

c.2 repl.

LBL-12942
c.2 repl.



Lawrence Berkeley Laboratory

UNIVERSITY OF CALIFORNIA

Accelerator & Fusion Research Division

RECEIVED
LAWRENCE
BERKELEY LABORATORY

AUG 17 1982

LIBRARY AND
DOCUMENTS SECTION

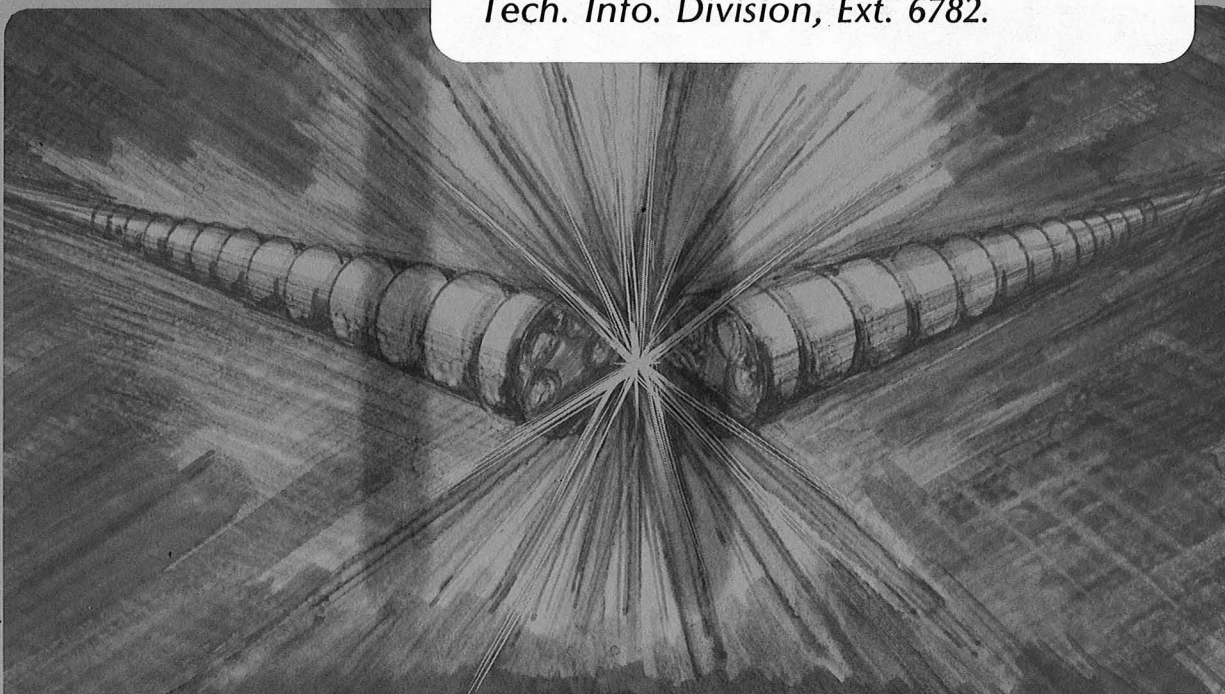
HAMILTONIAN THEORY OF GUIDING CENTER MOTION

Robert Grayson Littlejohn
(Ph.D. thesis)

May 1980

TWO-WEEK LOAN COPY

*This is a Library Circulating Copy
which may be borrowed for two weeks.
For a personal retention copy, call
Tech. Info. Division, Ext. 6782.*



LBL-12942
c.2 repl.

LEGAL NOTICE

This book was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States

Lawrence Berkeley Laboratory Library
University of California, Berkeley

Lawrence Berkeley Laboratory is an equal opportunity employer.

HAMILTONIAN THEORY OF GUIDING CENTER MOTION

Robert Grayson Littlejohn

Ph.D. Thesis

May 1980

Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

This work was supported by the Director, Office of Energy Research, Office of Fusion Energy, Division of Applied Plasma Physics of the U.S. Department of Energy under Contract No. W-7405-ENG-48.

Hamiltonian Theory of Guiding Center Motion

Robert Grayson Littlejohn

Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

ABSTRACT

With problems involving inhomogeneous magnetic fields in plasma physics it is common to call upon the guiding center approximation. Unfortunately, this approximation can lead to lengthy calculations, especially when carried beyond lowest order. Hamiltonian methods have proved to be an effective way of organizing and condensing such long calculations for many other perturbation problems in plasma physics, both in regard to single particle motion and collective effects. It has hitherto been difficult to apply similar Hamiltonian methods for the guiding center approximation because the demand for canonical variables in Hamiltonian mechanics has seemed to force one to use field line coordinates. These coordinates

are often not convenient in practical applications, and most research in the past has been carried out with non-Hamiltonian methods in rectangular coordinates.

This work presents a theory of guiding center motion which circumvents these problems by using noncanonical coordinates in phase space and rectangular coordinates in physical space. Methods are developed for carrying out Hamiltonian perturbation theory in noncanonical coordinates, and these preserve all the computational advantages of the more traditional Hamiltonian methods. Close attention is paid to the Poisson brackets of the coordinates among themselves. Darboux's theorem is used to create a set of coordinates which satisfy certain Poisson bracket relations and which simultaneously have a dynamical significance in the perturbation expansion. Lie transforms are used to carry out the perturbation expansion itself.

The theory is applied to the motion of a nonrelativistic particle moving in a magnetostatic field. The guiding center variables are the following. $\underline{X}$ is the position of the guiding center in rectangular coordinates; U is the parallel velocity of the guiding center; θ is the gyrophase; and M is the magnetic moment. Although these variables are well-known in guiding center theory and have simple physical interpretation, nevertheless they were not selected on that basis. Instead they emerge as a natural consequence of the transformation theory, as the solutions to a set of Hamiltonian differential equations. The guiding center variables are represented as formal power series in the adiabatic parameter ϵ of functions of the particle's position $\underline{x}$ and velocity $\underline{v}$. The guiding center variables exactly satisfy (i.e. to all orders) the following Poisson bracket relations:

$$\{\theta, M\} = 1/\epsilon$$

$$\{\theta, \underline{X}\} = \{\theta, U\} = \{M, \underline{X}\} = \{M, U\} = 0$$

$$\{\underline{X}, U\} = \underline{\Omega}^*/\Omega^*$$

$$\{X_i, X_j\} = \epsilon e_{ikj} \hat{b}_k / \Omega^*$$

where e_{ikj} is the Levi-Civita symbol, $\hat{b}$ is the unit vector along the magnetic field $\underline{B}$, $\underline{\Omega}^*$ is the modified gyrofrequency vector $e\underline{B}/mc + \epsilon UV \times \hat{b}$, and where $\Omega^* = \hat{b} \cdot \underline{\Omega}^*$.

Successive applications of the underlying transformation theory, including Darboux's theorem and Lie transforms, produce a Hamiltonian formulation of guiding center bounce and drift motions. Averaged equations of motion are derived as well as the series for the adiabatic invariants through (ϵ) .

DEDICATION

The completion of the Ph. D. dissertation is more than the end of a large task. It is a turning point, a milestone, marking the end of one phase of one's life and the beginning of another. For me it is a time for reflection on seven years of involvement with physics, the kind of involvement that is so intense that the boundary can no longer be distinguished between the object of one's work and the personality which drives it.

When I first plunged into this passionate and often disorienting world of physics, I found a friend with whom to share all my experience, intellectual and personal. We were of the same age, both of us had been out of school for a number of years, had never been married, had known the same haunts in our lives before Berkeley. We understood one another instinctively, without explanation. The generosity of his concern and respect for me never faltered.

Tom Kommers died shortly after his thirtieth birthday. I still cannot understand why it happened, even in terms of my own mortality. But here and now I would like to pause and dedicate to him, not this dissertation or this piece of scientific work, but rather this period of my life. I cannot think about it without thinking about him.

ACKNOWLEDGMENTS

On many accounts I would like to thank my thesis adviser, Allan Kaufman, but most especially for having created the working environment that I have enjoyed for the last three years. It is hard to imagine any other research group in which I could have enjoyed as much freedom for individual initiative and creativity as I have found here. This work would have been impossible without Prof. Kaufman's faith in the value of basic research and the guidance of his instincts for the important questions. His interests have in many ways become my own, and in one form or another I expect to be working on them for many years to come.

Of all the graduate students I have worked with, I would especially like to thank John Cary, from whom I learned more than I can easily reckon. John's intelligence and good common sense were guidance, support and inspiration to me on many occasions. In addition, I would like to thank Steve McDonald, Celso Gregobi, Harry Mynick, and Jim Hammer for many enjoyable discussions and exchanges about physics and other matters.

Most of all, I would like to thank my mother, Mary T Littlejohn, for having taught me that intellectual activity is an ennobling human endeavor in its own right, one which needs no moral justification beyond the spiritual satisfaction it brings.

TABLE OF CONTENTS

I. Introduction	1
1. The revolution in classical mechanics	2
2. The development of the method of the Darboux transformation	4
3. Synopsis of the thesis	10
4. References	13
II. Darboux's theorem and two-dimensional guiding center motion	14
1. Introduction	15
2. A covariant formulation of Hamiltonian mechanics	20
3. Darboux's theorem	29
4. Application of Darboux's theorem to the guiding center problem	38
5. The guiding center Hamiltonian	55
6. Discussion and conclusions	60
7. Acknowledgements	65
8. References	66
9. Figure captions	69
10. Figures	70
III. Three-dimensional guiding center motion	75
1. Introduction	77
2. Physical particle variables	82
3. Properties of the unit vectors	89
4. The Darboux transformation: formal properties	97
5. The Darboux transformation: explicit expressions	107

6. The guiding center Hamiltonian	115
7. Conclusions	133
8. Acknowledgements	136
9. Appendix A	137
10. References	141
11. Tables	143
12. Figure captions	144
13. Figures	145
IV. Guiding center bounce motion	149
1. Introduction	151
2. The Kruskal-Darboux-Lie method for nearly periodic systems	154
3. The preparatory transformations	161
4. The Darboux transformation	177
5. The averaging transformation	194
6. Conclusions	203
7. Acknowledgements	205
8. References	206
9. Tables	207
10. Figure captions	210
11. Figures	211
V. Concluding remarks	215
1. Improvements on the Darboux transformation method	216
2. Extensions of the Darboux transformation method	219
3. Applications of the guiding center Hamiltonian	222

CHAPTER I

INTRODUCTION

1. THE REVOLUTION IN CLASSICAL MECHANICS

Classical mechanics, the mainstay of theoretical physics in the nineteenth century, has lain dormant for most of the first half of this century. During this time the subject seemed to have solidified, and textbooks were written with a perspective based on the old research problems of the nineteenth century, especially celestial mechanics. These books became a part of every physicist's educational experience, and the prevalent sentiment that classical mechanics was a finished subject has no doubt been enhanced by this old perspective.

The last twenty years have witnessed a remarkable revival of interest in classical mechanics under the stimulus of a whole host of new problems. Among these are the practical problems of particle accelerators and controlled fusion devices, as well as theoretical problems in statistical mechanics, plasma physics and quantum mechanics. Under the impetus of the space program, even celestial mechanics has been revived.

In the last few years it has been discovered, to name but a few examples, that unstable motion is generic, even in systems with a small number of degrees of freedom (the Henon-Heiles system); that stability does not imply integrability (the KAM theorem); and that there are hitherto unsuspected dimensions to integrability (the Korteweg-de Vries equation).

The old and established methods of classical mechanics, on which it was thought for so long that the last word had been written, proved to be far from adequate to handle these new problems. As a result the whole subject is in a state of revolution as new methods are developed to handle new problems, and entirely new questions, as yet unresolved,

have moved to the forefront of research activity. Undoubtedly there will come a time for assimilation, when sufficiently large areas of research will have quieted down and the textbooks can be rewritten, but that time is not yet.

One aspect is already clear, however, about the modern developments in classical mechanics, and that is the predominantly mathematical point of view which permeates the new research. On the one hand, it is perhaps not surprising that the old mathematical methods, based on the intuitive but logically inadequate concept of infinitesimal, should give way to the differential-topological methods of today. But this change is more than simply a matter of tidying up a few logical shortcomings. One's very ability to conceptualize and to articulate the right questions, if not to find the right answers, is enhanced by the global, topological point of view of modern mathematics.

The three papers which form the main body of this thesis concern the guiding center problem, i.e. the perturbative solution of the equations of motion for a charged particle in a given electromagnetic field. Stated in these terms, the goal is quite traditional, and would have been appreciated by any physicist of the nineteenth century. There is no consideration given here to questions of a global nature, such as integrability or stability, and the whole perturbative approach is based on an unquestioned assumption that it can be useful to expand decidedly nonanalytic functions in power series. Nevertheless, this work does form a small part of the revolution in classical mechanics, for the simple reason that the traditional Hamilton-Jacobi methods of classical mechanics do not

work for the guiding center problem.

In working out the perturbation methods to be presented here, one cannot help but be impressed with the power of thinking in terms of the calculus of Cartan on differentiable manifolds, even though the three papers presented herein are written in terms of the index calculus familiar to the largest number of physicists. This was an unfortunate necessity; in fact, the very effort required to transcribe the ideas involved into the more familiar language is convincing proof of the superiority of the abstract, coordinate-free point of view.

From the standpoint of classical mechanics, an important aspect of this work is that it shows that a simple description of the dynamics of a mechanical system, i.e. a convenient set of coordinates, is not always easily compatible with a simple description of the symplectic structure on phase space. That it was thought otherwise in the past is due to the dominance of the perspective based on celestial mechanics, where the transformation $p=mv$ seems a small price to pay to make both the dynamics and the symplectic structure look simple. As a result, even an awareness of the symplectic structure of phase space has in the past been subdued.

2. THE DEVELOPMENT OF THE METHOD OF THE DARBOUX TRANSFORMATION

It may be helpful to the reader in following the remainder of this thesis if I give a brief history of the considerations which led to the ideas to be presented here.

The original motivation for looking at the guiding center problem was to find a simple, well-known, and physically interesting example

of a perturbation problem in classical mechanics on which to try the Poincaré-von Zeipel method, which I had just learned, for treating perturbations in Hamiltonian systems. (The Poincaré-von Zeipel method is the standard method of classical Hamiltonian perturbation theory,¹ in which undesirable terms in the Hamiltonian, usually the oscillatory ones, are transformed away by means of successive canonical transformations.) At first I did not have the slightest idea that the guiding center problem would be anything more than a straightforward application of this method. But I quickly discovered that things were not going to be so easy.

The reason is quite simple. All the books which discuss Hamiltonian perturbation theory do so in the context of a Hamiltonian of the form

$$H(\underline{q}, \underline{p}) = H_0(\underline{q}, \underline{p}) + \epsilon H_1(\underline{q}, \underline{p}) \quad (1)$$

Historically this model is taken from celestial mechanics, and, indeed, it is of common occurrence in other applications as well. But it is not universal. Consider the Hamiltonian for a charged particle in a static magnetic field:

$$H(\underline{q}, \underline{p}) = \frac{1}{2m} \left[\underline{p} - \frac{e}{c} \underline{A}(\underline{q}) \right]^2 \quad (2)$$

The physical meaning of the guiding center approximation, in terms of the ratio of the gyroradius to the scale length of the magnetic field, is well-known, but it is not clear how this approximation is to be built into the Hamiltonian (2). The Hamiltonian (2) simply does not look at all like the perturbation paradigm shown in (1).

The two leading ideas for introducing the guiding center

approximation were these. According to one idea, one replaces $\underline{q}$ with $\epsilon \underline{q}$, to indicate that the field is slowly varying. But once this is done, one does not know what to do next. In the other idea, one expands in powers of $1/\epsilon$, or equivalently, replaces e by e/ϵ and expands in ϵ . A curious aspect of this procedure is that it gives a Hamiltonian which appears to have a leading term which is $O(\epsilon^{-2})$:

$$H(\underline{q}, \underline{p}) = \frac{e^2}{2\epsilon^2 mc^2} \underline{A}^2 - \frac{e}{\epsilon mc} \underline{A} \cdot \underline{p} + \frac{1}{2m} \underline{p}^2 \quad (3)$$

The leading term is integrable, because it depends only on $\underline{q}$, but it does not give periodic orbits, as one expects for the guiding center problem. It is also not gauge invariant, and so it seemed that (3) was wrong. As it finally turned out, (3) is correct, but the ordering is not as indicated, since $\underline{p}$ itself is $O(\epsilon^{-1})$. Thus, the Hamiltonian (2) consists of two $O(\epsilon^{-1})$ terms in the parentheses which nearly cancel one another, leaving an $O(1)$ result for H . But even if I had recognized this fact at the time, it would still not have been clear how to proceed.

In the meantime it seemed prudent to find out how the guiding center expansion was known at all. This turned out to be quite a rich subject; there are at least four distinguishable methods for deriving the guiding center expansion which were published before 1967.

One method is based on a physical picture, a little algebra, and some intuitive notions about ordering and averaging. This is the method found in the plasma physics textbooks,² and also in Northrop's³ book. Perhaps the clearest and simplest application of this method was made by Baños. But this method did not satisfy me,

because it is not systematic to any order, and because it gave no insight into the Hamiltonian (2).

Another method is based on Kruskal's ansatz,⁵ which is a kind of WKB approximation. This method is not Hamiltonian, but it is systematic to any order. A discussion of it may also be found in Northrop's³ book. This method has in recent years been used by Northrop and Rome⁶ to carry the guiding center expansion to one order beyond the classic, well-known drift formulas. By their own account, the amount of algebra was brutal. But this approach also did not satisfy me, because it gave no insight into the failure of the Poincaré-von Zeipel method on the Hamiltonian (2).

A third method was both systematic to any order and Hamiltonian. This is the method of Gardner,⁷ which begins by transforming the Hamiltonian (2) to field line coordinates, and follows with a sequence of canonical transformations. This method is certainly logically complete, but it failed to satisfy me for a number of reasons. In the first place, it was not clear what connection this method had to the Poincaré-von Zeipel method, and it did not explain the failure of the latter on the guiding center problem. Second, the method seemed to involve more labor than one would expect on the basis of experience with the Poincaré-von Zeipel method. And third, the use of field line coordinates seemed a strange artifice, and a rather high price to pay to use Hamiltonian mechanics. The use of field line coordinates entails curvilinear coordinate systems in physical space, and these lead to such things as the metric tensor and covariant derivatives. It seemed that such complications should really not be necessary in Euclidean space, especially when non-Hamiltonian treatments of guiding

center motion can be carried out in rectangular coordinates.

The fourth method proved to be the most interesting of all to me, even though it is non-Hamiltonian. This method was called "the method of rapidly rotating phase" by Bogoliubov and Mitropolski,⁸ who seem to be its inventors. It is also often called "the method of averaging." This method is applicable to systems of ordinary differential equations whose solutions are nearly periodic, which includes the guiding center problem, but which, most interestingly, also includes other systems, some of which can be analyzed by the method of Poincaré-von Zeipel. Furthermore, the procedure one goes through in executing the method of averaging has a parallel, step for step, in the method of Poincaré-von Zeipel. The method of averaging was carefully analyzed in an important paper by Kruskal,⁹ in which he showed how adiabatic invariants could be extracted once the perturbation expansions themselves were carried out with non-Hamiltonian means. Thus, a partial link had already been established between the method of averaging and traditional Hamiltonian methods, and there were suggestions of other links as well.

It seemed that by paying close attention to the method of averaging as applied to the guiding center problem, one might be able to transcribe each step taken in the non-Hamiltonian procedure into an equivalent step in a Hamiltonian analysis. But this idea ran into trouble on the very first step. In the non-Hamiltonian method of averaging, one begins by introducing the instantaneous gyrophase θ . This is done in a preparatory transformation, the purpose of which is to make the unperturbed system look simple. In a textbook example of the Poincaré-von Zeipel method, one does something very similar, which is to introduce action/angle variables for the unperturbed system.

But since the Hamiltonian (2) does not neatly fall into an unperturbed part and a perturbation, it was not clear what to do.

This much was clear, however: it was the restriction to canonical transformations in Hamiltonian mechanics which was causing the difficulty. For indeed, if arbitrary transformations were allowed in Hamiltonian mechanics, then one could use the same preparatory transformation for a Hamiltonian analysis as for a non-Hamiltonian analysis. This line of reasoning led to a realization that it is either an unexplained phenomenon of some significance, or else a minor miracle, that the restricted class of canonical transformations are so often useful in practice. It also caused an anxiety that a canonical transformation which would satisfy the preparatory desiderata for the guiding center problem might not even exist.

What followed was a long and frustrating search for the right canonical transformation, one which would prepare the Hamiltonian (2) for a perturbation analysis. After much work the inadequacy of mixed variable generating functions for the task at hand became apparent. Considerably more work showed a similar inadequacy of Lie transforms, which are extremely awkward for expressing finite canonical transformations.

It was an act of desperation to pose the following question: What is the most general form of a canonical transformation $(q,p) \rightarrow (Q,P)$ in which one of the new variables is the instantaneous gyrophase θ ? This seemed to be a question of last resort, because the Poisson bracket relations in which one would formulate this question give a system of nonlinear, over-determined partial differential equations.

Nevertheless, it was immediately apparent that the question was

an interesting one, because it quickly produced the formulas shown in Eqs. (4.20) and (4.28) of Chapter II. Verifying that these formulas satisfied the nonlinear equations proved to be a more formidable task, because the demonstration involved infinite series of noncommuting linear operators. In the case of the relation $\{J, X\}=0$ (Eq. (4.13) of Chapter II), I was unable at first to prove the desired result to all orders, but a brute force substitution and calculation produced a proof of the result through fourth order, in a miraculous mutual annihilation of scores of terms. It was only later that the proof based on the commutativity of Hamiltonian flows, which forms the essence of Darboux's theorem, became apparent.

The rest of the theory fell quickly into place, and it is adequately explained in the papers which follow.

3. SYNOPSIS OF THE THESIS

The three research papers comprising the main body of this thesis concern the application of Darboux's theorem in various aspects of guiding center theory. Of the three, the first is the most important, since it gives an exposition of the new methods, whereas the other two papers simply apply the methods to cases of practical interest.

The first paper, entitled "A guiding center Hamiltonian: A new approach," comprises Chapter II of this thesis. This paper divides roughly into two parts. The first part, consisting of Secs. 1-3, presents a covariant formulation of Hamiltonian mechanics using the index calculus familiar to physicists. Included in this part is a discussion of Darboux's theorem in the context in which it has been understood in the past, i.e. concerning the existence of canonical

coordinates. The second part of the paper, consisting of Secs. 4-6, applies Darboux's theorem to guiding center motion. As a model to illustrate the method, particle motion in the two-dimensional magnetic field $\underline{B} = B(x,y)\hat{z}$ is studied. In the application of Darboux's theorem not just any canonical (or semicanonical) coordinate system is sought, but rather one which has a dynamical significance, based on physical considerations of the unperturbed motion. One result is a fascinating interplay between the dynamics and the symplectic structure. After the Darboux transformation has been carried out, the Hamiltonian is subjected to an averaging transformation which is effected by Lie transforms. One interesting aspect of this procedure is the use of scalar Lie generating functions with noncanonical variables.

The second paper, entitled "Hamiltonian formulation of guiding center motion," comprises Chapter III. This paper is a straightforward application of the Hamiltonian methods of the first paper to guiding center motion in three-dimensional magnetic fields $\underline{B} = B(\underline{x})$, with $\underline{E} = 0$. Although there is little that is new here from the standpoint of Hamiltonian mechanics, the three-dimensional guiding center problem is a substantial exercise in perturbation theory, even with the new methods. In this paper the drift equations are carried to one order beyond their classic, well-known forms. It turns out that if one is willing to have the guiding center position depend on the perpendicular unit vectors (these define the origin of gyrophase), then the $O(\epsilon^2)$ correction to the perpendicular drift equations takes on a remarkably simple form (see Eq. (6.55) of Chapter III). Perhaps the most striking result of this paper is the set of

Poisson bracket relations in Eqs. (4.1)-(4.5) and (4.31)-(4.32).

It is an interesting consequence of the theory underlying the Darboux transformation that these Poisson bracket relations are exact expressions, good to all orders, rather than truncated power series.

The third paper, entitled "Hamiltonian theory of guiding center bounce motion," comprises Chapter IV of the thesis. Although this paper, like the previous one, is an application of the Darboux transformation to a practical problem, it includes in addition some improvements on the Hamiltonian methods of Chapter II. In particular, it is found that the Lagrange tensor is simpler to deal with, for many purposes, than the Poisson tensor. In this paper a treatment of the bounce motion is given which parallels that given for the gyromotion in the previous paper. The averaged Hamiltonian for the bounce motion is derived, and this gives in turn the averaged equations of motion. In addition, the adiabatic invariant series is carried through $O(\epsilon)$, and Poisson bracket relations are derived for the averaged variables. Like the Poisson brackets derived in the previous paper, these are also exact.

Finally, in Chapter V, I suggest ways to improve on the methods which have been presented, and I discuss applications and directions for further research.

REFERENCES

1. A. H. Nayfeh, Perturbation Methods (Wiley, New York, 1973).
2. F. F. Chen, Introduction to Plasma Physics (Plenum, New York, 1974).
3. T. G. Northrop, The Adiabatic Motion of Charged Particles (Interscience, New York, 1963).
4. A. Baños, J. Plasma Phys. 1, 305(1967).
5. M. D. Kruskal, "Elementary Orbit and Drift Theory," in Plasma Physics (International Atomic Energy Agency, Vienna, 1965).
6. T. G. Northrop and J. A. Rome, Phys. Fluids 21, 384(1978).
7. C. S. Gardner, Phys. Rev. 115, 791(1959).
8. N. N. Bogoliubov and Y. A. Mitropolski, Asymptotic Methods in the Theory of Non-Linear Oscillations (Gordon and Breach, New York, 1961).
9. M. D. Kruskal, J. Math. Phys. 3, 806(1962).

CHAPTER II

DARBOUX'S THEOREM AND TWO-DIMENSIONAL

GUIDING CENTER MOTION

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 $\underline{B} = B(x,y)\hat{z}$ 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 formulation of the guiding center problem, an approach which leads to a remarkably deep insight into the formal structure of classical Hamiltonian mechanics. 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 studied by mathematicians.

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

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 and Marsden,¹ Vinogradov and Kupersmidt,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 applications of interest in physics, even quite theoretical ones, a description of Hamiltonian mechanics which focuses on the differential geometry of phase space may be deemed to be unnecessarily 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 physics 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 controversy 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 invariants 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 Hamiltonian 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 Sec. 4 below, cannot be easily

analyzed by the standard methods of Poincaré. The reason is that the relation between the canonical momentum $\underline{p}$ and the physical variables $\underline{x}$ and $\underline{y}$ describing the motion of the particle involves the use of the magnetic vector potential $\underline{A}$. 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 $\underline{p}$ from $\underline{x}$ and $\underline{y}$ 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 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 notational 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 to 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 $\underline{B}(\underline{x}) = B(x,y)\hat{z}$. 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 $\underline{z}$ or $\underline{z}^i$, 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 canonical coordinates, and refer to a canonical coordinate system. In this section, when we refer to canonical coordinates we will decompose the $2N$ coordinates $\underline{z}^i$ into q 's and p 's as follows:

$$\underline{z}^i = (q_1, \dots, q_N, p_1, \dots, p_N) \quad (2.1)$$

Canonical coordinates are to be regarded as a special case, and unless we state the contrary, the coordinates $\underline{z}^i$ 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 $\underline{z}$ represents a set of canonical coordinates, and if $\bar{\underline{z}}$ represents a set of $2N$ independent functions of $\underline{z}$, then $\bar{\underline{z}}$ 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

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

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

$$\gamma = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \quad (2.3)$$

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

$$\bar{\omega}_{ij} = \frac{\partial z^k}{\partial \bar{z}^i} \gamma_{k\ell} \frac{\partial z^\ell}{\partial \bar{z}^j} \quad (2.4)$$

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

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

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

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

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 $\bar{z}$ is arbitrary, i.e. not necessarily canonical, or when no distinction need be 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 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,

$$\omega_{ij} \sigma^{kj} = \delta_i^k \quad (2.7)$$

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

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

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:

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

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

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

The 2-form ω is nondegenerate, meaning

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

It is also closed, meaning $d\omega = 0$, or

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

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

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

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0 \quad (2.13)$$

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

$$\frac{\partial \omega_{ij}}{\partial t} = 0 \quad (2.14)$$

From a practical point of view, this means that most time-dependent transformations $\underline{z} = \underline{z}(\underline{q}, \underline{p}, t)$, taking us from a canonical coordinate system to an arbitrary system, must be excluded. Time-dependent canonical transformations 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 variables 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 $\underline{B}(\underline{x})$. The usual canonical coordinates $(\underline{q}, \underline{p})$ for the phase space of this system are given in terms of the particle's position $\underline{x}$ and velocity $\underline{v}$ by

$$\begin{aligned} \underline{q} &= \underline{x} \\ \underline{p} &= m\underline{v} + \frac{e}{c} \underline{A}(\underline{x}) \end{aligned} \quad (2.15)$$

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

$$\{x_i, x_j\} = 0 \quad (2.16a)$$

$$\{x_i, v_j\} = -\{v_i, x_j\} = \frac{1}{m} \delta_{ij} \quad (2.16b)$$

$$\{v_i, v_j\} = \frac{e}{m c} B_{ij} \quad (2.16c)$$

where

$$B_{ij} = \epsilon_{ijk} B_k \quad (2.17)$$

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

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

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

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

Observe that the closedness of ω implies the Maxwell equation $\nabla \cdot \underline{B} = 0$.

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

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

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 $(\underline{x}, \underline{v})$ coordinates used in Eqs. (2.15)-(2.19).

The Hamiltonian in the $(\underline{q}, \underline{p})$ coordinates is

$$H(\underline{q}, \underline{p}) = \frac{1}{2m} \left(\underline{p} - \frac{e}{c} \underline{A}(\underline{q}) \right)^2 \quad (2.21)$$

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

$$H(\underline{x}, \underline{v}) = \frac{1}{2} m \underline{v}^2 \quad (2.22)$$

Then the equations of motion are

$$\frac{d}{dt} \begin{pmatrix} \underline{x} \\ \underline{v} \end{pmatrix} = \frac{1}{m} \begin{pmatrix} 0 & I \\ -I & \frac{e}{mc} \underline{B} \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial H}{\partial \underline{x}} \\ \frac{\partial H}{\partial \underline{v}} \end{pmatrix} = \begin{pmatrix} \underline{v} \\ \frac{e}{mc} \underline{v} \times \underline{B} \end{pmatrix} \quad (2.23)$$

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

Let us now return to Hamilton's equations of motion and replace the parameter t , describing the trajectories in phase space, by 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 $\underline{S}(\underline{z}_0, \lambda)$ be the solution to Hamilton's equations which satisfies $\underline{z} = \underline{z}_0$ at $\lambda = 0$. That is, $\underline{S}(\underline{z}_0, \lambda)$ satisfies

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

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

$$\underline{S}(\underline{S}(\underline{z}_0, \lambda_1), \lambda_2) = \underline{S}(\underline{z}_0, \lambda_1 + \lambda_2) \quad (2.25)$$

for all $\underline{z}_0, \lambda_1, \lambda_2$. This is an elementary result from the theory of ordinary differential equations,²⁵ and it gives rise to an interpretation of the solution $\underline{S}$ 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 $\underline{z}$, 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 $\underline{z} \rightarrow \bar{\underline{z}}$ with $\bar{\underline{z}} = \underline{S}(\underline{z}, \lambda)$; 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

$$(T(\lambda)f)(z) = f(S(z, \lambda)) \quad (2.26)$$

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

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

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

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

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

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

with solution

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

It is well-known that the solutions of 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 transformations are the proper generalizations 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 a number of 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 $\underline{z}$, with respect to which ω has components $\omega_{ij}(\underline{z})$, which are to be regarded as definite functions of $\underline{z}$. Under an arbitrary change of coordinates $\underline{z} \rightarrow \bar{\underline{z}}$ the components of ω go into $\bar{\omega}_{ij}(\bar{\underline{z}})$, which we consider to be functions of the new coordinates $\bar{\underline{z}}$, according to the usual rule for covariant tensors:

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

However, if the transformation $\underline{z} \rightarrow \bar{\underline{z}}$ is a symplectic transformation, then the invariance of ω means $\bar{\omega}_{ij}(\bar{\underline{z}}) = \omega_{ij}(\bar{\underline{z}})$, for all $\bar{\underline{z}}$. 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 approach 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 $\underline{z}$ of canonical coordinates into q 's and p 's in the following order:

$$\underline{z} = (q_1, p_1, \dots, q_N, p_N) \quad (3.1)$$

Corresponding to this ordering, the matrix γ has the form

$$\gamma_{ij} = \begin{pmatrix} \begin{array}{cc|cc} 0 & 1 & & \\ -1 & 0 & & \\ \hline & & 0 & 1 \\ & & -1 & 0 \end{array} & & 0 \\ & \ddots & \\ 0 & & \begin{array}{cc|cc} & & 0 & 1 \\ & & -1 & 0 \end{array} \end{pmatrix} \quad (3.2)$$

This ordering differs from that used in Sec. 2.

We shall denote phase space by ϕ , representing a $2N$ dimensional manifold. The construction of canonical coordinates 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, nondegenerate 2-form ω on ϕ and a coordinate system $\underline{z}$ with respect to which ω has components ω_{ij} . Then there exists a coordinate transformation $\underline{z} \rightarrow \bar{\underline{z}}$ 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 $\underline{z}$, can be chosen at will.

We remark that if the original coordinate system $\underline{z}$ is canonical itself, then the constructive proof of Darboux's theorem gives a method of determining a canonical transformation $\underline{z} \rightarrow \bar{\underline{z}}$ in which one of the new coordinates $\bar{z}^i(\underline{z})$ 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 $\underline{z} \rightarrow \bar{\underline{z}}$ such that the components $\bar{\omega}_{ij}$ of ω with respect to the new coordinates $\bar{\underline{z}}$ have the form

$$\bar{\omega}_{ij} = \begin{pmatrix} \Omega_{ij} & 0 \\ 0 & \begin{array}{cc|cc} 0 & 1 \\ -1 & 0 \end{array} \end{pmatrix} \quad (3.3)$$

where Ω_{ij} represents a $(2N-2) \times (2N-2)$ matrix. 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:

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

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 $(2N-2) \times (2N-2)$ 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

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

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

$$\frac{\partial \Omega_{ij}}{\partial q} = \frac{\partial \Omega_{ij}}{\partial p} = 0 \quad (3.6)$$

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

$$\frac{\partial \Omega_{ij}}{\partial Z^k} + \frac{\partial \Omega_{jk}}{\partial Z^i} + \frac{\partial \Omega_{ki}}{\partial Z^j} = 0 \quad (3.7)$$

In Eqs. (3.6)-(3.7), the indices i,j,k run over the numbers $1, \dots, 2N-2$, corresponding to the coordinates $\underline{z}$.

The result of these corollaries is that the quantities Ω_{ij} are the components with respect to the coordinate system $\underline{z}$ 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 $\underline{z} \rightarrow \bar{\underline{z}}$, taking the components Ω_{ij} into $\bar{\Omega}_{ij}$, 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 made. 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 $\bar{\underline{z}} = (\underline{z}, q, p)$ the Darboux algorithm.

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 $\underline{z} \rightarrow \bar{\underline{z}}$, the components $\bar{\sigma}^{ij}$ of the σ tensor with respect to the new coordinates $\bar{\underline{z}}$ are the Poisson Brackets of the new coordinates among themselves. With the definition $\bar{\underline{z}} = (\underline{z}, q, p)$, we demand the following form for these Poisson brackets:

$$\{q, p\} = 1 \quad (3.8)$$

$$\{Z^i, q\} = 0 \quad (3.9)$$

$$\{Z^i, p\} = 0 \quad (3.10)$$

$$\{Z^i, Z^j\} = \Sigma^{ij} \quad (3.11)$$

The precise form of the quantities Σ^{ij} immaterial for the purpose of proving the lemma, although these quantities will automatically be the components of a $(2N-2) \times (2N-2)$, antisymmetric, invertible matrix, since the form of $\bar{\sigma}^{ij}$ is given by

$$\bar{\sigma}^{ij} = \left(\begin{array}{cc|cc} \Sigma^{ij} & & & 0 \\ & & & \\ \hline & & 0 & 1 \\ & & -1 & 0 \end{array} \right) \quad (3.12)$$

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 $q(\underline{z})$ on ϕ for one of the new coordinates; the other $2N-1$ functions, $p(\underline{z})$ and $Z^i(\underline{z})$, will then be constrained by Eqs. (3.8)-(3.10). In terms of the given function $q(\underline{z})$, Eq. (3.8) is a first-order, linear inhomogeneous partial differential equation for the unknown function $p(\underline{z})$. 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 $\underline{z} = \underline{z}(\lambda)$ which are the solutions to the following set of ordinary differential equations:

$$\frac{dz^i}{d\lambda} = \{z^i, q\} \quad (3.13)$$

These characteristics are the trajectories which result upon treating $q(\underline{z})$ 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

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

A picture of the solution $p(\underline{z})$ 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 $p(\underline{z})$, we choose a $2N-1$ dimensional manifold P_0 , cutting all the Q surfaces. P_0 is arbitrary, except that it must be nowhere tangent to any Q surface, since that would result in $dq \wedge dp = 0$ 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 $p(\underline{z})$; for example, it is convenient to take $p(\underline{z})=0$ for $\underline{z} \in P_0$. For $\underline{z} \notin P_0$, $p(\underline{z})$ is defined as the negative of the elapsed λ parameter, relative to P_0 , of the q -characteristic passing through $\underline{z}$. From Eq. (3.14) it then follows that

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

and Eq. (3.8) is satisfied.

Next we want to solve Eq. (3.9) for $2N-2$ function $Z^i(\underline{z})$ which are independent of each other and also of q and p . Considering q as given and p and $\underline{z}$ 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:

$$\frac{dZ^i}{d\lambda} = 0 \quad (3.16)$$

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

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

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

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

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) and (3.8)-(3.9):

$$\frac{d}{d\lambda} \{Z^i, p\} = \{\{Z^i, p\}, q\} = \{\{q, p\}, Z^i\} + \{\{Z^i, q\}, p\} = 0 \quad (3.19)$$

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 closedness of ω , 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 equations 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 is 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 $\epsilon \rightarrow 0$, it being understood that the particle variables $\underline{x}$ and $\underline{v}$ and the fields $\underline{A}$ and $\underline{B}$ are to be held fixed in this limiting process. For example, the gyroradius $m v_{\perp} c / e B$ is $O(\epsilon)$, and the gyrofrequency $e B / m c$ is $O(\epsilon^{-1})$. The physical meaning of the limit $\epsilon \rightarrow 0$ 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

$$H(\underline{q}, \underline{p}) = \frac{1}{2} (\underline{p} - \frac{1}{\epsilon} \underline{A}(\underline{q}))^2 \quad (4.1)$$

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

$$\begin{aligned} \underline{x} &= \underline{q} \\ \underline{v} &= \underline{p} - \frac{1}{\epsilon} \underline{A}(\underline{q}) \end{aligned} \quad (4.2)$$

The third step is to restrict consideration to magnetic fields of the form $\underline{B}(\underline{x}) = B(x, y) \hat{z}$, and furthermore to consider only particle motion in the x - y plane. The problem thereby becomes two-dimensional, and we write $\underline{x} = (x, y)$, $\underline{v} = (v_x, v_y)$, 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 (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 transformation $(q,p) \rightarrow (x,v)$, the Hamiltonian becomes

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

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 :

$$\{f,g\} = \frac{\partial f}{\partial x} \frac{\partial g}{\partial v} - \frac{\partial f}{\partial v} \frac{\partial g}{\partial x} + \frac{1}{\epsilon} B \cdot \left(\frac{\partial f}{\partial v} \times \frac{\partial g}{\partial v} \right) \quad (4.4)$$

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 $\epsilon \rightarrow 0$. 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 $\hat{b}$ along the magnetic field B . According to previous conventions, we have $\hat{b} = \hat{z}$. Next we define a unit vector $\hat{c}$ in the direction of the particle's velocity:

$$\underline{v} = \hat{c} \quad (4.5)$$

Finally, we define a unit vector $\hat{a}$ by $\hat{a} = \hat{b} \times \hat{c}$. Thus the triad $(\hat{a}, \hat{b}, \hat{c})$ forms a right-handed set. Note that for a uniform magnetic field $\hat{a}$ is in the direction of the gyroradius vector $\tilde{r}$, which is the displacement between the guiding center position $\tilde{X}$ and the position of the particle $\tilde{x}$:

$$\tilde{x} = \tilde{X} + \tilde{r} \quad (4.6)$$

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

$$\tilde{r} = \frac{\epsilon v}{B} \hat{a} \quad (4.7)$$

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

$$\begin{aligned} \hat{c} &= -\sin\theta \hat{x} - \cos\theta \hat{y} \\ \hat{a} &= \cos\theta \hat{x} - \sin\theta \hat{y} \\ \hat{b} &= \hat{z} \end{aligned} \quad (4.8)$$

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, and indicated here by the components of the σ tensor:

$$\begin{aligned} \{x_i, x_j\} &= 0 \\ \{x, v\} &= \hat{c} \\ \{x, \theta\} &= -\hat{a}/v \\ \{\theta, v\} &= \frac{B}{\epsilon v} \end{aligned} \quad (4.9)$$

4.3. The Darboux algorithm

The third coordinate transformation is not trivial, and requires some motivation. Consider a Hamiltonian $H(\underline{q}, \underline{p})$. A typical strategy in Hamiltonian perturbation theory is to find a canonical transformation $(\underline{q}, \underline{p}) \rightarrow (\bar{\underline{q}}, \bar{\underline{p}})$ such that the new Hamiltonian K is independent of one or more (perhaps all) of the new generalized coordinates $\bar{\underline{q}}$. 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 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 depends on $\underline{z}$. 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:

$$\begin{aligned}\frac{d\underline{x}}{dt} &= v\hat{c} \\ \frac{dv}{dt} &= 0 \\ \frac{d\theta}{dt} &= \frac{B(\underline{x})}{\epsilon}\end{aligned}\tag{4.10}$$

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 $(Z_1, \dots, Z_{2N-2}, q, p)$. 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 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 $\tilde{X} = (X, Y)$ corresponding to the $(2N-2)$ -vector $\tilde{Z}$. Actually, it is desirable to modify the form of Eqs. (3.8)-(3.10) slightly, and ask for solutions $J, \tilde{X}$ to the set

$$\{\theta, J\} = 1/\epsilon \quad (4.11)$$

$$\{\tilde{X}, \theta\} = 0 \quad (4.12)$$

$$\{\tilde{X}, J\} = 0 \quad (4.13)$$

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 $d/d\lambda = \{ \cdot, \theta \}$ and use Eq. (4.9) to get the following differential equations for the θ -characteristics:

$$\frac{dx}{d\lambda} = - \frac{\hat{a}}{v} \quad (4.14)$$

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

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

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

$$\frac{d\chi}{d\lambda} = 0 \quad (4.17)$$

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 surface $\theta = \text{constant}$ 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 variables is permissible, and it gives

$$\frac{dx}{dv} = \frac{\epsilon}{B} \hat{a} \quad (4.18)$$

$$\frac{dJ}{dv} = \frac{v}{B} \quad (4.19)$$

Although these equations depend upon the unspecified function $B(x)$ 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

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

where L is the Lie operator defined by

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

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

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

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 X

We proceed with the construction of a simultaneous solution to Eqs. (4.12)-(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 $d/d\mu = \{ \cdot, J \}$. 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

$$\frac{dx}{d\mu} = 0 \quad (4.23)$$

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

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

$$\underline{X}(\underline{x}, v=0, \theta) = \underline{x} \quad (4.25)$$

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

$$\{\underline{X}, J\} = 0 \quad (4.26)$$

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 $\underline{X}$ are now 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 $\underline{X}$ at any given phase point $\underline{z} = (\underline{x}, \theta, v)$ is found by following the θ -characteristic passing through $\underline{z}$ 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

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

Exactly as was done in Section 3, we can prove that $d/d\lambda \{\underline{X}, J\} = 0$,

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 $\tilde{X}$ given below.

At this point we find an explicit expression for the function $\tilde{X}(x, \theta, v)$. 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

$$\tilde{X}(x, \theta, v) = \exp(-\epsilon v L) \tilde{x} \quad (4.28)$$

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

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

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 $\tilde{X}$ 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 series²⁸ will be of use to us here. For example, Eq. (4.28) may be inverted to solve for $\tilde{x}$:

$$\tilde{x}(\tilde{X}, \theta, v) = \exp(+\epsilon v L) \tilde{X} \quad (4.30)$$

In this equation the Lie operator L is given by

$$L = \frac{1}{B(\underline{X})} \hat{a} \cdot \frac{\partial}{\partial \underline{X}} \quad (4.31)$$

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 $(\underline{x}, \theta, v) \rightarrow (\underline{X}, \theta, J)$, 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 $\underline{X}$, i.e. not on θ or J .

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

$$\frac{d}{d\lambda}\{X, Y\} = \frac{d}{d\mu}\{X, Y\} = 0 \quad (4.32)$$

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 $\underline{z} = (\underline{X}, \theta, J)$, and the result will be the same as at $\underline{z}$ itself. Clearly, the most convenient point to make such an evaluation is on V_0 .

In order to find $\{X, Y\}$ on V_0 is it necessary to compute $\{X, Y\}$ in the neighborhood of V_0 and then to let $v \rightarrow 0$. 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

$$\begin{aligned} X &= x - \frac{\epsilon v}{B} \cos \theta + O(v^2) \\ Y &= y + \frac{\epsilon v}{B} \sin \theta + O(v^2) \end{aligned} \quad (4.33)$$

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

$$\{X, Y\} = -\frac{\epsilon}{B(\underline{x})} + O(v) \quad (4.34)$$

But when we let $v \rightarrow 0$, $\underline{x}$ becomes identical with $\underline{X}$, and we obtain

$$\{X, Y\} = -\frac{\epsilon}{B(\underline{X})} \quad (4.35)$$

By the arguments above, this is valid at any point $(\underline{X}, \theta, J)$ of phase space. As predicted, $\{X, Y\}$ depends only on $\underline{X}$.

Altogether, in the coordinate system (X, Y, θ, J) the components of the matrix σ are

$$\sigma^{ij} = \begin{pmatrix} 0 & \frac{-\epsilon}{B(\tilde{X})} & 0 & 0 \\ \frac{+\epsilon}{B(\tilde{X})} & 0 & 0 & 0 \\ 0 & 0 & 0 & +\frac{1}{\epsilon} \\ 0 & 0 & -\frac{1}{\epsilon} & 0 \end{pmatrix} \quad (4.36)$$

That is, we can write the Poisson Bracket of two functions f and g in terms of the coordinates $(\tilde{X}, \theta, J)$ as follows:

$$\{f, g\} = \frac{\epsilon}{B(\tilde{X})} \left(\frac{\partial f}{\partial Y} \frac{\partial g}{\partial X} - \frac{\partial f}{\partial X} \frac{\partial g}{\partial Y} \right) + \frac{1}{\epsilon} \left(\frac{\partial f}{\partial \theta} \frac{\partial g}{\partial J} - \frac{\partial f}{\partial J} \frac{\partial g}{\partial \theta} \right) \quad (4.37)$$

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 $(X, Y, \theta, J) \rightarrow (Q, P, \theta, J)$, which would bring the σ tensor into the form $\sigma^{ij} = \gamma_{ij}/\epsilon$. Except for the factor $1/\epsilon$, which is a minor consideration, we would then have constructed, by means of a number of noncanonical intermediaries, an overall canonical transformation $(q_x, q_y, p_x, p_y) \rightarrow (Q, P, \theta, J)$. According to the theory in Section 3, the new coordinates Q and P would be functions of $\tilde{X}$ alone, and they would satisfy $\{Q, P\} = 1/\epsilon$.

The functions Q and P of $\tilde{X}$ which are produced by a second iteration of the Darboux algorithm cannot be constructed perturbatively, as were $\tilde{X}$ 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 :

$$Q(\underline{X}) = \beta(\underline{X})/\epsilon \quad (4.38)$$

$$P(\underline{X}) = \alpha(\underline{X})/\epsilon$$

The functions α and β satisfy

$$\nabla\alpha \times \nabla\beta = \underline{B} \quad (4.39)$$

which in our two-dimensional field configuration becomes

$$B(X,Y) = \frac{\partial\alpha}{\partial X} \frac{\partial\beta}{\partial Y} - \frac{\partial\alpha}{\partial Y} \frac{\partial\beta}{\partial X} \quad (4.40)$$

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 $\underline{X}$ 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 $\underline{A}$ is nonphysical. On the other hand, Eq. (4.37) shows that the Poisson Bracket in the $(\underline{X},\theta,J)$ 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 $(\underline{x},\theta,v) \rightarrow (\underline{X},\theta,J)$, 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 $(\underline{x},\theta,v)$. Using Eq. (4.30), J may be expressed as a function of $(\underline{X},\theta,v)$. In the process of eliminating $\underline{x}$ in favor of $\underline{X}$, there results a double infinite series involving the operator L . This can be collapsed back into a single series, yielding finally

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

Next, we invert this series to obtain v as a function of $(\underline{X},\theta,J)$.

Carried out through second order, this gives

$$\begin{aligned}
v(\underline{X}, \theta, J) = & (2BJ)^{1/2} + \epsilon \frac{(2BJ)}{3B^2} (\hat{a} \cdot \nabla B) \\
& + \epsilon^2 \frac{(2BJ)^{3/2}}{72 B^4} [9B(\hat{a}\hat{a} : \nabla \nabla B) - 7(\hat{a} \cdot \nabla B)^2] \quad (4.42)
\end{aligned}$$

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

$$\begin{aligned}
\underline{x}(\underline{X}, \theta, J) = & \underline{X} + \epsilon \frac{(2BJ)^{1/2}}{B} \hat{a} - \epsilon^2 \frac{(2BJ)}{6B^3} (\hat{a} \cdot \nabla B) \hat{a} \\
& + \epsilon^3 \frac{(2BJ)^{3/2}}{72 B^5} [-3B(\hat{a}\hat{a} : \nabla \nabla B) + 5(\hat{a} \cdot \nabla B)^2] \hat{a} \quad (4.43)
\end{aligned}$$

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

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

$$\begin{aligned}
H(\underline{X}, \theta, J) = & BJ + \epsilon \frac{(2BJ)^{3/2}}{3B^2} (\hat{a} \cdot \nabla B) \\
& + \epsilon^2 \frac{(2BJ)^2}{24B^4} [3B(\hat{a}\hat{a} : \nabla \nabla B) - (\hat{a} \cdot \nabla B)^2] + O(\epsilon^3) \quad (4.44)
\end{aligned}$$

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

$$L_n f = \epsilon \{w_n, f\} \quad (5.1)$$

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 $O(\epsilon^{-1})$.

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

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

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 :

$$T = \dots T_3 T_2 T_1 \quad (5.3)$$

$$T^{-1} = T_1^{-1} T_2^{-1} T_3^{-1} \dots \quad (5.4)$$

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 $O(\epsilon^2)$ part, according to Eq. (4.37). Therefore we define two more series of operators, as follows:

$$M_n f = \frac{\partial w_n}{\partial \theta} \frac{\partial f}{\partial J} - \frac{\partial w_n}{\partial J} \frac{\partial f}{\partial \theta} \quad (5.5)$$

and

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

so that

$$L_n = M_n + \epsilon^2 N_{n+2} \quad (5.7)$$

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 ,

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

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

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

$$\begin{aligned} \bar{z} &= Tz \\ z &= T^{-1}\bar{z} \end{aligned} \quad (5.10)$$

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

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

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

$$H = \sum_{n=0}^{\infty} \epsilon^n H_n \quad (5.13)$$

$$K = \sum_{n=0}^{\infty} \epsilon^n K_n \quad (5.14)$$

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

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

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

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

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

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

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

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

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

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

$$K(\tilde{X}, \tilde{J}) = B\tilde{J} + \epsilon^2 \frac{\tilde{J}^2}{4B^2} [B\nabla^2 B - 3(\nabla B)^2] + O(\epsilon^3) \quad (5.21)$$

where B means $B(\tilde{X})$ and where ∇ means $\partial/\partial \tilde{X}$.

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

$$\frac{d\tilde{X}}{dt} = \frac{\hat{e}b}{B} \times \left\{ \tilde{J} \nabla B + \epsilon^2 \frac{\tilde{J}^2}{4} \nabla \left[\frac{\nabla^2 B}{B} - \frac{3(\nabla B)^2}{B^2} \right] \right\} + O(\epsilon^5) \quad (5.22)$$

$$\frac{d\tilde{\theta}}{dt} = \frac{B}{\epsilon} + \epsilon \frac{\tilde{J}}{2B^2} [B\nabla^2 B - 3(\nabla B)^2] + O(\epsilon^3) \quad (5.23)$$

$$\frac{d\tilde{J}}{dt} = 0 \quad (5.24)$$

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

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

$$\tilde{X} = \tilde{X} + \frac{\epsilon^3}{B} \hat{b} \times \nabla \left[\frac{(2BJ)^{3/2}}{3B^3} (\hat{c} \cdot \nabla B) \right] + O(\epsilon^4) \quad (5.25)$$

$$\begin{aligned} \tilde{\theta} = \theta + \epsilon \frac{(2BJ)^{1/2}}{B^2} (\hat{c} \cdot \nabla B) + \epsilon^2 \frac{(2BJ)}{12B^4} \hat{a} \hat{c} : (3B \nabla \nabla B \\ - 5 \nabla B \nabla B) + O(\epsilon^3) \end{aligned} \quad (5.26)$$

$$\begin{aligned} \bar{J} = J + \epsilon \frac{(2BJ)^{3/2}}{3B^3}(\hat{a} \cdot \nabla B) + \epsilon^2 \frac{(2BJ)^2}{48B^5}[(7\hat{a}\hat{a} + 9\hat{c}\hat{c}) : \nabla B \nabla B \\ + 3B(\hat{a}\hat{a} - \hat{c}\hat{c}) : \nabla \nabla B] + O(\epsilon^3) \end{aligned} \quad (5.27)$$

In all cases these formulas have been carried 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 $(\bar{X}, \bar{\theta}, \bar{J})$ can be expressed in terms of (x, y) . 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 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 Kruskal¹⁷ 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, y) . 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 that 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 treatment 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 charged 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.

ACKNOWLEDGEMENTS

I would like to thank A. N. Kaurman for his patient encouragement of this work, and to acknowledge several helpful discussions with R. K. Sachs and J. R. Cary. In addition, critical readings of the manuscript by C. Grebogi, J. A. Rome, A. Ogawa, and A. Mine are greatly appreciated.

REFERENCES.

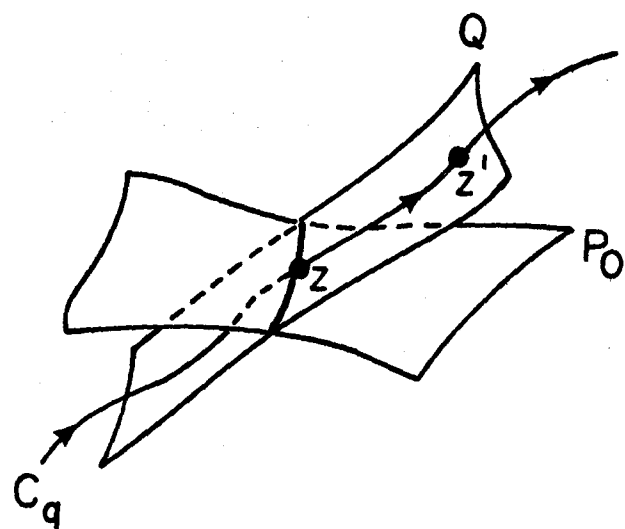
1. R. Abraham, Foundations of Mechanics (Benjamin, New York, 1967).
2. A. M. Vinogradov and B. A. Kupersmidt, Russ. Math. Surveys 32:4, 177(1977).
3. V. I. Arnold, Mathematical Methods of Classical Mechanics (Springer, New York, 1978).
4. H. Goldstein, Classical Mechanics (Addison-Wesley, Reading Massachusetts, 1950).
5. E. Saletan and A. Cromer, Theoretical Mechanics (Wiley, New York, 1971).
6. E. C. G. Sudershan and N. Mukunda, Classical Dynamics: A Modern Approach (Wiley, New York, 1974).
7. H. Alfven, 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. 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. 115, 791(1959).
13. T. G. Northrop and J. A. Rome, Phys. Fluids 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. Poincaré, 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., Ansterdam, 1964).
20. Y. Hagihara, Celestial Mechanics (MIT Press, Cambridge, Mass., 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. A9, 2043(1976).
25. V. I. Arnold, Ordinary Differential Equations (MIT Press, Cambridge, Mass., 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, Phys. Rev. Lett.
43, 1668(1979).

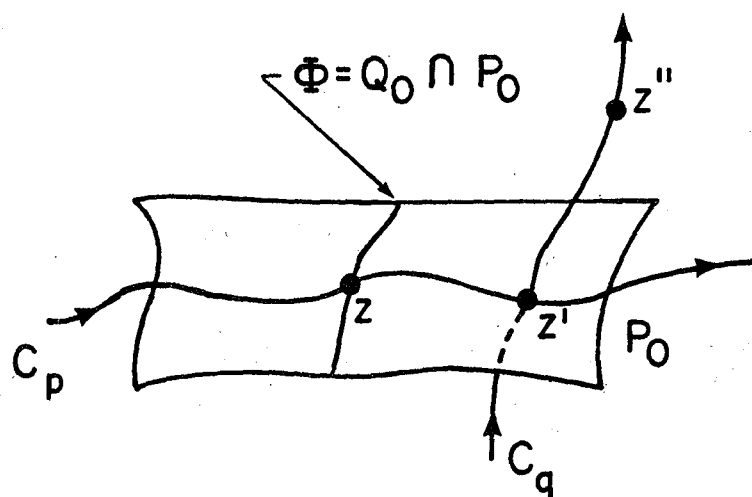
FIGURE CAPTIONS.

- Fig. 1. The q -characteristics and the construction of the functions $p(z)$ and $\tilde{Z}(z)$.
- Fig. 2. The construction of the functions $\tilde{Z}(z)$ 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 $\tilde{X}(x, \theta, v)$. The figure shows a θ -characteristic moving toward the surface $v=0$.



XBL793-972

Fig. 1



XBL 793 - 973

Fig. 2

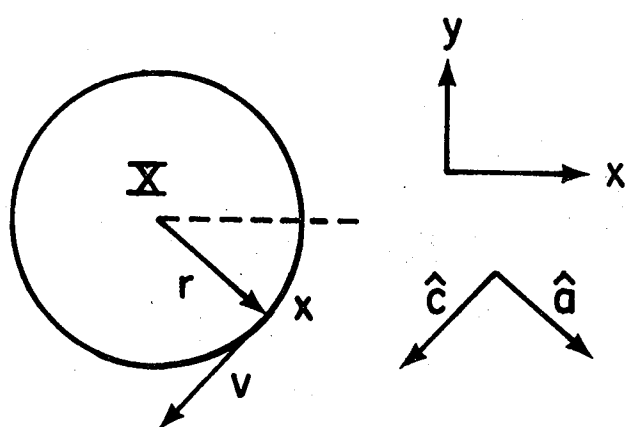
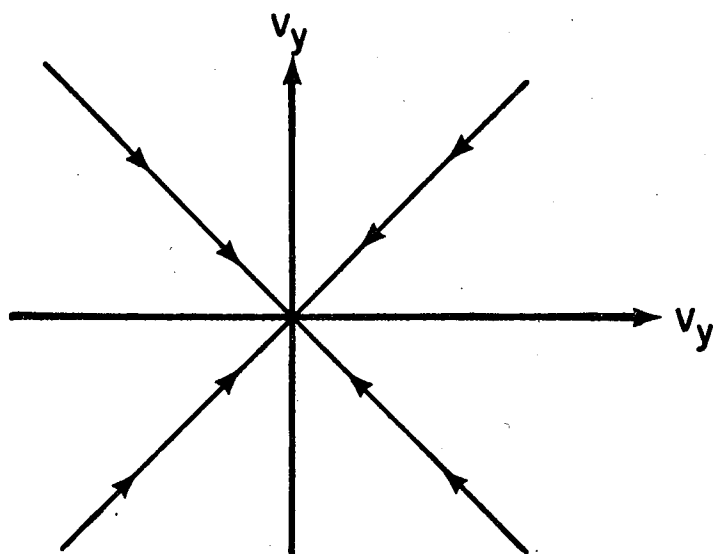
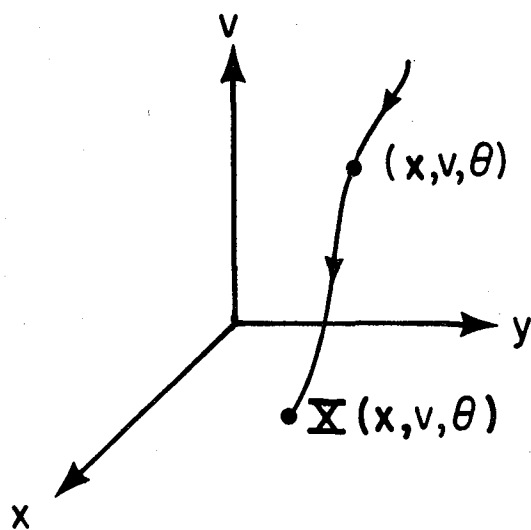


Fig. 3.



XBL 793-971

Fig. 4.



XBL 793-969

Fig. 5.

CHAPTER III

THREE-DIMENSIONAL GUIDING CENTER MOTION

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 $B=B(x)$, with $E=0$, 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 paper¹ 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 $\mathbf{B} = B(x,y)\hat{z}$. 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 $\mathbf{B} = \mathbf{B}(\mathbf{x})$, with the electric field $\mathbf{E} = 0$. 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 Mynick¹² 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 Meier¹³ has shown how the guiding center problem can be treated without using either mixed-variable generating functions or Lie transforms. Instead, Meier has developed canonical transformations by appealing directly to the defining Poisson bracket relations. Furthermore, Meier has avoided the use of field line coordinates. Meier'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 Rome⁸ 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 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 a 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 $(\underline{q}, \underline{p})$ 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 $\underline{B}(\underline{x})$ with $\underline{E}=0$ may be described by the Hamiltonian

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

where $\underline{A}$ is the magnetic vector potential satisfying $\underline{B} = \nabla \times \underline{A}$. 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 $(\underline{q}, \underline{p})$ are related to the particle's position $\underline{x}$ and velocity $\underline{v}$ by

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

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 $\underline{p}$ has dimensions of velocity.

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

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

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

$$H(\underline{q}, \underline{p}, \epsilon) = \frac{1}{2} \left[\underline{p} - \frac{e}{\epsilon mc} \underline{A}(\underline{q}) \right]^2 \quad (2.4)$$

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

$$\begin{aligned} \underline{x} &= \underline{q} \\ \underline{v} &= \underline{p} - \frac{e}{\epsilon mc} \underline{A}(\underline{q}) \end{aligned} \quad (2.5)$$

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 $\epsilon=1$ is the physical system. The order of an expression is determined by its behavior as $\epsilon \rightarrow 0$, while the position $\underline{x}$, the velocity $\underline{v}$, and the fields $\underline{A}$ and $\underline{B}$ are held fixed. Thus the Hamiltonian is $O(1)$, and the canonical momentum $\underline{p}$ is $O(\epsilon^{-1})$.

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

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

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

$$B_{ij} = \epsilon_{ijk} B_k \quad (2.7)$$

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 $(\underline{x}, \underline{v})$ coordinates:

$$H(\underline{x}, \underline{v}) = \frac{1}{2} v^2 \quad (2.8)$$

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 $(\underline{x}, \underline{v})$, then their Poisson bracket $\{F, G\}$ is given by

$$\{F, G\} = \frac{\partial F}{\partial \underline{x}} \cdot \frac{\partial G}{\partial \underline{v}} - \frac{\partial F}{\partial \underline{v}} \cdot \frac{\partial G}{\partial \underline{x}} + \frac{1}{e} \underline{\Omega} \cdot \left(\frac{\partial F}{\partial \underline{v}} \times \frac{\partial G}{\partial \underline{v}} \right) \quad (2.9)$$

where the vector $\underline{\Omega}$ is defined by

$$\underline{\Omega} = \frac{e}{mc} \underline{B} \quad (2.10)$$

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 $B(\underline{x}) = |\underline{B}(\underline{x})|$,

and define the unit vector $\hat{b}(\underline{x})$ by

$$\hat{b}(\underline{x}) = \frac{\underline{B}(\underline{x})}{B(\underline{x})} \quad (2.11)$$

It is convenient to assume that $B(\underline{x})$ is bounded away from zero in the spatial region of interest. Not only does this guarantee that the vector $\hat{b}(\underline{x})$ 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 $\underline{v}$ and the vector $\hat{b}$. These are u and w , the instantaneous parallel and perpendicular velocities, respectively, and they are given by

$$u = \underline{v} \cdot \hat{b}(\underline{x}) \quad (2.12)$$

$$w = (v^2 - u^2)^{1/2} \quad (2.13)$$

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

$$\hat{t}_1 \cdot \hat{t}_1 = \hat{t}_2 \cdot \hat{t}_2 = \hat{b} \cdot \hat{b} = 1 \quad (2.14)$$

$$\hat{t}_1 \cdot \hat{t}_2 = \hat{t}_1 \cdot \hat{b} = \hat{t}_2 \cdot \hat{b} = 0 \quad (2.15)$$

$$\hat{b} = \hat{t}_1 \times \hat{t}_2 \quad (2.16)$$

For the time being, we may assume that $\hat{t}_1$ and $\hat{t}_2$ 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 $\underline{v}_\perp$ by $\underline{v}_\perp = \underline{v} - u\hat{b}$. Next, the gyroradius vector $\underline{r}$ is given by $\underline{r} = \epsilon(\hat{b} \times \underline{v})/\Omega$, where Ω is the signed gyrofrequency:

$$\Omega = \frac{eB}{mc} = \hat{b} \cdot \underline{\Omega} = \text{sign}(e) |\underline{\Omega}| \quad (2.17)$$

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

$$\underline{v} = u\hat{b} + w\hat{c} \quad (2.18)$$

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

$$\hat{a} = \hat{b} \times \hat{c} \quad (2.19)$$

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

$$\hat{a} = \cos\theta \hat{\tau}_1 - \sin\theta \hat{\tau}_2 \quad (2.20)$$

$$\hat{c} = -\sin\theta \hat{\tau}_1 - \cos\theta \hat{\tau}_2$$

Our third coordinate system in phase space consists of the six physical particle variables $(\underline{x}, u, \theta, w)$ 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 best be taken as closed-form, algebraic relations specifying a variable transformation $(\underline{x}, \underline{v}) \rightarrow (\underline{x}, u, \theta, w)$.

Nevertheless, in the special case of a uniform magnetic field, these variables do have a 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 position is given by $\underline{X} = \underline{x} - \underline{r}$ exactly, and it is the precise center of the circle of motion. Later we will discuss ways in which the definition of $\underline{X}$ 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 $(\underline{x}, u, \theta, w)$ coordinate system, we need the Hamiltonian and the Poisson tensor. The former is easy to obtain:

$$H(\underline{x}, u, \theta, w) = \frac{1}{2} (w^2 + u^2) \quad (2.21)$$

As for the Poisson tensor, it may be obtained from Eq. (2.9) and

the relation $\{z^i, z^j\} = \sigma^{ij}$, with $\underline{z} = (x, u, \theta, w)$ being taken as the six-dimensional coordinate vector. A little calculation gives the components of the Poisson tensor in the following form:

$$\{x, x\} = 0 \quad (2.22a)$$

$$\{x, u\} = \hat{b} \quad (2.22b)$$

$$\{x, \theta\} = -\frac{\hat{a}}{w} \quad (2.22c)$$

$$\{x, w\} = \hat{c} \quad (2.22d)$$

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

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

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

Two notational conventions have been used in these equations and should be mentioned. First, for any pair of vectors $\underline{Y}$ and $\underline{Z}$, $\nabla \underline{Y} \cdot \underline{Z}$ means $(\nabla \underline{Y}) \cdot \underline{Z}$ and not $\nabla(\underline{Y} \cdot \underline{Z})$. 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 $\underline{y}$. This convention is followed whenever we are expressing any relation in the (x, u, θ, 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:

$$\hat{\tau}_1 \hat{\tau}_1 + \hat{\tau}_2 \hat{\tau}_2 + \hat{b} \hat{b} = \hat{a} \hat{a} + \hat{b} \hat{b} + \hat{c} \hat{c} = I \quad (3.1)$$

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

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

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

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

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

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

$$\underline{Y} \underline{Z} : \nabla \nabla \hat{b} \cdot \hat{b} = -(\underline{Y} \cdot \nabla \hat{b}) \cdot (\underline{Z} \cdot \nabla \hat{b}) \quad (3.6)$$

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

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

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

relation, on account of the Maxwell equation $\nabla \cdot \mathbf{B} = 0$:

$$\nabla \cdot \hat{\mathbf{b}} = - \frac{\hat{\mathbf{b}} \cdot \nabla \mathbf{B}}{B} \quad (3.8)$$

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

$$\underline{\mathbf{R}} = \nabla \hat{\mathbf{t}}_1 \cdot \hat{\mathbf{t}}_2 = \nabla \hat{\mathbf{c}} \cdot \hat{\mathbf{a}} \quad (3.9)$$

The vector $\underline{\mathbf{R}}$ has the following geometrical interpretation. The vectors $\hat{\mathbf{t}}_1$ and $\hat{\mathbf{t}}_2$, which define the perpendicular plane, are a function of position $\underline{\mathbf{x}}$ and hence vary from point to point. This variation is partly due to the variation in the vector $\hat{\mathbf{b}}$, to which $\hat{\mathbf{t}}_1$ and $\hat{\mathbf{t}}_2$ are orthogonal, and partly due to an arbitrariness in the definition of $\hat{\mathbf{t}}_1$ and $\hat{\mathbf{t}}_2$, which at this point in the work we are allowing for. Therefore if we examine the vectors $\hat{\mathbf{t}}_1$ and $\hat{\mathbf{t}}_2$ 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 $\hat{\mathbf{t}}_1$ and $\hat{\mathbf{t}}_2$ 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 $\hat{\mathbf{t}}_1$ and $\hat{\mathbf{t}}_2$ at P , and the angle $\Delta\psi$ will, for small separations, be proportional to the distance between P and P' . Indeed, if we let $\Delta\underline{\mathbf{x}}$ be the displacement vector between P and P' , then we have $\Delta\psi = \Delta\underline{\mathbf{x}} \cdot \underline{\mathbf{R}}$. In particular, the quantity $\hat{\mathbf{b}} \cdot \underline{\mathbf{R}}$ represents the rate (in terms of radians per unit length) at which the vectors $\hat{\mathbf{t}}_1$ and $\hat{\mathbf{t}}_2$ "twist" as one moves along a magnetic field line.

These considerations are important when we consider the arbitrariness in the definition of $\hat{\mathbf{t}}_1$ and $\hat{\mathbf{t}}_2$. 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 $\hat{\tau}_1(x)$ and $\hat{\tau}_2(x)$, another pair of perpendicular unit vector fields $\hat{\tau}'_1(x)$ and $\hat{\tau}'_2(x)$. 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 $\phi = \phi(x)$ 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 $\phi(x)$, 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 $\phi(x)$. 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, on the other hand, changes by the amount ϕ , since θ' is the gyrophase relative to the $\hat{\tau}'_1$ direction:

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

Therefore of the coordinates (x, u, θ, 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 $\underline{v}$, are naturally invariant under the transformation indicated by Eq. (3.10). Nevertheless, the vector $\underline{R}$, which can be expressed in terms of the gradients of $\hat{a}$ and $\hat{c}$ by $\underline{R} = \nabla \hat{c} \cdot \hat{a}$, is not invariant:

$$\underline{R}' = \nabla \hat{\tau}_1' \cdot \hat{\tau}_2' = \underline{R} - \nabla \phi \quad (3.12)$$

In view of the geometrical interpretation of the vector $\underline{R}$ 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 $\underline{R} = \nabla \hat{c} \cdot \hat{a}$ is not, we recall that the operator ∇ in the expression for $\underline{R}$ 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 $\underline{A}$.

Let us now ask ourselves to what extent the vector $\underline{R}$ 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 $\underline{R}=0$. The answer, as may be seen from Eq. (3.12), is no, because in general $\nabla \times \underline{R} \neq 0$. Nevertheless, this line of reasoning raises an interesting point, namely that the curl of $\underline{R}$ is invariant under a change of perpendicular unit vectors: $\nabla \times \underline{R} = \nabla \times \underline{R}'$. This in turn suggests that the vector $\nabla \times \underline{R}$ can be expressed purely in terms of $\hat{b}$. Some algebra shows that this is indeed the case:

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

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 used in computing the second order drifts appearing in Appendix A.

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

$$\phi(\underline{x}) = \int^{\underline{x}} \underline{R}(\underline{x}') \cdot d\underline{x}' \quad (3.14)$$

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

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

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

with the same integration conventions as in Eq. (3.14). Then $\hat{b} \cdot \nabla \phi = \hat{b} \cdot \underline{R} - \psi$, 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 $\underline{x}$ 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:

$$\hat{\tau}_1 = \frac{\hat{b} \cdot \nabla \hat{b}}{|\hat{b} \cdot \nabla \hat{b}|} \quad (3.16)$$

$$\hat{\tau}_2 = \hat{b} \times \hat{\tau}_1 \quad (3.17)$$

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:

$$\hat{a}\hat{a} + \hat{c}\hat{c} = I - \hat{b}\hat{b} \quad (3.18)$$

$$\hat{a}\hat{c} - \hat{c}\hat{a} = \hat{b} \times I \quad (3.19)$$

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

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

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

$$Z_2 = \hat{b} \cdot \nabla \hat{c} \cdot \hat{a} = \hat{b} \cdot \underline{R} \quad (3.22)$$

The symbol Z is a mnemonic for "zeroth harmonic." Observe that Z_0 vanishes in a current free region of space, i.e. where $\nabla \times \underline{B} = 0$, 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":

$$F_0 = \hat{b} \cdot \nabla \hat{b} \cdot \hat{a} \quad (3.23)$$

$$F_1 = \hat{b} \cdot \nabla \hat{b} \cdot \hat{c} \quad (3.24)$$

$$F_2 = \hat{a} \cdot \nabla \hat{c} \cdot \hat{a} = \hat{a} \cdot \underline{\hat{R}} \quad (3.25)$$

$$F_3 = \hat{c} \cdot \nabla \hat{c} \cdot \hat{a} = \hat{c} \cdot \underline{\hat{R}} \quad (3.26)$$

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":

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

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

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:

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

$$\{u, w\} = w \left(\frac{1}{2} Z_1 - S_1 \right) + u F_1 \quad (3.29b)$$

$$\{\theta, w\} = \frac{\Omega}{\epsilon w} + \frac{u}{w} Z_0 + F_3 \quad (3.29c)$$

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 (x,u,θ,w) coordinate system in phase space to a new system, denoted by (X,U,θ,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 these expressions. Most importantly, we will derive in this section the components of the Poisson tensor with respect to the (X,U,θ,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 (x,u,θ,w) , namely X , U , and J , which will satisfy the following Poisson bracket relations:

$$\{\theta, J\} = 1/\epsilon \quad (4.1)$$

$$\{\theta, X\} = 0 \quad (4.2)$$

$$\{\theta, U\} = 0 \quad (4.3)$$

$$\{J, X\} = 0 \quad (4.4)$$

$$\{J, U\} = 0 \quad (4.5)$$

The solution of these equations for the five unknown functions (X,U,J) will produce a "semicanonical" coordinate system in phase space, namely (X,U,θ,J) , in which the variables θ and J are, one

might say, "canonically decoupled" from the other four variables $(\underline{X}, U)$. The reason for choosing the symbols $\underline{X}$ 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 $1/\epsilon$ instead of 1 so that J will be $O(1)$ instead of $O(\epsilon)$.

The transformation $(\underline{x}, u, \theta, w) \rightarrow (\underline{X}, U, \theta, J)$ 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 $(\underline{X}, U, \theta, J)$ 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

$$d/d\lambda = \{ \cdot, \theta \} \quad (4.6)$$

$$d/d\mu = \{ \cdot, J \} \quad (4.7)$$

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:

$$dJ/d\lambda = -1/\epsilon \quad (4.8)$$

$$d\tilde{x}/d\lambda = 0 \quad (4.9)$$

$$dU/d\lambda = 0 \quad (4.10)$$

$$d\tilde{x}/d\mu = 0 \quad (4.11)$$

$$dU/d\mu = 0 \quad (4.12)$$

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

$$d\tilde{z}/d\lambda = \{\tilde{z}, \theta\} \quad (4.13)$$

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 $\tilde{z} = \tilde{z}(\mu)$ satisfying

$$d\tilde{z}/d\mu = \{\tilde{z}, J\} \quad (4.14)$$

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

$$\frac{d\tilde{x}}{d\lambda} = -\frac{\hat{a}}{w} \quad (4.15)$$

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

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

For ϵ sufficiently small, the right hand side of Eq. (4.17) is dominated by the term $-\Omega/\epsilon w$, 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:

$$\frac{dx}{dw} = \frac{\varepsilon \hat{a}}{D} \quad (4.18)$$

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

$$\frac{dJ}{dw} = \frac{w}{D} \quad (4.20)$$

Here the denominator D is given by

$$D = \Omega + \varepsilon(uZ_0 + wF_3) \quad (4.21)$$

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

4.3. Geometrical interpretation of the coordinates (X, U, J)

Let us give a geometrical interpretation to the θ -characteristics, and also to the functions (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 $\theta=\text{constant}$, 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 behavior for the θ -characteristics was observed in Ref. 1 and discussed there in greater detail.

Every point $\underline{z} = (\underline{x}, u, \theta, w)$ 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 $\underline{x}$ and u coordinates take on certain values, which can be considered functions of the original point $\underline{z}$. We will call these functions $\underline{X}(\underline{x}, u, \theta, w)$ and $U(\underline{x}, u, \theta, w)$; they have the property that when $w=0$, $\underline{X}=\underline{x}$ and $U=u$. Effectively, the functions $\underline{X}$ 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 $\underline{X}$ and U elsewhere in phase space are found by propagating these functions along θ -characteristics, i.e. by assigning the same values of $\underline{X}$ and U to any two points $\underline{z}$ and $\underline{z}'$ which lie on the same θ -characteristic. Clearly, the functions $\underline{X}$ and U so constructed are constants of the θ -characteristics, and hence satisfy Eqs. (4.9) and (4.10).

As for the function $J(\underline{x}, u, \theta, w)$, we define it to be $-1/\epsilon$ times the elapsed λ parameter between the point $\underline{z} = (\underline{x}, u, \theta, w)$ and the $w=0$ point on the θ -characteristic passing through $\underline{z}$. 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 $(\underline{x}, U)$ 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 $J(\underline{x}, u, \theta, w)$ 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

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

where

$$D_0 = \Omega + \epsilon u Z_0 \quad (4.23)$$

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

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

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

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

The right hand sides of all three of these equations go to zero as $w \rightarrow 0$, so that the J-characteristics on the surface $w=0$ consist of immobile points. Hence the functions $(\underline{x}, U)$, which take on the values $(\underline{x}, u)$ 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 $(\underline{X}, U, J)$, 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 $(\underline{X}, U, \theta, J)$ coordinate system

Of the 15 independent components of the Poisson tensor in the $(\underline{X}, U, \theta, J)$ coordinate system, nine are given by Eqs. (4.1)-(4.5). The remaining six components, i.e. the Poisson brackets of the coordinates $(\underline{X}, U)$ 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 $(\underline{X}, U)$ 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 $(\underline{X}, U)$, 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 $(\underline{X}, U)$ 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)-(4.19) as a power series in w . Again, to lowest order, the results can be written down practically by inspection:

$$\underline{X}(\underline{x}, u, \theta, w) = \underline{x} - \frac{\epsilon w}{D_0} \hat{a} + O(w^2) \quad (4.27)$$

$$U(\underline{x}, u, \theta, w) = u - \frac{\epsilon w u F_0}{D_0} + O(w^2) \quad (4.28)$$

Taking the Poisson brackets of these quantities with themselves

and keeping track of the w -ordering gives, after some algebra,

$$\{X_i, X_j\} = \frac{\epsilon}{D_0} (a_i c_j - a_j c_i) + O(w) \quad (4.29)$$

$$\{X_i, U\} = b_i + \frac{\epsilon U}{D_0} (a_i F_1 - c_i F_0) + O(w) \quad (4.30)$$

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

$$\{\tilde{X}, \tilde{X}\} = \frac{\epsilon}{\Omega + \epsilon U Z_0} \hat{b} \times I \quad (4.31)$$

$$\{\tilde{X}, U\} = \hat{b} + \frac{\epsilon U}{\Omega + \epsilon U Z_0} \hat{b} \times (\hat{b} \cdot \nabla \hat{b}) \quad (4.32)$$

In these equations all fields are evaluated at $\tilde{X}$, e.g. $\hat{b}$ means $\hat{b}(\tilde{X})$, and ∇ means $\partial / \partial \tilde{X}$. Eqs. (4.31) and (4.32), along with Eqs. (4.1)-(4.5), completely specify the Poisson tensor in the (X, U, θ, 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)-(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 (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 $(\underline{X}, \underline{U}) = (X, Y, Z, U)$ for the four phase space coordinates, and define, for the purposes of this demonstration, two vectors $\underline{M}$ and $\underline{N}$ by

$$\underline{M} = \hat{\epsilon} \underline{b} \quad (4.33)$$

$$\underline{N} = \underline{\Omega} + \epsilon \underline{U} \nabla \times \underline{b} \quad (4.34)$$

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

$$\{\underline{X}, \underline{X}\} = \underline{M} \times \underline{I} / (\hat{\underline{b}} \cdot \underline{N}) \quad (4.35)$$

$$\{\underline{X}, \underline{U}\} = \underline{N} / (\hat{\underline{b}} \cdot \underline{N}) \quad (4.36)$$

and hence the Poisson tensor has the form

$$\sigma_{(4)}^{ij} = \frac{1}{\hat{\underline{b}} \cdot \underline{N}} \begin{pmatrix} 0 & -M_z & M_y & N_x \\ M_z & 0 & -M_x & N_y \\ -M_y & M_x & 0 & N_z \\ -N_x & -N_y & -N_z & 0 \end{pmatrix} \quad (4.37)$$

Here the subscript 4 has been appended to the symbol σ to indicate that we are dealing with the reduced 4x4 Poisson tensor in the variables $(\underline{X}, \underline{U})$.

On taking the negative of the inverse of $\sigma_{(4)}^{ij}$ we obtain the 4x4 Lagrange tensor $\omega_{(4)ij}$:

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

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 $\hat{b} \cdot \tilde{N}$. The tensor $\omega_{(4)}$ is closed, i.e. it satisfies

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

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

$$\nabla \cdot \tilde{N} = 0 \quad (4.40)$$

$$\nabla \times \tilde{M} = \partial \tilde{N} / \partial U \quad (4.41)$$

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 $\tilde{z}=(X,U,\theta,J)$, and write ω_{ij} for the 6×6 Lagrange tensor. Then in accordance with Eq. (4.1)-(4.5) we have

$$\omega_{ij} = \left(\begin{array}{cc|cc} \omega_{(4)} & & & 0 \\ & & & \\ \hline 0 & & 0 & \epsilon \\ & & -\epsilon & 0 \end{array} \right) \quad (4.42)$$

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

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

On taking the determinant of this relation we obtain

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

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

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

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

$$|\Delta| = |\Omega + \epsilon U Z_0| \quad (4.46)$$

Therefore we have

$$d^3 \underline{q} d^3 \underline{p} = |\Omega + \epsilon U Z_0| d^3 \underline{X} dU d\theta dJ \quad (4.47)$$

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

5. THE DARBOUX TRANSFORMATION: EXPLICIT EXPRESSIONS

In this section we will give explicit formulas for the Darboux transformation $(\underline{x}, \underline{u}, \theta, w) \rightarrow (\underline{X}, \underline{U}, \theta, J)$ 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 $(\underline{X}, \underline{U}, \theta, J)$ 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 $z_i = (x_i, u_i, \theta_i, w_i)$ and $z_f = (x_f, u_f, \theta_f, w_f)$, 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 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 , 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 x_f , u_f , and J_f . The quantities x_i , u_i and J_i are to be thought of as initial conditions for the functions 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 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

$$x(w) = x_0 + \epsilon x_1 + \epsilon^2 x_2 + O(\epsilon^3) \quad (5.1)$$

$$u(w) = u_0 + \epsilon u_1 + O(\epsilon^2) \quad (5.2)$$

$$J(w) = J_0 + \epsilon J_1 + O(\epsilon^2) \quad (5.3)$$

in which the quantities x_0, 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

$$\Omega(x) = \Omega(x_0) + \epsilon x_1 \cdot \nabla \Omega(x_0) + O(\epsilon^2) \quad (5.4)$$

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 z_i and z_f on a θ -characteristic:

$$x_f = x_i + \frac{\epsilon \hat{a}}{\Omega} (w_f - w_i) + \epsilon^2 \left\{ \frac{1}{2\Omega} (w_f - w_i)^2 \hat{a} \cdot \nabla \left(\frac{\hat{a}}{\Omega} \right) - \frac{\hat{a}}{2\Omega^2} \left[(w_f^2 - w_i^2) F_3 + 2u_i (w_f - w_i) Z_0 \right] \right\} + O(\epsilon^3) \quad (5.5)$$

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

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

In these formulas, all fields on the right hand side are evaluated at x_i .

5.2. The Darboux transformation and its inverse

Let us specialize the formulas above so as to obtain X, U , and J as functions of (x, u, θ, w) . To do this we identify z_i with z and z_f

with the $w=0$ point on the θ -characteristic passing through $\underline{z}$. That is, we set $\underline{x}_i = \underline{x}$, $\underline{u}_i = \underline{u}$, $\underline{w}_i = \underline{w}$, and $\underline{J}_i = \underline{J}$, and also $\underline{x}_f = \underline{X}$, $\underline{u}_f = \underline{U}$, $\underline{w}_f = 0$, and $\underline{J}_f = 0$. These substitutions are in accordance with the definition and initial value properties of the functions $\underline{X}$, $\underline{U}$, and $\underline{J}$, as described in Sec. 4, and they give the following:

$$\begin{aligned} \underline{X}(\underline{x}, \underline{u}, \theta, \underline{w}) = \underline{x} - \frac{\epsilon \underline{w} \hat{\underline{a}}}{\Omega} + \epsilon^2 \left\{ \frac{\underline{w}^2}{2\Omega^2} \left[\Omega \hat{\underline{a}} \cdot \nabla \left(\frac{\hat{\underline{a}}}{\Omega} \right) + F_3 \hat{\underline{a}} \right] \right. \\ \left. + \frac{\underline{u} \underline{w}}{\Omega^2} Z_0 \hat{\underline{a}} \right\} + O(\epsilon^3) \end{aligned} \quad (5.8)$$

$$\underline{U}(\underline{x}, \underline{u}, \theta, \underline{w}) = \underline{u} - \frac{\epsilon}{2\Omega} \left[\underline{w}^2 (Z_2 - \frac{1}{2} Z_0 + S_0) + 2 \underline{u} \underline{w} F_0 \right] + O(\epsilon^2) \quad (5.9)$$

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

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

Note that through the $O(\epsilon)$ term the quantity $\underline{X}$ corresponds with the usual definition of the guiding center. Alternatively, we might say that $\underline{X}$ coincides with the exact guiding center for a uniform magnetic field. It is on these grounds that we will call the variables $(\underline{X}, \underline{U}, \theta, \underline{J})$ "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 $\underline{U}$ needs no interpretation; it is the instantaneous parallel velocity. And the first term of the expression for $\underline{J}$ is, of course, proportional to the magnetic moment to lowest order. Note that $\underline{J}$ 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 z_i and z_f . This will allow us to determine $\tilde{x}$, u , and J as functions of $\tilde{X}$, U , and w . That is, we set $x_i = \tilde{X}$, $u_i = U$, $w_i = 0$, and $J_i = 0$, and also $x_f = \tilde{x}$, $u_f = u$, $w_f = w$, and $J_f = J$. Doing so, we obtain

$$\begin{aligned} \tilde{x}(\tilde{X}, U, \theta, w) = \tilde{X} + \frac{\epsilon w \hat{a}}{\Omega} + \epsilon^2 \left\{ \frac{w^2}{2\Omega^2} \left[\Omega \hat{a} \cdot \nabla \left(\frac{\hat{a}}{\Omega} \right) - F_3 \hat{a} \right] \right. \\ \left. - \frac{Uw}{\Omega^2} Z_0 \hat{a} \right\} + O(\epsilon^3) \end{aligned} \quad (5.11)$$

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

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

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

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 (x, u, θ, w) in terms of the intermediate guiding center variables (X, U, θ, J) . To do this, we first invert the series in Eq. (5.13) to find w as a function of (X, U, θ, J) . To the order given this series inversion is trivial, and it gives

$$w(\underline{X}, U, \theta, J) = (2\Omega J)^{1/2} + \epsilon \left[\frac{(2\Omega J)}{3\Omega^2} (\hat{\mathbf{a}} \cdot \nabla \Omega + \Omega F_3) + \frac{(2\Omega J)^{1/2} U}{2\Omega} z_0 \right] + O(\epsilon^2) \quad (5.14)$$

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

$$\begin{aligned} \underline{x}(\underline{X}, U, \theta, J) = \underline{X} + \frac{\epsilon(2\Omega J)^{1/2}}{\Omega} \hat{\mathbf{a}} + \epsilon^2 \left\{ \frac{(2\Omega J)}{6\Omega^3} \left[3\Omega \hat{\mathbf{a}} \cdot \nabla \hat{\mathbf{a}} - \Omega F_3 \hat{\mathbf{a}} \right. \right. \\ \left. \left. - (\hat{\mathbf{a}} \cdot \nabla \Omega) \hat{\mathbf{a}} \right] - \frac{(2\Omega J)^{1/2} U}{2\Omega^2} z_0 \hat{\mathbf{a}} \right\} + O(\epsilon^3) \end{aligned} \quad (5.15)$$

$$\begin{aligned} u(\underline{X}, U, \theta, J) = U + \epsilon \left[\frac{(2\Omega J)}{2\Omega} (z_2 - \frac{1}{2} z_0 + S_0) + \frac{(2\Omega J)^{1/2} U}{\Omega} F_0 \right] + O(\epsilon^2) \end{aligned} \quad (5.16)$$

Again, all fields on the right hand side are evaluated at $\underline{X}$. 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 $(\underline{X}, U, \theta, J)$. It is obtained by simply substituting Eqs. (5.14) and (5.16) into (2.21), and this gives

$$\begin{aligned} H(\underline{X}, U, \theta, J) = \Omega J + \frac{1}{2} U^2 + \epsilon \left[\frac{(2\Omega J)^{3/2}}{3\Omega^2} (\Omega F_3 + \hat{\mathbf{a}} \cdot \nabla \Omega) \right. \\ \left. + \frac{(2\Omega J) U}{2\Omega} (\frac{1}{2} z_0 + z_2 + S_0) + \frac{(2\Omega J)^{1/2} U^2}{\Omega} F_0 \right] + O(\epsilon^2) \end{aligned} \quad (5.17)$$

We should not expect this Hamiltonian to be independent of θ , and

indeed, there are θ -dependent terms in the $O(\epsilon)$ term of Eq. (5.17). The angle θ is a well-defined function of the physical particle variables (x, 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 $\epsilon \rightarrow 0$.) This requirement makes the unperturbed system "recurrent," in Kruskal's¹⁶ 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

$$H(\tilde{x}, U, \theta, J) = \Omega J + \frac{1}{2} U^2 + \epsilon H_1(\tilde{x}, U, \theta, J) + O(\epsilon^2) \quad (5.18)$$

and then use the Poisson bracket relations, given in Eqs. (4.1)-(4.5) and (4.31)-(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)-(4.32) are to be expanded into a power series in ϵ in this process.

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

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

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

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

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

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

Mirroring effects are displayed by computing dU/dt :

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

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

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

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

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 $O(\epsilon)$. 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 $O(\epsilon)$ 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 by 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 $(X, U, \theta, J) \rightarrow (\bar{X}, \bar{U}, \bar{\theta}, \bar{J})$ 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 $g_1, g_2, \dots$, which we will call the generators of the transformation. The generators are associated with a sequence $L_1, L_2, \dots$ of "Lie operators," defined by

$$L_n = \epsilon \{g_n, \} \quad (6.1)$$

The factor ϵ has been inserted into this definition in order to cancel the factor $1/\epsilon$ in Eq. (4.1), so that the Lie operators L_n are $O(1)$. The Lie operators are in turn associated with a sequence $T_1, T_2, \dots$ of symplectic transformation operators, according to the rule

$$T_n = \exp(-\epsilon^n L_n / n) \quad (6.2)$$

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

$$T = \dots T_3 T_2 T_1 \quad (6.3)$$

$$T^{-1} = T_1^{-1} T_2^{-1} T_3^{-1} \dots \quad (6.4)$$

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

$$\bar{z} = T z \quad (6.5)$$

$$z = T^{-1} \bar{z} \quad (6.6)$$

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

$$K = T^{-1}H \quad (6.7)$$

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 a power series in ϵ , and write

$$L_n = L_{n0} + \epsilon L_{n1} + \epsilon^2 L_{n2} + \dots \quad (6.8)$$

where

$$L_{n0} = \frac{\partial g_n}{\partial \theta} \frac{\partial}{\partial J} - \frac{\partial g_n}{\partial J} \frac{\partial}{\partial \theta} \quad (6.9)$$

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

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

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

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

Finally, we write

$$H = \sum_{n=0}^{\infty} \epsilon^n H_n \quad (6.13)$$

$$K = \sum_{n=0}^{\infty} \epsilon^n K_n \quad (6.14)$$

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

$$K_0 = H_0 \quad (6.15)$$

$$L_{10} H_0 = K_1 - H_1 \quad (6.16)$$

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

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

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

$$K_1(\bar{X}, \bar{U}, \bar{J}) = \bar{J} \bar{U} \left(\frac{1}{2} Z_0 + Z_2 \right) \quad (6.18)$$

The oscillatory part gives a differential equation for g_1 :

$$\begin{aligned} \Omega \frac{\partial g_1}{\partial \theta} = & - \frac{(2\Omega J)^{3/2}}{3\Omega^2} (\Omega F_3 + \hat{a} \cdot \nabla \Omega) - \frac{(2\Omega J)U}{2\Omega} S_0 \\ & - \frac{(2\Omega J)^{1/2} U^2}{\Omega} F_0 \end{aligned} \quad (6.19)$$

Using Table I, this is easily integrated, yielding

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

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

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

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

6.3. The averaging transformation: explicit formulas

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

$$\bar{X}(X, U, \theta, J) = X + \frac{\epsilon \hat{b}^2}{4\Omega^2} \left[-(2\Omega J) S_1 + 8(2\Omega J)^{1/2} U F_1 \right] + O(\epsilon^3) \quad (6.22)$$

$$\bar{U}(X, U, \theta, J) = U + O(\epsilon^2) \quad (6.23)$$

$$\begin{aligned} \bar{J}(X, U, \theta, J) = J + \frac{\epsilon}{6\Omega^3} \left[2(2\Omega J)^{3/2} (\Omega F_3 + \hat{a} \cdot \nabla \Omega) + 3\Omega(2\Omega J) U S_0 \right. \\ \left. + 6\Omega(2\Omega J)^{1/2} U^2 F_0 \right] + O(\epsilon^2) \end{aligned} \quad (6.24)$$

$$\begin{aligned} \bar{\theta}(X, U, \theta, J) = \theta + \frac{\epsilon}{2\Omega^2} \left[2(2\Omega J)^{1/2} (-\Omega F_2 + \hat{c} \cdot \nabla \Omega) - \Omega U S_1 \right. \\ \left. + 2\Omega(2\Omega J)^{-1/2} U^2 F_1 \right] + O(\epsilon^2) \end{aligned} \quad (6.25)$$

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 $\underline{z}$ and $\bar{\underline{z}}$ 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 composing Eqs. (6.22)-(6.25) with Eqs. (5.8)-(5.10). The result is

$$\begin{aligned} \bar{\underline{x}}(\underline{x}, u, \theta, w) = \underline{x} - \frac{\epsilon w \hat{a}}{\Omega} + \epsilon^2 \left\{ \frac{w^2}{4\Omega^3} \left[-3\Omega S_1 \hat{b} + 2(\hat{b} \cdot \nabla \Omega) \hat{b} + 2\Omega(\hat{b} \times \underline{R}) \right. \right. \\ \left. \left. - \nabla \Omega - (\hat{a}\hat{a} - \hat{c}\hat{c}) \cdot \nabla \Omega \right] + \frac{wu}{\Omega^2} (Z_0 \hat{a} + 2F_1 \hat{b}) \right\} + O(\epsilon^3) \end{aligned} \quad (6.26)$$

$$\bar{u}(\underline{x}, u, \theta, w) = u - \frac{\epsilon}{2\Omega} \left[w^2 (Z_2 - \frac{1}{2} Z_0 + S_0) + 2uwF_0 \right] + O(\epsilon^2) \quad (6.27)$$

$$\begin{aligned} \bar{\theta}(\underline{x}, u, \theta, w) = \theta + \frac{\epsilon}{2\Omega^2} \left[2w(-\Omega F_2 + \hat{c} \cdot \nabla \Omega) - \Omega u S_1 + 2\Omega \frac{u^2}{w} F_1 \right] \\ + O(\epsilon^2) \end{aligned} \quad (6.28)$$

$$\begin{aligned} \bar{J}(\underline{x}, u, \theta, w) = \frac{w^2}{2\Omega} + \frac{\epsilon}{2\Omega^3} \left[w^3 \hat{a} \cdot \nabla \Omega + \Omega w^2 u (S_0 - Z_0) + 2\Omega w u^2 F_0 \right] \\ + O(\epsilon^2) \end{aligned} \quad (6.29)$$

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

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

$$\begin{aligned}
x(\bar{X}, \bar{U}, \bar{\theta}, \bar{J}) = & \bar{X} + \epsilon \frac{(2\Omega\bar{J})^{1/2}}{\Omega} \hat{a} + \epsilon^2 \left\{ \frac{(2\Omega\bar{J})}{4\Omega^3} \left[-\Omega S_1 \hat{b} + 4(\hat{b} \cdot \nabla \Omega) \hat{b} \right. \right. \\
& - 2\Omega(\hat{b} \times \mathbf{R}) - 3\nabla \Omega + (\hat{a}\hat{a} - \hat{c}\hat{c}) \cdot \nabla \Omega \left. \right] + \frac{(2\Omega\bar{J})^{1/2} \bar{U}}{4\Omega^2} \left[-8F_1 \hat{b} \right. \\
& \left. \left. - 2Z_0 \hat{a} - \hat{b} \times (\hat{a} \cdot \nabla \hat{b}) - \hat{c} \cdot \nabla \hat{b} \right] - \frac{\bar{U}^2}{\Omega^2} \hat{b} \cdot \nabla \hat{b} \right\} + O(\epsilon^3)
\end{aligned} \tag{6.30}$$

$$\begin{aligned}
u(\bar{X}, \bar{U}, \bar{\theta}, \bar{J}) = & \bar{U} + \frac{\epsilon}{2\Omega} \left[(2\Omega\bar{J}) (Z_2 - \frac{1}{2}Z_0 + S_0) + 2(2\Omega\bar{J})^{1/2} \bar{U} F_0 \right] \\
& + O(\epsilon^2)
\end{aligned} \tag{6.31}$$

$$\begin{aligned}
\theta(\bar{X}, \bar{U}, \bar{\theta}, \bar{J}) = & \bar{\theta} + \frac{\epsilon}{2\Omega^2} \left[2(2\Omega\bar{J})^{1/2} (\Omega F_2 - \hat{c} \cdot \nabla \Omega) + \Omega \bar{U} S_1 \right. \\
& \left. - 2\Omega(2\Omega\bar{J})^{-1/2} \bar{U}^2 F_1 \right] + O(\epsilon^2)
\end{aligned} \tag{6.32}$$

$$\begin{aligned}
w(\bar{X}, \bar{U}, \bar{\theta}, \bar{J}) = & (2\Omega\bar{J})^{1/2} + \frac{\epsilon}{2\Omega} \left[(2\Omega\bar{J})^{1/2} \bar{U} (Z_0 - S_0) - 2\bar{U}^2 F_0 \right] \\
& + O(\epsilon^2)
\end{aligned} \tag{6.33}$$

In these expressions, all fields on the right hand side are evaluated at the averaged guiding center position $\bar{X}$, 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 has been worked out by Hastie, Taylor, and Haas.¹⁷

6.4. Uniqueness of the averaged guiding center variables

It is important to ask to what extent the variables $(\bar{X}, \bar{U}, \bar{\theta}, \bar{J})$, 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's¹⁶ 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 $(\bar{X}, \bar{U}, \bar{\theta}, \bar{J})$ consists of two, namely $\bar{\theta}$ and $\bar{J}$, which are (apart from the factor $1/\epsilon$) canonically conjugate, plus four more, namely $\bar{X}$ and $\bar{U}$, which have vanishing Poisson brackets with $\bar{\theta}$ and $\bar{J}$. Let us write $\bar{\zeta}$ for the four variables $\bar{X}$ and $\bar{U}$ collectively. Then the semicanonical requirement can be written as $\{\bar{\theta}, \bar{J}\} = 1/\epsilon$ and $\{\bar{\zeta}, \bar{\theta}\} = \{\bar{\zeta}, \bar{J}\} = 0$. With the given identification for $\bar{\zeta}$, this requirement is equivalent to the Poisson bracket relations in Eqs. (4.1)-(4.5).

Third and finally, the four variables $\bar{\zeta}$ 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 $L = \epsilon\{g, \}$ 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

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

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

$$\bar{J}' = \bar{J} \quad (6.37)$$

Since we are assuming that $\partial g / \partial \theta = 0$, 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

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

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

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

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 $(\bar{X}, \bar{U}, \bar{J})$, on passing from Eq. (6.19) to Eq. (6.20), then the effect would be the same as the transformation above, with $g = \epsilon c$ and c being the constant of integration.

Let us now suspend the third requirement, and ask for the general form of a transformation $(\bar{\xi}, \bar{\theta}, \bar{J}) \rightarrow (\bar{\xi}'', \bar{\theta}'', \bar{J}'')$, such that the double primed variables are nice, and such that the variables $\bar{\theta}''$ and $\bar{J}''$ are canonically decoupled from the variables $\bar{\xi}''$ in the manner shown by Eqs. (4.1)-(4.5), but where the Poisson brackets of the variables $\bar{\xi}''$ 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 (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

$$\bar{\xi}'' = Z(\bar{\xi}') \quad (6.41)$$

$$\bar{\theta}'' = \bar{\theta}' \quad (6.42)$$

$$\bar{J}'' = \bar{J}' \quad (6.43)$$

where 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 $\bar{\xi}'$ 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 Z would be the transformation which leaves $\bar{U}$ unchanged but which converts $\bar{X}$ 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 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 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 $\bar{\xi}$ for the five nice variables $(\bar{X}, \bar{U}, \bar{J})$ collectively, and ask for the most general variable transformation which preserves niceness. Kruskal has answered this question; it is

$$\bar{\theta}' = \bar{\theta} + f(\bar{\xi}) \quad (6.44)$$

$$\bar{\xi}' = \bar{\xi} \quad (6.45)$$

where f is an arbitrary function and where $\underline{\Xi}$ 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 $\underline{R}$. In addition, the Hamiltonian K , shown in Eq. (6.21), depends on $\underline{R}$ 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 $\underline{R}$ through $Z_2 = \hat{\underline{b}} \cdot \underline{R}$ 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 $\hat{\underline{b}} \cdot \underline{R}$ equal to any scalar field we like. In particular, if we take

$$\hat{\underline{b}} \cdot \underline{R} = -\frac{1}{2} \hat{\underline{b}} \cdot (\nabla \times \hat{\underline{b}}) \quad (6.46)$$

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

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

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

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

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

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

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

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

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

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 $\bar{U}$:

$$\bar{U} = \text{Avg}(u) + \epsilon \bar{J} \hat{\underline{b}} \cdot (\nabla \times \hat{\underline{b}}) + O(\epsilon^2) \quad (6.52)$$

The variable $\bar{U}$ agrees with the variable $v_{\parallel}$ used by Northrop and Rome⁸ through the order given. A different choice of perpendicular unit vectors could have been made which would cause $\bar{U}$ to be identical with $\text{Avg}(u)$, although it would also cause the Hamiltonian K to be more complicated. The effect of the $O(\epsilon)$ 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 $\tilde{X}$, given in Eq. (6.26), of its dependence on $\underline{R}$, which is through the term $\hat{b} \cdot \underline{R}$. (The only exception is the case that $\nabla \times \underline{R}$, 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 $d\tilde{X}/dt$. We have

$$\frac{d\tilde{X}}{dt} = \{\tilde{X}, \tilde{X}\} \cdot [\bar{J} \nabla \Omega + O(\epsilon^2)] + \{\tilde{X}, \bar{U}\} [\bar{U} + O(\epsilon^2)] \quad (6.53)$$

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

$$\left(\frac{d\tilde{X}}{dt} \right)_{\parallel} = \hat{b} \bar{U} + O(\epsilon^2) \quad (6.54)$$

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

Eq. (6.54) shows that $\bar{U}$ is actually the parallel velocity of the guiding center. Eq. (6.56) shows that the $O(\epsilon^2)$ correction to the

perpendicular drifts is proportional to the $O(\epsilon)$ 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 $\bar{\mathbf{X}}$. Of course, Eq. (6.55) is easily expanded properly into a power series in ϵ .

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

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

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

Finally, we have the evolution of the gyrophase:

$$\frac{d\bar{\theta}}{dt} = \frac{\Omega(\bar{\mathbf{X}})}{\epsilon} + O(\epsilon) \quad (6.57)$$

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

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

One's intuition says that the guiding center position $\bar{\mathbf{X}}$ should not depend on the choice of perpendicular unit vectors, and hence that the appearance of the term $\hat{\mathbf{b}}\times\bar{\mathbf{R}}$ 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 $\bar{\mathbf{X}}$. It is here that we call upon the discussion of subsection 6.4.

Any alternate definition for $\bar{\mathbf{X}}$ 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 $\tilde{\mathbf{X}}$. There are many ways to define a variable $\tilde{\mathbf{X}}'$ which agrees with our $\tilde{\mathbf{X}}$ in any number of leading terms of Eq. (6.26) and which is also nice, and at the $O(\epsilon^2)$ 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

$$\tilde{\mathbf{X}}' = \tilde{\mathbf{X}} - \epsilon^2 \frac{\tilde{\mathbf{J}}}{\Omega} \hat{\mathbf{b}} \times \mathbf{R} \quad (6.58)$$

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

$$\bar{\mathbf{U}}' = \bar{\mathbf{U}} + \epsilon \tilde{\mathbf{J}} [\hat{\mathbf{b}} \cdot \mathbf{R} + \frac{1}{2} \hat{\mathbf{b}} \cdot (\nabla \times \hat{\mathbf{b}})] \quad (6.59)$$

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

Unfortunately, the variable $\tilde{\mathbf{X}}'$ 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.¹⁸ 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.²⁰⁻²⁶ 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 $(\bar{X}, U, J)$ 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 $\bar{X}$ is identical to $\bar{X}'$ of Eq. (6.58).

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

$$\begin{aligned}
\tilde{x} = x - \frac{\epsilon}{\Omega} \hat{b} \times \tilde{v} + \epsilon^2 \left\{ \frac{1}{8\Omega^3} \left[4v_{\perp} (v_{\perp} \cdot \nabla \Omega) + 9v_{\perp}^2 \hat{b} (\hat{b} \cdot \nabla \Omega) - 4v_{\perp}^2 \nabla \Omega \right. \right. \\
\left. \left. + 6\Omega \hat{b} (v_{\perp} \cdot \nabla \hat{b} \cdot v_{\perp}) \right] + \frac{u}{2} \left[\hat{b} \cdot (\nabla \times \hat{b}) \hat{b} \times \tilde{v} + 2\hat{b} (\hat{b} \cdot \nabla \hat{b} \cdot v_{\perp}) \right] \right\} \\
+ O(\epsilon^3)
\end{aligned} \tag{A.1}$$

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:

$$\begin{aligned}
U = u - \frac{\epsilon}{4\Omega} \left[2(\hat{b} \times \tilde{v}) \cdot \nabla \hat{b} \cdot v_{\perp} - v_{\perp}^2 \hat{b} \cdot (\nabla \times \hat{b}) + 4u \hat{b} \cdot \nabla \hat{b} \cdot (\hat{b} \times \tilde{v}) \right] \\
+ O(\epsilon^2)
\end{aligned} \tag{A.2}$$

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

$$\begin{aligned}
J = \frac{v_{\perp}^2}{2\Omega} + \frac{\epsilon}{4\Omega^3} \left\{ 2v_{\perp}^2 (\hat{b} \times \tilde{v}) \cdot \nabla \Omega + \Omega u \left[2(\hat{b} \times \tilde{v}) \cdot \nabla \hat{b} \cdot v_{\perp} - v_{\perp}^2 \hat{b} \cdot (\nabla \times \hat{b}) \right] \right. \\
\left. + 4\Omega u^2 \hat{b} \cdot \nabla \hat{b} \cdot (\hat{b} \times \tilde{v}) \right\} + O(\epsilon^2)
\end{aligned} \tag{A.3}$$

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 $\underline{x}$ and $\underline{v}$. Therefore all fields on the right hand sides, such as Ω and $\hat{b}$, are evaluated at the instantaneous particle position $\underline{x}$. 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 $K = \frac{1}{2} m v^2$ of the particle can be expressed in terms of the guiding center variables. The relation is

$$K = m[\Omega(\underline{X})J + \frac{1}{2} U^2] + O(\epsilon^2) \quad (A.4)$$

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

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

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 $d\underline{X}/dt$ which is in the direction $\hat{b}(\underline{X})$. This is

$$\left(\frac{d\underline{X}}{dt} \right)_{\parallel} = \hat{b}U + O(\epsilon^2) \quad (A.6)$$

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

$$\begin{aligned}
\left(\frac{d\tilde{\chi}}{dt}\right)_{\perp} = & \frac{\epsilon}{\Omega} \hat{b} \times (J \nabla \Omega + U^2 \hat{b} \cdot \nabla \hat{b}) + \epsilon^2 \hat{b} \times \left\{ -\frac{U^3}{\Omega} (\hat{b} \cdot \nabla \times \hat{b}) \hat{b} \cdot \nabla \hat{b} \right. \\
& + \frac{JU}{\Omega} \left[-\frac{\nabla \Omega}{\Omega} (\hat{b} \cdot \nabla \times \hat{b}) - \frac{1}{2} (\hat{b} \cdot \nabla \times \hat{b}) \hat{b} \cdot \nabla \hat{b} + \frac{1}{2} \nabla (\hat{b} \cdot \nabla \times \hat{b}) \right. \\
& \left. \left. - \hat{b} \times (\hat{b} \cdot \nabla \hat{b} \cdot \nabla \hat{b}) + (\nabla \cdot \hat{b}) \hat{b} \times (\hat{b} \cdot \nabla \hat{b}) \right] \right\} + O(\epsilon^3) \quad (A.7)
\end{aligned}$$

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

$$\begin{aligned}
\frac{JU}{\Omega} \hat{b} \times \left[\dots \right] = & \frac{JU}{\Omega} \left[\frac{\nabla \Omega \times (\nabla \times \hat{b})}{\Omega} - \frac{5}{2} (\nabla \cdot \hat{b}) \hat{b} \cdot \nabla \hat{b} - \frac{1}{2} \hat{b} \hat{b} : \nabla \nabla \hat{b} \right. \\
& \left. + \hat{b} \cdot \nabla \hat{b} \cdot \nabla \hat{b} - \frac{1}{2} \nabla (\nabla \cdot \hat{b}) + \frac{1}{2} \nabla^2 \hat{b} \right]_{\perp} \quad (A.8)
\end{aligned}$$

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

REFERENCES

1. R. G. Littlejohn, J. Math. Phys. 20, 2445(1979).
2. H. Alfven, 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. Meier (ORNL), private communication.
14. A. I. Morozov and L. S. Solov'ev, in Reviews of Plasma Physics, edited by M. A. Leontovich (Consultants Bureau, New York, 1966), vol. 2, p. 201.
15. A. H. Nayfeh, Perturbation Methods (Wiley, New York, 1973).
16. M. D. Kruskal, J. Math. Phys. 3, 806(1962).
17. R. J. Hastie, J. B. Taylor and F. A. Haas, Ann. Phys. 41, 302(1967).

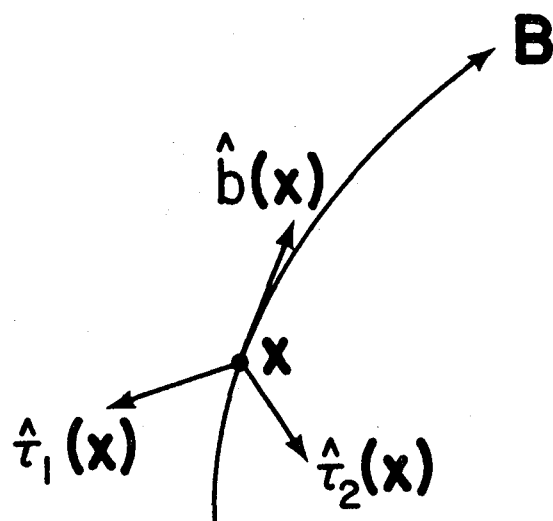
18. C. Grebogi, A. N. Kaufman, and R. G. Littlejohn, Phys. Rev. Lett. 43, 1668(1979).
19. T. G. Northrop, M. D. Kruskal, and C. S. Liu, Phys. Fluids 9, 1503(1966).
20. R. Dewar, Phys. Fluids 16, 1102(1973).
21. S. Johnston, Phys. Fluids 19, 93(1976).
22. R. Dewar, J. Phys. A9, 2043(1976).
23. J. R. Cary and A. N. Kaufman, Phys. Rev. Lett. 39, 402(1977).
24. S. Johnston and A. N. Kaufman, Phys. Rev. Lett. 40, 1266(1978).
25. S. Johnston, A. N. Kaufman, and G. L. Johnston, J. Plasma Phys. 20, 365(1978).
26. J. R. Cary, Ph. D. Thesis, University of California, Berkeley, 1979.
Available as Lawrence Berkeley Laboratory Report No. LBL-8155.

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	$2S_0$	$-\frac{1}{2} S_0$

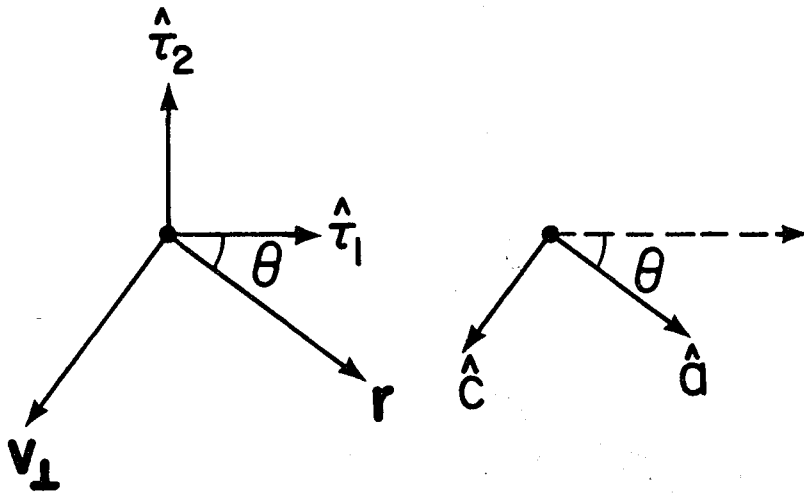
FIGURE CAPTIONS.

- Fig. 1. The three unit vectors $\hat{b}$, $\hat{r}_1$, and $\hat{r}_2$.
- 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. $\tilde{X}$ is the guiding center position.
- Fig. 4. A schematic illustration of a θ -characteristic. The $w=0$ "plane" in the diagram actually represents a four-dimensional surface in phase space.



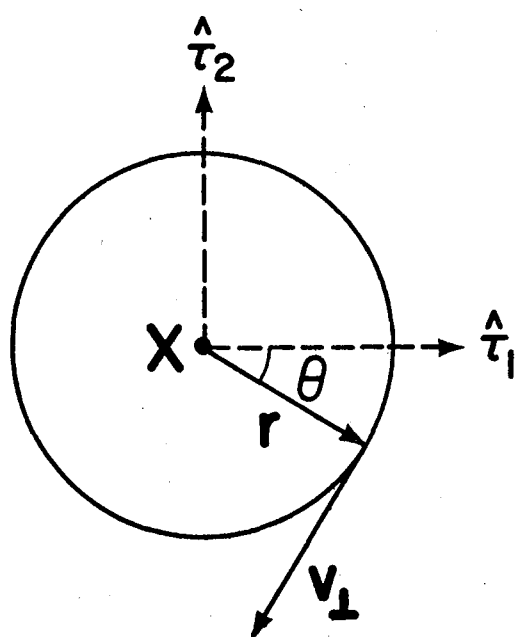
XBL 795-1522

Fig. 1.



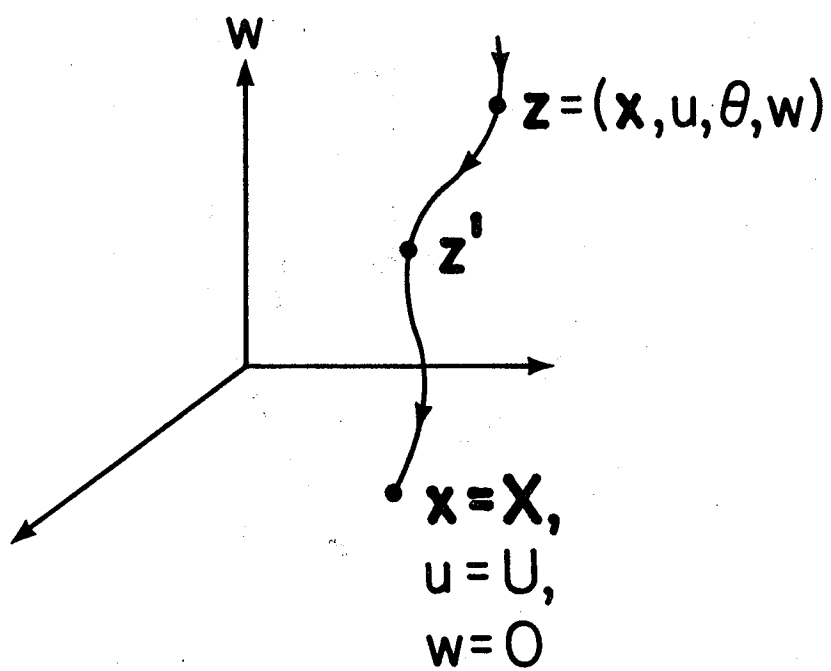
XBL 795-1521

Fig. 2.



XBL 795-1523

Fig. 3.



XBL 795-1524

Fig. 4.

CHAPTER IV

GUIDING CENTER BOUNCE MOTION

HAMILTONIAN THEORY OF GUIDING CENTER
BOUNCE MOTION*

Robert G. Littlejohn

Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

ABSTRACT

Guiding center bounce motion is analyzed with Hamiltonian methods, which include the use of Darboux's theorem and Lie transforms. The system studied is a nonrelativistic particle moving in a static magnetic field $\underline{B} = \underline{B}(x)$ with $\underline{E} = 0$. The averaged equations of motion and adiabatic invariant series are derived.

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

1. INTRODUCTION

In two previous papers^{1,2} I have given a Hamiltonian treatment of particle gyro-motion using Darboux's theorem and Lie transforms. In this paper the same methods are used to analyze the longitudinal oscillations of guiding center motion (the bounce motion); a brief Hamiltonian treatment of the drift motion comes quite easily at the end. As in Refs. 1-2, the problem definition includes the restriction to a nonrelativistic particle moving in a static magnetic field $B=B(x)$ with $E=0$.

This work parallels quite closely some previous work on bounce motion, especially that of Northrop, Liu, and Kruskal.³ The results derived here are identical to the results of those authors, although the methods of derivation are quite different. In particular, there is complete agreement on the averaged equations of field line motion and on the first correction to the bounce invariant. This correction was also worked out by Hastie, Taylor, and Haas,⁴ using an approach based on the Vlasov equation rather than single particle equations of motion. All three sets of results are in agreement.

The original treatment of bounce motion was given by Northrop and Teller,⁵ and Northrop⁶ has given an excellent review of the whole subject. Throughout this paper a familiarity with Northrop's review is assumed, and several techniques are drawn from this source without explicit reference.

The unique feature of this work concerns the Hamiltonian methods used to carry out the perturbation expansion. These are the methods developed in Ref. 1-2, and a familiarity with them is assumed here. Nevertheless, in several places improvements on the methods of Refs. 1-2

were called upon for this work, and they are described in detail.

The new Hamiltonian methods are a replacement for Kruskal's⁷ systematic adiabatic theory. In an actual calculation, they have the effect of replacing large amounts of mechanical algebra with a small amount of simple calculations, requiring nevertheless careful concentration.

Although the formalism of differential geometry is called upon from time to time in this work, it is for the sake of enrichment value, and is not a critical element in any proof or demonstration. Therefore this paper can be read quite satisfactorily without any knowledge of differential geometry. A physicist's standard knowledge of tensor calculus is, however, required.

For ease of comparison of this work with Refs. 2 and 3, Table I has been prepared showing notational differences.

A pervasive notational problem with this work concerns partial derivatives. There are at least six coordinate systems used in this paper, and a partial differential operator such as $\partial/\partial E$ can have a different meaning in different coordinate systems, depending on which other variables besides E are held fixed. This problem is solved in thermodynamics by explicitly indicating the variables which are to be held fixed, but such a solution would lead to extremely cumbersome formulas for this work. Therefore in this paper the variables to be held fixed are determined from the context. For example, when we are discussing the coordinate system (y, ψ, E) in Sec. 3.5, the operator $\partial/\partial E$ is taken at fixed (y, ψ) . In cases of possible confusion the variables to be held fixed are indicated in the text.

A similar problem concerns the components of covariant vectors and tensors. It is not enough, for example, to ask for the ψ -component of the vector ρ . One must also indicate what the other coordinates are, whether they be (y, E) or (Y, J) , for example. Therefore in reading this paper one must be careful to note the coordinate system involved, especially when dealing with the 1-form ρ and the 2-form ω .

Sec. 2 of this paper places the Hamiltonian method used here, which we call the Kruskal-Darboux-Lie method, in the context of general Hamiltonian systems. Sec. 3 analyzes the unperturbed bounce motion (i.e. in the limit $\epsilon \rightarrow 0$), and establishes several important coordinate systems and their properties. Sec. 4 carries out the Darboux transformation, and derives the Poisson tensor in the new coordinates. As in Refs. 1-2, the Poisson tensor appears as an exact expression, i.e. it is not a power series in ϵ . In Sec. 5 we carry out the averaging transformation using Lie transforms, and we derive the averaged equations of motion and the adiabatic invariant. Sec. 5 also gives a brief discussion of the drift motion. Finally, in Sec. 6 we discuss the significance of this work and suggest some extensions. Most importantly, this paper lays the groundwork for the study of resonances among the different degrees of freedom of charged particle motion, and ways of carrying this out are suggested.

2. THE KRUSKAL-DARBOUX-LIE METHOD FOR NEARLY PERIODIC SYSTEMS

2.1. Introduction

This section summarizes and coordinates some material from Refs. 1-2 which will describe an explicit program for the Hamiltonian treatment of nearly periodic systems. This program is a kind of Hamiltonian extension of Kruskal's theory,⁷ using Darboux's theorem to find an appropriate set of variables and Lie transforms to carry out the averaging transformation. Refs. 1-2 have already illustrated the program with the example of the nearly periodic motion of particle gyration, but at the risk of some repetition it is useful to summarize it here before proceeding with an analysis of the guiding center bounce motion.

There are two reasons for this. First, although Kruskal⁷ has given a clear exposition of his theory, the Hamiltonian extensions to it can be found in Refs. 1-2 only intertwined with the details of their application to the gyro-motion. Therefore a description of the new method is in order which emphasizes its generality. Second, when applying the Hamiltonian program to the case of guiding center bounce motion, it is easy to become confused by the algebraic details of the application, unless one keeps clearly in mind the general structure and purpose of the system of variable transformations. This was not the case for the gyro-motion, because for that problem the unperturbed oscillator is harmonic. For the bounce motion the unperturbed oscillator is in general anharmonic, and it depends on the unspecified function $B(x)$ for its description. As a result the bounce motion is potentially more confusing.

2.2. Kruskal's theory

It is appropriate to begin with a brief description of Kruskal's theory, the structure of which may be seen in Table II. Although Kruskal's theory works equally well for Hamiltonian and non-Hamiltonian systems, it is assumed here that the nearly periodic system in question is a Hamiltonian system with N degrees of freedom. Thus the variable set $(\underline{q}, \underline{p})$ in Table II represents two N -vectors of canonical coordinates in a phase space of dimensionality $2N$.

The notation used in Table II is based on that of Kruskal's original paper,⁷ which employs symbols in quite a different manner from either Refs. 1-2 or the main body of this paper. Kruskal's notation appears in this paper only in this section and in Table II and column (a) of Table III. Some care may be required to avoid confusion.

The first preparatory transformation, $(\underline{q}, \underline{p}) \rightarrow \underline{x}$, is simply a matter of convenience that may be useful in certain cases. The $2N$ -vector $\underline{x}$ is some set of phase space coordinates, which are possibly non-canonical. A non-Hamiltonian system begins with coordinates of this type, rather than $(\underline{q}, \underline{p})$.

The second preparatory transformation takes account of the nearly periodic nature of the system. The expression "nearly periodic" means that the equations of motion can be written in the form

$$\frac{d\underline{x}}{dt} = \underline{F}_0(\underline{x}) + \epsilon \underline{F}_1(\underline{x}) + \dots \quad (2.1)$$

and that the equations of motion for the unperturbed system

$$\frac{d\underline{x}}{dt} = \underline{F}_0(\underline{x}) \quad (2.2)$$

are solvable and yield only periodic orbits in phase space. Generically speaking, the phase space solutions to Eq. (2.1) look like the orbit shown in Fig. 1(a), while the solutions to Eq. (2.2) are topologically equivalent to circles, as illustrated by the two orbits shown in Fig. 1(b). These orbits are called loops by Kruskal.

The unperturbed phase space orbits provide a set of phase space coordinates which are descriptive of the unperturbed system in a natural way. Since a finite region (possibly all) of $2N$ -dimensional phase space is filled up by a $(2N-1)$ -parameter family of loops, it is natural to make some choice of $2N-1$ variables $\tilde{y}$ which select a particular loop out of the family. Regarded as functions of the $2N$ quantities $\tilde{x}$, the $2N-1$ quantities $\tilde{y}$ are constant along any given loop, and their time derivatives vanish with respect to the unperturbed system. Likewise, position along a given loop can be specified by some angle-like variable θ , which can be taken to run monotonically from 0 to 2π around the loop. The variables $(\tilde{y}, \theta)$ form the third coordinate system in Kruskal's scheme, as shown in Table II.

The equations of motion of the unperturbed system in the variables $(\tilde{y}, \theta)$ have the form

$$\begin{aligned}\frac{d\tilde{y}}{dt} &= 0 \\ \frac{d\theta}{dt} &= \psi_0(\tilde{y})\end{aligned}\tag{2.3}$$

Clearly, the $2N-1$ quantities $\tilde{y}$ are constants of the unperturbed motion. Since in general the frequency of the unperturbed motion differs from one loop to another, the function ψ_0 in Eq. (2.3) depends on $\tilde{y}$. The

fact that ψ_0 is independent of θ is not automatic, but depends on the definition of the angle θ , as discussed by Kruskal. Note that the angle θ evolves linearly in time with respect to the unperturbed system, with a period $2\pi/\psi_0(y)$. It may be seen that carrying out the second preparatory transformation of Table II involves solving the unperturbed system, finding all the constants of motion, and finding an angle whose time evolution is linear in the unperturbed system. By hypothesis, this can always be done.

When the full equations of motion (2.1) are expressed in the coordinates (y, θ) , there results a set of the form

$$\frac{dy}{d\tilde{t}} = 0 + \epsilon G_1(y, \theta) + \dots \quad (2.4)$$

$$\frac{d\theta}{d\tilde{t}} = \psi_0(y) + \epsilon \psi_1(y, \theta) + \dots$$

Because of the higher order terms which have appeared, the quantities y are not constants of the true motion. Also, because of the θ -dependence of these terms, the y - and θ -evolutions do not decouple from one another. In these circumstances it is often said that y and θ have "rapid oscillations" in their time evolution.

The next variable transformation in Kruskal's theory, indicated as the "averaging transformation" in Table II, is a near-identity transformation of the form $(y, \theta) \rightarrow (\bar{y}, \bar{\theta})$, or, more explicitly,

$$\begin{aligned} \bar{y}(y, \theta) &= y + \epsilon Y_1(y, \theta) + \dots \\ \bar{\theta}(y, \theta) &= \theta + \epsilon \Theta_1(y, \theta) + \dots \end{aligned} \quad (2.5)$$

The functions $\bar{Y}_1, \bar{\Theta}_1$, etc., are chosen so that two criteria are fulfilled. First, the transformation itself is periodic in $\bar{\theta}$, i.e. it is free of secular terms. Second, the equations of motion for the overbarred variables have the form

$$\begin{aligned}\frac{d\bar{y}}{dt} &= 0 + \epsilon H_1(\bar{y}) + \dots \\ \frac{d\bar{\theta}}{dt} &= \omega_0(\bar{y}) + \epsilon \omega_1(\bar{y}) + \dots\end{aligned}\tag{2.6}$$

where the essential point is that the right hand sides are independent of the angle $\bar{\theta}$. Kruskal's theory gives an iterative algorithm for the determination of the functions $\bar{Y}_n, \bar{\Theta}_n, \bar{H}_n, \bar{\omega}_n$, etc. It may be very laborious to carry out in practice.

An important result of the averaging transformation is that the time evolutions of $\bar{y}$ and $\bar{\theta}$ are decoupled, since the functions $\bar{H}_n$ and $\bar{\omega}_n$ in Eq. (2.6) are independent of $\bar{\theta}$. On account of this property, the new variables are said to be "free of rapid oscillations"; in this paper they will be called averaged. Kruskal called them nice variables. The overbar notation is a reminder that the variables in question are averaged.

An important mathematical advantage of using the averaged variables $\bar{y}$ is that their time evolution is governed by a system of only $2N-1$ differential equations, rather than $2N$ for the earlier variables x . Once the evolution of the variables $\bar{y}$ has been determined, that of $\bar{\theta}$ follows by a simple quadrature.

Actually, assuming that the original system is Hamiltonian, it is possible to reduce the system by two variables instead of just one.

To do this, one first expresses the original coordinates (q, p) as functions of the averaged variables $(\bar{y}, \bar{\theta})$. Then the quantity $\bar{J}$ is computed as a function of $\bar{y}$:

$$\bar{J}(\bar{y}) = \frac{1}{2\pi} \int_0^{2\pi} p(\bar{y}, \bar{\theta}) \cdot \frac{\partial}{\partial \bar{\theta}} q(\bar{y}, \bar{\theta}) d\bar{\theta} \quad (2.7)$$

This integral can be considered to be a line integral in phase space taken around the closed phase space curve $\bar{y} = \text{constant}$, which is called a ring by Kruskal. Kruskal's ring encircles one direction of the torus or cylinder-like surface in phase space on which the orbit lies. Further details concerning the topological properties of this surface may be found in Abraham and Marsden.⁸ Actually performing the ring integral can be quite tedious, as may be seen in the work of Northrop, Liu, and Kruskal.³

Kruskal has shown that the quantity $\bar{J}$ is a constant of the motion, and this fact allows the system to be reduced by one more variable. The last variable transformation shown in Table II is completed by finding $2N-2$ other functions $\bar{z}$ of $\bar{y}$, in addition to $\bar{J}(\bar{y})$. The variables $\bar{z}$ evolve according to a set of $2N-2$ differential equations.

2.3. The Darboux-Lie extension to Kruskal's theory

The structure of the Darboux-Lie extension to Kruskal's theory may be seen in column (a) of Table III. Again, a warning is in order concerning notation; the notation used here for a description of the general case is based on that of Ref. 7, and is independent of the remainder of this paper. By comparing column (a) of Table III with Table II it may be seen that the Darboux-Lie extension to Kruskal's

theory involves, roughly speaking, performing the last two variable transformations in reverse order.

Column (b) of Table III exemplifies the Kruskal-Darboux-Lie method with the problem of particle gyro-motion, the subject of Ref. 2, and may be compared with column (a) in order to see more clearly the meaning of the transformations. Column (c) shows the application to guiding center bounce motion, which will be discussed in later sections of this paper.

The first two preparatory transformations have exactly the same form and purpose in the Darboux-Lie extension to Kruskal's theory as they have in the original theory. Referring to column (b) of Table III, the purpose of the first preparatory transformation $(q,p) \rightarrow (\tilde{x},\tilde{v})$ may be seen to be the creation of a set of gauge independent phase coordinates. This is merely a matter of convenience. Likewise, the second preparatory transformation is concerned with the properties of the unperturbed orbits. For example, the $(2N-1)$ -vector $\tilde{y}$ in the general case corresponds to the five quantities $(\tilde{x},\tilde{u},\tilde{w})$ for the case of the gyro-motion, and these quantities are constants of the motion at lowest order.

The Darboux transformation, shown in column (a) of Table III as $(y,\theta) \rightarrow (\tilde{z},\theta,J)$, is designed to create two variables, θ and J , which are canonically conjugate and which are "canonically decoupled," one might say, from the remaining $2N-2$ variables $\tilde{z}$. More precisely, one demands that the following Poisson bracket relations be satisfied:

$$\begin{aligned} \{\theta,J\} &= \text{constant} \\ \{\theta,\tilde{z}\} &= \{J,\tilde{z}\} = 0 \end{aligned} \tag{2.8}$$

The attention which is paid to Poisson bracket relations is a characteristic feature of the Darboux-Lie extension to Kruskal's theory. Clearly, the $2N-2$ variables $\underline{z}$ in the general case correspond to the four variables (X,U) in the case of the gyro-motion.

Finally, the last variable transformation shown in Table III is the averaging transformation. This is a near-identity, symplectic transformation which is specified by its scalar Lie generators, in the manner illustrated in Refs. 1-2.

Let us now apply the Kruskal-Darboux-Lie method to the case of guiding center bounce motion.

3. THE PREPARATORY TRANSFORMATIONS

3.1. The guiding center variables

It is well known^{3,5,6} that for certain magnetic field configurations $B(x)$ guiding center motion exhibits a bouncing behavior, which is a kind of nearly periodic motion. In addition, guiding center motion can be represented by a Hamiltonian dynamical system, as shown in Ref. 2. Therefore guiding center motion is a suitable candidate for the Kruskal-Darboux-Lie method, as outlined above in Sec. 2. An equivalent analysis of this system, using Kruskal's non-Hamiltonian method, has been carried out by Northrop, Liu, and Kruskal.³

The guiding center variables will be denoted by (X,U,θ,M) in this and remaining sections of this paper. These are the averaged, not the intermediate, guiding center variables, in the terminology of Ref.2; the overbars have been dropped for notational convenience. Note that the adiabatic invariant of gyration is denoted here by M , rather than J , as indicated in Table I. The latter symbol is reserved

in this paper for the longitudinal invariant, as is customary in plasma physics.

The guiding center variables satisfy the following Poisson bracket relations, which specify the symplectic structure of the six-dimensional phase space of which they are coordinates:

$$\{\theta, M\} = 1/\epsilon \quad (3.1)$$

$$\{\theta, \underline{X}\} = \{\theta, U\} = \{M, \underline{X}\} = \{M, U\} = 0 \quad (3.2)$$

$$\{\underline{X}, U\} = \frac{\Omega^*}{\Omega} = \hat{b} + \frac{\epsilon U \hat{b} \times (\hat{b} \cdot \nabla \hat{b})}{\Omega^*} \quad (3.3)$$

$$\{\underline{X}, \underline{X}\} = \frac{\epsilon \hat{b} \times I}{\Omega^*} \quad (3.4)$$

In these equations $\hat{b}$ is the unit vector along the magnetic field B , I is the unit tensor, Ω is the signed gyrofrequency eB/mc , $\underline{\Omega}$ is the gyrofrequency vector $eB/mc = \hat{b}\Omega$, and Ω^* and $\underline{\Omega}^*$ are the Morozov and Solov'ev⁹ variables:

$$\Omega^* = \Omega + \epsilon U \nabla \times \hat{b} \quad (3.5)$$

$$\underline{\Omega}^* = \hat{b} \cdot \underline{\Omega}^* = \Omega + \epsilon U \hat{b} \cdot (\nabla \times \hat{b}) \quad (3.6)$$

All fields in Eqs. (3.1)-(3.4) are evaluated at the guiding center position $\underline{X}$, and ∇ means $\partial/\partial \underline{X}$. Finally, ϵ is a dimensionless parameter indicating order in the guiding center approximation. Physical results correspond to $\epsilon=1$.

Because the right hand sides of Eqs. (3.3) and (3.4) are independent of θ and M , these equations define a symplectic structure on a reduced phase space of four dimensions, which has the four variables

$(\underline{X}, U)$ for coordinates. This property is not an accident, but follows from the use of the Darboux algorithm, as explained in Ref. 1. It is the reduced phase space which will be of interest in this paper.

The time evolution of the guiding center variables is governed by the following Hamiltonian, in which the $O(\epsilon)$ term has been made to vanish by an appropriate choice for the origin of gyrophase, as explained in Ref. 2:

$$H(\underline{X}, U, M) = M\Omega(\underline{X}) + \frac{1}{2} U^2 + O(\epsilon^2) \quad (3.7)$$

The Hamiltonian is independent of θ , implying that M is a constant of the motion. Because of this property, the Hamiltonian can be considered to be a function on the reduced phase space of four dimensions described by the variables $(\underline{X}, U)$, with the constant M being taken as a parameter.

Sometimes it is useful to imagine the Hamiltonian in its formal limit as an infinite power series in ϵ , whereas for practical purposes it must be truncated after a finite number of terms. In either case the Hamiltonian can be thought of as a function on the reduced phase space parametrized by $(\underline{X}, U)$, and it gives rise to nearly periodic motion on that space. The unperturbed system, in the sense of Eqs. (2.1)-(2.2), depends only on the leading term in the Hamiltonian, so that the preparatory transformations indicated in Table III do not depend on how the Hamiltonian is truncated. Using the Poisson bracket relations, Eqs. (3.3)-(3.4), it is trivial to write out the differential equations of the unperturbed system in the variables $(\underline{X}, U)$:

$$\frac{d\tilde{x}}{dt} = \hat{b}U$$

(3.8)

$$\frac{dU}{dt} = - \hat{M}\hat{b} \cdot \nabla \Omega$$

For certain magnetic fields $\tilde{B}$ and in certain regions of phase space, this system is periodic, and it defines Kruskal's loops.

3.2. Transforming the Poisson tensor

In a Hamiltonian theory it is necessary to know in each coordinate system used, at least implicitly, the Poisson brackets of the coordinates among themselves. These Poisson brackets are the components σ^{ij} of the contravariant Poisson tensor with respect to the given coordinate system, and a straightforward computation of the new Poisson brackets under a change of coordinates is equivalent to the execution of the usual transformation law for contravariant tensors.

In the straightforward method of transforming the Poisson tensor, the Poisson brackets of the new coordinates among themselves are computed first. These emerge initially as functions of the old coordinates, and the second step is to express them as functions of the new coordinates. Finally, any partial derivative expressions appearing in the Poisson brackets, such as the operator ∇ in Eqs. (3.3)-(3.4), must be transformed to the new coordinates.

The straightforward method sufficed for the purposes of Refs. 1-2, but it quickly leads to very laborious calculations when applied to guiding center bounce motion. The following is a presentation of a short-cut method, which could have been used to advantage in Refs. 1-2, but which is nearly essential here. Instead of dealing with the Poisson tensor σ^{ij} , the short-cut method focuses on the Lagrange

tensor ω_{ij} , and even more so on the covariant vector ρ_i , of which ω_{ij} is the exterior derivative.

A brief review of the properties of the vector ρ_i and the tensors ω_{ij} and σ^{ij} is in order. If the phase space coordinates are z^i , then the Poisson tensor σ^{ij} is defined by $\sigma^{ij} = \{z^i, z^j\}$, with the usual definition for the Poisson bracket. Next, the Lagrange tensor ω_{ij} has components which form a matrix which is inverse to the matrix σ^{ij} :

$$\sum_{k=1}^{2N} \sigma^{ik} \omega_{kj} = \delta_j^i \quad (3.9)$$

This definition for the components ω_{ij} differs by a sign from that employed in Refs. 1-2, and it causes the quantity ω_{ij} to be the negative of the Lagrange bracket of z^i with z^j , according to the usual definition of the Lagrange bracket which is given in mechanics texts. The present sign convention is better, however, because it is more in accordance with the theory of differential forms as applied to mechanics.^{8,10}

The Lagrange tensor is closed, which means that it satisfies the following differential equation:

$$\frac{\partial \omega_{ij}}{\partial z^k} + \frac{\partial \omega_{jk}}{\partial z^i} + \frac{\partial \omega_{ki}}{\partial z^j} = 0 \quad (3.10)$$

The closedness of ω is equivalent to the Jacobi identity. Eq. (3.10) will prove to be a theorem of great power and utility in Secs. 4 and 5 of this paper. Because ω is closed, it can, according to a theorem called Poincaré's lemma, be written in terms of the derivatives of a covariant vector ρ_i , which will be called the distinguished 1-form:

$$\omega_{ij} = \frac{\partial \rho_j}{\partial z^i} - \frac{\partial \rho_i}{\partial z^j} \quad (3.11)$$

The essence of the short-cut method for finding the Poisson tensor under a change of coordinates is to transform first the components of the distinguished 1-form. Next, the Lagrange tensor is computed from Eq. (3.11), and then the Poisson tensor is computed from Eq. (3.9), i.e. by matrix inversion. Since ρ_i is a first rank tensor it is easier to transform than σ^{ij} , which is a second rank tensor, and in practice this advantage more than compensates for the necessity for matrix inversion. Actually, as will be shown below, the distinguished 1-form ρ_i can be written as a linear combination of the gradients of phase space scalars. Because the transformation properties of scalars are trivial, the result is an almost immediate way of writing down the components of the Poisson tensor in any coordinate system.

For later reference it is useful to tabulate here the relations between the quantities ω_{ij} and σ^{ij} which follow from Eq. (3.9), i.e. from matrix inversion. We let D represent an arbitrary invertible antisymmetric 4x4 matrix, and we let G be its inverse. Then

$$\begin{aligned} G_{12} &= -D_{34}/\Gamma \\ G_{13} &= +D_{24}/\Gamma \\ G_{14} &= -D_{23}/\Gamma \\ G_{23} &= -D_{14}/\Gamma \\ G_{24} &= +D_{13}/\Gamma \\ G_{34} &= -D_{12}/\Gamma \end{aligned} \quad (3.12)$$

where

$$\Gamma = D_{12}D_{34} - D_{13}D_{24} + D_{14}D_{23} \quad (3.13)$$

It is interesting to note that if ω is identified with D then the quantity Γ is proportional to the one independent component of the 4-form $\omega \wedge \omega$, i.e. the Liouville volume element.

It is interesting to examine the Lagrange tensor ω_{ij} and the distinguished 1-form ρ_i in the (X, U) coordinate system. It is also useful to do so, because before ρ_i can be found in an arbitrary coordinate system, it must be known in a given system.

The Lagrange tensor is found from Eqs. (3.3)-(3.4) and (3.12)-(3.13). To display the results it is convenient to let the indices i, j run over the numbers 1, 2, 3, corresponding to the components of X , and to let the index 4 correspond to the coordinate U . Then we have

$$\omega_{ij} = \frac{1}{\epsilon} e_{ijk} \Omega_k^* \quad (3.14)$$

$$\omega_{i4} = -b_i$$

where e_{ijk} is the Levi-Civita symbol.

Let us now find the distinguished 1-form ρ . This is a problem which is very similar to finding a vector potential A corresponding to a magnetic field B . A satisfactory solution is given by

$$\rho_i = \frac{e}{\epsilon mc} A_i + U b_i \quad (3.15)$$

$$\rho_4 = 0$$

The first three components of ρ are suggestive of a modified vector potential A^* , such as that introduced by Morozov and Solov'ev⁹:

$$\tilde{A}^* = \tilde{A} + \frac{\epsilon mc}{e} U \hat{b} \quad (3.16)$$

Likewise, the vanishing fourth component of ρ can be associated with a modified electrostatic potential ϕ^* by $0 = \rho_4 = -(e/\epsilon mc)\phi^*$. Then the verification of Eq. (3.11), using Eqs.(3.14)-(3.15), reduces to the following identities:

$$\begin{aligned} \tilde{\Omega}^* &= \frac{e}{mc} \nabla \times \tilde{A}^* \\ -\epsilon \hat{b} &= \frac{e}{mc} \left(-\frac{\partial \tilde{A}^*}{\partial U} - \nabla \phi^* \right) \end{aligned} \quad (3.17)$$

It is also interesting to write out the closedness property of ω for the specific form given in Eq. (3.14). The result is

$$\begin{aligned} \nabla \cdot \tilde{\Omega}^* &= 0 \\ \frac{\partial \tilde{\Omega}^*}{\partial U} &= \epsilon \nabla \times \hat{b} \end{aligned} \quad (3.18)$$

The obvious analogy between these equations and the well-known formulas of electromagnetic theory is both interesting and suggestive. It comes about because the electromagnetic field tensor $F_{\mu\nu}$, like ω_{ij} , represents a closed 2-form on a space of four dimensions. In the case of $F_{\mu\nu}$, the space in question is space-time, with coordinates (x, t) , and in the case of ω_{ij} the space is our four dimensional phase space, with coordinates (X, U) . Evidently, it would be suggestive to write $\tilde{E}^* = -(\epsilon mc/e)\hat{b}$ in Eqs. (3.17)-(3.18).

3.3. Field line coordinates; the transformation $(\underline{X}, U) \rightarrow (\underline{y}, s, U)$

According to the discussion of Sec. 2, we now seek a transformation to a set of phase space coordinates corresponding to the set $(\underline{y}, \theta)$ of column (a) of Table III. These variables will consist of three quantities which are constants of the unperturbed motion, and an angle evolving linearly in time. In practice, it is convenient to carry out this transformation in a number of steps, of which the first is a transformation to field line coordinates. By "field line coordinates" we mean a curvilinear coordinate system in physical space in which two of the three coordinates label field lines. Transforming to such a system has the advantage of immediately creating two constants of the unperturbed motion, since this motion, according to Eq. (3.8), is always parallel to field lines.

Two of the field line coordinates will be denoted by the 2-vector $\underline{y} = (y_1, y_2)$, which is related to the familiar Euler potentials¹¹ (α, β) of the guiding center position $\underline{X}$ by

$$\begin{aligned} y_1 &= \beta(\underline{X}) \\ y_2 &= \frac{e}{mc} \alpha(\underline{X}) \end{aligned} \quad (3.19)$$

Since $\underline{B} = \nabla\alpha \times \nabla\beta$, it follows that

$$\underline{\Omega}(\underline{X}) = \frac{\partial y_2}{\partial \underline{X}} \times \frac{\partial y_1}{\partial \underline{X}} \quad (3.20)$$

and that

$$\hat{\underline{b}} \cdot \frac{\partial \underline{y}_a}{\partial \underline{X}} = 0 \quad (3.21)$$

In this paper the letters a,b,c,d will always be used for indices which run over the numbers 1,2.

In association with the coordinates y it is convenient to employ the two-dimensional Levi-Civita tensor γ_{ab} , with $\gamma_{11}=\gamma_{22}=0$ and $\gamma_{12}=-\gamma_{21}=1$. For example, if the gauge $A = (1/2)(\alpha\nabla\beta - \beta\nabla\alpha)$ is chosen, then

$$\frac{e}{mc} \tilde{A}(\tilde{X}) = \frac{1}{2} \gamma_{ab} y_b \frac{\partial y_a}{\partial \tilde{X}} \quad (3.22)$$

Here and throughout this paper the summation convention is used, unless greater clarity is called for. The tensor γ_{ab} satisfies the following identities, which are of use later:

$$\begin{aligned} \gamma_{ab} \gamma_{cd} &= \delta_{ac} \delta_{bd} - \delta_{ad} \delta_{bc} \\ \gamma_{ac} \gamma_{cb} &= -\delta_{ab} \end{aligned} \quad (3.23)$$

A third spatial coordinate is conveniently chosen to be s , the distance along a field line relative to an arbitrary initial value surface:

$$s(\tilde{X}) = \int^{\tilde{X}} \hat{\tilde{b}}(\tilde{x}') \cdot d\tilde{x}' \quad (3.24)$$

It follows from this definition that s satisfies

$$\begin{aligned} \hat{\tilde{b}} \cdot \frac{\partial s}{\partial \tilde{X}} &= 1 \\ \hat{\tilde{b}} \cdot \frac{\partial \tilde{X}}{\partial s} &= 1 \end{aligned} \quad (3.25)$$

To complete the coordinate transformation $(\tilde{X}, U) \rightarrow (y, s, U)$ we first

work out the equations of motion of the unperturbed system. From Eq. (3.8), these are

$$\frac{dy}{dt} = 0$$

$$\frac{ds}{dt} = U \quad (3.26)$$

$$\frac{dU}{dt} = -M \frac{\partial \Omega}{\partial s}$$

Second, the distinguished 1-form is transformed to the new coordinates. This transformation can be carried out using the usual transformation law for covariant vectors, but it is easier to observe that in the gauge given by Eq. (3.22), the distinguished 1-form is a linear combination of phase space gradients. That is, with $z^i = (X, U)$, Eqs. (3.15) and (3.22) are equivalent to

$$\rho_i = \frac{1}{2\epsilon} \gamma_{ab} y_b \frac{\partial y_a}{\partial z^i} + U \hat{b} \cdot \frac{\partial X}{\partial z^i} \quad (3.27)$$

where $i=1,2,3,4$. This is a manifestly covariant equation, in which the five quantities $(y_1, y_2, X_1, X_2, X_3)$ are being treated as phase space scalars, and hence it is valid in any coordinate system z^i .

In particular, in the coordinate system (y, s, U) , Eq. (3.27) gives

$$\begin{aligned} \rho_a &= \frac{1}{2\epsilon} \gamma_{ab} y_b + U \hat{b} \cdot \frac{\partial X}{\partial y_a} \\ \rho_s &= U \\ \rho_U &= 0 \end{aligned} \quad (3.28)$$

In the first of these equations, ρ_a stands for the y_a component of ρ . This notation avoids the use of multi-level subscripts, and it will be followed throughout this paper. Eq. (3.25) has been used to simplify the component ρ_s . If the Poisson brackets were needed in these coordinates, it would be easy to derive them from the components of ρ given above.

3.4. The variable E ; the transformation $(y,s,U) \rightarrow (y,s,E)$

In the unperturbed system of Eq. (3.26) the variables y are constants of the motion and the variables s and U form a potential system of one degree of freedom. Therefore this system possesses an energy-like integral, supplying a third constant of the unperturbed motion:

$$E(y,s,U) = \frac{1}{2} U^2 + M\Omega(y,s) \quad (3.29)$$

The quantity E is not a constant of the true, perturbed motion, and it is to be treated here as a variable or new phase space coordinate. The three variables (y,E) correspond to the variables y of column (a) of Table III.

Eq. (3.29) may be used to eliminate U in favor of E , thus effecting the transformation $(y,s,U) \rightarrow (y,s,E)$. The coordinate transformation is not one-to-one, because of the double root for U :

$$U(y,s,E) = \pm \{2[E - M\Omega(y,s)]\}^{1/2} \quad (3.30)$$

It is convenient to imagine that the one coordinate system (y,s,U) breaks up into two patches when variables are changed to the (y,s,E) system. The upper sign in Eq. (3.30) corresponds to the "upper

patch," meaning $U > 0$, and similarly for the lower patch. The same convention will be followed throughout this paper wherever a double sign appears. If an equation is quoted without a double sign, then it may be presumed to be valid in both patches.

Let us assume that for some value of y the quantity $M\Omega(y, s)$, considered as a function of s , has the form shown schematically in Fig. 2. In particular, we assume that at $s = s_m$, $M\Omega$ has a minimum. Recall from Ref. 2 that $\text{sign}(M) = \text{sign}(\Omega) = \text{sign}(e)$, so that a minimum in $M\Omega$ is also a minimum in the magnetic field strength $B(y, s)$. The quantity s_m is to be regarded as a function of y , and we will write $\Omega_m(y)$ for $\Omega(y, s_m(y))$. It follows that E must be greater than $M\Omega_m(y)$, and if there are two turning points of the unperturbed motion, as shown in Fig. 2, then this motion is periodic. It will not, however, be harmonic in general, so we are obliged to deal with a nonlinear oscillator. Note that there are other ways that the unperturbed motion can be periodic, for example if $\Omega(y, s)$ is periodic in s . These cases will not be considered here, however.

The turning points s_i , $i=0,1$, are the roots of

$$E = M\Omega(y, s_i) \quad (3.31)$$

It is convenient to order the roots so that $s_0 < s_1$. The quantities s_i are regarded as functions of (y, E) , and in the unperturbed motion $U=0$ when $s = s_i(y, E)$. In this paper we will only consider a range of parameters such that Eq. (3.31) has two roots.

3.5. The variable ψ ; the transformation $(y, s, E) \rightarrow (y, \psi, E)$

Fig. 3 shows the s - U plane for the unperturbed system, for some

specific field line with coordinates $\tilde{y}$. The unperturbed motion takes place on the closed curves $E=\text{constant}$, two of which are shown in Fig. 3. These curves are symmetric about the line $U=0$, and they form a family which is topologically equivalent to a set of concentric circles. The point $s=s_m(\tilde{y})$, $U=0$ is a fixed point of the unperturbed motion, and it forms the center of the family of curves $E=\text{constant}$.

It is convenient to define ψ as the phase of the unperturbed oscillator, taking on the value 0 at the point $s=s_0(\tilde{y}, E)$, $U=0$, the value π at the point $s=s_1(\tilde{y}, E)$, $U=0$, and approaching 2π as the unperturbed phase point returns to $s=s_0(\tilde{y}, E)$, $U=0$. Two contour lines of ψ are shown in Fig. 3; considered as a function of (s, U) or (s, E) , ψ has a branch point at $s=s_m$, $U=0$.

Let us denote the frequency of the unperturbed oscillator by $\omega_0(\tilde{y}, E)$. Then it follows from Eq. (3.26) that

$$\frac{1}{\omega_0(\tilde{y}, E)} = \pm \frac{1}{\pi} \int_{s_0(\tilde{y}, E)}^{s_1(\tilde{y}, E)} \frac{ds'}{U(\tilde{y}, s', E)} \quad (3.32)$$

Similarly, ψ itself is given by

$$\psi(\tilde{y}, s, E) = \omega_0(\tilde{y}, E) \int_{s_0(\tilde{y}, E)}^s \frac{ds'}{U(\tilde{y}, s', E)} \quad (3.33)$$

in the upper patch, and 2π plus the expression above in the lower patch. Eq. (3.33) can be used, at least in principle, to eliminate s in favor of ψ , thus bringing about the transformation $(\tilde{y}, s, E) \rightarrow (\tilde{y}, \psi, E)$. The following identities follow immediately from Eq. (3.33), and are useful in carrying out this transformation:

$$\frac{\partial \psi}{\partial s} = \frac{\omega_0(y, E)}{U(y, s, E)} \quad (3.34)$$

$$\frac{\partial s}{\partial \psi} = \frac{U(y, s, E)}{\omega_0(y, E)} \quad (3.35)$$

In addition, it is useful to note that although the guiding center position $\tilde{X}$ depends on both ψ and E in the new coordinates, it does so only through the function $s(y, \psi, E)$. Therefore we have

$$\frac{\partial \tilde{X}}{\partial \psi} = \hat{b} \frac{\partial s}{\partial \psi} = \hat{b} \frac{U}{\omega_0} \quad (3.36)$$

$$\frac{\partial \tilde{X}}{\partial E} = \hat{b} \frac{\partial s}{\partial E} \quad (3.37)$$

In a similar manner, if f and g are any two functions depending only on $\tilde{X}$, then they satisfy the differential equation

$$\frac{\partial f}{\partial \psi} \frac{\partial g}{\partial E} = \frac{\partial f}{\partial E} \frac{\partial g}{\partial \psi} \quad (3.38)$$

Using these relations, it is easy to work out the components of the distinguished 1-form in the coordinates (y, ψ, E) . Here we call upon Eq. (3.27). The result is

$$\rho_a = \frac{1}{2\epsilon} \gamma_{ab} y_b + U \hat{b} \cdot \frac{\partial \tilde{X}}{\partial y_a}$$

$$\rho_\psi = U \frac{\partial s}{\partial \psi} = \frac{U^2}{\omega_0} \quad (3.39)$$

$$\rho_E = U \frac{\partial s}{\partial E}$$

The derivatives in these expressions are taken with respect to the (y, ψ, E) coordinates.

Finally, there is a relation of importance connecting the action of the unperturbed oscillator with its frequency. We define the action $I(y, E)$ as follows:

$$I(y, E) = \pm \frac{1}{\pi} \int_{s_0(y, E)}^{s_1(y, E)} U(y, s', E) ds' \quad (3.40)$$

Then it follows, using Eqs. (3.30) and (3.32), that

$$\frac{\partial I}{\partial E}(y, E) = \frac{1}{\omega_0(y, E)} \quad (3.41)$$

This result is not as obvious as it looks, because the integrand of Eq. (3.40) is not differentiable at the endpoints. Nevertheless it is true and it can be justified rather easily.

The unperturbed system (3.8) possesses an important symmetry, which finds simple expression in the coordinates (y, ψ, E) . If the original coordinates (X, U) are expressed as functions of (y, ψ, E) , then we have

$$X(y, -\psi, E) = +X(y, \psi, E) \quad (3.42)$$

$$U(y, -\psi, E) = -U(y, \psi, E) \quad (3.43)$$

This is obviously a kind of time-reversal symmetry, since ψ evolves linearly in time in the unperturbed system. It is a curious fact that this system possesses time-reversal symmetry, since neither the perturbed guiding center equations, described by Eqs. (3.1)-(3.4)

and (3.7), nor the original charged particle motion from which the guiding center equations were derived possesses such symmetry. Important use of Eqs. (3.42)-(3.43) will be made as we proceed.

Referring to Table III, column (c), it may be seen that the introduction of the angle ψ completes the second preparatory transformation, taking us altogether from the original coordinates (X, U) to the coordinates (y, ψ, E) . Let us now turn to the Darboux transformation.

4. THE DARBOUX TRANSFORMATION

4.1. The Poisson tensor in the coordinates (y, ψ, E)

The Darboux transformation will follow from treating the angle ψ as a Hamiltonian and determining the phase space trajectories (the ψ -characteristics) which result. The Poisson tensor in the (y, ψ, E) coordinates is required for this analysis, since Poisson brackets are used to determine equations of motion.

According to the discussion of Sec. 3, we will proceed from the distinguished 1-form ρ to the Lagrange tensor ω to the Poisson tensor σ . The Lagrange tensor ω has a linear dependence on the distinguished 1-form ρ , and the latter consists of an $O(1/\epsilon)$ term and an $O(1)$ term, as can be seen in Eq. (3.39). Therefore ω also breaks up into two such terms when ordered in ϵ , so that it is convenient to write

$$\omega_{ij} = \frac{1}{\epsilon} \mu_{ij} + \nu_{ij} \quad (4.1)$$

The tensors μ and ν are both closed (see Eq. (3.10)), and important consequences will follow from this fact as we proceed.

By drawing on Eq. (3.11), the components of μ and ν can be written down immediately. The only non-vanishing component of μ is

$$\mu_{ab} = -\gamma_{ab} \quad (4.2)$$

The notation here is like that of Eq. (3.28); μ_{ab} represents the (y_a, y_b) component of μ . It is curious to note that although μ is closed, it is not symplectic, since $\det(\mu_{ij})=0$.

The components of v can be found similarly. They are

$$v_{ab} = \frac{\partial}{\partial y_a} \left(\hat{U}b \cdot \frac{\partial X}{\partial y_b} \right) - \frac{\partial}{\partial y_b} \left(\hat{U}b \cdot \frac{\partial X}{\partial y_a} \right) \quad (4.3a)$$

$$v_{a\psi} = \frac{\partial}{\partial y_a} \left(U \frac{\partial s}{\partial \psi} \right) - \frac{\partial}{\partial \psi} \left(\hat{U}b \cdot \frac{\partial X}{\partial y_a} \right) \quad (4.3b)$$

$$v_{aE} = \frac{\partial}{\partial y_a} \left(U \frac{\partial s}{\partial E} \right) - \frac{\partial}{\partial E} \left(\hat{U}b \cdot \frac{\partial X}{\partial y_a} \right) \quad (4.3c)$$

$$v_{\psi E} = \frac{\partial}{\partial \psi} \left(U \frac{\partial s}{\partial E} \right) - \frac{\partial}{\partial E} \left(U \frac{\partial s}{\partial \psi} \right) \quad (4.3d)$$

Of these, a significant and important simplification occurs in (4.3d). Expanding out the derivatives and using Eqs. (3.30), (3.38), (3.35), and (3.41), we have

$$v_{\psi E} = -\frac{1}{\omega_0} = -\frac{\partial I}{\partial E} \quad (4.4)$$

The components of the tensor v have the following symmetry properties:

$$\begin{aligned} v_{ab}(-\psi) &= -v_{ab}(\psi) \\ v_{a\psi}(-\psi) &= +v_{a\psi}(\psi) \\ v_{aE}(-\psi) &= -v_{aE}(\psi) \\ v_{\psi E}(-\psi) &= +v_{\psi E}(\psi) \end{aligned} \quad (4.5)$$

These follow simply from Eqs. (3.42)-(3.43). Recall that y and s are functions only of $\underline{x}$, and hence are even under $\psi \rightarrow -\psi$.

To find the Poisson tensor we may return to Eqs. (3.12) and identify ω with the matrix D and σ with the matrix G . Then the equation for Γ , Eq. (3.13), may be written in terms of the tensor γ_{ab} :

$$\Gamma = \gamma_{ab} \left(\frac{1}{2} \omega_{ab} \omega_{\psi E} + \omega_{aE} \omega_{b\psi} \right) \quad (4.6)$$

Substituting Eqs. (4.1)-(4.4) into this and simplifying gives

$$\Gamma = \frac{1}{\epsilon \omega_0} (1 + \epsilon \Delta) \quad (4.7)$$

where

$$\Delta = \gamma_{ab} (\omega_0 v_{aE} v_{b\psi} - \frac{1}{2} v_{ab}) \quad (4.8)$$

With the help of Eq. (3.12) the Poisson tensor may now be computed.

The result is

$$\begin{aligned} \{y_a, y_b\} &= \epsilon \gamma_{ab} / (1 + \epsilon \Delta) \\ \{y_a, \psi\} &= \epsilon \omega_0 \gamma_{ab} v_{bE} / (1 + \epsilon \Delta) \\ \{y_a, E\} &= -\epsilon \omega_0 \gamma_{ab} v_{b\psi} / (1 + \epsilon \Delta) \\ \{\psi, E\} &= \omega_0 \left(1 - \frac{1}{2} \epsilon \gamma_{ab} v_{ab} \right) / (1 + \epsilon \Delta) \end{aligned} \quad (4.9)$$

4.2. The ψ -characteristics

The ψ -characteristics result from treating ψ as a Hamiltonian. We are also interested in the function $J(y, \psi, E)$, which is to satisfy $\{\psi, J\} = 1$. Setting $d/d\lambda = \{ \cdot, \psi \}$, we have

$$\frac{dy_a}{d\lambda} = \frac{\epsilon \omega_0 \gamma_{ab} v_b E}{1 + \epsilon \Delta} \quad (4.10a)$$

$$\frac{dE}{d\lambda} = - \frac{\omega_0 (1 - \frac{1}{2} \epsilon \gamma_{ab} v_{ab})}{1 + \epsilon \Delta} \quad (4.10b)$$

$$\frac{dJ}{d\lambda} = -1 \quad (4.10c)$$

Before proceeding to an analytic treatment of these equations, it is valuable to fix in mind a qualitative picture of the ψ -characteristics, which always lie in contour surfaces of the function ψ . These are three-dimensional surfaces in four-dimensional phase space, and their general nature can be seen in Fig. 3, in which the two coordinates y are suppressed. The two-dimensional surface $s = s_m(y)$, $U=0$, which appears as a single point in Fig. 3, will be called the "initial value surface." This surface can also be characterized by the equation $E = M\Omega_m(y)$, so that it is not a surface of constant E . As the parameter λ increases, the ψ -characteristics converge inward toward the initial value surface, passing through decreasing values of E . This conclusion follows from the equation above for $dE/d\lambda$, Eq. (4.10b), which shows that for ϵ small, E is a monotonically decreasing function of λ . When viewing Fig. 3 it is important to remember that this figure is a cross-cut taken at constant y , and that the quantities y suffer an $O(\epsilon)$ evolution along the ψ -characteristics, according to Eq. (4.10a). Therefore the lines $\psi = \text{constant}$ shown in Fig. 3, representing three-dimensional surfaces in phase space, differ from the exact ψ -characteristics by $O(\epsilon)$.

Another picture of a ψ -characteristic is given in Fig. 4, this

time in the coordinates (y, ψ, E) . The coordinate ψ is suppressed in the drawing, since this coordinate does not change along ψ -characteristics, and only the three coordinates (y, E) are shown.

Fig. 4 shows why the initial value surface is so called. The quantities $\underline{Y}$ are the values y take on when λ is such that the ψ -characteristic is on the initial value surface. That is, $\underline{Y}$ are the initial values for the functions $y(\lambda)$. Similarly, it is convenient to let the function $J(\lambda)$ take on the value zero on the initial value surface. This is a reasonable convention, since, as may be seen in Fig. 3, the action I of the unperturbed oscillator vanishes on the initial value surface.

The analogy between these definitions and those of Ref. 2 should be clear. In particular, the relation between the variables y and $\underline{Y}$ here is the same as that between (x, u) and (X, U) in Ref. 2.

4.3. The Poisson tensor in the coordinates $(\underline{Y}, \psi, J)$

The solution to Eqs. (4.10), subject to the given initial conditions, gives the Darboux transformation $(y, \psi, E) \rightarrow (\underline{Y}, \psi, J)$. Before obtaining explicit formulas for the Darboux transformation, let us evaluate the Poisson tensor in the new coordinates. As in Refs. 1-2, the results will be exact, i.e. they are represented by formulas in closed form, rather than by infinite power series.

In Ref. 2 the Poisson tensor in the coordinates (X, U, θ, M) was evaluated by examining the behavior of the θ -characteristics near the initial value surface. The analogous strategy does not work well for this problem, because there is no natural expansion parameter to describe the ψ -characteristics near the initial value

surface. Such a parameter can be introduced artificially, but then one is led through a calculation which is complicated by numerous details which are specific to the parameter chosen. These details show their superfluity by vanishing from consideration once the final result is obtained. Therefore in the following we will present an alternate method of deriving the Poisson tensor, a method which could have been used to advantage in Refs. 1-2. Once again, the key is to focus on the Lagrange tensor, rather than the Poisson tensor.

Of the six independent components of the Poisson tensor in the (Y, ψ, J) coordinates, three are built into the construction of the ψ -characteristics. These are $\{Y, \psi\}=0$ and $\{\psi, J\}=1$. Using these known Poisson brackets, Eqs. (3.12)-(3.13) allow us to write down the following equations for the components of the Lagrange tensor in the coordinate system (Y, ψ, J) :

$$\begin{aligned}\omega_{12} &= -1/\{Y_1, Y_2\} \\ \omega_{a\psi} &= \gamma_{ab}\{Y_b, J\}/\{Y_1, Y_2\} \\ \omega_{aJ} &= 0 \\ \omega_{\psi J} &= -1\end{aligned}\tag{4.11}$$

Here the indices 1,2,a refer to the components of Y , not y . In the following we will determine first the three unknown components of the Lagrange tensor, ω_{12} and $\omega_{a\psi}$, and then complete the Poisson tensor, using the equations above.

Using Poisson's theorem, it is easy to show that all the

components of the Poisson tensor in the coordinates (Y, ψ, J) are constant along ψ -characteristics. The proof of this fact was given in Ref. 1, and need not be repeated here. Since the components of the Lagrange tensor are functions of the components of the Poisson tensor, these also, with respect to the coordinates (Y, ψ, J) , are constant along ψ -characteristics. Therefore the unknown components ω_{12} and $\omega_{a\psi}$ may be found at any point of phase space by determining their values on the initial value surface.

Of the four components of the distinguished 1-form ρ in the coordinates (Y, ψ, J) , the J -component is not needed in order to determine the unknown components of the Lagrange tensor. By using Eq. (3.27), the other three components of ρ may be written down:

$$\begin{aligned}\rho_a &= \frac{1}{2\varepsilon} \gamma_{bc} y_c \frac{\partial y_b}{\partial Y_a} + \hat{U} b \cdot \frac{\partial \hat{X}}{\partial Y_a} \\ \rho_\psi &= \frac{1}{2\varepsilon} \gamma_{bc} y_c \frac{\partial y_b}{\partial \psi} + \hat{U} b \cdot \frac{\partial \hat{X}}{\partial \psi}\end{aligned}\tag{4.12}$$

These expressions simplify on the initial value surface. First, we have $U=0$ on the initial value surface. Next, note that the differential operators $\partial/\partial\psi$ and $\partial/\partial Y_a$, taken in the coordinate system (Y, ψ, J) , imply that J is held constant. Since the initial value surface is $J=0$, these operators, when evaluated on the initial value surface, represent directional derivatives which lie in the initial value surface. Therefore if a function is known on the initial value surface, but not necessarily anywhere else, it is possible to evaluate the action of the operators $\partial/\partial\psi$ and $\partial/\partial Y_a$ on this function at points on the initial value surface. Since

$\underline{y}=\underline{Y}$ on the initial value surface, these arguments imply

$$\rho_a = \frac{1}{2\epsilon} \gamma_{ab} Y_b \quad (4.13)$$

$$\rho_\psi = 0$$

on the initial value surface.

This in turn gives

$$\omega_{ab} = -\frac{1}{\epsilon} \gamma_{ab} \quad (4.14)$$

$$\omega_{a\psi} = 0$$

These are valid on the initial value surface, and hence everywhere else in phase space as well.

Finally, by using Eq. (4.11) all these results may be assembled to give the complete Poisson tensor in the coordinates $(\underline{Y}, \psi, J)$:

$$\{Y_a, Y_b\} = \epsilon \gamma_{ab} \quad (4.15a)$$

$$\{Y, \psi\} = \{Y, J\} = 0 \quad (4.15b)$$

$$\{\psi, J\} = 1 \quad (4.15c)$$

The especially simple form of Eq. (4.15a) comes about because of the simple relation between $\underline{y}$ and the Euler Potentials (α, β) , which satisfy $\underline{B} = \nabla\alpha \times \nabla\beta$. If $\underline{y}$ had been defined in terms of some arbitrary pair of field line labels (what Stern¹¹ has called "unmatched" Euler Potentials), then the right hand side of Eq. (4.15a) would have been some nonconstant function of $\underline{Y}$.

4.4. Explicit formulas for the Darboux transformation

Let us return to Eqs. (4.10) and develop an explicit solution representing the ψ -characteristics. As in Refs. 1-2, the solution will be developed as a power series in ϵ . In Eqs. (4.10), the independent variable λ may be eliminated in favor of E , since the two are monotonic functions of one another. Expanding the resulting differential equations in powers of ϵ gives

$$\frac{dy_a}{dE} = -\epsilon \gamma_{ab} v_{bE} + O(\epsilon^2) \quad (4.16)$$

$$\frac{dJ}{dE} = \frac{\partial I}{\partial E} + \epsilon \gamma_{ab} v_{aE} v_{b\psi} + O(\epsilon^2)$$

Eqs. (3.41) and (4.8) have been used in writing the results in this form.

Through $O(\epsilon)$, Eq. (4.16) may be integrated immediately. Much less sophistication is called for here than in the integration of the Darboux equations in Ref. 2. It is convenient to introduce the two-vector F , defined by

$$F_a(\underline{y}, \psi, E) = \int_{M\Omega_m(\underline{y})}^E v_{aE}(\underline{y}, \psi, E') dE' \quad (4.18)$$

In terms of F , the solution to Eq. (4.16) can be written

$$y_a(\underline{y}, \psi, E) = Y_a - \epsilon \gamma_{ab} F_b(\underline{y}, \psi, E) + O(\epsilon^2) \quad (4.19)$$

or

$$Y_a(\underline{y}, \psi, E) = y_a + \epsilon \gamma_{ab} F_b(\underline{y}, \psi, E) + O(\epsilon^2) \quad (4.20)$$

The quantities F satisfy the following useful identities:

$$\frac{\partial F_a}{\partial E} = v_{aE} \quad (4.21)$$

$$\frac{\partial F_a}{\partial \psi} = v_{a\psi} - \frac{\partial I}{\partial y_a} \quad (4.22)$$

$$\frac{\partial F_a}{\partial y_b} - \frac{\partial F_b}{\partial y_a} = v_{ab} \quad (4.23)$$

The proof of these identities will be deferred for a moment.

In Eqs. (4.17) all the terms shown are evaluated at (y, ψ, E) . Since y is not a constant along ψ -characteristics, we may use Eq. (4.19) to eliminate y in favor of Y , which is constant.

$$\frac{dJ}{dE} = \frac{\partial I}{\partial E} + \epsilon \gamma_{ab} \left(F_a \frac{\partial^2 I}{\partial E \partial Y_b} + v_{aE} v_{b\psi} \right) + O(\epsilon^2) \quad (4.24)$$

In this equation, all terms are evaluated at (Y, ψ, E) , so that integration is immediate. An integration by parts may be performed on the term in F , and by using Eqs. (4.21)-(4.22) the result can be written as

$$J(Y, \psi, E) = I(Y, E) + \epsilon \left(\gamma_{ab} F_a \frac{\partial I}{\partial Y_b} + G(Y, \psi, E) \right) + O(\epsilon^2) \quad (4.25)$$

where

$$G(Y, \psi, E) = \gamma_{ab} \int_{M\Omega_m(Y)}^E \frac{\partial F_a}{\partial E} \frac{\partial F_b}{\partial \psi} dE' \quad (4.26)$$

Finally, using Eq. (4.20) to revert to the variables (y, ψ, E) , we have

$$J(y, \psi, E) = I(y, E) + \epsilon G(y, \psi, E) + O(\epsilon^2) \quad (4.27)$$

Along with Eq. (4.20), this equation specifies the Darboux transformation, $(y, \psi, E) \rightarrow (Y, \psi, J)$.

The symmetry properties of F and G are important. From their definitions and from Eqs. (4.5) we have

$$\begin{aligned} F(y, -\psi, E) &= -F(y, \psi, E) \\ G(y, -\psi, E) &= -G(y, \psi, E) \end{aligned} \quad (4.28)$$

To obtain the inverse of the Darboux transformation, it is convenient to introduce the function $W(y, I)$, defined by

$$E = W(y, I(y, E)) \quad (4.29)$$

W is merely the function which gives the energy of the unperturbed oscillator in terms of its action. It satisfies

$$\frac{\partial W}{\partial I}(y, I) = \omega_0 \quad (4.30)$$

$$\frac{\partial W}{\partial y}(y, I) = -\omega_0 \frac{\partial I}{\partial y}(y, E) \quad (4.31)$$

where the independent variables given indicate which variables are to be held fixed in the differentiation process. Using these relations, Eq. (4.25) may be inverted to give

$$E(Y, \psi, J) = W(Y, J) - \epsilon \omega_0 \left(\gamma_{ab} F_a \frac{\partial I}{\partial Y_b} + G \right) + O(\epsilon^2) \quad (4.32)$$

In this equation the E argument of the $O(\epsilon)$ terms, e.g. $\omega_0(\underline{Y}, E)$, is evaluated at $W(\underline{Y}, J)$. Along with Eq. (4.19), this equation gives the inverse of the Darboux transformation. It also gives the Hamiltonian, through $O(\epsilon)$, in the variables $(\underline{Y}, \psi, J)$, because of Eq. (3.7):

$$H(\underline{y}, \psi, E) = E + O(\epsilon^2) \quad (4.33)$$

$$H(\underline{Y}, \psi, J) = W(\underline{Y}, J) - \epsilon \omega_0 \left(\gamma_{ab} F_a \frac{\partial I}{\partial Y_b} + G \right) + O(\epsilon^2) \quad (4.34)$$

4.5. A proof of Eqs. (4.21)-(4.23)

Let us return to Eqs. (4.21)-(4.23) and supply their proofs, as promised. Eq. (4.21) follows immediately from the definition of F , Eq. (4.18). Eq. (4.22) also follows from the definition, but it requires more work. Directly differentiating Eq. (4.18) with respect to ψ and using the closedness of v (see Eq. (3.10)) give

$$\frac{\partial F_a}{\partial \psi} = \int_{M\Omega_m(\underline{y})}^E \frac{\partial}{\partial E} \left(v_{a\psi} - \frac{\partial I}{\partial y_a} \right) dE' \quad (4.35)$$

The quoted result follows if it can be shown that the quantity in the parentheses vanishes at the lower limit $E=M\Omega_m(\underline{y})$, which represents, of course, the initial value surface.

To see this, it is convenient to change from the variable E to the new variable E_1 , given by

$$E_1 = E - M\Omega_m(\underline{y}) \quad (4.36)$$

The advantage of the variable set $(\underline{y}, \psi, E_1)$ is that in this set, the

operators $\partial/\partial y$ and $\partial/\partial \psi$, when evaluated on the initial value surface, are directional derivatives lying in this surface. This follows because the initial value surface is $E_1=0$, and the indicated operators are taken at fixed E_1 . To transform partial derivatives from the variable set (y, ψ, E) to the set (y, ψ, E_1) , the following substitutions may be used:

$$\begin{aligned}\frac{\partial}{\partial y} &\rightarrow \frac{\partial}{\partial y} - M \frac{\partial \Omega_m}{\partial y} \frac{\partial}{\partial E_1} \\ \frac{\partial}{\partial \psi} &\rightarrow \frac{\partial}{\partial \psi} \\ \frac{\partial}{\partial E} &\rightarrow \frac{\partial}{\partial E_1}\end{aligned}\tag{4.37}$$

By using Eq. (4.3), the term in parentheses in Eq. (4.35) can be written in the original variables (y, ψ, E) as

$$(\dots) = \frac{\partial}{\partial y_a} (\hat{U}b) \cdot \frac{\partial X}{\partial \psi} - \frac{\partial}{\partial \psi} (\hat{U}b) \cdot \frac{\partial X}{\partial y_a} - \frac{\partial I}{\partial y_a}\tag{4.38}$$

Since U and I both vanish on the initial value surface, the derivatives of these quantities with respect to y or ψ , in the new variables (y, ψ, E_1) , also vanish. Transforming Eq. (4.38) to the new variables and taking advantage of this fact gives

$$(\dots) = -M \frac{\partial \Omega_m}{\partial y_a} \frac{\partial}{\partial E_1} (\hat{U}b) \cdot \frac{\partial X}{\partial \psi} + \frac{M}{\omega_0} \frac{\partial \Omega_m}{\partial y_a}\tag{4.39}$$

plus terms which vanish on the initial value surface. Next, with the help of Eqs. (3.30) and (3.36) and the fact that $\hat{b}$ is a unit vector,

this becomes

$$(\dots) = \frac{M^2}{\omega_0} \frac{\partial \Omega_m}{\partial y_a} \frac{\partial \Omega}{\partial s} \frac{\partial s}{\partial E_1} \quad (4.40)$$

Finally, this vanishes on the initial value surface because $\partial \Omega / \partial s$ vanishes at $s=s_m$.

The proof of Eq. (4.23) is similar, but somewhat simpler.

4.6. The equations of motion in the coordinates (y, ψ, E)

The equations of motion in the coordinates (y, ψ, E) are interesting on account of the physical picture they provide and because of their relation to the averaged equations of motion to be derived later.

These equations can be derived from the Poisson bracket relations

in Eqs. (4.9) and the Hamiltonian (4.33). Expanding the equations

of motion in powers of ϵ gives, through $O(\epsilon)$,

$$\frac{dy_a}{dt} = -\epsilon \omega_0 \gamma_{ab}^v v_{b\psi} + O(\epsilon^2) \quad (4.41a)$$

$$\frac{d\psi}{dt} = \omega_0 - \epsilon \omega_0^2 \gamma_{ab}^v v_{aE} v_{b\psi} + O(\epsilon^2) \quad (4.41b)$$

$$\frac{dE}{dt} = O(\epsilon^2) \quad (4.41c)$$

Eq. (4.41a) is especially interesting. This equation is nothing more than the usual perpendicular drift equation expressed in field line coordinates. As the guiding center bounces between its mirror points, the slow drift across field lines gradually accumulates. This is because the instantaneous drift, given by Eq. (4.41a), consists of a part which is oscillatory in the bounce cycle, and a part which

is secular. One advantage of the coordinates (y, ψ, E) is that it is relatively easy to separate terms into their averaged and oscillatory parts.

To this end, we define the linear averaging operator, Avg , which takes a function f of (y, ψ, E) and produces a function $\text{Avg}(f)$, depending on (y, E) :

$$\text{Avg}(f) = \frac{1}{2\pi} \int_0^{2\pi} f(y, \psi, E) d\psi \quad (4.42)$$

The complementary operator is Osc , defined by

$$\text{Osc}(f) = f - \text{Avg}(f) \quad (4.43)$$

The variables (y, s, E) are more physically immediate than (y, ψ, E) , and the operators Avg and Osc can be expressed in these coordinates.

If f is a function which is even in U , then

$$\text{Avg}(f) = \pm \frac{\omega_0(y, E)}{\pi} \int_{s_0(y, E)}^{s_1(y, E)} \frac{f(y, s', E)}{U(y, s', E)} ds' \quad (4.44)$$

If f is odd in U , then $\text{Avg}(f)=0$. Note that the integrand is singular at the endpoints.

To lowest order Eq. (4.41a) may be averaged to obtain a kind of drift kinetic equation. This new equation represents the motion of the guiding center across field lines, with the oscillatory behavior on the bounce time scale having been removed. To obtain this equation, note that Eq. (4.22) implies

$$\text{Avg}(v_{a\psi}) = \frac{\partial I}{\partial y_a} \quad (4.45)$$

$$\text{Osc}(v_{a\psi}) = \frac{\partial F_a}{\partial \psi} \quad (4.46)$$

Then from Eq. (4.41a),

$$\text{Avg}\left(\frac{dy_a}{dt}\right) = -\epsilon\omega_0\gamma_{ab} \frac{\partial I}{\partial y_b} + O(\epsilon^2) \quad (4.47)$$

This is a well-known result, and more conventional proofs of it may be found elsewhere.^{5,6}

The $O(\epsilon^2)$ term in Eq. (4.41a) could easily have been worked out, because the Hamiltonian is known through $O(\epsilon)$. This term represents the second order perpendicular drifts of the guiding center, which were first derived, in rectangular coordinates, by Northrop and Rome.¹² Nevertheless, the simple averaging procedure used above to derive Eq. (4.47) is valid only through lowest order, and the averaged equations at $O(\epsilon^2)$ must be derived by performing an averaging transformation. This will be done in Sec. 5.

4.7. The equations of motion in the coordinates (Y, ψ, J)

The variables (Y, ψ, J) are not free of rapid oscillations to all orders, because no attention has been paid as yet to the higher order terms in the Hamiltonian. Nevertheless, they are free of rapid oscillations to one higher order than the variables (y, ψ, E) , as will be shown below. An analogous behavior was observed in Ref. 2 in relation to the sets of variables (X, U, θ, M) and (x, u, θ, w) .

By using the Poisson brackets in Eq. (4.4) and the lowest order term of the Hamiltonian in Eq. (4.34), it is easy to derive the equations of motion to lowest order:

$$\frac{dY_a}{dt} = \epsilon \gamma_{ab} \frac{\partial W}{\partial Y_b} + O(\epsilon^2) \quad (4.48a)$$

$$\frac{d\psi}{dt} = \frac{\partial W}{\partial J} + O(\epsilon) \quad (4.48b)$$

$$\frac{dJ}{dt} = O(\epsilon) \quad (4.48c)$$

By using Eqs. (4.30)-(4.31), these may be seen to be the averaged versions of Eqs. (4.41), as anticipated.

From this observation it follows that the $O(\epsilon)$ term in the Darboux transformation for the variables y and Y , given by Eqs. (4.19)-(4.20), must be responsible for removing the rapid oscillations to lowest order.

An alternate form for the functions F will show this property more clearly. Because of Eq. (4.28) we have $F(y, \psi=0, E)=0$. With the use of this equation as initial conditions, Eq. (4.46) may be integrated to give

$$F_a(y, \psi, E) = \int_0^\psi \text{Osc}(v_{a\psi}) d\psi \quad (4.49)$$

This equation allows a convenient numerical determination of the functions F , since by Eq. (4.44) it is equivalent to an integration along field lines.

Another form of this equation follows from Eq. (4.41):

$$F_a(y, \psi, E) = \frac{1}{\epsilon \omega_0} \gamma_{ab} \int_0^\psi \left[\frac{dy_b}{dt} - \text{Avg} \left(\frac{dy_b}{dt} \right) \right] d\psi \quad (4.50)$$

This form shows clearly how the term in F in Eqs. (4.19)-(4.20) causes the rapid oscillations to be removed. It should also be compared with Eq. (23) of Ref. 4, where the same expression was derived in a non-Hamiltonian context.

5. THE AVERAGING TRANSFORMATION

Although averaged equations of motion can be obtained at lowest order by throwing away oscillatory terms, as we did in Sec. 4, nevertheless it is necessary to employ an averaging transformation to obtain averaged equations of motion beyond lowest order. This we shall do in this section.

5.1. The Lie transform procedure

The averaging transformation is a near-identity, symplectic transformation of the form $(Y, \psi, J) \rightarrow (\bar{Y}, \bar{\psi}, \bar{J})$, which preserves the functional form of the Poisson brackets in Eqs. (4.15). Its purpose and method of execution are so similar to those of the averaging transformations in Refs. 1-2 that only the barest outlines need be given here.

One aspect that is slightly different, however, concerns the expansion of the Poisson bracket as a power series in ϵ . If f and g are any two functions of (Y, ψ, J) , then, according to Eqs. (4.15), their Poisson bracket is given by

$$\{f, g\} = \frac{\partial f}{\partial \psi} \frac{\partial g}{\partial J} - \frac{\partial f}{\partial J} \frac{\partial g}{\partial \psi} + \epsilon \gamma_{ab} \frac{\partial f}{\partial Y_a} \frac{\partial g}{\partial Y_b} \quad (5.1)$$

Note that the Poisson bracket consists of an $O(1)$ term and an $O(\epsilon)$

term.

We let g_n represent the n -th Lie generator, and we define two corresponding first-order, partial differential operators (i.e. tangent vectors) M_n and N_{n+1} by

$$M_n = \frac{\partial g_n}{\partial \psi} \frac{\partial}{\partial J} - \frac{\partial g_n}{\partial J} \frac{\partial}{\partial \psi} \quad (5.2a)$$

$$N_{n+1} = \gamma_{ab} \frac{\partial g_n}{\partial Y_a} \frac{\partial}{\partial Y_b} \quad (5.2b)$$

Then the operator L_n , defined by

$$L_n = \{g_n, \} \quad (5.3)$$

can be written as

$$L_n = M_n + \epsilon N_{n+1} \quad (5.4)$$

The rest of the details of the use of Lie transforms may be summarized by the following list of formulas:

$$T_n = \exp\left(-\frac{\epsilon^n L_n}{n}\right) \quad (5.5)$$

$$T = \dots T_3 T_2 T_1 \quad (5.6)$$

$$T^{-1} = T_1^{-1} T_2^{-1} T_3^{-1} \dots \quad (5.7)$$

$$\bar{z} = T z \quad (5.8a)$$

$$z = T^{-1} \bar{z} \quad (5.8b)$$

$$K = T^{-1} H \quad (5.9)$$

$$T = I - \epsilon M_1 + \frac{1}{2} \epsilon^2 (M_1^2 - M_2 - 2N_2) + O(\epsilon^3) \quad (5.10)$$

$$T^{-1} = I + \epsilon M_1 + \frac{1}{2} \epsilon^2 (M_1^2 + M_2 + 2N_2) + O(\epsilon^3) \quad (5.11)$$

$$H(Y, \psi, J) = \sum_{n=0}^{\infty} \epsilon^n H_n(Y, \psi, J) \quad (5.12)$$

$$K(\bar{Y}, \bar{J}) = \sum_{n=0}^{\infty} \epsilon^n K_n(\bar{Y}, \bar{J}) \quad (5.13)$$

$$K_0 = H_0 \quad (5.14a)$$

$$M_1 H_0 = K_1 - H_1 \quad (5.14b)$$

5.2. The Lie generator g_1

To the order we are working, we need only the first Lie generator g_1 . Calling on Eq. (4.34), we have

$$H_0 = W(Y, J) \quad (5.15)$$

$$H_1 = -\omega_0 \left(\gamma_{ab} F_a \frac{\partial I}{\partial Y_b} + G \right) \quad (5.16)$$

where the E argument of the terms in H_1 is evaluated at $E=W(Y, J)$.

Then Eqs. (5.2), (5.14b) and (4.30) give the following equation for g_1 :

$$\frac{\partial g_1}{\partial \psi} = K_1 + \gamma_{ab} F_a \frac{\partial I}{\partial Y_b} + G \quad (5.17)$$

According to the usual procedure we break this equation into its averaged and oscillatory parts. Since both F and G are odd in ψ (see Eq. (4.28)), they are purely oscillatory, and we have

$$K_1 = 0 \quad (5.18)$$

and

$$g_1 = \int_0^\psi d\psi' \gamma_{ab} \left[\frac{\partial I}{\partial Y_b} F_a(\underline{Y}, \psi', E) + G(\underline{Y}, \psi', E) \right] d\psi' \quad (5.19)$$

where the integrand is evaluated at $E=W(\underline{Y}, J)$.

5.3. The averaged variables $(\bar{Y}, \bar{\psi}, \bar{J})$

Using the Lie generator g_1 , we may work out the averaging transformation through $O(\epsilon)$. The transformation for the variables $\underline{Y}$ is easiest. From Eqs. (5.6)-(5.8) we have

$$\bar{\underline{Y}} = \underline{Y} + O(\epsilon^2) \quad (5.20)$$

Actually, the knowledge of g_1 would allow us to determine the $O(\epsilon^2)$ term in Eq. (5.20), but since the Darboux transformation was carried only through $O(\epsilon)$ it is consistent to leave this equation at the same order.

The transformation for the variable ψ is

$$\bar{\psi} = \psi + \epsilon \frac{\partial g_1}{\partial J} + O(\epsilon^2) \quad (5.21)$$

The quantity $\partial g_1 / \partial J$ can be written in a number of forms, using Eqs. (3.41), (4.21)-(4.22), and (4.30). One of the most interesting is

$$\bar{\psi} = \psi + \epsilon \omega_0 \gamma_{ab} \int_0^\psi d\psi' \left[F_a \frac{\partial}{\partial Y_b} \left(\frac{1}{\omega_0} \right) + v_a E_b v_\psi \right] + O(\epsilon^2) \quad (5.22)$$

In a sense, the first term in this integral is a correction for the fact that ω_0 changes as the guiding center drifts from one field line to another, while the second term is a correction for the $O(\epsilon)$ term in the equation for $d\psi/dt$ (see Eq. (4.41b)). The result is the variable $\bar{\psi}$ which is free of rapid oscillations.

The transformation for the variable J is

$$\bar{J} = J - \epsilon \frac{\partial g_1}{\partial \psi} + O(\epsilon^2) \quad (5.23)$$

or, using Eq. (5.17),

$$\bar{J} = J - \epsilon \left(\gamma_{ab} F_a \frac{\partial I}{\partial Y_b} + G \right) + O(\epsilon^2) \quad (5.24)$$

This gives $\bar{J}$ as a function of (Y, ψ, J) . Of perhaps greater interest is the composition of this transformation with the Darboux transformation, which will give $\bar{J}$ as a function of (y, ψ, E) . From Eq. (4.27), this is

$$\bar{J}(y, \psi, E) = I(y, E) - \epsilon \gamma_{ab} F_a \frac{\partial I}{\partial y_b} + O(\epsilon^2) \quad (5.25)$$

This result has been obtained previously by Northrop, Liu, and Kruskal³ and independently by Hastie, Taylor, and Haas.⁴ To show the equivalence of the results we may use Eqs. (4.47) and (4.50):

$$\bar{J} = I + \frac{1}{\epsilon \omega_0} \gamma_{ab} \int_0^\psi d\psi \text{Avg} \left(\frac{dy_a}{dt} \right) \frac{dy_b}{dt} + O(\epsilon^2) \quad (5.26)$$

This may be compared to Eq. (41) of Ref. 3, or Eq. (5.13) of Ref. 4.

5.4. The averaged equations of motion

The averaged Hamiltonian K was obtained above:

$$K(\bar{Y}, \bar{J}) = W(\bar{Y}, \bar{J}) + O(\epsilon^2) \quad (5.27)$$

The independence of K on $\bar{\psi}$ holds to any order to which one has the endurance to carry out the required Darboux transformation and Lie transforms. In Eq. (5.27) we imagine that $\bar{\psi}$ has been eliminated to all orders, i.e. that K represents a formal power series.

From the averaged Hamiltonian and the Poisson bracket relations in Eq. (4.15) the equations of motion are immediate:

$$\frac{d\bar{Y}_a}{dt} = \epsilon \gamma_{ab} \frac{\partial W}{\partial \bar{Y}_b} + O(\epsilon^3) \quad (5.28)$$

$$\frac{d\bar{\psi}}{dt} = \omega_0 + O(\epsilon^2) \quad (5.29)$$

$$\frac{d\bar{J}}{dt} = 0 \text{ to all orders} \quad (5.30)$$

The quantities on the right hand sides of these equations are evaluated at $(\bar{Y}, E=W(\bar{Y}, \bar{J}))$. These should be compared to Eqs. (4.41) and (4.48). Note in particular that the averaged equations for $\bar{Y}$ and $\bar{\psi}$ hold to one higher order than the equations for Y and ψ , which still contain rapid oscillations.

Eq. (5.30) is especially interesting, because it gives a constant

of the motion. In practice, of course, one must deal with a truncated series for $\bar{J}$, so it is relevant to examine the properties of the time derivative of these truncated series.

Suppose to be definite that $\bar{J}$ is expressed in terms of the variables (y, ψ, E) as in Eq. (5.25), so as to give a unique specification of a sequence of functions $\bar{J}_n$:

$$\bar{J}(y, \psi, E) = \sum_{n=0}^{\infty} \epsilon^n \bar{J}_n(y, \psi, E) \quad (5.31)$$

Let us define $\bar{J}_{[N]}$ as the truncated version of this series:

$$\bar{J}_{[N]} = \sum_{n=0}^N \epsilon^n \bar{J}_n(y, \psi, E) \quad (5.32)$$

Then we have

$$\bar{J}_{[N]} = \bar{J} - \sum_{n=N+1}^{\infty} \epsilon^n \bar{J}_n \quad (5.33)$$

or, on taking the time derivative and using Eqs. (5.30) and (4.41),

$$\frac{d}{dt} \bar{J}_{[N]} = -\epsilon^{N+1} \omega_0 \frac{\partial \bar{J}_{N+1}}{\partial \psi} + O(\epsilon^{N+2}) \quad (5.34)$$

Thus, $\bar{J}_{[N]}$ is a constant through $O(\epsilon^N)$. More importantly, however, the average of the time derivative of $\bar{J}_{[N]}$ vanishes to one higher order:

$$\text{Avg} \left(\frac{d}{dt} \bar{J}_{[N]} \right) = O(\epsilon^{N+2}) \quad (5.35)$$

Eqs. (5.25)-(5.26) effectively give $\bar{J}_{[1]}$. Northrop, Liu, and Kruskal³ showed that this quantity is a constant through $O(\epsilon)$, in accordance with Eq. (5.34). It was correctly pointed out by Hastie, Taylor, and Haas,⁴ however, that the average of the time derivative of $\bar{J}_{[1]}$ must vanish through $O(\epsilon^2)$, in accordance with Eq. (5.35), and that any demonstration of the constancy of $\bar{J}_{[1]}$ which does not prove Eq. (5.35) is incomplete. A lengthy but straightforward calculation, which calls on many of the identities proved in Sec. 4, shows that $\bar{J}_{[1]}$ does indeed satisfy Eq. (5.35).

5.5. The drift motion

The Hamiltonian (5.27) is essentially of one degree of freedom, since the constant $\bar{J}$ can be treated as a parameter. The two-dimensional phase space for the drift motion can be identified with the initial value surface described in Sec. 4, although its only essential characteristic is that points of this phase space must be in one-to-one correspondence with magnetic field lines. From an abstract point of view, the phase space of drift motion is the differentiable manifold consisting of field lines.

If we write $Q = \bar{Y}_1$ and $P = \bar{Y}_2$ and revert to the symbol H for the Hamiltonian, then the Hamiltonian for the drift motion is

$$H(Q, P; \bar{J}) = W(\bar{Y}, \bar{J}) + O(\epsilon^2) \quad (5.36)$$

The variables Q and P satisfy the Poisson bracket relation

$$\{Q, P\} = \epsilon \quad (5.37)$$

which comes from Eq. (4.15a). Because this system is of one degree

freedom, it is integrable, and a transformation to action/angle variables is possible with the traditional Hamilton-Jacobi method. It is not necessary to use the Darboux algorithm. We assume here that the drift motion is periodic, so that action/angle variables exist.

The action variable Φ is the integral

$$\Phi = \frac{1}{2\pi} \oint PdQ \quad (5.38)$$

which is taken around a contour of $H(Q,P;\bar{J}) = \text{constant}$. This contour is, of course, the trajectory in the phase space of drift motion, since it is the curve of constant energy. The constancy of Φ is trivial, since Φ is a function only of the energy. It is well-known that an expression similar to Eq. (5.38) gives a constant of the motion in the case of time-dependent fields, and that in this more general case the constancy is not trivial. Unfortunately, it is not possible to analyze here the case of time-dependent fields.

The representation of Φ as a power series in ϵ comes about through the power series representation of the contour of integration, given by Eq. (5.36). Thus, as it stands, Eq. (5.38) is valid to all orders. If instead the contour of integration were taken around $W(Q,P,\bar{J}) = \text{constant}$, then Eq. (5.38) would be valid through $O(\epsilon)$. This is a matter of practical importance, because in practice one possesses only a finite series for $H(Q,P;\bar{J})$.

If the frequency of the drift motion is denoted by ω_f , then we have

$$\frac{1}{\omega_f} = \frac{1}{2\pi} \oint \left(\frac{\partial H}{\partial P} \right)^{-1} dQ \quad (5.39)$$

Since the contour of integration depends only on the energy, so also does ω_f . Alternatively, ω_f may be considered a function of Φ .

The drift angle ϕ is given by

$$\phi(Q,P) = \omega_f \int^{(Q,P)} \left(\frac{\partial H}{\partial P} \right)^{-1} dQ \quad (5.40)$$

with the integral, once again, taken along a contour of constant H . Because the transformation $(Q,P) \rightarrow (\phi,\Phi)$ is a canonical transformation, we have

$$\{\phi, \Phi\} = \epsilon \quad (5.41)$$

The variables (ϕ, Φ) are action/angle variables for the drift motion.

6. CONCLUSIONS

The major obstacle to the Hamiltonian treatment of guiding center motion is the difficulty of finding an appropriate set of canonical variables in which to carry out the perturbation expansion and to express the results. The standard Hamilton-Jacobi method does not work for this problem, and it was for this reason that the approach based on Darboux's theorem was developed in Ref. 1. The result of the first application of Darboux's theorem is a semicanonical coordinate system which is not only adequate for an analysis of the gyro-motion but even preferable, for many purposes, to a fully canonical set of variables.

Nevertheless, the original goal of finding a set of canonical variables for the guiding center problem is a question both of academic interest and of practical importance for the description

of the bounce and drift motion. This paper has gone a long way toward the completion of this goal, since we have constructed here a set of variables (dropping the overbars) $(\theta, M, \psi, J, \phi, \Phi)$ which satisfy the Poisson bracket relations

$$\begin{aligned}\{\theta, M\} &= 1/\epsilon \\ \{\psi, J\} &= 1 \\ \{\phi, \Phi\} &= \epsilon\end{aligned}\tag{6.1}$$

The system is still not completely canonical, because in addition to these Poisson brackets, the Poisson brackets $\{\theta, \psi\}$, $\{\theta, \phi\}$, and $\{\psi, \phi\}$ are non-zero.

Nevertheless, the coordinates $(\theta, M, \psi, J, \phi, \Phi)$ are capable of giving a complete and systematic (to any desired order in ϵ) Hamiltonian description of the three degrees of freedom of charged particle motion. Such a description is quite new; although many aspects of the Hamiltonian structure of guiding center motion have been noticed in the past, these came about as observations based on the averaged equations of motion at lowest order, rather than as a result of a systematic Hamiltonian theory.

The most promising application for the methods of this paper may well be a Hamiltonian formalism for the study of resonances among the three degrees of freedom indicated by Eq. (6.1). This is currently an active area of research, and the results of these investigations will be reported on in future publications.

ACKNOWLEDGEMENTS

I would like to thank C. Grebogi, J. A. Rome, and H. E. Mynick for useful discussions.

REFERENCES

1. R. G. Littlejohn, J. Math. Phys. 20, 2445(1979).
2. R. G. Littlejohn, Lawrence Berkeley Laboratory Report No. LBL-9141 (1979). To be published in Phys. Fluids.
3. T. G. Northrop, C. S. Liu, and M. D. Kruskal, Phys. Fluids 9, 1503(1966).
4. R. J. Hastie, J. B. Taylor, and F. A. Haas, Ann. Phys. 41, 302(1967).
5. T. G. Northrop and E. Teller, Phys. Rev. 117, 215(1960).
6. T. G. Northrop, The Adiabatic Motion of Charged Particles (Interscience, New York, 1963).
7. M. D. Kruskal, J. Math. Phys. 3, 806(1962).
8. R. Abraham and J. E. Marsden, Foundations of Mechanics (Benjamin/Cummings, Reading, Mass., 1978), 2nd Ed.
9. A. I. Morozov and L. S. Solov'ev, in Reviews of Plasma Physics, edited by M. A. Leontovich (Consultants Bureau, New York, 1966), vol. 2, p. 201.
10. V. I. Arnold, Mathematical Methods of Classical Mechanics (Springer-Verlag, New York, 1978).
11. D. P. Stern, Am. J. Phys. 38, 494(1970).
12. T. G. Northrop and J. A. Rome, Phys. Fluids 21, 384(1978).

Table I. Symbol equivalences. A formula from this paper or for Ref. 2 can be translated into the notation of Ref. 3 by making the indicated symbol substitutions.

	This Paper	Ref. 2	Ref. 3
1.	ϵ	ϵ	$(e/m)\epsilon$
2.	----- ^a	$\tilde{x}$	$\tilde{r}$
3.	$\hat{b}$	$\hat{b}$	$\hat{L}$
4.	-----	$\tilde{x}^b$	$\tilde{\rho}^b$
5.	-----	w	σ
6.	-----	u	η
7.	-----	θ	$2\pi(u + \frac{1}{2})$
8.	$\tilde{x}$	$\tilde{x}^b$	$\tilde{p}^b$
9.	U	$\bar{U}^c$	H^c
10.	θ	$\bar{\theta}$	$2\pi(\phi + \frac{1}{2})$
11.	M	$\bar{J}$	$(mc/e)M$
12.	ψ	-----	$2\pi v'$
13.	E^c	-----	K^c
14.	ω_0	-----	$2\pi/T$
15.	I	-----	$mJ_0/2\pi$
16.	$\bar{J}$	-----	$mJ/2\pi$

^aThe omitted variables do not appear in the indicated paper.

^bEqual through $O(\epsilon)$.

^cEqual through $O(\epsilon^2)$.

Table II. The structure of variable transformations in Kruskal's theory. The notation is based on that of Ref. 7. The variables $\underline{q}$, $\underline{p}$ are N-vectors, $\underline{x}$ is a 2N-vector, $\underline{y}$ and $\bar{\underline{y}}$ are (2N-1)-vectors, and $\bar{\underline{z}}$ is a (2N-2)-vector.

Transformation		Variables
1.		$(\underline{q}, \underline{p})^a$
	Preparatory	↓
2.		$\underline{x}^a$
	Preparatory	↓
3.		$(\underline{y}, \theta)$
	Averaging	↓
4.		$(\bar{\underline{y}}, \bar{\theta})$
	Ring Integral ^b	↓
5.		$(\bar{\underline{z}}, \bar{\theta}, \bar{J})^b$

^aThe method begins with the variables $\underline{x}$ instead of $(\underline{q}, \underline{p})$ if the system is non-Hamiltonian.

^bThe variable $\bar{J}$ can be computed only for a Hamiltonian system.

Table III. The structure of variable transformations in the Kruskal-Darboux-Lie method. In (a), the general case; the notation is based on Ref. 7 and is independent of this paper; $\underline{q}, \underline{p}$ are N-vectors, $\underline{x}$ is a 2N-vector, $\underline{y}$ is a (2N-1)-vector, and $\underline{z}$ and $\underline{\bar{z}}$ are (2N-2)-vectors. In (b), the application to particle gyro-motion, as explained in Ref. 2; $\underline{q}, \underline{p}, \underline{x}, \underline{v}, \underline{X}, \underline{\bar{X}}$ are 3-vectors. In (c), the application to bounce motion, the subject of this paper; $\underline{X}$ is a 3-vector, and $\underline{y}, \underline{Y}, \underline{\bar{Y}}$ are 2-vectors.

	(a)	(b)	(c)
Transformation	General Case	Gyro-Motion	Bounce Motion
1.	$(\underline{q}, \underline{p})$	$(\underline{q}, \underline{p})$	-----
Preparatory	↓	↓	
2.	$\underline{x}$	$(\underline{x}, \underline{v})$	$(\underline{X}, \underline{U})^a$
Preparatory	↓	↓	↓
3.	$(\underline{y}, \theta)$	$(\underline{x}, \underline{u}, \theta, w)$	$(\underline{y}, \psi, E)^b$
Darboux	↓	↓	↓
4.	$(\underline{z}, \theta, J)$	$(\underline{X}, \underline{U}, \theta, M)^c$	$(\underline{Y}, \psi, J)$
Averaging/Lie	↓	↓	↓
5.	$(\underline{\bar{z}}, \bar{\theta}, \bar{J})$	$(\underline{\bar{X}}, \underline{\bar{U}}, \bar{\theta}, \bar{M})^c$	$(\underline{\bar{Y}}, \bar{\psi}, \bar{J})$

^aThese variables are identified with the overbarred variables $(\underline{\bar{X}}, \underline{\bar{U}})$ of column b.

^bThe transformation $(\underline{X}, \underline{U}) \rightarrow (\underline{y}, \psi, E)$ passes through several intermediate steps; see Sec. 3.

^cThe symbol J was used instead of M in Ref. 2.

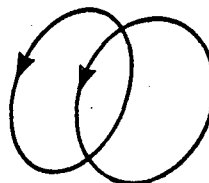
FIGURE CAPTIONS.

- Fig. 1. In (a), a qualitative picture of nearly periodic motion, i.e. the exact, perturbed motion. In (b), the unperturbed motion. These are Kruskal's loops.
- Fig. 2. Potential energy diagram for the unperturbed bounce motion, which takes place at fixed y . s_0 and s_1 are the turning points, and s_m is the minimum of the potential well.
- Fig. 3. The s - U phase plane for the unperturbed bounce motion. The motion follows the contours of constant E in the direction of the arrows. Two contours of constant ψ are shown.
- Fig. 4. A ψ -characteristic in phase space. The surface $E = M\Omega_m(y)$, or $J=0$, is the initial value surface, which is two-dimensional.

(a)

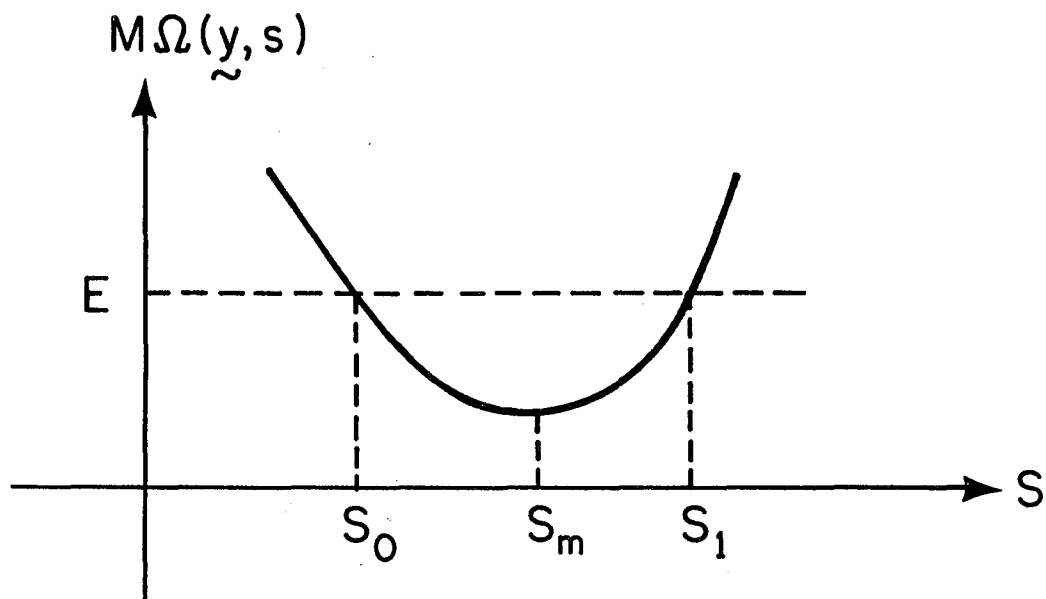


(b)



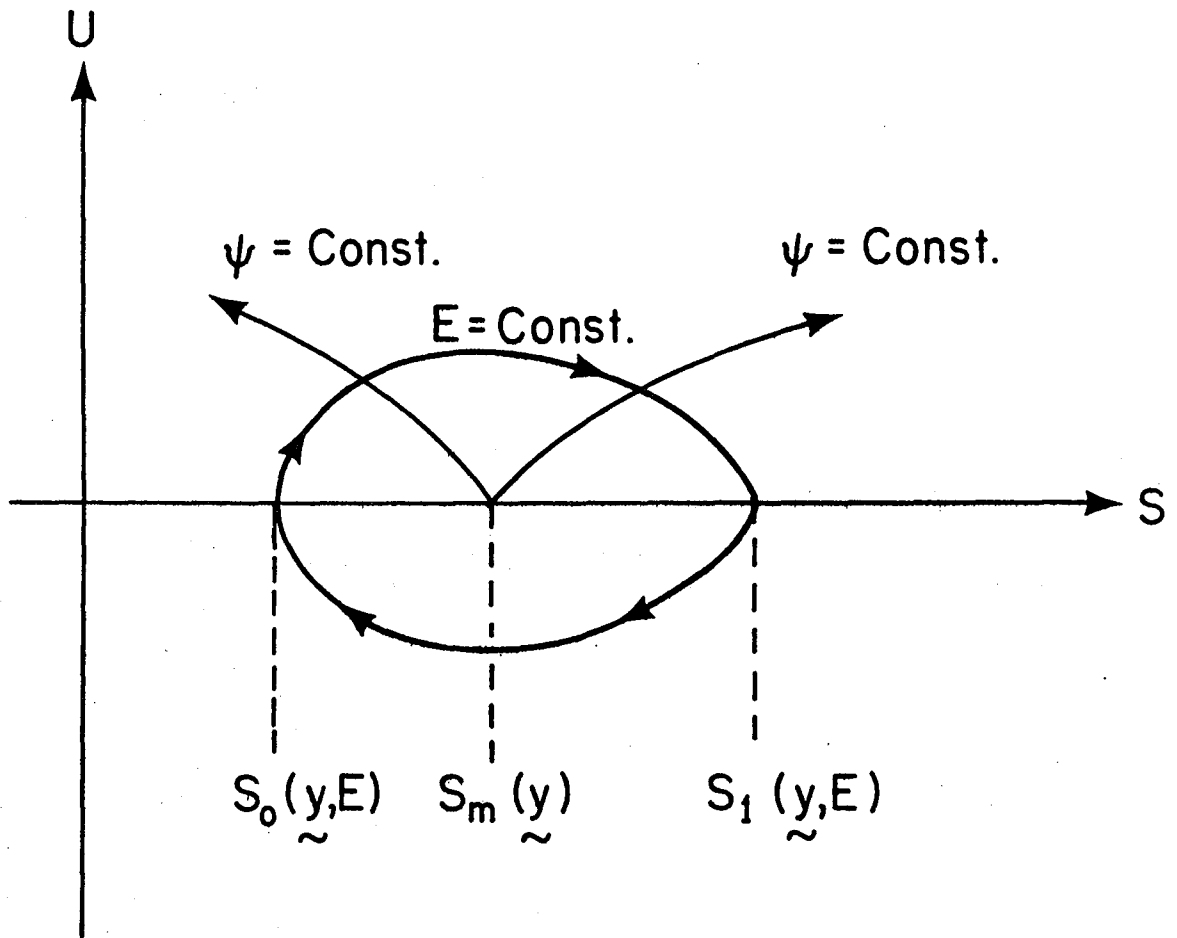
XBL 805-958

Fig. 1.



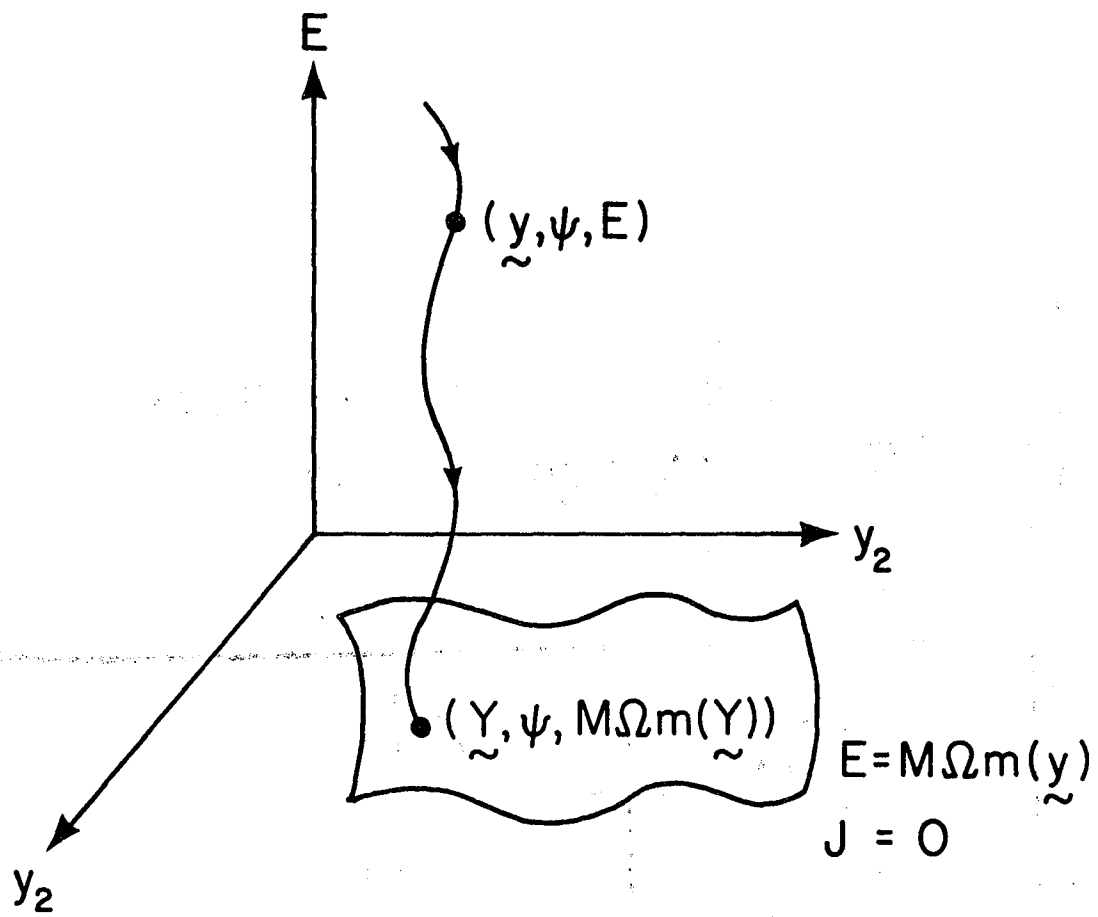
XBL805-959

Fig. 2.



XBL 805-957

Fig. 3.



XBL 805-960

Fig. 4.

CHAPTER V

CONCLUDING REMARKS

1. IMPROVEMENTS ON THE DARBOUX TRANSFORMATION METHOD

The method of the Darboux transformation which was developed in Chapter II and applied in Chapters III and IV is undoubtedly capable of improvement. Some suggestions for this may be offered here, and any careful reader will certainly think of others.

The most obvious drawback to the method is the requirement for carrying out two transformations, the Darboux transformation and the averaging transformation, both of which are expressed as infinite series. If these two transformations could be merged there would perhaps result a simplification. One way to do this would be to carry out the Darboux transformation using a modified gyrophase ϕ :

$$\phi(\underline{x}, \underline{v}) = \theta(\underline{x}, \underline{v}) + \varepsilon \phi_1(\underline{x}, \underline{v}) + \dots \quad (1)$$

where $\theta(\underline{x}, \underline{v})$ is the instantaneous gyrophase, exactly as defined in Chapters II or III, and where $\phi_1, \phi_2, \dots$ are initially unknown functions. The resulting Darboux transformation would then be parametrized by the functions $\phi_1, \phi_2, \dots$, which would be determined by the demand that the Hamiltonian be independent of ϕ .

Another approach is based on the appreciation, gathered in Chapter IV and elsewhere, that the Lagrange tensor ω and the distinguished 1-form ρ are easier to deal with than the Poisson tensor. In reference to the problem of gyromotion, for example, one might proceed as follows. First, observe that in any phase space coordinate system $\underline{z}$, the components ρ_i of the distinguished 1-form ρ can be written

$$\rho_i(\underline{z}) = \underline{p}(\underline{z}) \cdot \frac{\partial \underline{q}(\underline{z})}{\partial \underline{z}^i} + \frac{\partial S(\underline{z})}{\partial \underline{z}^i} \quad (2)$$

where the index i runs from 1 to 6, where $\underline{p}$ and $\underline{q}$ are the usual Cartesian canonical coordinates, and where S is an arbitrary scalar function on phase space which specifies a "gauge transformation" on the 1-form ρ . Of course, $\omega = d\rho$ is not affected by the term in S , since $ddS=0$.

Next, one would consider a set of coordinates which is equal, to lowest order, to the final guiding center variables one wants. The $(\underline{x}, \underline{u}, \theta, w)$ coordinate system of Chapter III is close to this, but since the variable w is not canonically conjugate to θ , even at lowest order, it should be replaced by $m = w^2/2B(\underline{x})$. Then, one would consider a near-identity coordinate transformation of the form $(\underline{x}, \underline{u}, \theta, m) \rightarrow (\underline{X}, \underline{U}, \Theta, M)$ such that the following criteria are fulfilled. First, the Hamiltonian in the new variables is to be independent of Θ . Second, the components of ρ corresponding to Θ and M are to have the values

$$\begin{aligned} \rho_{\Theta} &= \epsilon M \\ \rho_M &= 0 \end{aligned} \quad (3)$$

And third, the other four components of ρ are to be independent of both M and Θ . These criteria give a hierarchy of underdetermined equations for the $O(\epsilon)$ and higher correction terms in the transformation $(\underline{x}, \underline{u}, \theta, m) \rightarrow (\underline{X}, \underline{U}, \Theta, M)$, and allow those terms to be determined.

Finally, when the new Hamiltonian, the components of ρ , and the

transformation itself are known, the Lagrange and Poisson tensors are computed from ρ , and the averaged equations of motion follow immediately.

There are many unresolved questions about this approach, but if it worked as indicated, it would completely bypass Darboux's theorem and it would be easier to understand than the method developed in Chapter II.

In working with the Lagrange tensor one is often faced with the problem of inverting a $2N \times 2N$ antisymmetric matrix in order to pass from the Lagrange tensor to the Poisson tensor and vice versa. In this connection. I have worked out the following relations for performing such a matrix inversion, which are helpful especially when N is three or greater. These relations are essentially simplifications to Cramer's rule which come about because of antisymmetry.

Let M be a $2N \times 2N$ antisymmetric matrix with components $M_{ij} = -M_{ji}$. Define the quantity Γ by

$$\Gamma = \sum_P S_P M_{i_1 j_1} M_{i_2 j_2} \dots M_{i_N j_N} \quad (4)$$

where the sum is taken over all permutations $(1, 2, \dots, 2N) \rightarrow (i_1, j_1, i_2, j_2, \dots, i_N, j_N)$ which satisfy $i_1 < i_2 < \dots < i_N$ and $i_k < j_k$, $k=1, \dots, N$, and where S_P is the parity of the permutation. The quantity Γ is the square root of the determinant of M , and it is computed in (4) with only $(2N-1)!!$ terms instead of the $(2N)!$ terms which would be required in a straightforward expansion of the determinant. The reduced number of terms is smaller by a factor of

$2^N N!$, which can be substantial. If the matrix M_{ij} is taken to be the component matrix of a 2-form ω , then Γ is $1/N!$ times the one independent component of the $2N$ -form $\omega \wedge \omega \wedge \dots \wedge \omega$ (N times), i.e. the Liouville volume element.

If $\Gamma \neq 0$, let K be the inverse of M . Then if $i < j$, the component K_{ij} is given by

$$K_{ij} = \frac{(-1)^{i+j}}{\Gamma} \sum_{P'} S_{P'} M_{i_1 j_1} M_{i_2 j_2} \dots M_{i_{N-1} j_{N-1}} \quad (5)$$

where the sum is taken over all permutations P' of the numbers $(1, 2, \dots, 2N)$, with i, j removed, to $(i_1, j_1, i_2, j_2, \dots, i_{N-1}, j_{N-1})$, where $i_1 < i_2 < \dots < i_{N-1}$ and $i_k < j_k$, $k=1, 2, \dots, N-1$. Again, $S_{P'}$ is the parity of the permutation. If $i > j$, then use $K_{ij} = -K_{ji}$. The sum (5) contains $(2N-3)!!$ terms.

2. EXTENSIONS OF THE DARBOUX TRANSFORM METHOD

An obvious shortcoming of the work presented in Chapters II, III and IV is that it does not allow for time-dependent fields. Time dependence cannot be added to these results as an afterthought, but rather must be built into the formalism from the outset.

It is shown in any textbook on classical mechanics that the form of Hamilton's equations of motion is preserved under time-dependent canonical transformations, although the Hamiltonian does not transform as a phase space scalar:

$$K(Q, P, t) = H(q, p, t) + \frac{\partial S}{\partial t}(q, P, t) \quad (6)$$

However, when one considers general coordinate transformations in phase space, it is easily shown that a new Hamiltonian does not in general exist. The reason for this is encapsulated in Eq. (2.14) of Chapter II.

Therefore in treating the time-dependent problem, using noncanonical coordinates in phase space, it seems best to employ an extended phase space of $N+1$ degrees of freedom in which time is made conjugate to a new variable h which is physically akin to the energy of the particle. For example, in the nonrelativistic problem, one could begin with the Hamiltonian

$$H(\underline{q}, \underline{t}, \underline{p}, h) = \frac{1}{2} \left[\underline{p} - \frac{1}{\epsilon} A(\underline{q}, \underline{t}) \right]^2 + \frac{1}{\epsilon} \phi(\underline{q}, \underline{t}) - h \quad (7)$$

where t and h are "anticonjugate," i.e. $\{t, h\} = -1$.

Instead of doing this, however, it seems better to treat the relativistic problem, which has the same number of degrees of freedom as the Hamiltonian (7). The nonrelativistic equations then follow as a limit. I have carried out part of this calculation, using the Hamiltonian

$$H(x^\mu, p_\mu) = -[-(p^\mu - \frac{1}{\epsilon} A^\mu)(p_\mu - \frac{1}{\epsilon} A_\mu)]^{1/2} \quad (8)$$

The independent parameter specifying phase space trajectories is, with this Hamiltonian, the proper time of the particle, and the Hamiltonian (7) may be directly derived as the nonrelativistic limit of (8).

It is not possible to give final results on the analysis of the Hamiltonian (8), but at least the Poisson brackets in the guiding center variables may be displayed. These variables are denoted

(X^μ, K, U, θ, M) , and they result from the Darboux transformation. In a simplified description of these variables, we may say that X^μ is the guiding center position (and time), K and U are the energy and parallel velocity of the guiding center (relative to the $\mathbf{E} \times \mathbf{B}$ drift in the nonrelativistic limit), and θ and M are the gyrophase and magnetic moment. We let $A^{*\mu}$ be a modified vector potential, given by

$$A^{*\mu} = A^\mu + \epsilon(Kd^\mu + Ub^\mu) \quad (9)$$

where d^μ is the time-like eigenvector of the electromagnetic stress-energy tensor, and where b^μ is a space-like eigenvector, having $(0, \hat{b})$ as its nonrelativistic limit. From $A^{*\mu}$ we define a modified field tensor $F^{*\mu\nu}$ and its dual, $G^*_{\mu\nu}$. In terms of these, the Poisson brackets are

$$\begin{aligned} \{X^\mu, X^\nu\} &= -\frac{\epsilon}{D} \epsilon^{\mu\nu\sigma\tau} d_\sigma b_\tau \\ \{X^\mu, K\} &= +\frac{1}{D} G^{*\mu\nu} b_\nu \\ \{X^\mu, U\} &= -\frac{1}{D} G^{*\mu\nu} d_\nu \\ \{K, U\} &= \frac{1}{4\epsilon D} F^{*\mu\nu} G^*_{\mu\nu} = \frac{1}{\epsilon D} (\mathbf{E}^* \cdot \mathbf{B}^*) \end{aligned} \quad (10)$$

where

$$D = G^{*\mu\nu} d_\mu b_\nu \quad (11)$$

In quite another sense, the method of the Darboux transformation can be extended to other systems, apart from the guiding center problem. It seems that the method is generally applicable to

Hamiltonian systems which display multiple time scales, and these are not easily treated by the standard perturbation techniques. The guiding center problem is a rather complicated example of such a system, and perhaps some features could be seen more easily with simpler systems.

3. APPLICATIONS OF THE GUIDING CENTER HAMILTONIAN

When carried only to the order of the classic, well-known drift equations, the formalism developed here may be of limited practical advantage for problems in plasma physics, in spite of its intrinsic interest. When carried beyond this lowest order, however, its advantages as a labor saving device become striking. This much is clear from the analysis of single particle motion presented in this thesis, but of course the same will be true in other applications.

Applications to single particle motion are not exhausted by the derivation of the guiding center equations of motion. For example, guiding center motion in the presence of an electromagnetic wave, with various possible relations between the wave frequency and the gyrofrequency, is relevant to plasma physics and may be treated, as a first approximation to reality, as a problem in single particle motion. This problem has already been analyzed to lowest order in ϵ by Grebogi, Kaufman, and Littlejohn,¹ who derive a ponderomotive Hamiltonian for this problem, i.e. an averaged Hamiltonian carried to second order in the wave amplitude.

The bounce and drift motions can also be analyzed with the Hamiltonian methods presented here, as shown in Chapter IV. These analyses should be extended to time-dependent phenomena, and they should also

be generalized to handle other cases of interest, such as trapped particles in a tokamak. There are a number of resonance phenomena in particle motion in tokamak fields which are of current interest, such as resonances caused by divertor coils, and these may be investigated with Hamiltonian means.

As was mentioned in Chapter IV, another area of application for the guiding center Hamiltonian is in the study of gyro-bounce-drift resonances in single particle motion. These resonances are intrinsically higher order effects in the parameter ϵ , because of the separation of time scales for the three types of near periodicities. Although much work has been done on this problem, a thorough and unified treatment has yet to be given.

Finally, the use of the guiding center Hamiltonian for self-consistent problems is a rich field that is as yet completely unexplored. Hamiltonian methods are most useful for the analysis of nonlinear phenomena, such as mode coupling and nonlinear wave packet evolution. It is likely that weak dissipation can also be included in such analyses. The guiding center Hamiltonian would certainly be useful for treating such phenomena in inhomogeneous magnetic fields, and, in the case of higher order effects in the gyroradius, it would be nearly indispensable.

REFERENCE

1. C. Grebogi, A. N. Kaufman, and R. G. Littlejohn, Phys. Rev. Lett. 43, 1668(1979).

This report was done with support from the Department of Energy. Any conclusions or opinions expressed in this report represent solely those of the author(s) and not necessarily those of The Regents of the University of California, the Lawrence Berkeley Laboratory or the Department of Energy.

Reference to a company or product name does not imply approval or recommendation of the product by the University of California or the U.S. Department of Energy to the exclusion of others that may be suitable.

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