a Hamiltonian and a non-Hamiltonian treatment of the guiding center problem.

In textbook problems on perturbation theory the unperturbed system is separated from the perturbation at the outset, and hence the separation, as a task in itself, is hardly recognized. In a non-Hamiltonian treatment of the guiding center problem it is nearly trivial to achieve this separation, as has been shown by Bogoliubov and Mitropolski.[16] It was on the basis of this non-Hamiltonian separation that the angle θ was chosen as a new coordinate in the construction of the Darboux transformation in Sec. 4, and this choice caused the desired separation in the Hamiltonian treatment as well.

Likewise, the choice of appropriate variables for the unperturbed system is often nearly unconscious in textbook examples. In Hamiltonian systems, this choice can be formalized by saying that one must solve the Hamilton-Jacobi equation for the unperturbed system before proceeding with a perturbation treatment, although often the required solution is obvious. In our example, the Darboux transformation automatically provides us with a set of coordinates appropriate to the unperturbed system, because the canonically conjugate variables θ and J are effectively action-angle variables for the unperturbed system.

The construction of the Darboux transformation, as it was given in Section 4, is not unique, in the sense that the selection of any phase function which differs from θ by terms of order ε or higher would satisfy the two goals discussed above equally as well as θ itself. The only reason for choosing θ is that it has a simple dependence on (x, v). Indeed, if $\bar{\theta}$, given by Eq. (5.26), were chosen, then not only would the unperturbed system separate from the perturbation, but also the entire Hamiltonian would decouple from $\bar{\theta}$. This consideration raises the possibility that the construction of the Darboux transformation in Section 4 and the perturbation treatment in Section 5 could be merged, although I have not yet investigated this question.

In this paper a Hamiltonian treatment of the guiding center problem has been achieved at the expense of the construction of the Darboux transformation. It may well be asked if the result is worth the price. There are several reasons to believe the answer is yes.

In the first place, even if the results are carried to lowest order, giving only the classic, well-known "drifts," the method provides, nonetheless, a Hamiltonian treatment of these lowest order results within the framework of a systematic ordering scheme.

Second, the method seems to give the shortest avenue to higher order results, in terms of the labor involved, although this may best be judged by those who have used other methods. The perturbation treatment in Section 5 is no worse than any standard Hamiltonian perturbation treatment, and enormously better than a non-Hamiltonian treatment. The Darboux transformation itself is perturbative, i.e. it is a power series in ε instead of a transformation in closed form, but it is based on a secular perturbation treatment which is quite simple. On balance, it seems that

a simple secular perturbation expansion plus a standard Hamiltonian perturbation expansion is much less laborious than a non-Hamiltonian expansion.

Third, a simple Hamiltonian treatment of the guiding center problem opens the door to the addition of other perturbations, such as electromagnetic waves, and to the study of, for example, the effects of these on adiabatic invariants. Some results along these lines have already been achieved by Grebogi, Kaufman and Littlejohn.[31]

Fourth, successive iterations of the Darboux algorithm give a simple means of exploring the other adiabatic invariants of guiding center motion, such as the longitudinal invariant and the flux invariant.[8]

Fifth, since the dynamics of statistical ensembles of charge particles in the Vlasov approximation can be described in Hamiltonian terms, the guiding center Hamiltonian can be used to treat nonuniform magnetic fields in a plasma, a case of great practical importance. The possible applications of a guiding center Hamiltonian to kinetic theory are too numerous to mention.

Several extensions of the results of the present paper have already been completed and will be reported upon in forthcoming publications. Two-dimensional, fully electromagnetic fields have been treated, as well as three-dimensional magnetostatic fields. The results are promising, and work is beginning on three-dimensional electromagnetic fields and relativistic treatments, as well as on applications in other directions.

REFERENCES

1. R. Abraham, _Foundations of Mechanics_ (Benjamin, New York, 1967).

2. A. M. Vinogradov and B. A. Kupershmidt, Russ. Math. Surveys $\underset{\sim\sim\sim\sim}{32:4}$, 177(1977).

3. V. I. Arnold, _Mathematical Methods of Classical Mechanics_ (Springer-Verlag, New York, 1978).

4. H. Goldstein, _Classical Mechanics_ (Addison-Wesley, Reading Massachusetts, 1950).

5. E. Saletan and A. Cromer, _Theoretical Mechanics_ (Wiley, New York, 1971).

6. E. C. G. Sudarshan and N. Mukunda, _Classical Dynamics: A Modern Approach_ (Wiley, New York, 1974).

7. H. Alfvén, _Cosmical Electrodynamics_ (Clarendon Press, Oxford, 1950).

8. T. G. Northrop, _The Adiabatic Motion of Charged Particles_ (Interscience, New York, 1963).

9. M. D. Kruskal, "Elementary Orbit and Drift Theory," in _Plasma Physics_ (International Atomic Energy Agency, Vienna, 1965).

10. A. Baños, J. Plasma Phys. $\underset{\sim}{1}$, 305(1967).

11. P. Martín, "Second Order Drift Velocities," preprint, Dpto. Física, Universidad Simón Bolívar, Caracas, Venezuela, 1977.

12. C. S. Gardner, Phys. Rev. $\underset{\sim\sim\sim}{115}$, 791(1959).

13. T. G. Northrop and J. A. Rome, Phys. Fluids $\underset{\sim\sim}{21}$, 384(1978).

14. D. P. Stern, "Hamiltonian Formulation of Guiding Center Motion," Goddard Space Flight Center Report No. X-641-71-56.

15. N. Krylov, N. Bogoliubov, _Introduction to Non-Linear Mechanics_ (Princeton University Press, Princeton, 1947).

16. N. N. Bogoliubov and Y. A. Mitropolski, _Asymptotic Methods in the Theory of Non-Linear Oscillations_ (Gordon and Breach, New York, 1961).

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

18. H. Poincare, _Les Methods Nouvelles de la Mechanique Celeste_, 1892 (Reprinted by Dover, New York, 1957).

19. D. ter Haar, _Elements of Hamiltonian Dynamics_ (North Holland Publ., Amsterdam, 1964).

20. Y. Hagihara, _Celestial Mechanics_ (MIT Press, Cambridge, Massachusettts, 1972).

21. A. H. Nayfeh, _Perturbation Methods_ (Wiley, New York, 1973).

22. G. Hori, Pub. Astron. Soc. Jap. 18, 287 (1966).

23. A. Deprit, Cel. Mech. 1, 12(1969).

24. R. L. Dewar, J. Phys. A 9, 2043(1976).

25. V. I. Arnold, _Ordinary Differential Equations_ (MIT Press, Cambridge, Massachusetts, 1978).

26. R. Courant and D. Hilbert, _Methods of Mathematical Physics_ (Interscience, New York, 1953).

27. M. D. Kruskal, "Asymptotology," in _Mathematical Models in Physical Sciences_ (Prentice-Hall, Englewood Cliffs, N. J., 1963).

28. W. Gröbner, _Die Lie-Reihen und Ihre Anwendungen_ (VEB Deutscher Verlag der Wissenschaften, Berlin, 1967).

29. D. P. Stern, Am. J. Phys. 38, 494(1970).

30. J. R. Cary, "Lie Transforms and Their Use in Hamiltonian Perturbation Theory," Lawrence Berkeley Laboratory Report No. LBL-6350, 1978.

31. C. Grebogi, A. N. Kaufman, and R. G. Littlejohn, "Ponderomotive Effects Due to an Electromagnetic Wave in a Nonuniform Magnetic Field," Lawrence Berkeley Laboratory Report No. LBL-8916, 1979.

FIGURE CAPTIONS

Fig. 1. The q-characteristics and the construction of the functions

$\underset{\sim}{p}(z)$ and $\underset{\sim}{Z}(\underset{\sim}{z})$.

Fig. 2. The construction of the functions $\underset{\sim}{Z}(\underset{\sim}{z})$ as simultaneous constants

of the q- and p-characteristics.

Fig. 3. Guiding center variables for a uniform magnetic field. The unit
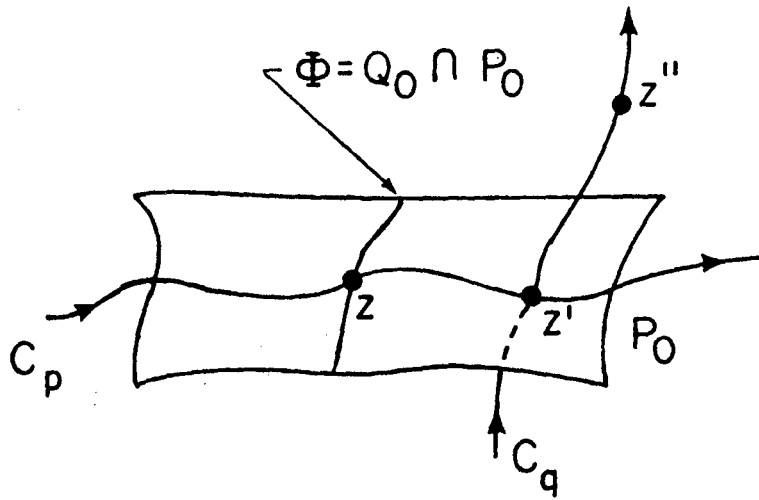
vectors $\hat{a}$, $\hat{c}$ rotate with the particle.

Fig. 4. The θ-characteristics converge on the surface $v = 0$.

Fig. 5. Geometrical meaning of the functions $X(x, \theta, v)$. The figure shows

a θ-characteristic moving toward the surface $v = 0$.
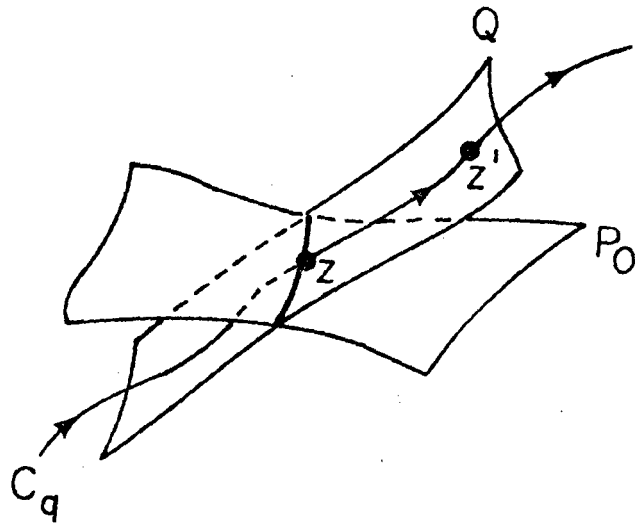
XBL 793-972

Fig. 1

XBL 793 - 973

Fig. 2

XBL 793-970

Fig. 3

XBL 793-971

Fig. 4

XBL 793-969

Fig. 5