REVIEWS OF MODERN PHYSICS, VOLUME 81, APRIL–JUNE 2009

Hamiltonian theory of guiding-center motion

John R. Cary

Center for Integrated Plasma Studies and Department of Physics, University of Colorado, Boulder, Colorado 80309-0390, USA and Tech-X Corporation, Boulder, Colorado 80303, USA

Alain J. Brizard

Department of Chemistry and Physics, Saint Michael's College, Colchester, Vermont 05439, USA

(Published 22 May 2009)

Guiding-center theory provides the reduced dynamical equations for the motion of charged particles in slowly varying electromagnetic fields, when the fields have weak variations over a gyration radius (or gyroradius) in space and a gyration period (or gyroperiod) in time. Canonical and noncanonical Hamiltonian formulations of guiding-center motion offer improvements over non-Hamiltonian formulations: Hamiltonian formulations possess Noether's theorem (hence invariants follow from symmetries), and they preserve the Poincaré invariants (so that spurious attractors are prevented from appearing in simulations of guiding-center dynamics). Hamiltonian guiding-center theory is guaranteed to have an energy conservation law for time-independent fields—something that is not true of non-Hamiltonian guiding-center theories. The use of the phase-space Lagrangian approach facilitates this development, as there is no need to transform *a priori* to canonical coordinates, such as flux coordinates, which have less physical meaning. The theory of Hamiltonian dynamics is reviewed, and is used to derive the noncanonical Hamiltonian theory of guiding-center motion. This theory is further explored within the context of magnetic flux coordinates, including the generic form along with those applicable to systems in which the magnetic fields lie on nested tori. It is shown how to return to canonical coordinates to arbitrary accuracy by the Hazeltine-Meiss method and by a perturbation theory applied to the phase-space Lagrangian. This noncanonical Hamiltonian theory is used to derive the higher-order corrections to the magnetic moment adiabatic invariant and to compute the longitudinal adiabatic invariant. Noncanonical guiding-center theory is also developed for relativistic dynamics, where covariant and noncovariant results are presented. The latter is important for computations in which it is convenient to use the ordinary time as the independent variable rather than the proper time. The final section uses noncanonical guiding-center theory to discuss the dynamics of particles in systems in which the magnetic-field lines lie on nested toroidal flux surfaces. A hierarchy in the extent to which particles move off of flux surfaces is established. This hierarchy extends from no motion off flux surfaces for any particle to no average motion off flux surfaces for particular types of particles. Future work in magnetically confined plasmas may make use of this hierarchy in designing systems that minimize transport losses.

DOI: 10.1103/RevModPhys.81.693 PACS number(s): 52.35.Mw, 52.35.Qz, 52.35.Ra

CONTENTS

I. INTRODUCTION

A charged particle in a constant magnetic field $\mathbf{B} = B\hat{\mathsf{b}}$ moves along a helix, while conserving its kinetic energy and, therefore, its speed $v = |\mathbf{v}|$. In addition, the motion parallel to the magnetic field is uniform, i.e., the velocity $v_{\|} = \mathbf{v}\cdot\hat{\mathsf{b}}$ parallel to the magnetic field is constant, and so the perpendicular speed $v_{\perp} \equiv |\mathbf{v}_{\perp}| = (v^2 - v_{\|}^2)^{1/2}$ is also constant. The perpendicular motion (or gyromotion) is confined to a circle, whose gyration center remains on the same magnetic-field line. The gyration frequency (or gyrofrequency) is given by $\Omega \equiv eB/mc$ and the gyration radius (or gyroradius) vector $\boldsymbol{\rho}(v_{\perp}, \zeta) \equiv \hat{\mathsf{b}} \times \mathbf{v}_{\perp}/\Omega$ depends explicitly on the gyration angle (or gyroangle) ζ.

Guiding-center theory gives the modifications to these results for motion in a magnetic field that is slowly varying rather than constant. Slowly varying means that the scale length L of the magnetic field is large compared with the gyroradius $\rho = v_{\perp}/\Omega$ and the distance $v_{\|}/\Omega$ traveled by the particle in one gyroperiod. In this case, the field is approximately constant, and so the results of the constant-field theory should be approximately correct: Drifts across field lines should be small, and some constant of motion in the constant magnetic-field case should be an adiabatic invariant (Kulsrud, 1957; Gardner, 1959) in the case of a slowly varying magnetic field. Alfvén (1940) showed that the magnetic moment

$$\mu \equiv \frac{e}{mc}\oint \frac{d\zeta}{2\pi}\left[m\mathbf{v} + \frac{e}{c}\mathbf{A}(\mathbf{X}+\boldsymbol{\rho})\right]\cdot\frac{\partial\boldsymbol{\rho}}{\partial\zeta} = \frac{mv_{\perp}^2}{2B} \quad (1.1)$$

is the adiabatic invariant associated with the fast gyromotion of a charged particle (with mass m and charge e) in a slowly varying magnetic field $\mathbf{B} = \nabla\times\mathbf{A}$ and the gyroaction $J \equiv (mc/e)\mu$ is canonically conjugate to the ignorable gyrophase angle ζ. From this adiabatic invariance and energy conservation, it follows that there must be a parallel force due to the gradient of μB (the perpendicular kinetic energy), taken with μ held constant. Alfvén (1940) also calculated the cross-field drifts due to the gradient of B and the magnetic-field-line curvature. His results showed that the cross-field drifts are smaller than $v_{\perp}$ by the ratio $\rho/L \equiv \epsilon$. These results and other early works are reviewed by Northrop (1963).

A. History

In the past two decades, researchers began to face the shortcomings of standard guiding-center theory (Northrop, 1963): it does not have an energy conservation law (for time-independent systems), it fails to provide equations of motion consistent with Liouville's theorem (Goldstein *et al.*, 2002), and it does not derive from a variational principle. These shortcomings had become increasingly important as applications of guiding-center theory were taken to their limits, especially in numerical studies of particle motion in complex magnetic geometries. A small amount of energy nonconservation is not critical for short-time integration, but over long times it can accumulate and cause numerical analyses to give unphysical results. Liouville's theorem is important for particle-in-cell simulation techniques (Birdsall and Langdon, 1985), in which each particle is advecting a part of phase space, and so the equations of motion need to conserve phase-space volume. Integration of guiding-center equations of motion having no Liouville property may yield unphysical attractors or repellers giving spurious loss. Further, without a variational form of the equations of motion, Noether's theorem may not be applicable, which would otherwise provide constants of motion for systems that exhibit symmetries.

All of these problems would ultimately be solved by the development of a Hamiltonian theory of guiding-center motion [Boozer (1980) showed that the guiding-center equations could be modified to have the Liouville property without necessarily being Hamiltonian]. Hamiltonian theories naturally have the Liouville property, possess a variational structure, and have an energy conservation law for time-independent systems. Moreover, having a Hamiltonian theory of guiding-center motion guarantees that further reductions, such as the introduction of the longitudinal invariant, are possible to all orders (Kruskal, 1962). Efforts were made in at least three separate directions to obtain a Hamiltonian theory of guiding-center motion.

Canonical perturbation theory was used to obtain guiding-center equations of motion by Gardner (1959) and Wong (1982). In these analyses, one begins by introducing flux coordinates α and β (Stern, 1970) for the magnetic field $\mathbf{B}=\nabla\alpha\times\nabla\beta$. With this (Euler-Clebsch) representation for the magnetic field, canonical coordinates are found. Wong (1982) proceeded perturbatively to introduce a transformation to new coordinates such that one canonical pair represents the fast gyromotion, while the remaining four coordinates evolve on slow time scales. Being canonical, these coordinates have a unit phase-space Jacobian. Moreover, the equations of motion are simply described by a Hamiltonian function and a canonical Poisson bracket. However, one must use nonphysical coordinates in this theory.

Boozer (1980) showed that, for magnetic fields consistent with scalar pressure magnetohydrodynamic (MHD) equilibrium, the guiding-center equations of motion can be modified to have the Liouville property and, further, if the fields are curl-free ($\boldsymbol{\nabla}\times\mathbf{B}=\mathbf{0}$), these equations can be derived from a Hamiltonian. This work yielded guiding-center equations of motion similar or identical to those that had been obtained previously (Morozov and Solov'ev, 1966; Rutherford, 1970; Dobrott and Frieman, 1971). Subsequently, White *et al.* (1982) showed that, for non-curl-free magnetic fields ($\boldsymbol{\nabla}\times\mathbf{B}\neq\mathbf{0}$) consistent with scalar pressure MHD equilibrium, the guiding-center equations of Boozer (1980) can be derived from a Hamiltonian. In fact, their results were somewhat more general, in that the magnetic field could be composed of two terms, one that has nested flux surfaces and another that corresponds to the primary terms for breaking those flux surfaces in perturbation theory. Later, White and Chance (1984) showed that, for systems with nested flux surfaces, one could introduce new variables uniformly close to the toroidal and poloidal angles, such that with the neglect of higher-order terms in guiding-center theory the guiding-center equations of motion are canonical. Indeed, White and Chance (1984) used these coordinates for analyzing the complicated guiding-center orbits in helical configurations. At nearly the same time, Boozer (1984) pointed out that these canonical equations of motion arise by neglecting a term important only for toroidal configurations with significant plasma pressure.

This review will be based on the results of Littlejohn, who, in a series of papers (Littlejohn, 1979, 1981, 1982a) culminating in the work of Littlejohn (1983), used noncanonical Hamiltonian mechanics to derive the guiding-center phase-space Lagrangian from first principles. The approach is to use Lie-transform perturbation theory (Cary, 1981a, 1981b) for noncanonical Hamiltonian mechanics (Cary and Littlejohn, 1983). The noncanonical formulation of mechanics follows from noting that the canonical equations of motion derive from requiring stationarity of the action integral

$$\mathcal{A}[\mathbf{x}]\equiv\int L(\mathbf{x},\dot{\mathbf{x}};t)dt, \tag{1.2}$$

where $L\equiv\mathbf{p}(\mathbf{x},\dot{\mathbf{x}},t)\cdot\dot{\mathbf{x}}-H(\mathbf{x},\dot{\mathbf{x}},t)$, with respect to virtual displacements $\delta\mathbf{x}$ in configuration space. Since this formulation is variational, arbitrary coordinates can be introduced. In particular, physical variables, such as Cartesian guiding-center coordinates, can be used without sacrificing the formal results of Hamiltonian theories: noncanonical Hamiltonian mechanics is variational (and is, therefore, amenable to the derivation of dynamical invariants from Noether symmetries), possesses the Liouville property, and possesses all Poincaré integral invariants (thereby preserving the structure of phase space). Furthermore, one is guaranteed the existence of an adiabatic invariant to all orders for systems with well-separated time scales.

Littlejohn (1983) introduced a perturbative transformation to new guiding-center coordinates. The rapid motion is contained in the gyrophase ζ, the magnetic moment (1.1) is an adiabatic invariant, and the remaining variables, the three spatial coordinates of the guiding center and the parallel (kinetic) momentum, are slowly varying. At each order at which the phase-space Lagrangian is calculated, there is an exact Liouville property and a Noether derivation of exact conservation laws. Moreover, this formulation, like other Hamiltonian formulations, can be the starting point for the derivation of other adiabatic invariants such as the longitudinal invariant and the drift invariant or the parallel invariant for motion in toroidal magnetic fields with ripple (Cary *et al.*, 1988). Additionally, when canonical variables are available, symplectic integration techniques (Forest and Ruth, 1990; Candy and Rozmus, 1991), with their enhanced numerical stability, may be used. Finally, because the transformation is known, one can show how to calculate the currents and densities from the guiding-center distribution (Brizard, 1989, 1992).

Our calculations proceed through first order in the ratio of the gyroradius to scale length. However, unlike previous work (Littlejohn, 1983), we do not have an ordering in which the electric field is small. Thus, the $E\times B$ drift velocity occurs at lowest order in the analysis. With this subsidiary ordering, we are able to obtain the polarization drift at the same order as the magnetic drift velocities.

B. Notation

The latin indices i, j, and k are used to denote components of covariant, contravariant, or mixed tensors in configuration space or momentum space and take values from 1 to N (N is the number of degrees of freedom). The greek indices α and β are used to denote phase-space components and take values from 1 to $2N$, while the greek indices μ and ν are used to denote space-time components of four-vectors or four-by-four tensors and take values from 0 to 3. The sans-serif latin indices a and b are used to denote components in eight-dimensional extended phase space.

C. Organization

We begin, in Sec. II, with a discussion of the phase-space Lagrangian formulation of mechanics. As we have noted, the phase-space Lagrangian formalism has the benefits of the Hamiltonian formalism (i.e., Liouville and Noether properties) without the drawback of having to use canonical coordinates. We show how the phase-space Lagrangian formalism possesses the Hamiltonian properties.

In Sec. III, we introduce the guiding-center phase-space Lagrangian. From the guiding-center phase-space Lagrangian, and the general formalism of phase-space Lagrangians, we immediately obtain the Hamiltonian guiding-center equations of motion. At this point we turn our attention to the derivation of the guiding-center Lagrangian in our more general ordering of large electric fields. Next we show how the transformation can be used to obtain the particle currents and densities in configuration space from the guiding-center distribution. Last, we consider two applications. We show how to obtain the canonical angular momentum conserved for systems with azimuthal symmetry, and we show how the Hamiltonian formulation of magnetic-field line flow can be obtained from the guiding-center Lagrangian. In this and all sections, we develop results only through first order in the ratios of gyroradius to scale length and characteristic frequencies to gyrofrequency, even though there has been a large amount of development of higher-order guiding-center equations (Northrop and Rome, 1978) and the problems of gyrogauge invariance (Littlejohn, 1984, 1988) that arise in this context.

In Sec. IV, we turn to the problem of finding canonical coordinates for the guiding-center Lagrangian. We approach this problem from the question of how one obtains canonical coordinates from a general phase-space Lagrangian. With this basis, we are able to show how the various canonical coordinates can be obtained by applying transformations to the guiding-center variables and Lagrangian of Sec. III. In particular, we show how flux coordinates are modified to obtain canonical variables.

In Sec. V, we discuss the longitudinal and drift adiabatic invariants of Hamiltonian guiding-center theory. First, we calculate these invariants to lowest order, as is valid for the guiding-center calculation developed here, and show how to obtain appropriate canonical action-angle coordinates. However, our calculation illustrates that the longitudinal adiabatic invariant exists to all orders for a guiding-center phase-space Lagrangian calculated only to finite order. This property is important for numerical analyses. Because an invariant corresponding to the longitudinal adiabatic invariant exists for a reasonably small ratio of gyroradius to scale length, one would like guiding-center dynamics also to have an invariant, so that numerical integrations yield long-time results consistent with such an invariant, e.g., the motion stays close to some surface in phase space. Truncated Hamiltonian guiding-center dynamics possesses an adiabatic invariant to all orders, while truncated non-Hamiltonian theories in general do not. For accurate, long-time results, it is therefore important to integrate Hamiltonian guiding-center equations of motion. First-order corrections for the magnetic moment (first adiabatic invariant) and the bounce action (second adiabatic invariant) are systematically derived by Lie-transform perturbation theory.

In Sec. VI, we discuss the covariant and noncovariant formulations of relativistic guiding-center Hamiltonian dynamics. While the covariant formulation is elegant, it is based on a covariant Hamiltonian that is not energy-like (since it is a Lorentz invariant), which makes it difficult to apply. The noncovariant formulation, on the other hand, treats time separately from spatial coordinates and uses an energylike relativistic Hamiltonian.

In Sec. VII, we use Hamiltonian guiding-center theory to discuss properties of configurations having reduced collisional transport. These concepts, isodynamism, guiding-center integrability, omnigeneity, and specific omnigeneity, are more easily discussed using the phase-space Lagrangian of Hamiltonian guiding-center dynamics, as it is easily transformed to the flux coordinates in which the properties of the equilibrium are naturally stated. Nührenberg and Zille (1988) relied crucially on the developments of Hamiltonian guiding-center dynamics to find their improved confinement configurations. Our discussion reviews the results in this area and shows how the concepts of specific omnigeneity, omnigeneity, guiding-center integrability, and isodynamism are successively more restrictive.

In Sec. VIII, we summarize our work and indicate possible areas of new research. In Appendix A, we show how the original guiding-center equations of motion of Northrop (1963), which lack Hamiltonian properties, can be modified to become the Northrop Hamiltonian guiding-center equations. The primary difference between the Northrop Hamiltonian guiding-center equations and the Hamiltonian guiding-center equations presented in Sec. III is that the polarization drift velocity is included in Northrop's guiding-center velocity while it is absent in the standard guiding-center velocity presented in Sec. III. In Appendix B, we review the derivation of several sets of coordinates for toroidal magnetic fields with nested flux surfaces. Last, in Appendix C, we present an introduction to the derivation of a Fokker-Planck collision operator in guiding-center phase space. Through the guiding-center phase-space transformation, the classical transport coefficients for spatial diffusion in a strongly magnetized plasma are recovered.

II. PHASE-SPACE LAGRANGIAN FORMULATION OF MECHANICS

Our analysis uses the phase-space Lagrangian formulation of mechanics. This formulation allows one to use arbitrary (noncanonical) coordinates in phase space, while retaining features of Hamiltonian mechanics, such as Noether's theorem and the Poincaré invariants. We begin this section with a review of Lagrangian and Hamiltonian mechanics. The phase-space Lagrangian follows immediately. Finally, we discuss how Noether's theorem and the Poincaré invariants appear in this formulation.

A. Lagrangian and Hamiltonian formulation

The Lagrangian formalism (Saletan and Cromer, 1971; Landau and Lifshitz, 1976; Arnold, 1989; Goldstein *et al.*, 2002) allows one to use arbitrary coordinates in configuration space. The Lagrangian for a set of coordinates $\mathbf{q}=(q^1,q^2,\ldots,q^N)$, where N denotes the number of degrees of freedom, is a function $L(\mathbf{q},\dot{\mathbf{q}},t)$ of the coordinates and their time derivatives. The equations of motion follow from requiring a trajectory's action (1.2) to be stationary with respect to variations of the trajectory $\mathbf{q}(t)$ in configuration space. This requirement yields the Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right)=\frac{\partial L}{\partial q^i} \tag{2.1}$$

for the trajectory. For charged-particle motion in an electromagnetic field, for example, the Lagrangian in Cartesian coordinates (for a particle of mass m and charge e) is

$$L(\mathbf{x},\dot{\mathbf{x}},t)=\frac{m}{2}|\dot{\mathbf{x}}|^2+\frac{e}{c}\dot{\mathbf{x}}\cdot\mathbf{A}(\mathbf{x},t)-e\Phi(\mathbf{x},t), \tag{2.2}$$

in terms of the scalar potential Φ and the vector potential $\mathbf{A}$, which give the electromagnetic field via $\mathbf{E}=-\nabla\Phi-c^{-1}\partial\mathbf{A}/\partial t$ and $\mathbf{B}=\nabla\times\mathbf{A}$.

The Lagrangian formalism is said to be *coordinate independent*. That is, the action integral (1.2) may be calculated in any coordinate system, and the statement $\delta\mathcal{A}=0$ that determines the trajectories may be stated without reference to any particular set of coordinates. In practice, this means that, as long as the time variable is not changed, the Lagrangian is a scalar. Hence, to transform the Lagrangian (2.2) to generalized coordinates $\mathbf{q}$, one substitutes $\mathbf{x}(\mathbf{q},t)$ and $\dot{\mathbf{x}}=\partial\mathbf{x}/\partial t+\Sigma\dot{q}^i\partial\mathbf{x}/\partial q^i$ into the Lagrangian (2.2), and one has the Lagrangian appropriate for the new variables. This property of the Lagrangian formalism allows one to find easily the dynamical equations (2.1) in new, more convenient coordinates, such as polar coordinates for the central force problem.

The Hamiltonian formalism (Arnold, 1989; Goldstein *et al.*, 2002) is one step more general than the Lagrangian formalism in that it places $\mathbf{q}$ and $\dot{\mathbf{q}}$ on an equal footing and allows more general transformations. We define the *canonical momentum* $\mathbf{p}$ with components

$$p_i=\frac{\partial L}{\partial \dot{q}^i}(\mathbf{q},\dot{\mathbf{q}},t). \tag{2.3}$$

The Hamiltonian formalism may be used whenever Eq. (2.3) may be inverted to find the velocities as functions of the canonical momenta, $\dot{q}^i(\mathbf{q},\mathbf{p},t)$. In this case, a point in *phase space* is determined by $\mathbf{q}$ and $\mathbf{p}$, rather than $\mathbf{q}$ and $\dot{\mathbf{q}}$. The equations of motion are Hamilton's equations,

$$\dot{q}^i=\frac{\partial H}{\partial p_i}\quad\text{and}\quad \dot{p}_i=-\frac{\partial H}{\partial q^i}, \tag{2.4}$$

where

$$H(\mathbf{q},\mathbf{p},t)=\mathbf{p}\cdot\dot{\mathbf{q}}(\mathbf{q},\mathbf{p},t)-L[\mathbf{q},\dot{\mathbf{q}}(\mathbf{q},\mathbf{p},t),t] \tag{2.5}$$

defines the Hamiltonian through the Legendre transformation (2.5). In Cartesian coordinates, the Hamiltonian corresponding to a charged particle moving in an electromagnetic field, from Eqs. (2.2) and (2.3), is

$$H(\mathbf{x},\mathbf{p},t)=\frac{1}{2m}\left|\mathbf{p}-\frac{e}{c}\mathbf{A}(\mathbf{x},t)\right|^2+e\Phi(\mathbf{x},t). \tag{2.6}$$

Here the canonical momenta (2.3) are

$$p_i=\frac{\partial L}{\partial \dot{x}^i}=m\dot{x}_i+\frac{e}{c}A_i.$$

The first term represents the kinetic momentum while the second term represents the magnetic part of the canonical momentum for a charged particle moving in a magnetic field.

While Hamiltonian mechanics treats the coordinates and momenta on equal footing and allows for a broader class of transformations, the set of transformations is yet restricted to *canonical* transformations. To define these transformations, it is convenient to denote the phase-space point by the $2N$-dimensional vector $\mathbf{z}=(q^1,\ldots,q^N,p_1,\ldots,p_N)$. The transformation to another set of coordinates $\mathbf{Z}(\mathbf{z},t)$ is canonical if the Jacobian matrix

$$D^\beta_\alpha=\frac{\partial Z^\beta}{\partial z^\alpha} \tag{2.7}$$

is symplectic, i.e., it satisfies $\mathsf{D}\cdot\sigma\cdot\mathsf{D}^\dagger=\sigma$, where σ is the fundamental symplectic $N\times N$ matrix

$$\boldsymbol{\sigma}=\begin{pmatrix}0 & \delta^{ij}\\ -\delta^{ij} & 0\end{pmatrix}, \tag{2.8}$$

and $\mathsf{D}^\dagger$ denotes the transpose of D.

The fundamental symplectic form (2.8) defines the Poisson brackets (denoted {,}) of the coordinates and the canonical-momenta among themselves,

$$\{z^\alpha,z^\beta\}=\sigma^{\alpha\beta}\quad\text{or}\quad\{q^i,p_j\}=\delta^i_j. \tag{2.9}$$

For any two functions f and g in phase space, the canonical Poisson bracket {,} is a bilinear antisymmetric differential operator defined as

$$\{f,g\}=\frac{\partial f}{\partial z^\alpha}\sigma^{\alpha\beta}\frac{\partial g}{\partial z^\beta}=\frac{\partial f}{\partial \mathbf{q}}\cdot\frac{\partial g}{\partial \mathbf{p}}-\frac{\partial f}{\partial \mathbf{p}}\cdot\frac{\partial g}{\partial \mathbf{q}}. \tag{2.10}$$

(Here and throughout summation over repeated indices is implied.) When expressed in terms of the canonical Poisson bracket (2.10), Hamilton's equations of motion can be written as

$$\frac{dz^\alpha}{dt} = \sigma^{\alpha\beta}\frac{\partial H}{\partial z^\beta}. \tag{2.11}$$

Canonical transformations are defined as those for which the Poisson bracket (2.10) remains form invariant, i.e., one could have used any set of canonical coordinates $\mathbf{Z}$ in place of $\mathbf{z}$ in Eq. (2.10). This and other properties of canonical transformations are discussed in standard textbooks on classical mechanics (Goldstein *et al.*, 2002).

Canonical transformations have the property of preserving the form of Hamilton's equations. That is, for any Hamiltonian $H(\mathbf{q},\mathbf{p},t)$ and any canonical transformation $(\mathbf{q},\mathbf{p})\to(\mathbf{Q},\mathbf{P})$ to new coordinates $\mathbf{Q}(\mathbf{q},\mathbf{p},t)$ and momenta $\mathbf{P}(\mathbf{q},\mathbf{p},t)$, there exists a new Hamiltonian $K(\mathbf{Q},\mathbf{P},t)$ giving the evolution of the new variables by Hamilton's equations $\dot{Q}^i \equiv \partial K/\partial P_i$ and $\dot{P}_i \equiv -\partial K/\partial Q^i$. Furthermore, the set of canonical transformations is the largest set such that a new Hamiltonian is guaranteed for any Hamiltonian H in the original variables.

B. Phase-space Lagrangian

A valuable feature of a Lagrangian formalism is that one can make arbitrary coordinate transformations. A phase-space Lagrangian (a Lagrangian that yields the correct equations of motion in phase space when all the phase-space coordinates are varied) should be easily transformed to an arbitrary (e.g., noncanonical) set of phase-space coordinates. In fact, the phase-space Lagrangian $\mathcal{L}$ is well known and equals the configuration-space Lagrangian L in value,

$$\mathcal{L}(\mathbf{q},\mathbf{p},\dot{\mathbf{q}},\dot{\mathbf{p}},t) \equiv \mathbf{p}\cdot\dot{\mathbf{q}} - H(\mathbf{q},\mathbf{p},t). \tag{2.12}$$

Hamilton's equations follow from the phase-space Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}^i}\right) = \frac{\partial \mathcal{L}}{\partial q^i} \to \dot{p}_i = -\frac{\partial H}{\partial q^i}$$

and

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{p}_i}\right) = \frac{\partial \mathcal{L}}{\partial p_i} \to 0 = \dot{q}^i - \frac{\partial H}{\partial p_i}.$$

The phase-space Lagrangian (2.12) that gives Hamilton's equations (2.11) has a very special form. The time derivatives of only half $(q^1,\ldots,q^N)$ of the phase-space variables $(q^1,\ldots,q^N;\, p_1,\ldots,p_N)$ are present. Further, these time derivatives are multiplied by the other half of the variables $(p_1,\ldots,p_N)$. We say that this phase-space Lagrangian is in canonical form. In the next section, we consider more general forms of the phase-space Lagrangian.

To illustrate the use of this formalism, we consider the motion of a particle in an electromagnetic field. Equations (2.6) and (2.12) together imply that the phase-space Lagrangian for this system is

$$\mathcal{L} = \mathbf{p}\cdot\dot{\mathbf{x}} - \frac{1}{2m}\left|\mathbf{p} - \frac{e}{c}\mathbf{A}(\mathbf{x},t)\right|^2 - e\Phi(\mathbf{x},t). \tag{2.13}$$

We may now choose to use the particle velocity

$$\mathbf{v} \equiv \frac{1}{m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right) \tag{2.14}$$

as a variable in place of the canonical momentum $\mathbf{p}$. The phase-space Lagrangian for the noncanonical variables $(\mathbf{x},\mathbf{v})$ is

$$\mathcal{L} = \left(m\mathbf{v} + \frac{e}{c}\mathbf{A}\right)\cdot\dot{\mathbf{x}} - \left(e\Phi + \frac{m}{2}|\mathbf{v}|^2\right). \tag{2.15}$$

The phase-space Euler-Lagrange equations for the velocity variables $\mathbf{v}$ yield $\partial\mathcal{L}/\partial\mathbf{v}=0$ (since $\partial\mathcal{L}/\partial\dot{\mathbf{v}}=0$), or

$$\dot{\mathbf{x}} = \mathbf{v}, \tag{2.16}$$

while for the coordinates $\mathbf{x}$ we obtain

$$\frac{d}{dt}\left[m\mathbf{v} + \frac{e}{c}\mathbf{A}(\mathbf{x},t)\right] = \nabla\left[\frac{e}{c}\mathbf{A}(\mathbf{x},t)\cdot\dot{\mathbf{x}} - e\Phi(\mathbf{x},t)\right],$$

which, when evaluated explicitly, yields

$$m\dot{\mathbf{v}} = e\mathbf{E} + \frac{e}{c}\dot{\mathbf{x}}\times\mathbf{B}. \tag{2.17}$$

This expression becomes the Lorentz force on a charged particle after the identification (2.16) is made.

C. Equations of motion for the phase-space Lagrangian

With a Lagrangian for the canonical phase-space variables $(\mathbf{q},\mathbf{p})$, we may transform to any $2N$ coordinates z^α $(\alpha=1,\ldots,2N)$ that parametrize the phase space by making the appropriate substitutions into the phase-space Lagrangian. In general, upon doing so, we no longer have the canonical form (2.12) for the phase-space Lagrangian, so we no longer have Hamilton's canonical equations. We now investigate the structure of the equations given by the general phase-space Lagrangian.

To calculate the new phase-space Lagrangian, we need the functions $q^i(\mathbf{z},t)$ and $p_i(\mathbf{z},t)$, which define the new parametrization of phase space. The total derivatives of the $\mathbf{q}$ coordinates become

$$\dot{q}^i = \frac{\partial q^i}{\partial t} + \dot{z}^\alpha\frac{\partial q^i}{\partial z^\alpha}. \tag{2.18}$$

Insertion of Eq. (2.18) into Eq. (2.12) yields the general form for a phase-space Lagrangian,

$$\mathcal{L} \equiv \Lambda_\alpha\dot{z}^\alpha - \mathcal{H}, \tag{2.19}$$

where

$$\Lambda_\alpha \equiv \mathbf{p}\cdot\frac{\partial\mathbf{q}}{\partial z^\alpha} \quad \text{and} \quad \mathcal{H} = H - \mathbf{p}\cdot\frac{\partial\mathbf{q}}{\partial t}. \tag{2.20}$$

The general phase-space Lagrangian (2.19) is written in terms of a *symplectic* part $(\Lambda_\alpha\dot{z}^\alpha)$, where time derivatives $\dot{z}^\alpha$ appear at first order only, and a Hamiltonian part $(-\mathcal{H})$. The notation (2.19) emphasizes that $\mathbf{\Lambda}$ is a covariant vector (or one-form) in the $2N$-dimensional phase space and the symplectic part is said to be canonical if the $2N$-dimensional covector $\mathbf{\Lambda}$ has N nonvanishing components. Note that, while the phase-space Lagrangian (2.19) represents a scalar field, it is sometimes useful to refer to its $2N+1$ components $(\Lambda_\alpha, -\mathcal{H})$. The action integral (1.2), therefore, becomes a path integral in phase-space-time $(\mathbf{z}, t)$,

$$\mathcal{A} = \int (\Lambda_\alpha dz^\alpha - \mathcal{H} dt). \tag{2.21}$$

This formalism allows one to express trajectories as occurring in the $(2N+1)$-dimensional geometry of phase-space-time. However, since we are not transforming time itself, it is sufficient to consider the $2N$-dimensional phase space with time-dependent transformations in what follows.

The addition of a total time derivative to the Lagrangian (2.19) does not affect the equations of motion, as its integral depends only on the end points, not the path. Addition of the total time derivative of a phase-space function $F(\mathbf{z}, t)$ to the Lagrangian $\mathcal{L} \to \mathcal{L} + dF/dt$ introduces the transformation

$$\Lambda_\alpha \to \Lambda_\alpha + \frac{\partial F}{\partial z^\alpha} \quad \text{and} \quad \mathcal{H} \to \mathcal{H} - \frac{\partial F}{\partial t}. \tag{2.22}$$

This is a type of gauge transformation and, hence, F is called a gauge function.

Variation of the Lagrangian (2.19) yields the phase-space Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{z}^\alpha}\right) = \frac{\partial \mathcal{L}}{\partial z^\alpha}. \tag{2.23}$$

Upon using Eq. (2.19), we find

$$\frac{\partial \mathcal{L}}{\partial \dot{z}^\alpha} = \Lambda_\alpha,$$

$$\frac{\partial \mathcal{L}}{\partial z^\alpha} = \frac{\partial \Lambda_\beta}{\partial z^\alpha}\dot{z}^\beta - \frac{\partial \mathcal{H}}{\partial z^\alpha},$$

so that Eq. (2.23) becomes

$$\frac{d\Lambda_\alpha}{dt} \equiv \frac{\partial \Lambda_\alpha}{\partial t} + \dot{z}^\beta \frac{\partial \Lambda_\alpha}{\partial z^\beta} = \frac{\partial \Lambda_\beta}{\partial z^\alpha}\dot{z}^\beta - \frac{\partial \mathcal{H}}{\partial z^\alpha}.$$

Next, upon rearranging terms, we finally obtain

$$\omega_{\alpha\beta}\frac{dz^\beta}{dt} \equiv \frac{\partial \mathcal{H}}{\partial z^\alpha} + \frac{\partial \Lambda_\alpha}{\partial t}, \tag{2.24}$$

where the Lagrange-bracket two-form $\boldsymbol{\omega}$ is the exterior derivative of the one-form $\mathbf{\Lambda}$ (Arnold, 1989),

$$\omega_{\alpha\beta} \equiv \frac{\partial \Lambda_\beta}{\partial z^\alpha} - \frac{\partial \Lambda_\alpha}{\partial z^\beta} = \frac{\partial \mathbf{p}}{\partial z^\alpha}\cdot\frac{\partial \mathbf{q}}{\partial z^\beta} - \frac{\partial \mathbf{p}}{\partial z^\beta}\cdot\frac{\partial \mathbf{q}}{\partial z^\alpha}. \tag{2.25}$$

The components of the two-form $\boldsymbol{\omega}$ can be used to construct a $2N \times 2N$ matrix known as the Lagrange matrix (Goldstein *et al.*, 2002). We note that the Lagrange matrix (2.25) and the right-hand side of Eq. (2.24) are both invariant with respect to the gauge transformation (2.22) and, thus, the equations of motion (2.24) are also gauge invariant.

Under the assumption of a regular Lagrangian (for which $\det \boldsymbol{\omega} \neq 0$), we define the Poisson matrix $\mathbf{\Pi}$ to be the inverse of the Lagrange matrix $\boldsymbol{\omega}$ (i.e., $\Pi^{\alpha\beta}\omega_{\beta\gamma} = \delta^\alpha_\gamma$). The equations of motion (2.24) can then be inverted to give

$$\begin{aligned}\frac{dz^\alpha}{dt} &= \Pi^{\alpha\beta}\left(\frac{\partial \mathcal{H}}{\partial z^\beta} + \frac{\partial \Lambda_\beta}{\partial t}\right) \\ &= \Pi^{\alpha\beta}\left[\frac{\partial H}{\partial z^\beta} + \left(\frac{\partial \mathbf{p}}{\partial t}\cdot\frac{\partial \mathbf{q}}{\partial z^\beta} - \frac{\partial \mathbf{p}}{\partial z^\beta}\cdot\frac{\partial \mathbf{q}}{\partial t}\right)\right].\end{aligned} \tag{2.26}$$

Thus, the $(2N+1)$-dimensional phase-space Lagrangian (2.19) is associated with Hamilton's equations (2.26) as follows: the Poisson matrix $\Pi^{\alpha\beta}$ is obtained from the $2N$-dimensional symplectic covector $\mathbf{\Lambda}$ and the Hamiltonian $\mathcal{H}$ is obtained from its Hamiltonian part. Hamilton's equations (2.26) can be used to derive the energy equation

$$\begin{aligned}\frac{d\mathcal{H}}{dt} &= \frac{\partial \mathcal{H}}{\partial t} + \frac{\partial \mathcal{H}}{\partial z^\alpha}\frac{dz^\alpha}{dt} \\ &= \frac{\partial \mathcal{H}}{\partial t} + \frac{\partial \mathcal{H}}{\partial z^\alpha}\Pi^{\alpha\beta}\left(\frac{\partial \mathcal{H}}{\partial z^\beta} + \frac{\partial \Lambda_\beta}{\partial t}\right) \\ &= \frac{\partial \mathcal{H}}{\partial t} - \frac{dz^\beta}{dt}\frac{\partial \Lambda_\beta}{\partial t},\end{aligned} \tag{2.27}$$

where we used the antisymmetry of the Poisson matrix. The energy equation (2.27) clearly shows that energy is conserved in time-independent $(\partial/\partial t \equiv 0)$ Hamiltonian systems. Note that a general Hamiltonian system based on general $2N$-dimensional phase-space coordinates $\mathbf{z}$ is determined by a Hamiltonian function $\mathcal{H}(\mathbf{z}, t)$ and a Poisson-bracket structure based on the Poisson tensor $\Pi^{\alpha\beta}(\mathbf{z}, t)$.

We now calculate the Jacobian matrix for the time-dependent transformation $\mathbf{Z} \to \mathbf{z}(\mathbf{Z}, t)$ from canonical coordinates $\mathbf{Z} = (\mathbf{q}, \mathbf{p})$ to noncanonical coordinates $\mathbf{z}$. From Eq. (2.7), we find $D^\beta_\alpha = \partial z^\beta / \partial Z^\alpha$. Thus, the inverse transformation $\mathbf{z} \to \mathbf{Z}(\mathbf{z}, t)$ has the Jacobian matrix $(\mathsf{D}^{-1})^\alpha_\beta = \partial Z^\alpha / \partial z^\beta$. In matrix notation, the Lagrange tensor (2.25) can be written as

$$\boldsymbol{\omega} = -(\mathsf{D}^{-1})^\dagger \cdot \boldsymbol{\sigma} \cdot \mathsf{D}^{-1}, \tag{2.28}$$

and, therefore, its inverse is given by

$$\mathbf{\Pi} = \mathsf{D} \cdot \boldsymbol{\sigma} \cdot \mathsf{D}^\dagger. \tag{2.29}$$

Hence, $\Pi^{\alpha\beta} = \{z^\alpha, z^\beta\}$ is the Poisson bracket of z^α with z^β; it equals $\boldsymbol{\sigma}$ for canonical coordinates. (Often the Poisson tensor is denoted by J, but we reserve this symbol for the action of the action-angle variables.)

For the case of time-independent transformations to noncanonical coordinates $(\mathcal{H} = H)$, the noncanonical equations of motion (2.26) become

$$\frac{dz^\alpha}{dt} = \Pi^{\alpha\beta}\frac{\partial H}{\partial z^\beta} \equiv \{z^\alpha, H\}. \tag{2.30}$$

The last form shows that the evolution can still be expressed in terms of a Poisson bracket, here generalized to

$$\{f,g\} \equiv \frac{\partial f}{\partial z^\alpha}\Pi^{\alpha\beta}\frac{\partial g}{\partial z^\beta} \tag{2.31}$$

for noncanonical coordinates. The noncanonical Poisson bracket (2.31) satisfies the Jacobi identity $\{f,\{g,h\}\}+\{g,\{h,f\}\}+\{h,\{f,g\}\}=0$ for three arbitrary functions f, g, and h, which can also be expressed as

$$\Pi^{\alpha\nu}\partial_\nu\Pi^{\beta\gamma} + \Pi^{\beta\nu}\partial_\nu\Pi^{\gamma\alpha} + \Pi^{\gamma\nu}\partial_\nu\Pi^{\alpha\beta} = 0. \tag{2.32}$$

That the Poisson bracket (2.31) satisfies the Jacobi identity (2.32) follows from the fact that $\mathbf{\Pi}$ is the inverse of the exterior derivative of a one-form (i.e., $\omega_{\alpha\gamma}\Pi^{\gamma\beta}=\delta^\beta_\alpha$). Indeed, the Jacobi identity (2.32) may be expressed in terms of components of the Lagrange tensor $\omega_{\alpha\beta}$ as

$$\partial_\sigma\omega_{\alpha\beta} + \partial_\alpha\omega_{\beta\sigma} + \partial_\beta\omega_{\sigma\alpha} = 0,$$

which is identically satisfied from the definition (2.25) of $\omega_{\alpha\beta}$. We note that in the canonical case, the Jacobi identity (2.32) is identically satisfied since the components of the Poisson tensor $\mathbf{\Pi}=\boldsymbol{\sigma}$ are constants (i.e., the components are either 0, +1, or −1). In the noncanonical case, however, the components $\Pi^{\alpha\beta}(\mathbf{z},t)$ of the Poisson tensor depend on phase-space-time coordinates and, therefore, the fact that our Poisson tensor is guaranteed to satisfied the Jacobi identity is a great advantage.

It is illustrative to evaluate these tensors for the phase-space Lagrangian (2.15), which uses the noncanonical coordinates $(\mathbf{x},\mathbf{v})$. We express these tensors in 3×3 block form as

$$\boldsymbol{\omega} = m\begin{pmatrix} \epsilon_{ijk}\Omega^k & -\delta_{ij} \\ \delta_{ij} & 0 \end{pmatrix} \tag{2.33}$$

for the Lagrange tensor (here, i, j, and k take values from 1 to 3 and $\Omega^k=eB^k/mc$). Inverting this matrix yields the Poisson tensor

$$\mathbf{\Pi} = m^{-1}\begin{pmatrix} 0 & \delta^{ij} \\ -\delta^{ij} & \epsilon^{ijk}\Omega_k \end{pmatrix}. \tag{2.34}$$

The noncanonical Poisson bracket for charged particle motion in an electromagnetic field is therefore

$$\{f,g\} = \frac{1}{m}\left(\nabla f\cdot\frac{\partial g}{\partial \mathbf{v}} - \frac{\partial f}{\partial \mathbf{v}}\cdot\nabla g\right) + \frac{e\mathbf{B}}{m^2c}\cdot\frac{\partial f}{\partial \mathbf{v}}\times\frac{\partial g}{\partial \mathbf{v}}. \tag{2.35}$$

Here we see that, up to a factor of mass m, the first two terms possess the canonical form. The last term, on the other hand, involves the magnetic field $\mathbf{B}=\nabla\times\mathbf{A}$ explicitly, while the noncanonical Hamiltonian

$$H(\mathbf{x},\mathbf{v},t) = \frac{m}{2}|\mathbf{v}|^2 + e\Phi(\mathbf{x},t) \tag{2.36}$$

now involves only the scalar potential. Last, note that the energy equation (2.27) becomes

$$\frac{dH}{dt} = e\frac{\partial\Phi}{\partial t} - \frac{e}{c}\frac{\partial\mathbf{A}}{\partial t}\cdot\dot{\mathbf{x}}, \tag{2.37}$$

which vanishes for time-independent potentials.

D. Noether's theorem

Noether's theorem states in general that where there is a continuous family of transformations for which the Lagrangian is invariant (i.e., the transformations correspond to a symmetry of the Lagrangian), then there exists a corresponding constant of motion. Such a symmetry exists when one of the coordinates is ignorable (i.e., only derivatives of the ignorable coordinate appear in the Lagrangian, not the coordinate itself). Then the symmetry is represented by the family of translations in the ignorable coordinate.

For the present case, we assume that one of the coordinates (say, z^β) does not appear in the Lagrangian, i.e., none of the one-form components Λ_α depend on z^β and $\Lambda_\beta=\partial\mathcal{L}/\partial\dot{z}^\beta\neq0$. The phase-space Euler-Lagrange equation (2.23) for $\alpha=\beta$ yields

$$\frac{d\Lambda_\beta}{dt} = \frac{\partial\mathcal{L}}{\partial z^\beta} = 0, \tag{2.38}$$

which shows that Λ_β is an invariant (Cary, 1977). For the case in which the Lagrangian does not depend on time, the same argument can be made to show that the Hamiltonian H is an invariant.

E. Liouville's theorem

The transformation Jacobian $\mathcal{J}=\det(\mathsf{D}^{-1})$ may not be constant when the phase-space transformation is noncanonical. In fact, for a time-dependent transformation, the Jacobian $\mathcal{J}(\mathbf{z},t)$ satisfies the divergence equation

$$\frac{\partial\mathcal{J}}{\partial t} + \frac{\partial}{\partial z^\alpha}\left(\mathcal{J}\frac{dz^\alpha}{dt}\right) = 0. \tag{2.39}$$

This equation implies that the equations of motion (2.26) satisfy the Liouville theorem, i.e., the Hamiltonian flow conserves the phase-space volume $d^3q\,d^3p=\mathcal{J}d^6z$.

For a time-independent transformation, the Liouville theorem (2.39) becomes

$$0 = \frac{\partial}{\partial z^\alpha}\left(\mathcal{J}\Pi^{\alpha\beta}\frac{\partial H}{\partial z^\beta}\right) = \frac{\partial}{\partial z^\alpha}(\mathcal{J}\Pi^{\alpha\beta})\frac{\partial H}{\partial z^\beta}, \tag{2.40}$$

which, for an arbitrary Hamiltonian H, yields the Liouville identities

$$\frac{\partial}{\partial z^\alpha}(\mathcal{J}\Pi^{\alpha\beta}) = 0. \tag{2.41}$$

These Liouville identities imply that we may write the noncanonical Poisson bracket (2.31) as a phase-space divergence,

$$\{f,g\} = \frac{1}{\mathcal{J}}\frac{\partial}{\partial z^\alpha}\left(f\mathcal{J}\Pi^{\alpha\beta}\frac{\partial g}{\partial z^\beta}\right), \quad (2.42)$$

so that $\int\{f,g\}\mathcal{J}d^6z \equiv 0$.

We note that the phase-space Lagrangian (2.19) contains within it the information needed to calculate the Jacobian $\mathcal{J}$. Using Eq. (2.28), we find $\det(\boldsymbol{\omega}) = \mathcal{J}^2$, so that $\det(\boldsymbol{\omega})$ is positive and, with the convention that $\mathcal{J}$ is positive,

$$\mathcal{J} = \sqrt{\det(\boldsymbol{\omega})}. \quad (2.43)$$

Hence, the Jacobian can be calculated directly from the (phase-space) Lagrangian matrix (2.25) rather than the Jacobian matrix (2.28). For example, from the Lagrange matrix (2.33), we find $\mathcal{J}=m^3$ and, thus, $d^3q d^3p = m^3 d^3x d^3v$. In addition, the Liouville identities associated with the noncanonical Poisson bracket (2.35) are

$$\frac{\partial}{\partial v^j}\left(\frac{em}{c}B_i\epsilon^{ijk}\right) = 0.$$

III. NONCANONICAL HAMILTONIAN GUIDING-CENTER THEORY

In the present section, we start by giving the guiding-center phase-space Lagrangian and show how the equations of motion are derived from it. From there, and the results of Sec. II, it is straightforward to derive the drift-kinetic equation, which determines the evolution of a distribution of guiding centers. From the Hamiltonian guiding-center theory, we are able to obtain the conserved volume and Hamiltonian drift-kinetic equation directly, in Sec. II.E, without reference to the transformation, as the volume element is contained in the phase-space Lagrangian. Only then do we step back and show how the guiding-center Lagrangian is derived (in Sec. III.C). This derivation is needed to understand the transformation from the usual variables to guiding-center variables as is needed and used in Sec. III.D to derive the currents due to a given guiding-center distribution. As an application of noncanonical Hamiltonian guiding-center theory, we show how to derive the guiding-center angular momentum conserved for axisymmetric systems in Sec. III.E. Finally, in Sec. III.F, we show how the Hamiltonian formulation of field line flow follows directly from noncanonical Hamiltonian guiding-center theory.

A. Guiding-center Lagrangian

The guiding-center phase space consists of the guiding-center position $\mathbf{X}$, essentially the center of the helix; the guiding-center parallel velocity variable $u \equiv \hat{\mathsf{b}}\cdot\dot{\mathbf{X}}$; the (lowest-order) magnetic moment,

$$\mu \equiv \frac{m|\mathbf{w}|^2}{2B(\mathbf{X},t)}, \quad (3.1)$$

where $\mathbf{w} \equiv \mathbf{v}_\perp - \mathbf{v}_E$ is the perpendicular velocity in the local frame moving with the $E\times B$ drift velocity $\mathbf{v}_E \equiv \mathbf{E}\times c\hat{\mathsf{b}}/B$; and the ignorable gyrophase ζ, which gives the location of the particle on the circle about the guiding center. As there are still six variables parametrizing phase space, there is no loss of information in making the guiding-center transformation $(\mathbf{x},\mathbf{v}) \rightarrow (\mathbf{X},u,\mu,\zeta)$. For the sake of simplicity of notation, we occasionally use the gyroaction variable $J \equiv (mc/e)\mu$ instead of the magnetic moment μ whenever we need to refer to the action-angle coordinates (J,ζ) associated with gyromotion.

The equations of motion for these variables are given by the guiding-center phase-space Lagrangian,

$$\mathcal{L}_{\rm gc}(\mathbf{X},u,\mu,\zeta;t) = \left[\frac{e}{c}\mathbf{A}(\mathbf{X},t) + mu\hat{\mathsf{b}}(\mathbf{X},t)\right]\cdot\dot{\mathbf{X}} + J\dot{\zeta} - H_{\rm gc}, \quad (3.2)$$

in which the guiding-center Hamiltonian is given by

$$H_{\rm gc}(\mathbf{X},u,\mu;t) = \frac{m}{2}u^2 + \mu B(\mathbf{X},t) + e\Phi(\mathbf{X},t) - \frac{m}{2}|\mathbf{v}_E(\mathbf{X},t)|^2. \quad (3.3)$$

The arguments are shown here to emphasize that, for example, the magnetic-field strength $B(\mathbf{X},t)$ is evaluated at the guiding-center position $\mathbf{X}$, not at the particle position $\mathbf{x}$. Here and throughout, the effects of a gravitational field may by found by adding the gravitational $m\Phi_G$ to the electrostatic potential energy $e\Phi$. In addition, we now drop the adjective phase space, as the guiding-center Lagrangian is always henceforth a phase-space Lagrangian.

The guiding-center Lagrangian (3.2) comes not simply from gyrophase averaging, but from a transformation from the physical (particle) variables $(\mathbf{x},\mathbf{v})$ to the guiding-center variables $(\mathbf{X},u,\mu,\zeta)$. The details of the gyrophase ζ definition will be presented later (Sec. III.C). The definitions of the parallel velocity u and magnetic moment μ have already been given. To complete the picture, we must give the relation

$$\mathbf{x} \equiv \mathbf{X} + \boldsymbol{\rho} \quad (3.4)$$

between the physical location $\mathbf{x}$ and the guiding-center position $\mathbf{X}$, where $\boldsymbol{\rho}$ denotes the (gyroradius) displacement vector in the frame drifting with the $E\times B$ velocity $\mathbf{v}_E$. Here we simply note that the displacement vector $\boldsymbol{\rho} \equiv \tilde{\boldsymbol{\rho}} + \bar{\boldsymbol{\rho}}$ has a part (denoted $\tilde{\boldsymbol{\rho}}$) that is explicitly gyrophase dependent and a part (denoted $\bar{\boldsymbol{\rho}}$) that is gyrophase independent. In what follows, we show that the latter part

$$\bar{\boldsymbol{\rho}} = \frac{\hat{\mathsf{b}}}{\Omega} \times \mathbf{v}_E = \frac{c\mathbf{E}_\perp}{\Omega B} \tag{3.5}$$

denotes the guiding-center polarization displacement (Kaufman, 1986). The guiding-center Lagrangian (3.2) has the noncanonical form (2.19) and the guiding-center Hamiltonian (3.3) is given as the sum of the guiding-center kinetic energy $(m/2)u^2+\mu B$ plus the effective electric potential energy $e\Phi-(m/2)|\mathbf{v}_E|^2$, which is obtained as a result of the finite-Larmor-radius (FLR) expansion of the sum $e\Phi(\mathbf{X}+\bar{\boldsymbol{\rho}})+(m/2)|\mathbf{v}_E|^2$. An alternate choice for the guiding-center Lagrangian and Hamiltonian associated with the choice $\bar{\boldsymbol{\rho}}\equiv 0$ is discussed in Appendix A.

Our guiding-center Lagrangian (3.2) is obtained from an ordering in which the scalar potential Φ is one order lower than the particle kinetic energy, unlike previous derivations of the Hamiltonian theory of guiding-center motion. In this ordering, the electric drift $\mathbf{v}_E$ is of the same order as the perpendicular velocity $\mathbf{w}$, as in some non-Hamiltonian calculations (Northrop, 1963). As we will see, this ordering allows us to obtain the polarization drift in the same order as the curvature and ∇B drifts, although it appears differently in the theory. However, for consistency the parallel electric field $E_\parallel=\mathbf{E}\cdot\hat{\mathsf{b}}$ must be smaller (by one order) than the perpendicular field $\mathbf{E}_\perp$.

The variables (J,ζ) appear in canonical form in the symplectic part of the guiding-center Lagrangian (3.2) as $J\dot{\zeta}$ while the guiding-center Hamiltonian $H_{\rm gc}$ depends on J (or μ) alone. The Hamilton equations for (J,ζ) are

$$\dot{J} = -\frac{\partial H_{\rm gc}}{\partial \zeta} \equiv 0, \tag{3.6}$$

$$\dot{\zeta} = \frac{\partial H_{\rm gc}}{\partial J} \equiv \Omega. \tag{3.7}$$

Equation (3.6) shows that the gyroaction (or magnetic moment) is conserved by the guiding-center equations of motion. This also follows from Noether's theorem (Sec. II.D) since the gyrophase ζ is an ignorable coordinate, i.e., only its time derivative appears in the guiding-center Lagrangian (3.2).

If one is concerned with only the motion of the guiding center and not the evolution of the gyrophase, the term linear in $\dot{\zeta}$ can be dropped from the guiding-center Lagrangian, as it does not affect the equations of motion of the other variables, $\mathbf{X}$ and u. In the evolution equations for $\mathbf{X}$ and u, the adiabatic invariant μ (or J) does appear but only as a guiding-center dynamical parameter.

Variation of the guiding-center Lagrangian (3.2) with respect to the variable u gives the Euler-Lagrange equation

$$0 = \frac{\partial \mathcal{L}}{\partial u} = m\hat{\mathsf{b}}\cdot\dot{\mathbf{X}} - \frac{\partial H_{\rm gc}}{\partial u},$$

which yields

$$u \equiv \hat{\mathsf{b}}(\mathbf{X},t)\cdot\dot{\mathbf{X}}. \tag{3.8}$$

Thus, the guiding-center Lagrangian (3.2) dictates that u is the velocity of the guiding center in the direction of the magnetic field at the guiding center. [In the present theory, we have not included the parallel drift of Baños (1967) and Hazeltine (1973), which is discussed by Northrop and Rome (1978). In the derivation in Sec. III.D, we discuss how these terms are obtained.]

Last, we vary the Lagrangian (3.2) with respect to the guiding-center position $\mathbf{X}$. With manipulations similar to those used to derive the Lorentz force from the Lagrangian (2.15), we obtain the Euler-Lagrange equation

$$\begin{aligned} m\dot{u}\hat{\mathsf{b}} &= e\mathbf{E} - \mu\nabla B + \frac{m}{2}\nabla|\mathbf{v}_E|^2 - mu\frac{\partial\hat{\mathsf{b}}}{\partial t} \\ &\quad + \dot{\mathbf{X}}\times\left(\frac{e}{c}\mathbf{B} + mu\nabla\times\hat{\mathsf{b}}\right) \\ &\equiv e\left(\mathbf{E}^* + \frac{1}{c}\dot{\mathbf{X}}\times\mathbf{B}^*\right), \end{aligned} \tag{3.9}$$

where the effective electromagnetic fields

$$\mathbf{E}^* \equiv -\nabla\Phi^* - \frac{1}{c}\frac{\partial\mathbf{A}^*}{\partial t} \quad \text{and} \quad \mathbf{B}^* \equiv \nabla\times\mathbf{A}^* \tag{3.10}$$

are defined in terms of the effective electromagnetic potentials

$$\begin{aligned} e\Phi^* &\equiv e\Phi + \mu B - (m/2)|\mathbf{v}_E|^2, \\ \mathbf{A}^* &\equiv \mathbf{A} + (mc/e)u\hat{\mathsf{b}}. \end{aligned} \tag{3.11}$$

The guiding-center canonical momentum is now simply expressed as $e\mathbf{A}^*/c$ and the guiding-center Hamiltonian is $e\Phi^*+mu^2/2$.

We obtain the rate of change of the variable u by taking the scalar product of Eq. (3.9) with the effective magnetic field $\mathbf{B}^*$,

$$\dot{u} = -\frac{\mathbf{B}^*}{mB^*_\parallel}\cdot\left(\nabla H_{\rm gc} + \frac{e}{c}\frac{\partial\mathbf{A}^*}{\partial t}\right) \equiv \frac{e}{m}\frac{\mathbf{B}^*}{B^*_\parallel}\cdot\mathbf{E}^*, \tag{3.12}$$

with $B^*_\parallel\equiv\hat{\mathsf{b}}\cdot\mathbf{B}^*$ the effective magnetic field in the parallel direction (see Sec. III.B for more details concerning $B^*_\parallel$). The time derivative (3.12) contains terms that are higher order (in gyroradius) compared with the dominant terms, which are all that is usually kept. These higher-order terms, however, are needed for energy conservation.

The guiding-center velocity $\dot{\mathbf{X}}$ comes from the vector product of Eq. (3.9) with $\hat{\mathsf{b}}$ which, using Eq. (3.8), yields

$$\dot{\mathbf{X}} = \frac{\mathbf{B}^*}{mB_\parallel^*}\frac{\partial H_{\rm gc}}{\partial u} + \frac{c\hat{\mathsf{b}}}{eB_\parallel^*}\times\left(\nabla H_{\rm gc} + \frac{e}{c}\frac{\partial \mathbf{A}^*}{\partial t}\right)$$
$$= u\frac{\mathbf{B}^*}{B_\parallel^*} + \mathbf{E}^*\times\frac{c\hat{\mathsf{b}}}{B_\parallel^*}. \tag{3.13}$$

If the effective fields (3.10) were replaced by the standard fields $(\mathbf{E},\mathbf{B})$, Eqs. (3.12) and (3.13) would be the equations of motion for a particle in straight, constant electric, and magnetic fields.

The guiding-center equations of motion (3.12) and (3.13) can be used to derive the guiding-center energy equation

$$\frac{dH_{\rm gc}}{dt} = \frac{\partial H_{\rm gc}}{\partial t} + \dot{\mathbf{X}}\cdot\nabla H_{\rm gc} + \dot{u}\frac{\partial H_{\rm gc}}{\partial u} = e\frac{\partial \Phi^*}{\partial t} - \frac{e}{c}\frac{\partial \mathbf{A}^*}{\partial t}\cdot\dot{\mathbf{X}}, \tag{3.14}$$

which implies that the guiding-center energy $E_{\rm gc} \equiv \frac{1}{2}mu^2 + e\Phi^*$ is a constant of the motion for time-independent fields.

Taylor (1964) obtained the following Lagrangian for guiding-center motion:

$$L_{\rm T} = \frac{m}{2}\dot{s}^2 + \frac{e}{c}\alpha\dot{\beta} - e\Phi - \mu B, \tag{3.15}$$

where the magnetic coordinates (α,β,s) are used to describe the magnetic field $\mathbf{B} = \nabla\alpha\times\nabla\beta = B\partial\mathbf{x}/\partial s$, and the kinetic energy associated with the drift motion has been omitted. The Euler-Lagrange equations obtained from the Taylor Lagrangian (3.15) are

$$(e/c)\dot{\alpha} = -e\partial_\beta\Phi - \mu\partial_\beta B,$$
$$(e/c)\dot{\beta} = e\partial_\alpha\Phi + \mu\partial_\alpha B,$$
$$m\ddot{s} = -e\partial_s\Phi - \mu\partial_s B. \tag{3.16}$$

It is immediately clear that Taylor's Lagrangian (3.15) is not a phase-space Lagrangian [see Eq. (2.19)] since the parallel velocity $\dot{s}$ appears quadratically. As a result, the Poisson-bracket structure of Taylor's Lagrangian and the Liouville properties of Eqs. (3.16) are unclear. [See Eq. (4.30) for the correct phase-space Lagrangian form for the Taylor Lagrangian.] In contrast, the guiding-center Lagrangian (3.2) fits naturally into the general formalism of phase-space Lagrangians presented in the previous section. Hence, the Hamiltonian (conserved for autonomous systems) is known, the volume element can be derived from the phase-space Lagrangian via Eq. (2.43), and for ignorable coordinates z^α the conjugate component Λ_α is a constant of motion.

An important remark must be made here concerning the polarization drift, which is absent from the guiding-center velocity (3.13). This drift, however, is critical for obtaining the dielectric response of a low-frequency plasma. Instead, it appears in the transformation (3.4) itself, i.e., the derivative of this relation gives

$$\dot{\mathbf{x}} = \dot{\mathbf{X}} + \dot{\tilde{\boldsymbol{\rho}}} + \mathbf{v}_{\rm pol}, \tag{3.17}$$

with

$$\mathbf{v}_{\rm pol} \equiv \frac{d\bar{\boldsymbol{\rho}}}{dt} \tag{3.18}$$

representing the polarization drift (Sosenko *et al.*, 2001; Brizard, 2008), and where $\dot{\tilde{\boldsymbol{\rho}}} = \dot{\zeta}\partial\tilde{\boldsymbol{\rho}}/\partial\zeta + \cdots$ consists of terms that oscillate on the gyroperiod time scale. The polarization drift is a pure derivative and, hence, can always be integrated.[1]

An alternate set of guiding-center equations of motion may be derived in which the polarization drift appears explicitly in the guiding-center velocity $\dot{\mathbf{X}}$ by choosing $\bar{\boldsymbol{\rho}} \equiv \mathbf{0}$ instead of Eq. (3.5). This alternate set is presented in Appendix A.

B. Guiding-center phase-space volume conservation law

We calculate the Jacobian $\mathcal{J}_{\rm gc}$ from the Lagrange tensor (2.33) via Eq. (2.43), thereby avoiding the transformation, which we have not introduced at this point. It follows from Eq. (2.33) that the antisymmetric guiding-center Lagrange tensor for the variables $(\mathbf{X},u,\mu,\zeta)$ is in block-diagonal form,

$$\boldsymbol{\omega}_{\rm gc} = \begin{pmatrix} \hat{\boldsymbol{\omega}}_{\rm gc} & 0 \\ 0 & \begin{pmatrix} 0 & -(mc/e) \\ (mc/e) & 0 \end{pmatrix} \end{pmatrix}, \tag{3.19}$$

where $\hat{\boldsymbol{\omega}}_{\rm gc}$ is the 4×4 $(\mathbf{X},u)$ part of the guiding-center Lagrange tensor. The components of $\hat{\boldsymbol{\omega}}_{\rm gc}$ are found from the guiding-center Lagrangian (3.2) by exterior differentiation of its phase-space part as in Eq. (2.25). For the $\mathbf{X}$-u and u-$\mathbf{X}$ parts, we find

$$\hat{\omega}_{ui} = \hat{\omega}_{iu} = mb_i. \tag{3.20}$$

For the $\mathbf{X}$-$\mathbf{X}$ part, we obtain

$$\hat{\omega}_{ij} = \frac{\partial}{\partial X^i}\left(\frac{e}{c}A_j + mub_j\right) - \frac{\partial}{\partial X^j}\left(\frac{e}{c}A_i + mub_i\right)$$
$$= \epsilon_{ijk}\frac{e}{c}B^{*k}, \tag{3.21}$$

where ϵ_{ijk} is the Levi-Civita symbol. From these expressions it follows that the 4×4 guiding-center Lagrange tensor is

[1] This implies, in particular, that the polarization drift cannot lead to diffusion even in a turbulent field. This is important, as the difference between the guiding center and the average location [found by dropping the second, oscillating term in Eq. (3.4)] is the polarization, the integral of the polarization drift. This difference must remain small or else the theory, which assumes that the particle remains close to $\mathbf{X}$, would break down.

TABLE I. Guiding-center ordering required for the existence of the magnetic-moment invariant.

Order	Dimensionless	Fields	Distances	Rates	Velocities
ϵ^{-1}		$\mathbf{B}$, $\mathbf{E}_\perp$		Ω	
1		$E_\parallel$	L	$\mathbf{v}/\mathrm{L}$, $\mathbf{v}_E/\mathrm{L}$, τ^{-1}	$\mathbf{v}$, $\mathbf{v}_E$
ϵ	ρ/L, $(\Omega\tau)^{-1}$		ρ	$\mathbf{v}_\nabla/L$, $\mathbf{v}_{\rm pol}/\mathrm{L}$, $\mathbf{v}_\kappa/\mathrm{L}$	$\mathbf{v}_\nabla$, $\mathbf{v}_\kappa$, $\mathbf{v}_{\rm pol}$

$$\hat{\boldsymbol{\omega}}_{\rm gc} = m\begin{pmatrix} 0 & \Omega_3^* & -\Omega_2^* & -b_1 \\ -\Omega_3^* & 0 & \Omega_1^* & -b_2 \\ \Omega_2^* & -\Omega_1^* & 0 & -b_3 \\ b_1 & b_2 & b_3 & 0 \end{pmatrix}, \tag{3.22}$$

where

$$\Omega_i^* \equiv \frac{eB_i^*}{mc}, \tag{3.23}$$

and so

$$\mathcal{J}_{\rm gc} = \sqrt{\det(\boldsymbol{\omega}_{\rm gc})} = m^2\Omega_\parallel^* \equiv (me/c)B_\parallel^* \tag{3.24}$$

is the Jacobian, where

$$B_\parallel^* \equiv \hat{\mathsf{b}}\cdot\mathbf{B}^* = B + (mc/e)u\hat{\mathsf{b}}\cdot\nabla\times\hat{\mathsf{b}}. \tag{3.25}$$

An important Hamiltonian property of the guiding-center equations of motion (3.12) and (3.13) is that they satisfy the guiding-center phase-space volume conservation law

$$\frac{\partial B_\parallel^*}{\partial t} + \nabla\cdot(B_\parallel^*\dot{\mathbf{X}}) + \frac{\partial}{\partial u}(B_\parallel^*\dot{u}) = 0. \tag{3.26}$$

The explicit proof of this important Hamiltonian conservation law is presented in Appendix A.

For later reference, we also provide the guiding-center Poisson tensor associated with Eq. (3.2). It has the block diagonal form

$$\boldsymbol{\Pi}_{\rm gc} = \begin{pmatrix} \hat{\boldsymbol{\Pi}}_{\rm gc} & 0 \\ 0 & \begin{pmatrix} 0 & (e/mc) \\ -(e/mc) & 0 \end{pmatrix} \end{pmatrix}, \tag{3.27}$$

where $\hat{\boldsymbol{\Pi}}_{\rm gc}$ is the 4×4 $(\mathbf{X},u)$ part of the guiding-center Poisson tensor,

$$\hat{\boldsymbol{\Pi}}_{\rm gc} = \frac{1}{m\Omega_\parallel^*}\begin{pmatrix} 0 & -b_3 & -b_2 & \Omega_1^* \\ b_3 & 0 & -b_1 & \Omega_2^* \\ -b_2 & b_1 & 0 & \Omega_3^* \\ -\Omega_1^* & -\Omega_2^* & -\Omega_3^* & 0 \end{pmatrix}. \tag{3.28}$$

With the guiding-center Poisson tensor (3.27), the guiding-center equations of motion (3.12) and (3.13) clearly have the form (2.26). The guiding-center Poisson bracket is thus expressed as

$$\{f,g\}_{\rm gc} = \epsilon^{-1}\frac{\Omega}{B}\left(\frac{\partial f}{\partial\theta}\frac{\partial g}{\partial\mu} - \frac{\partial f}{\partial\mu}\frac{\partial g}{\partial\theta}\right) + \frac{\mathbf{B}^*}{mB_\parallel^*}\cdot\left(\nabla f\frac{\partial g}{\partial u} - \frac{\partial f}{\partial u}\nabla g\right) - \epsilon\frac{c\hat{\mathsf{b}}}{eB_\parallel^*}\cdot\nabla f\times\nabla g, \tag{3.29}$$

where the ϵ scaling of each term is shown explicitly. The first term (with ϵ^{-1} ordering) represents the fast gyromotion dynamics, the second term (with ϵ^0 ordering) represents the intermediate bounce-motion dynamics along magnetic-field lines, and the third term (with ϵ ordering) represents the slow drift-motion dynamics across magnetic-field lines.

C. Derivation of the guiding-center Lagrangian

To obtain the guiding-center Lagrangian (3.2), we seek a transformation to alternative phase-space variables in which the new, guiding-center Lagrangian has a simple form. The simple form is one in which the degree of freedom corresponding to gyromotion is absent from the equations of motion. This means that one of the variables, in the present case the gyrophase ζ, is ignorable: it appears in the symplectic part of the Lagrangian only linearly through its first derivative, as in Eq. (3.2). As a consequence, ζ does not appear in the equations of motion of the remaining variables, and its conjugate (i.e., its factor in the Lagrangian) is a constant of motion [see Eq. (2.38)]. Thus, the equations of motion have only two degrees of freedom.

To accomplish this transformation, one relies on the small-gyroradius, slowly-varying-field approximation. We introduce this transformation by inserting an ordering parameter ϵ into the Lagrangian (2.15). Simultaneously, we carry out this derivation with units such that $c=1$ and $e/m\equiv\epsilon^{-1}=\Omega/B$ (we restore units at the end of the calculation). The resulting Lagrangian is

$$\mathcal{L}(\mathbf{x},\mathbf{v};t) = [\epsilon^{-1}\mathbf{A}(\mathbf{x},t)+\mathbf{v}]\cdot\dot{\mathbf{x}} - [\tfrac{1}{2}|\mathbf{v}|^2 + \epsilon^{-1}\Phi(\mathbf{x},t)]. \tag{3.30}$$

This parameter ϵ is only an ordering device that allows us to collect terms of similar size. The relative orders of the terms in the Lagrangian were chosen by noting that in the limit of small ϵ the electromagnetic field should dominate while the remaining terms, essentially kinetic energy, should be of the same order. This ordering (summarized in Table I) assures that the gyroradius is relatively small, but it allows for the electric field and the electric drift to be of order unity.

In fact, the ordering specification implied by Eq. (3.30) is not sufficient for completing the calculation. In general, ordering in Lagrangian theory is difficult, as one must know at the outset the relative sizes of the time derivatives of the variables. Furthermore, a given term will contain parts of various orders. For example, the time derivative of the gyroradius vector $\tilde{\boldsymbol{\rho}}$ contains the convective time derivative of the gyroradius $(\mathbf{v}\cdot\nabla\tilde{\rho})$, which is small, and the time derivative of the gyrophase $(\dot{\zeta}\partial\tilde{\boldsymbol{\rho}}/\partial\zeta)$, which is large. With this in mind, we will have to further discuss the ordering in reference to Table I as we proceed with the derivation.

The lowest-order motion is found by keeping the $O(\epsilon^{-1})$ terms in the Lagrangian (3.30): $\mathcal{L}_{-1}\equiv\mathbf{A}\cdot\dot{\mathbf{x}}-\Phi$. The resulting Euler-Lagrange equation for $\mathbf{x}$ is

$$\frac{d\mathbf{A}}{dt}=\frac{\partial\mathbf{A}}{\partial t}+\dot{\mathbf{x}}\cdot\nabla\mathbf{A}=\nabla\mathbf{A}\cdot\dot{\mathbf{x}}-\nabla\Phi, \tag{3.31}$$

which becomes

$$\dot{\mathbf{x}}\times\mathbf{B}+\mathbf{E}=\mathbf{0}. \tag{3.32}$$

Thus, the Lagrangian $\mathcal{L}_{-1}$ does not determine the parallel velocity $(\hat{\mathsf{b}}\cdot\dot{\mathbf{x}})$. It determines only that the perpendicular velocity is the electric drift,

$$\dot{\mathbf{x}}_\perp=\mathbf{v}_E. \tag{3.33}$$

Indeed, Eq. (3.32) implies that, to lowest order in ϵ, the parallel electric field $(E_\parallel\equiv\hat{\mathsf{b}}\cdot\mathbf{E})$ must vanish for our perturbation analysis to be consistent.

The lowest-order Lagrangian $\mathcal{L}_{-1}$ is said to be singular. Indeed Eq. (3.32) cannot be solved for the rates of change of all variables, which implies that the Lagrange tensor cannot be inverted (see Sec. II.C) since there is no equation for the parallel velocity. The Lagrange tensor will be invertible when we obtain its $O(1)$ corrections. In fact, because the Lagrange tensor consists of an $O(\epsilon^{-1})$ part that is singular with an additional $O(1)$ part that allows inversion, the Poisson tensor (the inverse of the Lagrange tensor) will be $O(1)$. This, indeed, is what motivates having the Lagrangian start with $O(\epsilon^{-1})$ terms.

To obtain the guiding-center Lagrangian to higher order, we must introduce a coordinate system. Following Littlejohn (1983), we introduce the *fixed-frame* unit vectors $\hat{1}$ and $\hat{2}$, which, together with the magnetic unit vector $\hat{\mathsf{b}}$, form a local right-handed set, $\hat{1}\times\hat{2}=\hat{\mathsf{b}}$. We write the particle velocity as a sum of its parallel, electric drift, and perpendicular parts,

$$\mathbf{v}=v_\parallel\hat{\mathsf{b}}+\mathbf{v}_E+w\hat{\mathsf{c}}, \tag{3.34}$$

with the rotating perpendicular velocity unit vector $\hat{\mathsf{c}}$ expressed in terms of the fixed-frame unit vectors by

$$\hat{\mathsf{c}}=-\sin(\zeta)\hat{1}-\cos(\zeta)\hat{2}. \tag{3.35}$$

In these equations, we make explicit the point that the unit vectors are evaluated at the guiding-center location. Correspondingly, we introduce the orthogonal unit vector perpendicular to the magnetic field,

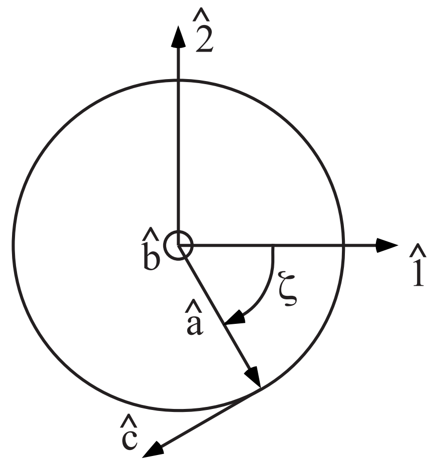

FIG. 1. Fixed-frame unit vectors $(\hat{1},\hat{2},\hat{\mathsf{b}})$ and rotating-frame unit vectors $(\hat{\mathsf{a}},\hat{\mathsf{b}},\hat{\mathsf{c}})$.

$$\hat{\mathsf{a}}=\hat{\mathsf{b}}\times\hat{\mathsf{c}}=\cos(\zeta)\hat{1}-\sin(\zeta)\hat{2}\equiv-\frac{\partial\hat{\mathsf{c}}}{\partial\zeta}. \tag{3.36}$$

This vector will prove to be the direction of the gyroradius. These vectors are sketched in Fig. 1.

As we have stated, our goal is to introduce a transformation, such that the Lagrangian for the guiding-center variables $(\mathbf{X},u,w,\zeta)$ has the gyrophase ζ ignorable. We expect that a transformation of the form

$$\mathbf{x}=\mathbf{X}+\epsilon\boldsymbol{\rho} \tag{3.37}$$

will work on the basis of knowing the solution for the case of constant magnetic field. To use the form (3.37) in the Lagrangian (2.15), we must know the derivatives, which we write in the form

$$\dot{\mathbf{x}}=\dot{\mathbf{X}}+\dot{\boldsymbol{\rho}}, \tag{3.38}$$

with

$$\dot{\boldsymbol{\rho}}=\dot{\zeta}\frac{\partial\tilde{\boldsymbol{\rho}}}{\partial\zeta}+\epsilon(\mathbf{v}\cdot\nabla\tilde{\boldsymbol{\rho}}+\dot{\tilde{\boldsymbol{\rho}}})=\frac{\mathbf{w}}{B}\dot{\zeta}+O(\epsilon), \tag{3.39}$$

where we note that the rate of variation of the gyroradius is $O(\epsilon^{-1})$ according to Table I (i.e., $\dot{\zeta}=\Omega$).

We now expand the field quantities using Eq. (3.37). Through zeroth order we obtain

$$\mathcal{L}=\epsilon^{-1}[\mathbf{A}\cdot(\dot{\mathbf{X}}+\dot{\boldsymbol{\rho}})-\Phi]+\boldsymbol{\rho}\cdot[\nabla\mathbf{A}\cdot(\dot{\mathbf{X}}+\dot{\boldsymbol{\rho}})-\nabla\Phi]$$
$$+\mathbf{v}\cdot(\dot{\mathbf{X}}+\dot{\boldsymbol{\rho}})-\frac{|\mathbf{v}|^2}{2}+O(\epsilon), \tag{3.40}$$

where $(\Phi,\mathbf{A})$ now denote potentials evaluated at the guiding-center position $\mathbf{X}$ in what follows and ∇ denotes a gradient with respect to $\mathbf{X}$.

The second term $\epsilon^{-1}\mathbf{A}\cdot\dot{\boldsymbol{\rho}}$ in the Lagrangian (3.40) should be pushed to higher order so that the guiding-center Lagrangian to lowest order has no gyrophase-dependent terms. To accomplish this, we write

$$\epsilon^{-1}\mathbf{A}\cdot\dot{\boldsymbol{\rho}} = \frac{d}{dt}(\mathbf{A}\cdot\boldsymbol{\rho}) - \left(\frac{\partial\mathbf{A}}{\partial t} + \dot{\mathbf{X}}\cdot\nabla\mathbf{A}\right)\cdot\boldsymbol{\rho} \tag{3.41}$$

so that, excluding the exact time derivative, Eq. (3.40) becomes

$$\mathcal{L} = \epsilon^{-1}(\mathbf{A}\cdot\dot{\mathbf{X}} - \Phi) + \boldsymbol{\rho}\cdot(\mathbf{E} + \dot{\mathbf{X}}\times\mathbf{B}) + \boldsymbol{\rho}\cdot\nabla\mathbf{A}\cdot\dot{\boldsymbol{\rho}} + \mathbf{v}\cdot(\dot{\mathbf{X}} + \dot{\boldsymbol{\rho}}) - \frac{|\mathbf{v}|^2}{2}, \tag{3.42}$$

where we henceforth exclude terms of order $O(\epsilon)$. Next, we use the exact derivative

$$\frac{\epsilon}{2}\frac{d}{dt}(\boldsymbol{\rho}\cdot\nabla\mathbf{A}\cdot\boldsymbol{\rho}) = \frac{1}{2}(\dot{\boldsymbol{\rho}}\cdot\nabla\mathbf{A}\cdot\boldsymbol{\rho} + \boldsymbol{\rho}\cdot\nabla\mathbf{A}\cdot\dot{\boldsymbol{\rho}}) + \frac{\epsilon}{2}\boldsymbol{\rho}\cdot\left(\frac{d\,\nabla\mathbf{A}}{dt}\right)\cdot\boldsymbol{\rho},$$

so that we may write

$$\boldsymbol{\rho}\cdot\nabla\mathbf{A}\cdot\dot{\boldsymbol{\rho}} = \tfrac{1}{2}(\boldsymbol{\rho}\cdot\nabla\mathbf{A}\cdot\dot{\boldsymbol{\rho}} - \dot{\boldsymbol{\rho}}\cdot\nabla\mathbf{A}\cdot\boldsymbol{\rho}) + \tfrac{1}{2}(\boldsymbol{\rho}\cdot\nabla\mathbf{A}\cdot\dot{\boldsymbol{\rho}} + \dot{\boldsymbol{\rho}}\cdot\nabla\mathbf{A}\cdot\boldsymbol{\rho}) = \tfrac{1}{2}\boldsymbol{\rho}\times\dot{\boldsymbol{\rho}}\cdot\mathbf{B} + O(\epsilon)$$

and the Lagrangian (3.42) is now written as

$$\mathcal{L} = \epsilon^{-1}(\mathbf{A}\cdot\dot{\mathbf{X}} - \Phi) + \boldsymbol{\rho}\cdot(\mathbf{E} + \dot{\mathbf{X}}\times\mathbf{B}) + \frac{1}{2}\boldsymbol{\rho}\times\dot{\boldsymbol{\rho}}\cdot\mathbf{B} + \mathbf{v}\cdot(\dot{\mathbf{X}} + \dot{\boldsymbol{\rho}}) - \frac{|\mathbf{v}|^2}{2}. \tag{3.43}$$

Note that the terms $\boldsymbol{\rho}\cdot(\mathbf{E}+\dot{\mathbf{X}}\times\mathbf{B})$ represent the electric-dipole and moving electric-dipole contributions to the guiding-center polarization and magnetization, respectively, while $\frac{1}{2}\boldsymbol{\rho}\times\dot{\boldsymbol{\rho}}\cdot\mathbf{B}$ represents the intrinsic magnetic-dipole contribution to the guiding-center magnetization.

To make further progress, we introduce the particular form of the transformation term $\boldsymbol{\rho}$,

$$\boldsymbol{\rho} \equiv \tilde{\boldsymbol{\rho}} + \bar{\boldsymbol{\rho}} = \hat{\mathsf{b}}\times\frac{\mathbf{w}}{B} + \bar{\boldsymbol{\rho}}, \tag{3.44}$$

with the quantity $\bar{\boldsymbol{\rho}}$ (to be determined below) independent of the gyrophase. Putting this into Eq. (3.43), with Eq. (3.34), gives the Lagrangian

$$\mathcal{L} = (\epsilon^{-1}\mathbf{A} + u\hat{\mathsf{b}})\cdot\dot{\mathbf{X}} + \left(\frac{|\mathbf{w}|^2}{2B} + \mathbf{v}_E\cdot\frac{\mathbf{w}}{B} - \frac{1}{2}\bar{\boldsymbol{\rho}}\cdot\tilde{\boldsymbol{\rho}}\right)\dot{\zeta} + \dot{\mathbf{X}}\cdot(\mathbf{v}_E - \bar{\boldsymbol{\rho}}\times\mathbf{B}) + \left(\bar{\boldsymbol{\rho}}\cdot\mathbf{E} - \frac{|\mathbf{v}_E|^2}{2}\right) - \left(\epsilon^{-1}\Phi + \frac{u^2}{2} + \frac{|\mathbf{w}|^2}{2}\right), \tag{3.45}$$

where we used $\tilde{\boldsymbol{\rho}}\cdot\mathbf{E} = \mathbf{w}\cdot\mathbf{v}_E$ and $u \equiv v_\| + O(\epsilon)$. We see that Eq. (3.45) will be simplified with the choice

$$\bar{\boldsymbol{\rho}} = \mathbf{B}\times\frac{\mathbf{v}_E}{B^2} = \frac{\mathbf{E}_\perp}{B^2}, \tag{3.46}$$

which is the guiding-center polarization displacement (3.5), where $\bar{\boldsymbol{\rho}}\cdot\mathbf{E} = |\mathbf{v}_E|^2$. Thus we arrive at

$$\mathcal{L} = (\epsilon^{-1}\mathbf{A} + u\hat{\mathsf{b}})\cdot\dot{\mathbf{X}} + \left(\frac{|\mathbf{w}|^2}{2B} + \mathbf{v}_E\cdot\frac{\mathbf{w}}{2B}\right)\dot{\zeta} - \left(\epsilon^{-1}\Phi + \frac{u^2}{2} + \frac{|\mathbf{w}|^2}{2} - \frac{|\mathbf{v}_E|^2}{2}\right). \tag{3.47}$$

An alternate choice for $\bar{\boldsymbol{\rho}}$ is provided by $\bar{\boldsymbol{\rho}}=\mathbf{0}$ and is discussed in Appendix A; this alternate choice leads to guiding-center equations of motion from which the standard equations of Northrop (1963) are obtained.

The last stage in this derivation is to eliminate the last $O(1)$, gyrophase-dependent term $\mathbf{v}_E\cdot(\mathbf{w}/2B)\dot{\zeta}$. This is done by subtracting the total time derivative

$$\frac{\epsilon}{2}\frac{d}{dt}(\mathbf{v}_E\cdot\tilde{\boldsymbol{\rho}}) = \left(\mathbf{v}_E\cdot\frac{\mathbf{w}}{2B}\right)\dot{\zeta} + \frac{\epsilon}{2}\frac{d\mathbf{v}_E}{dt}\cdot\tilde{\boldsymbol{\rho}}$$

from the Lagrangian (3.47) and omitting terms of order ϵ. We obtain, finally, the dimensionless form of the guiding-center Lagrangian,

$$\mathcal{L} = (\epsilon^{-1}\mathbf{A} + u\hat{\mathsf{b}})\cdot\dot{\mathbf{X}} + J\dot{\zeta} - H_{\rm gc} \equiv \mathcal{L}_{\rm gc}(\mathbf{X},u,\mu,\zeta;t), \tag{3.48}$$

where $J \equiv w^2/2B$ denotes the gyroaction and the guiding-center Hamiltonian is

$$H_{\rm gc}(\mathbf{X},u,\mu;t) = \epsilon^{-1}\Phi + J\Omega + \frac{u^2}{2} - \frac{|\mathbf{v}_E|^2}{2}. \tag{3.49}$$

There are no terms in $\hat{\mathsf{a}}$ or $\hat{\mathsf{c}}$ in this equation and, thus, there is no gyrophase dependence in these equations, so that the quantity J conjugate to ζ is a constant of motion.

To understand the accuracy of our calculation, it is useful to consider the orders appearing in the (dimensionless) equations of motion from the Lagrangian (3.48). Equation (3.12) gives a parallel velocity change that is $O(\epsilon^{-1})$ in the absence of any relative ordering of $E_\|$ and $\mathbf{E}_\perp$. Thus, the energy change of a particle in a gyroperiod would be order unity, and, hence, one could not assume that u is slowly changing on a gyroperiod. The resolution of this problem is to assume that the parallel electric field is $O(1)$, not $O(\epsilon^{-1})$ as is the perpendicular electric field.

Equation (3.13) for the rate of change of the guiding-center location simply states that this quantity is $O(1)$, with $\dot{\mathbf{X}}_\perp$ being $O(\epsilon)$. Closer examination shows that the perpendicular guiding-center velocity has an $O(1)$ piece due to the perpendicular electric field and an $O(\epsilon)$ piece due to the magnetic drifts. This raises the question of whether the calculation has been carried to sufficiently high order to allow one to keep these magnetic drifts, as they are $O(\epsilon)$, while the Hamiltonian is accurate only through $O(1)$.

To answer this question, we must consider how the equations of motion would change due to the addition of $O(\epsilon)$ terms in the Hamiltonian or other terms of the one-form. Such changes affect the time derivative of the coordinates only after multiplication by the Poisson tensor, as in Eq. (2.26). Inspection of the Poisson tensor

(3.27) shows that the part that leads to $\dot{\mathbf{X}}_\perp$ is $O(\epsilon)$. Hence, $O(\epsilon)$ changes to the Hamiltonian (3.49) [or the one-form (3.48)] that are $O(\epsilon)$ will change $\dot{\mathbf{X}}_\perp$ only by $O(\epsilon^2)$, and it is legitimate to keep the magnetic drifts. In contrast, the $O(\epsilon)$ terms in the parallel velocity and the parallel acceleration, while necessary for keeping the Hamiltonian structure, are not complete. Other $O(\epsilon)$ terms, such as the parallel drift (Baños, 1967; Hazeltine, 1973), would arise if the Hamiltonian and/or the one-form were calculated to $O(\epsilon)$, as the part of the Poisson tensor leading to these terms is $O(1)$.

D. Guiding-center currents

For finding the self-consistent electromagnetic fields driven by charged particles and for determining observables, we need to know the density and current in physical space. As we shall see, the magnetic gradient drifts do not appear in the formulas for the physical density and current.

We begin with the general moment integral of the particle phase-space function χ,

$$\begin{aligned} n[\chi] &\equiv \int d^3p\,\chi f \\ &= \int d^6z\,\chi\,\delta^3(\mathbf{x}-\mathbf{r})f \\ &= \int d^6Z_{\rm gc}\mathsf{T}_{\rm gc}^{-1}\chi\,\delta^3(\mathbf{X}+\boldsymbol{\rho}-\mathbf{r})g, \end{aligned} \tag{3.50}$$

where n is the particle density and $[\chi]$ denotes the particle velocity-space average of χ with respect to the Vlasov distribution f. In the last line of Eq. (3.50), $\mathsf{T}_{\rm gc}^{-1}\chi$ represents the guiding-center transformation of the function χ, the guiding center distribution is $g(\mathbf{X},u,\mu)$, the guiding-center volume element is $d^6Z_{\rm gc} \equiv m^2B_\|^*(\mathbf{X})d^3X du\,d\mu\,d\zeta$, and we have inserted the relation (3.4) between the particle position $\mathbf{x}\equiv\mathbf{r}$ and the guiding-center position $\mathbf{X}$. By expanding the delta function and integrating by parts, we obtain the multipole expansion

$$\begin{aligned} n[\chi] &= N[\mathsf{T}_{\rm gc}^{-1}\chi]_{\rm gc} - \nabla\cdot(N[\boldsymbol{\rho}\mathsf{T}_{\rm gc}^{-1}\chi]_{\rm gc}) \\ &\quad + \nabla\nabla{:}\left(N\left[\frac{\boldsymbol{\rho}\boldsymbol{\rho}}{2}\mathsf{T}_{\rm gc}^{-1}\chi\right]_{\rm gc}\right) + \cdots , \end{aligned} \tag{3.51}$$

where N is the guiding-center density and $[\cdots]_{\rm gc}$ denotes the guiding-center velocity-space average with respect to the guiding-center distribution g.

For the case $\chi=1$, the relation between the particle density n and the guiding-center density N is

$$n = N - \nabla\cdot\left(\frac{N}{e}\boldsymbol{\pi}_{\rm gc} + \cdots\right), \tag{3.52}$$

where the guiding-center dipole moment $\boldsymbol{\pi}_{\rm gc}\equiv e\bar{\boldsymbol{\rho}}$ is the guiding-center average of the gyrophase-independent (polarization) part (3.46), which survives the gyrophase integration in $[\cdots]_{\rm gc}$. The particle density n is thus expressed as the sum of the guiding-center density N and a polarization density that includes a dipole contribution (shown) as well as higher-order multipole contributions (not shown).

For the case $\chi=\mathbf{v}$, where $\mathsf{T}_{\rm gc}^{-1}\mathbf{v}=\dot{\mathbf{X}}_{\rm gc}+\dot{\boldsymbol{\rho}}_{\rm gc}$ is expressed as the sum of the guiding-center velocity $\dot{\mathbf{X}}_{\rm gc}$ and the guiding-center displacement velocity $\dot{\boldsymbol{\rho}}_{\rm gc}$ (Brizard and Hahm, 2007; Brizard, 2008), the relation between the particle flux $n[\mathbf{v}]\equiv n\mathbf{u}$ and the guiding-center flux $N[\dot{\mathbf{X}}_{\rm gc}]_{\rm gc}\equiv N\mathbf{U}_{\rm gc}$ is

$$n\mathbf{u} = N\mathbf{U}_{\rm gc} + \frac{\partial}{\partial t}\left(\frac{N}{e}\boldsymbol{\pi}_{\rm gc}\right) + \nabla\times\left(\frac{cN}{e}\boldsymbol{\mu}_{\rm gc}\right), \tag{3.53}$$

where the guiding-center magnetic moment

$$\boldsymbol{\mu}_{\rm gc} \equiv -\hat{\mathsf{b}}[\mu]_{\rm gc} + \frac{1}{c}\boldsymbol{\pi}_{\rm gc}\times\mathbf{U}_{\rm gc} \tag{3.54}$$

is expressed as the sum of the intrinsic magnetic-dipole moment and the moving electric-dipole moment contributions.

We note that the guiding-center continuity equation $\partial N/\partial t+\nabla\cdot(N\mathbf{U}_{\rm gc})=0$ is consistent with the particle continuity equation $\partial n/\partial t+\nabla\cdot(n\mathbf{u})=0$ since

$$\begin{aligned} 0 = \frac{\partial n}{\partial t} + \nabla\cdot(n\mathbf{u}) &= \frac{\partial}{\partial t}\left[N - \nabla\cdot\left(\frac{N}{e}\boldsymbol{\pi}_{\rm gc}\right)\right] + \nabla\cdot(N\mathbf{U}_{\rm gc}) \\ &\quad + \nabla\cdot\left[\frac{\partial}{\partial t}\left(\frac{N}{e}\boldsymbol{\pi}_{\rm gc}\right)\right] \\ &= \frac{\partial N}{\partial t} + \nabla\cdot(N\mathbf{U}_{\rm gc}), \end{aligned} \tag{3.55}$$

where $\nabla\cdot\nabla\times(N\boldsymbol{\mu}_{\rm gc})\equiv 0$. Hence, while the particle flux $n\mathbf{u}$ is not equal to the guiding-center flux $N\mathbf{U}_{\rm gc}$, the two fluid formulations are consistent with each other. We also note that the addition of the intrinsic guiding-center magnetization flux (with $N[\mu]_{\rm gc}\equiv p_\perp/B$) to the perpendicular guiding-center flux

$$N\mathbf{U}_{\perp\rm gc} = \frac{c\hat{\mathsf{b}}}{eB}\times(p_\perp\nabla\ln B + p_\|\hat{\mathsf{b}}\cdot\nabla\hat{\mathsf{b}})$$

yields the perpendicular particle flux $n\mathbf{u}_\perp\equiv(c\hat{\mathsf{b}}/eB)\times\nabla\cdot\mathsf{P}$, which is the lowest-order solution to the fluid equation of motion

$$mn\left(\frac{d\mathbf{u}}{dt} - \Omega\mathbf{u}\times\hat{\mathsf{b}}\right) \simeq -mn\Omega\mathbf{u}\times\hat{\mathsf{b}} = -\nabla\cdot\mathsf{P},$$

where $\mathsf{P}\equiv(p_\|-p_\perp)\hat{\mathsf{b}}\hat{\mathsf{b}}+p_\perp\mathbf{I}$ denotes the Chew-Goldberger-Low pressure tensor.

By defining the guiding-center polarization and magnetization vectors

$$\begin{pmatrix}\mathbf{P}_{\rm gc}\\ \mathbf{M}_{\rm gc}\end{pmatrix} \equiv \sum N\begin{pmatrix}\boldsymbol{\pi}_{\rm gc}\\ \boldsymbol{\mu}_{\rm gc}\end{pmatrix}, \tag{3.56}$$

the guiding-center Maxwell's equations for the macroscopic electromagnetic fields

$$\begin{pmatrix} \mathbf{D}_{\rm gc} \\ \mathbf{H}_{\rm gc} \end{pmatrix} \equiv \begin{pmatrix} \mathbf{E} + 4\pi \mathbf{P}_{\rm gc} \\ \mathbf{B} = 4\pi \mathbf{M}_{\rm gc} \end{pmatrix} \tag{3.57}$$

are expressed as

$$\nabla \cdot \mathbf{D}_{\rm gc} = 4\pi \rho_{\rm gc} \tag{3.58}$$

and

$$\nabla \times \mathbf{H}_{\rm gc} - \frac{1}{c}\frac{\partial \mathbf{D}_{\rm gc}}{\partial t} = \frac{4\pi}{c}\mathbf{J}_{\rm gc}, \tag{3.59}$$

where the guiding-center charge and current densities are, respectively, $\rho_{\rm gc} = \Sigma eN$ and $\mathbf{J}_{\rm gc} = \Sigma eN\mathbf{U}_{\rm gc}$.

E. Guiding-center angular momentum for azimuthally symmetric systems

As an example of the use of Noether's theorem for finding invariants of the guiding-center Lagrangian, we consider the case of a magnetized plasma with azimuthal symmetry. That is, we consider cylindrical coordinates $X^i \equiv (R, \varphi, Z)$ and assume that no quantities depend on toroidal angle φ. Hence the guiding-center Lagrangian (3.2) takes the form

$$\mathcal{L}_{\rm gc} = \left(\frac{e}{c}A_i + mub_i\right)dX^i + J\dot{\zeta} - H_{\rm gc}, \tag{3.60}$$

in which $A_i \equiv \mathbf{A} \cdot \partial \mathbf{X}/\partial X^i$ are the covariant components of the vector potential $\mathbf{A}$, $b_i \equiv \hat{\mathsf{b}} \cdot \partial \mathbf{X}/\partial X^i$ are the covariant components of the unit magnetic field $\hat{\mathsf{b}} \equiv \mathbf{B}/B$, and the Hamiltonian (3.3) has the form

$$H_{\rm gc} = \frac{m}{2}[u^2 - |\mathbf{v}_{\rm E}|^2(R,Z,t)] + \mu B(R,Z,t) + e\Phi(R,Z,t). \tag{3.61}$$

In this Lagrangian, there is no explicit dependence on the toroidal angle φ. Hence, its canonically conjugate momentum

$$p_\varphi \equiv \frac{e}{c}A_\varphi + mub_\varphi \tag{3.62}$$

is a constant of motion. This derivation does not require one to have nested flux surfaces or flux variables for the toroidal magnetic field or for the system to be time independent.

F. Hamiltonian formulation of field line flow

As noted in the derivation of the guiding-center Lagrangian, keeping only the lowest-order terms results in a singular Lagrangian; the parallel velocity cannot be determined, and, hence, one cannot determine the trajectory. However, if the electrostatic field is ignored, so that only the vector potential remains, and the vector potential is static, the Lagrangian does determine the spatial trajectories of the field lines. In this section, we develop this Hamiltonian formalism, and obtain canonical equations of motion for the field lines (Boozer, 1983; Cary and Littlejohn, 1983; Littlejohn, 1985).

According to our discussion, the action for magnetic-field line flow is

$$\mathcal{A}_B = \int \mathbf{A} \cdot d\mathbf{x}. \tag{3.63}$$

Variation of this equation gives $d\mathbf{x} \times \mathbf{B} = \mathbf{0}$, which states that the trajectories follow magnetic-field lines. The configuration space in which the field lines flow is three-dimensional. Hence, one can think of a trajectory in terms of two of the variables being given as functions of the third, at least locally. Specifically, we consider cylindrical coordinates (R, φ, Z) and take φ to be the independent coordinate. Thus, we write the Lagrangian as

$$\mathcal{A}_B = \int \left(A_R\frac{dR}{d\varphi} + A_Z\frac{dZ}{d\varphi} + A_\varphi\right)d\varphi. \tag{3.64}$$

Variation of this action gives the rates of change $dR/d\varphi$ and $dZ/d\varphi$.

This system can be put into Hamiltonian form by a gauge transformation. We introduce the function

$$\Psi(R,\varphi,Z,t) \equiv \int_0^Z dZ' A_Z(R,\varphi,Z',t), \tag{3.65}$$

whose gradient we subtract from the vector potential: $\mathbf{A} \rightarrow \mathbf{A} - \nabla\Psi$. This is equivalent to adding a total derivative to the Lagrangian. The resulting action is

$$\mathcal{A}_B = \int \left[\left(A_R - \frac{\partial \Psi}{\partial R}\right)\frac{dR}{d\varphi} + \left(A_\varphi - \frac{\partial \Psi}{\partial \varphi}\right)\right]d\varphi, \tag{3.66}$$

in which the Lagrangian has the form of $p\dot{q} - H$. Hence, we have found the canonical form for magnetic-field line flow. More explicitly, we define the magnetic momentum

$$P \equiv \left(A_R - \frac{\partial \Psi}{\partial R}\right), \tag{3.67}$$

and the magnetic Hamiltonian,

$$H_M \equiv -A_\varphi + \frac{\partial \Psi}{\partial \varphi}. \tag{3.68}$$

Then, because the Lagrangian is in canonical form, Hamilton's equations

$$\left(\frac{dP}{d\varphi}, \frac{dR}{d\varphi}\right) = \left(-\frac{\partial H_M}{\partial R}, \frac{\partial H_M}{\partial P}\right) \tag{3.69}$$

give the field lines.

IV. CANONICAL GUIDING-CENTER THEORY

The noncanonical, Hamiltonian guiding-center formalism developed in Sec. III has formal advantages, such as having a conserved phase-space volume and Noether's theorem. However, there are times when it is advantageous to use canonical coordinates. For example, analytical progress can be eased by having the dynamics encapsulated in a single function, the Hamiltonian, rather than spread through the $2N+1$ components (symplectic plus Hamiltonian) of the phase-space Lagrangian (2.19). Symplectic integration algorithms are known for the integration of trajectories (Kang, 1986; Forest and Ruth, 1990; Yoshida, 1990; Candy and Rozmus, 1991; Qin and Guan, 2008) and for particle simulations of plasma (Cary and Doxas, 1993). The goal of this section is to show how canonical coordinates are obtained in a number of cases.

The guiding-center Lagrangian (3.2) is linear in the three spatial time derivatives as well as the time derivative of the gyrophase. Thus, the four-dimensional guiding-center position and parallel velocity part contains three time derivatives, while to be in canonical form, as is the phase-space Lagrangian (2.13), there should be only two time derivatives. To eliminate the third time derivative, either one seeks purely spatial coordinates (Meiss and Hazeltine, 1992) such that both magnetic fields $\mathbf{A}$ and $\mathbf{B}$ have only the same two nonzero covariant components or one introduces coordinates mixing the guiding-center coordinates with the parallel velocity such that only two time derivatives appear in the guiding-center Lagrangian. These methods give different canonical coordinates. This should not be surprising, as given one set of canonical coordinates, there are arbitrary other sets, all related by canonical transforms.

We begin with the general case, and start this analysis from the framework of flux coordinates (Stern, 1970). General flux coordinates are defined such that two of the coordinates are constant along the magnetic-field lines, and so that the differential flux is simply the product of the differentials of the two coordinates. We also choose these coordinates so that the third variable measures the distance from some reference surface along the field line. To illustrate the two methods for finding canonical coordinates, we apply both to this case. We apply the method of Meiss and Hazeltine (1992) to find coordinates in which the vector potential and magnetic field have only the same two covariant components nonzero. We then show that, by adding terms linear in the parallel velocity to the flux variables, we obtain guiding-center canonical coordinates to relevant order. The latter method of obtaining canonical coordinates generalizes to the case of toroidally nested flux surface easily without the need for patches, as we show in the last part of this section, and will be discussed in Appendix B.

A. General magnetic coordinates

1. Magnetic-flux coordinates

A general magnetic field can be expressed either in terms of covariant components B_i,

$$\mathbf{B} = B_i \nabla \psi^i, \tag{4.1}$$

or contravariant components B^i,

$$\mathbf{B} = B^i \frac{\partial \mathbf{x}}{\partial \psi^i}, \tag{4.2}$$

where ψ^i denote general curvilinear coordinates. Using the orthogonality relations

$$\nabla \psi^i \cdot \frac{\partial \mathbf{x}}{\partial \psi^j} = \delta^i_j, \tag{4.3}$$

we see that the covariant and contravariant components of $\mathbf{B}$ are, respectively, defined as

$$B_i \equiv \mathbf{B} \cdot \frac{\partial \mathbf{x}}{\partial \psi^i} \quad \text{and} \quad B^i \equiv \mathbf{B} \cdot \nabla \psi^i. \tag{4.4}$$

Because of its relationship with the magnetic field $\mathbf{B} = \nabla \times \mathbf{A}$, the vector potential $\mathbf{A}$ is preferably written in terms of its covariant components (up to a gauge term)

$$\mathbf{A} = A_i \nabla \psi^i, \tag{4.5}$$

and, hence, we find the contravariant representation

$$\mathbf{B} = \nabla A_i \times \nabla \psi^i. \tag{4.6}$$

If we now introduce the Jacobian $\mathcal{V}$ of the transformation $\mathbf{x} \to \psi^i(\mathbf{x})$, defined by

$$\frac{\partial \mathbf{x}}{\partial \psi^i} \cdot \frac{\partial \mathbf{x}}{\partial \psi^j} \times \frac{\partial \mathbf{x}}{\partial \psi^k} = \epsilon_{ijk} \mathcal{V} \tag{4.7}$$

or

$$\nabla \psi^i \cdot \nabla \psi^j \times \nabla \psi^k = \mathcal{V}^{-1} \epsilon^{ijk}, \tag{4.8}$$

we obtain the contravariant components

$$B^i = \nabla \psi^i \cdot \nabla \times \mathbf{A} = \frac{\epsilon^{ijk}}{\mathcal{V}} \frac{\partial A_k}{\partial \psi^j}. \tag{4.9}$$

The covariant representation (4.1) is useful to calculate the current density $\mathbf{J} = \nabla \times (c\mathbf{B}/4\pi) = \nabla (cB_i/4\pi) \times \nabla \psi^i$, while the contravariant representation (4.6) is manifestly divergenceless.

Last, the differential equation for the magnetic-field lines is expressed as

$$\frac{d\mathbf{x}}{ds} = \frac{\mathbf{B}}{B}, \tag{4.10}$$

where s represents the position along a given magnetic-field line. Using

$$\frac{d\mathbf{x}}{ds} = \frac{d\psi^j}{ds} \frac{\partial \mathbf{x}}{\partial \psi^j}$$

in Eq. (4.10) and substituting the identity (4.3) and Eq. (4.9), we find the magnetic differential equations

$$\frac{d\psi^i}{ds} = \frac{B^i}{B} = \frac{\epsilon^{ijk}}{\mathcal{V}B} \frac{\partial A_k}{\partial \psi^j}. \tag{4.11}$$

These equations express how the curvilinear coordinates ψ^i vary along a magnetic-field line.

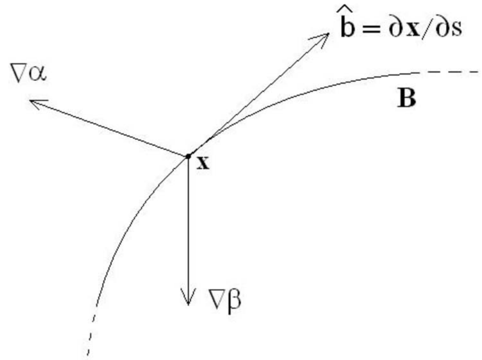

FIG. 2. Magnetic flux coordinates (α,β,s) for general magnetic fields.

2. Simple magnetic coordinates

One of the simplest representations for the magnetic field is in terms of the magnetic coordinates (α,β,s),

$$\mathbf{A} = \alpha \nabla \beta, \tag{4.12}$$

$$\mathbf{B} = \nabla\alpha \times \nabla\beta = B(\alpha,\beta,s)\frac{\partial \mathbf{x}}{\partial s}, \tag{4.13}$$

where the Jacobian $\mathcal{V}$ has the simple form

$$\mathcal{V}^{-1} = \nabla\alpha \times \nabla\beta \cdot \nabla s \equiv B. \tag{4.14}$$

The magnetic-field line labels α and β are known as Euler (or Clebsch) potentials (with $\mathbf{B}\cdot\nabla\alpha=\mathbf{0}=\mathbf{B}\cdot\nabla\beta$) and the third flux coordinate s is the parallel coordinate measuring position along a single magnetic-field line (Fig. 2 illustrates these flux coordinates). In this representation, the contravariant components of $\mathbf{B}$ are simply

$$B^\alpha = 0 = B^\beta \quad \text{and} \quad B^s = B \tag{4.15}$$

while the covariant components $B_i = Bb_i$ can be expressed with $b_\alpha = \hat{b}\cdot\partial\mathbf{x}/\partial\alpha$, $b_\beta=\hat{b}\cdot\partial\mathbf{x}/\partial\beta$, and $b_s=1$.

The values of the coordinates α and β throughout space are determined by requiring them to be constant on field lines. Conservation of flux by the flow of magnetic-field lines then assures that the relation (4.13) holds throughout space. From this construction it follows that magnetic flux coordinates must be patched between regions not connected by the magnetic field, and that they are singular where the magnetic field vanishes. They can also be multivalued when the magnetic field returns through the original, defining surface.

Regardless, we can now write the guiding-center Lagrangian (3.48) in terms of the magnetic flux coordinates (α,β,s),

$$\mathcal{L}(\alpha,\beta,s,u,\mu;t) = \frac{e}{c}\alpha\dot{\beta} + mu(b_\alpha\dot{\alpha} + b_\beta\dot{\beta} + \dot{s}) - H(\alpha,\beta,s,u,\mu;t), \tag{4.16}$$

where H denotes the standard guiding-center Hamiltonian (3.49) and we omitted the term $J\dot{\zeta}$ in the symplectic part of Eq. (4.16) since it does not enter into the guiding-center equations for (α,β,s) and u. The guiding-center Lagrangian (4.16) describes guiding-center motion in a four-dimensional phase space with coordinates (α,β,s,u), where the magnetic moment μ appears as a dynamically invariant label. We note, however, that since the symplectic part of Eq. (4.16) exhibits all three velocities $(\dot{\alpha},\dot{\beta},\dot{s})$, the Lagrangian (4.16) does not have the canonical form. In the next two sections, we discuss the methods used to reduce the number of independent velocities to two and, thus, obtain a canonical guiding-center phase-space Lagrangian.

3. Application of the Meiss-Hazeltine method

Meiss and Hazeltine (1992) noted that a further coordinate transformation, to replace the variable s, can eliminate one covariant component of the magnetic field and so lead to a Lagrangian with only two time derivatives. Thus one obtains canonical variables. To find the Meiss-Hazeltine coordinates, we introduce a new variable σ in place of the field line variable s via the function

$$\sigma = \hat{\sigma}(\alpha,\beta,s), \tag{4.17}$$

which has for its inverse

$$s = \hat{s}(\alpha,\beta,\sigma). \tag{4.18}$$

With this transformation, the covariant representation of the magnetic field becomes

$$\begin{aligned}\mathbf{B} &= B_\alpha \nabla\alpha + B_\beta \nabla\beta + B\nabla s \\ &= \left(B_\alpha + B\frac{\partial\hat{s}}{\partial\alpha}\right)\nabla\alpha \\ &\quad + \left(B_\beta + B\frac{\partial\hat{s}}{\partial\beta}\right)\nabla\beta + B\frac{\partial\hat{s}}{\partial\sigma}\nabla\sigma.\end{aligned} \tag{4.19}$$

The covariant components of the vector potential are unchanged by this transformation as the vector potential has no s component. Thus, the vector potential and the magnetic field will have only β and σ covariant components provided one can find a function $\hat{s}$ such that

$$B_\alpha + B\frac{\partial\hat{s}}{\partial\alpha} = 0. \tag{4.20}$$

To better understand the requirements on the transformation, we seek instead the function $\hat{\sigma}$. In terms of this function, Eq. (4.20) becomes

$$\frac{\partial\hat{\sigma}}{\partial s}B_\alpha - B\frac{\partial\hat{\sigma}}{\partial\alpha} = 0. \tag{4.21}$$

This equation says that the variable σ is constant along trajectories defined by

$$\frac{ds}{d\alpha} = b_\alpha. \tag{4.22}$$

Hence, for each value of β and for every initial condition in the α-s plane for that value of β, one can integrate Eq. (4.22) to obtain trajectories. These trajectories cannot cross as long as the flow (4.22) is nonsingular. To complete the definition of the new coordinates, a value of σ, such as the value of s where this trajectory crosses the s axis, is assigned to each trajectory.

In these coordinates neither the vector potential nor the magnetic field has an α covariant component, and so $\dot{\alpha}$ does not appear in the guiding-center Lagrangian,

$$\mathcal{L} = \left[\frac{e\alpha}{c} + mu\left(b_\beta + \frac{\partial \hat{s}}{\partial \beta} \right) \right] \dot{\beta} + mu \frac{\partial \hat{s}}{\partial \sigma} \dot{\sigma} - H. \tag{4.23}$$

Hence, the two coordinates are β and σ with their conjugates p_β and p_σ being the factors of their time derivatives in the Lagrangian (4.23): $\mathcal{L} \equiv p_\beta \dot{\beta} + p_\sigma \dot{\sigma} - H$.

These variables, while canonical, no longer have the simple interpretation of one of them being the distance along a field line. Moreover, the deviation of the new variable σ from the original variable s increases arbitrarily as one moves away from some reference value of α, according to Eq. (4.20). Additionally, according to Eq. (4.22), $\hat{s}(\alpha)$ has local extrema where b_α vanishes. Near these points σ is multivalued, and so patches will be needed. Last, White and Zakharov (2003) pointed out that, while the Meiss-Hazeltine method can be rigorously applied to construct a canonical guiding-center Lagrangian, it suffers from serious difficulties in its numerical implementation for many magnetic equilibria.

4. Canonical flux-based coordinates

Another way to obtain canonical, flux-based coordinates is to remove one of the temporal derivatives through appropriate addition of total time derivatives to the Lagrangian. In this section, we use dimensionless units ($c=1$ and $e/m=\epsilon^{-1}$) since this simplifies the algebra and allows us to see easily how to drop higher-order terms (we reinsert the units and remove the formal adiabatic ordering parameter in appropriate summary equations). Now the guiding-center Lagrangian (4.16) takes the form

$$\mathcal{L} = \epsilon^{-1} \alpha \dot{\beta} + u(b_\alpha \dot{\alpha} + b_\beta \dot{\beta} + \dot{s}) - H, \tag{4.24}$$

where H denotes the standard guiding-center Hamiltonian (3.49).

The reduced guiding-center phase space associated with the phase-space Lagrangian (4.24) is described in terms of four coordinates (α, β, s, u). Since the symplectic covector in Eq. (4.24) has three nonvanishing components, the phase-space Lagrangian (4.24) is noncanonical. To obtain canonical variables, we must transform Eq. (4.24) into one in which time derivatives of only two coordinates appear, and then the corresponding multipliers are the canonical variables. For example, one can group the coefficients of the time derivative of β to obtain

$$\mathcal{L} = \epsilon^{-1}(\alpha + \epsilon u b_\beta)\dot{\beta} + u\dot{s} + u b_\alpha \dot{\alpha} - H, \tag{4.25}$$

but there remains the problematic term proportional to $\dot{\alpha}$. However, we can move this problem to higher order by subtracting the total time derivative,

$$\frac{d}{dt}(u b_\alpha \alpha) = (u b_\alpha)\dot{\alpha} + \alpha \frac{d(u b_\alpha)}{dt},$$

to obtain the Lagrangian

$$\mathcal{L} = \epsilon^{-1} \alpha^{(1)} \dot{\beta}^{(1)} + u\dot{s} + \epsilon \left[u b_\beta \frac{d}{dt}(u b_\alpha) \right] - H, \tag{4.26}$$

where the first-order corrected (denoted by the superscript) canonical flux variables are

$$\alpha^{(1)} \equiv \alpha + \epsilon u b_\beta,$$

$$\beta^{(1)} \equiv \beta - \epsilon u b_\alpha. \tag{4.27}$$

Thus, neglecting ϵ-order terms, we obtain the desired canonical form

$$\mathcal{L} = \epsilon^{-1} \alpha^{(1)} \dot{\beta}^{(1)} + u\dot{s} - H. \tag{4.28}$$

From this form we see that the flux variables $(\alpha^{(1)}, \beta^{(1)})$ are canonically conjugate, and that the parallel velocity is conjugate to the distance along a field line.

For reference, we restore ordinary units, and we eliminate the formal ordering parameter. We find the canonical coordinates

$$\alpha^{(1)} \equiv \alpha + (uB/\Omega) b_\beta,$$

$$\beta^{(1)} \equiv \beta - (uB/\Omega) b_\alpha \tag{4.29}$$

and we note that $\mathbf{B}^* = \nabla \alpha^{(1)} \times \nabla \beta^{(1)} = \mathbf{B} + (uB/\Omega)\hat{\mathsf{b}} \cdot \nabla \times \hat{\mathsf{b}} + \cdots$. The Lagrangian for these variables is

$$\mathcal{L} = \frac{e}{c} \alpha^{(1)} \dot{\beta}^{(1)} + mu\dot{s} - H, \tag{4.30}$$

and the equations of motion are simply Hamilton's equations

$$\begin{pmatrix} \dot{s} \\ \dot{u} \end{pmatrix} = m^{-1} \begin{pmatrix} \partial H/\partial u \\ -\partial H/\partial s \end{pmatrix} \tag{4.31}$$

and

$$\begin{pmatrix} \dot{\alpha}^{(1)} \\ \dot{\beta}^{(1)} \end{pmatrix} = (c/e) \begin{pmatrix} -\partial H/\partial \beta^{(1)} \\ \partial H/\partial \alpha^{(1)} \end{pmatrix}. \tag{4.32}$$

In these (essentially canonical) coordinates $(\alpha^{(1)}, \beta^{(1)}, s, u)$, the phase-space Jacobian, found from Eq. (2.43), is constant: $\mathcal{J} = em/c$.

One can compute these coordinates to higher order as well. The basic idea is that, in each step, one eliminates the time derivatives of the parallel velocity by adding a total time derivative to the Lagrangian, and then one proceeds much as in first order, collecting the derivatives of the flux coordinates with the addition of other total time derivatives. In brief, subtraction of the total time derivative

$$\epsilon \frac{d}{dt}(u^2 b_\alpha b_\beta) \tag{4.33}$$

gives the new Lagrangian

$$\mathcal{L} = \epsilon^{-1}\alpha^{(1)}\dot{\beta}^{(1)} + u\dot{s} + \frac{\epsilon}{2}u^2\left(b_\beta\frac{\partial b_\alpha}{\partial\alpha} - b_\alpha\frac{\partial b_\beta}{\partial\alpha}\right)\dot{\alpha} + \frac{\epsilon}{2}u^2\left(b_\beta\frac{\partial b_\alpha}{\partial\beta} - b_\alpha\frac{\partial b_\beta}{\partial\beta}\right)\dot{\beta} + \frac{\epsilon}{2}u^2\left(b_\beta\frac{\partial b_\alpha}{\partial s} - b_\alpha\frac{\partial b_\beta}{\partial s}\right)\dot{s} - H. \tag{4.34}$$

At this point, the Lagrangian is similar to Eq. (4.24) and so a procedure similar to that leading to Eq. (4.28) can be used. The two new terms in the time derivatives of the flux variables can be replaced by the corrected variables to lowest order, and the terms in $\dot{s}$ can be collected. The result is the canonical Lagrangian

$$\mathcal{L} = \epsilon^{-1}\alpha^{(2)}\dot{\beta}^{(2)} + u^{(1)}\dot{s} - H + O(\epsilon^2), \tag{4.35}$$

where

$$\alpha^{(2)} = \alpha^{(1)} + \frac{\epsilon^2}{2}u^2\left(b_\beta\frac{\partial b_\alpha}{\partial\beta} - b_\alpha\frac{\partial b_\beta}{\partial\beta}\right), \tag{4.36}$$

$$\beta^{(2)} = \beta^{(1)} - \frac{\epsilon^2}{2}u^2\left(b_\beta\frac{\partial b_\alpha}{\partial\alpha} - b_\alpha\frac{\partial b_\beta}{\partial\alpha}\right), \tag{4.37}$$

$$u^{(1)} = u + \frac{\epsilon}{2}u^2\left(b_\beta\frac{\partial b_\alpha}{\partial s} - b_\alpha\frac{\partial b_\beta}{\partial s}\right). \tag{4.38}$$

One can continue this procedure to higher order, but doing so might not be useful, as the guiding-center Lagrangian is valid only through the lowest order. Thus, at best, it would seem appropriate to use only the first-order corrected canonical variables (4.29).

B. Toroidal magnetic fields with nested flux surfaces

For magnetic fields that lie on toroidal, nested flux surfaces, there are many sets of convenient coordinates having the convenient property that two of the coordinates are proper angles, i.e., they increase by 2π for circulation around the torus in the appropriate way. In this section, we concentrate on flux coordinates and canonical coordinates that can be obtained by allowing the angles to depend on the parallel velocity. In flux coordinates, the vector potential has a particularly simple form, and in which field lines are straight, while in canonical coordinates, the equations of evolution have the usual canonical form.

1. Toroidal flux coordinates

For magnetic fields that lie on nested tori, one coordinate T is chosen to label the tori (see Fig. 3). Two other coordinates, a poloidal angle θ_0 and a toroidal angle φ_0, then specify a point on the torus. Then, because the magnetic lines lie in surfaces of constant T, the magnetic field can be written in Clebsch representation,

$$\mathbf{B} = B_{T\theta_0}(T,\theta_0,\varphi_0)\,\nabla T \times \nabla\theta_0 + B_{\varphi_0 T}(T,\theta_0,\varphi_0)\,\nabla\varphi_0 \times \nabla T. \tag{4.39}$$

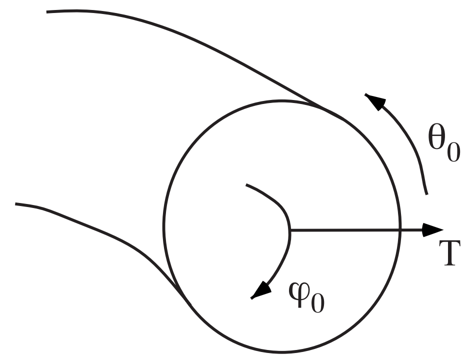

FIG. 3. Coordinates for magnetic fields lying on nested toroidal surfaces.

Flux coordinates $(\Psi,\theta_F,\varphi_F)$ are chosen so that the magnetic field has a particularly simple representation

$$\mathbf{B} = \nabla\Psi \times \nabla\theta_F + \iota(\Psi)\,\nabla\varphi_F \times \nabla\Psi. \tag{4.40}$$

These coordinates have the convenient property that in them the magnetic-field lines are straight,

$$\frac{d\theta_F}{d\varphi_F} = \frac{\mathbf{B}\cdot\nabla\theta_F}{\mathbf{B}\cdot\nabla\varphi_F} = \iota(\Psi). \tag{4.41}$$

This quantity ι is known as the rotational transform. The vector potential has the particularly simple form

$$\mathbf{A} = \Psi\,\nabla\,\theta_F + A_\varphi(\Psi)\,\nabla\,\varphi_F, \tag{4.42}$$

in which its covariant components are functions of the surface alone. Here

$$\frac{dA_\varphi}{d\Psi} = \iota(\Psi). \tag{4.43}$$

In these coordinates the differential toroidal flux,

$$\mathbf{B}\cdot\left(\frac{\partial\mathbf{r}}{\partial\Psi} \times \frac{\partial\mathbf{r}}{\partial\theta_F}\right) = 1, \tag{4.44}$$

is unity. Hence, $2\pi(\Psi_2 - \Psi_1)$ is the toroidal flux between toroidal surfaces of label Ψ_1 and Ψ_2. This is shown as the gray area (annulus at constant φ_F) in Fig. 4. Similarly, by integrating $\mathbf{A}\cdot d\mathbf{x}$ around a loop in φ_F at constant θ_F and

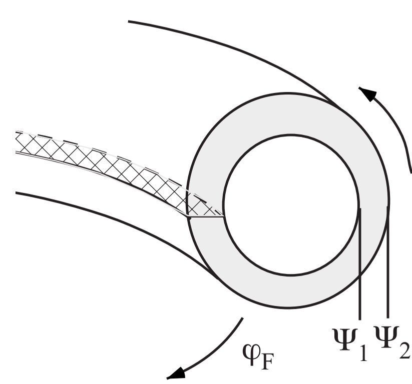

FIG. 4. Toroidal flux coordinates for magnetic fields lying on nested toroidal surfaces showing the toroidal and poloidal flux between surfaces.

Ψ, one can show that $2\pi[A_\varphi(\Psi_1)-A_\varphi(\Psi_2)]$ is *minus* the flux in the poloidal direction between toroidal surfaces of label Ψ_1 and Ψ_2. This is shown as the cross-hatched area of Fig. 4.

To find flux coordinates, we begin by supposing that they exist, then use the relation between the two representations, Eqs. (4.39) and (4.40), to determine the transformation between the original variables (T,θ_0,φ_0) and the flux variables $(\Psi,\theta_F,\varphi_F)$. We assume that this transformation leaves the toroidal angle unchanged, $\varphi_F=\varphi_0$, and that the new flux variable $\Psi(T)$ is independent of the angles and, hence, still a flux function, so that the poloidal angle is changed only by the addition of a function periodic in the angles,

$$\theta_F = \theta_0 + \hat{\theta}_F(T,\theta_0,\varphi_0). \tag{4.45}$$

With these assumptions, the magnetic field (4.40) can be put into the basis of Eq. (4.39). Equating the components gives

$$\frac{\partial\Psi}{\partial T}\left(1+\frac{\partial\hat{\theta}_F}{\partial\theta_0}\right) = B_{T\theta_0} \tag{4.46}$$

and

$$\frac{\partial\Psi}{\partial T}\left(\iota+\frac{\partial\hat{\theta}_F}{\partial\varphi_0}\right) = B_{\varphi_0 T}. \tag{4.47}$$

The first step in solving these equations is to determine the conditions imposed by averaging over the angles θ_0 and φ_0. The resulting equations imply

$$\frac{d\Psi}{dT} = \bar{B}_{T\theta_0} \tag{4.48}$$

and

$$\iota = -\frac{\bar{B}_{\varphi_0 T}}{\bar{B}_{T\theta_0}}, \tag{4.49}$$

where the overbar denotes the average over the original angles. The first equation determines the transformation from original toroidal surface label T to the new label Ψ. It is conventional to take the value of the flux Ψ to vanish on the magnetic axis, the degenerate single field going the long way around the torus at the center of the torus. We denote the value of the flux at the edge of the plasma by Ψ_{edge}. The second equation determines the rotational transform.

The transformation function $\hat{\theta}_F$ must then satisfy the following two equations:

$$\frac{\partial\hat{\theta}_F}{\partial\theta_0} = \frac{B_{T\theta_0}}{\bar{B}_{T\theta_0}} - 1 \tag{4.50}$$

and

$$\frac{\partial\hat{\theta}_F}{\partial\varphi_0} = \iota - \frac{B_{\varphi_0 T}}{\bar{B}_{\varphi_0 T}}. \tag{4.51}$$

That these equations can be solved locally follows from showing that the mixed second partial derivatives are equal. This follows from the vanishing of the divergence of the magnetic field,

$$\nabla\cdot\mathbf{B} = \frac{\partial B_{T\theta_0}}{\partial\varphi_0} + \frac{\partial B_{\varphi_0 T}}{\partial\theta_0} = 0. \tag{4.52}$$

In fact, it is relatively straightforward to show by Fourier transforming in the angles that Eqs. (4.50) and (4.51) can be solved globally. Hence, we can always find flux variables when magnetic fields lie in nested toroidal surfaces.

At this point, the toroidal angle is arbitrary. To determine the extent of arbitrariness in the flux coordinates, we analyze the restrictions placed on a transformation to other variables. The flux variable must remain invariant for it to equal the toroidal flux enclosed by a surface divided by 2π. Hence, we consider a transformation to new angles q_1 and q_2,

$$\theta_F = q_1 + f(\Psi,q_1,q_2) \tag{4.53}$$

and

$$\varphi_F = q_2 + g(\Psi,q_1,q_2), \tag{4.54}$$

where f and g are period functions of the new poloidal angle q_1 and the new toroidal angle q_2.

Inserting these transformations into the magnetic-field representation (4.39) gives

$$\mathbf{B} = \left(1+\frac{\partial f}{\partial q_1}-\iota\frac{\partial g}{\partial q_1}\right)\nabla\Psi\times\nabla q_1 + \left(\iota+\frac{\partial f}{\partial q_2}-\iota\frac{\partial g}{\partial q_2}\right)\nabla q_2\times\nabla\Psi. \tag{4.55}$$

Hence, provided the ratio of f, up to a function h of Ψ alone, and g is the rotational transform

$$f = \iota g + \bar{f}(\Psi), \tag{4.56}$$

any transformation of the form (4.53) and (4.54) is a transformation from one set of flux variables to another.

2. Canonical toroidal flux coordinates

The above flux coordinates (including the variations described in Sec. IV.B.1) are convenient for calculations because they reflect the dynamics (parallel motion is within a flux surface), and because they have the usual angular periodicity (i.e., $0\leqslant\theta<2\pi$ and $0\leqslant\varphi<2\pi$). In this section, we show how the associated canonical coordinates arise in guiding-center Lagrangian theory. For this analysis, there remains significant arbitrariness, as no preliminary transformations are needed. Hence, one need not transform to Boozer coordinates as a preliminary step (White and Chance, 1984; White and Zakharov, 2003).

For this analysis, we return to dimensionless units $m=c=1$ and $e=\epsilon^{-1}$ such that the adiabatic ordering parameter ϵ is shown explicitly in the Lagrangian. The guiding-center Lagrangian is

$$\mathcal{L} = (\epsilon^{-1}\Psi + ub_{\theta_F})\dot{\theta}_F + (\epsilon^{-1}A_\varphi + ub_{\varphi_F})\dot{\varphi}_F + (ub_\Psi)\dot{\Psi} - H \tag{4.57}$$

with the Hamiltonian H given by Eq. (3.49). In its current form, the guiding-center Lagrangian (4.57) is noncanonical because all three time derivatives $(\dot{\Psi},\dot{\theta}_F,\dot{\varphi}_F)$ appear explicitly. The derivation of a canonical guiding-center Lagrangian, which focuses on the term $(ub_\Psi)\,\dot{\Psi}$ in the symplectic part of Eq. (4.57), can proceed by either a redefinition of the poloidal angle $\theta_F \to \theta_c$ (White and Chance, 1984), a redefinition of the toroidal angle $\varphi_F \to \varphi_c$ (White and Zakharov, 2003), or a redefinition of the parallel velocity (White, 1990). We now show how a redefinition of the poloidal angle leads to a canonical guiding-center Lagrangian.

Subtracting the total derivative

$$\frac{d}{dt}(ub_\Psi\Psi) = ub_\Psi\dot{\Psi} + \Psi\frac{d(ub_\Psi)}{dt} \tag{4.58}$$

from the guiding-center Lagrangian (4.57) gives the Lagrangian

$$\mathcal{L} = (\epsilon^{-1}\Psi + ub_{\theta_F})\dot{\theta}_c + (\epsilon^{-1}A_\varphi + ub_{\varphi_F})\dot{\varphi}_F - H + O(\epsilon), \tag{4.59}$$

where the new canonical poloidal angle (White, 1990, 2008)

$$\theta_c \equiv \theta_F - \epsilon u b_\Psi \tag{4.60}$$

differs from the old by a bounded term that is $O(\epsilon)$. From the symplectic part of the Lagrangian (4.59), it is clear that the canonical coordinates are θ_c and φ_F with conjugate momenta given by their factors in the Lagrangian (4.59). White (2008) pointed out that the omission of the b_Ψ term in the guiding-center Lagrangian (4.57) amounts to a simple redefinition of the guiding-center position, which allows the retention of the angle coordinates (θ_F,φ_F) as canonical variables.

We now summarize our results in standard units. The canonical toroidal angle is any flux-variable consistent toroidal angle, and its conjugate momentum is

$$p_\varphi = \frac{e}{c}A_\varphi + mub_{\varphi_F}. \tag{4.61}$$

The canonical poloidal angle is

$$\theta_c \equiv \theta_F - \frac{u}{\Omega}B_\Psi \tag{4.62}$$

and its canonical momentum is

$$p_\theta = \frac{e}{c}\Psi + mub_{\varphi_F}. \tag{4.63}$$

The guiding-center Lagrangian (4.57) is given by

$$\mathcal{L} = p_\theta\dot{\theta}_c + p_\varphi\dot{\varphi}_F - H, \tag{4.64}$$

for which the Hamiltonian is given by Eq. (3.49).

This definition of the canonical variables does mix in the parallel velocity with the poloidal angle, but it retains the desired periodicity. For any fixed value of u, the increase of θ_c for one poloidal circuit is 2π because the increase in θ_F was 2π, and the difference $(u/\Omega)\,B_\Psi$ is a periodic function of θ_F, and so has no change in a poloidal transit. As discussed in Appendix B, this continues to be true for any of the specialized coordinates, such as Hamada coordinates and Boozer coordinates.

To obtain an explicit form for the Hamiltonian, Eqs. (4.62) and (4.63) must be solved for Ψ and u as functions of canonical momenta. This is not usually done. Instead, one of two methods is followed. If a nonsymplectic integrator is used, it is applied directly to the equations of motion in convenient variables. If a symplectic integrator is used, the implicit equations to be solved are set up also as functions of the convenient variables.

V. HIGHER-ORDER ADIABATIC INVARIANTS

The magnetic confinement of charged particles implies the existence of orbits enclosed within a compact volume in space, which in turn generically allows the existence of three orbital frequencies (Northrop, 1963). The first of these orbital frequencies, called the gyrofrequency (denoted $\omega_g \equiv \Omega$), exists even in uniform (unconfining) magnetic fields and describes the gyration of a charged particle about a single magnetic-field line. The second orbital frequency, called the bounce frequency (denoted ω_b), requires longitudinal confinement along magnetic-field lines (due to nonuniformity parallel to the field lines) and describes oscillations in the parallel component of the particle's velocity which vanishes at turning points along the trapped-particle orbit. Although certain magnetic geometries, e.g., axisymmetric tokamak geometry allow for the existence of confined, untrapped (or circulating) charged particles whose parallel velocity exhibits oscillatory behavior about a nonvanishing value, we focus our attention only on trapped-particle orbits in the present section. The third orbital frequency, called the drift-precession frequency (denoted ω_d), describes the periodic drift motion across magnetic-field lines (e.g., due to magnetic curvature). In general, these three orbital frequencies are widely separated (for a single particle species), with $\omega_g \gg \omega_b \gg \omega_d$,

$$(\mathbf{x},\mathbf{v};w,t) \xrightarrow{g} \begin{cases} (\mathbf{X},p_\parallel;W,t) \xrightarrow{b} \begin{cases} (\alpha,\beta;k,t) \xrightarrow{d} \begin{cases} (K,t) \\ (J_d,\theta_d) \end{cases} \\ (J_b,\theta_b) \end{cases} \\ (J_g,\theta_g) \end{cases} . \tag{5.1}$$

Equation (5.1) shows the hierarchy of adiabatic invariants associated with the guiding-center (g), bounce-

TABLE II. Bounce ordering required for the existence of the longitudinal adiabatic invariant.

Order	Dimensionless	Fields	Distances	Rates	Velocities
$1/\epsilon$		$\mathbf{B}$		Ω	
1		$E_\parallel$, $\mathbf{E}_\perp$	L	$\mathbf{v}/L$	$\mathbf{v}$
ϵ	ρ/L, $(\Omega\tau)^{-1}$		ρ	τ^{-1}, $\mathbf{v}_E$/L, $\mathbf{v}_\nabla$/L, $\mathbf{v}_\kappa$	$\mathbf{v}_E$, $\mathbf{v}_\nabla$, $\mathbf{v}_\kappa$

center (b), and drift-center (d) dynamical reductions (Tao *et al.*, 2007). Each reduction, which is carried out by a transformation on extended phase space (which includes the time-energy canonical pair), involves the elimination of a fast orbital angle $\theta_j=(\theta_g,\theta_b,\theta_d)$ and the construction of a corresponding adiabatic invariant $J_j=(J_g,J_b,J_d)$.

When the characteristic time scale of interest τ is much longer than the gyroperiod (i.e., when the particle has executed many gyration cycles during time τ), the fast gyration angle θ_g can be asymptotically removed from the particle's orbital dynamics and a corresponding adiabatic invariant $J_g=(mc/e)\mu$ (the magnetic moment μ) can be constructed (Northrop, 1963). The resulting guiding-center dynamics takes place in a reduced six-dimensional phase space with noncanonical coordinates $(\mathbf{X},p_\parallel;W,t)$, where $\mathbf{X}$ denotes the particle's guiding-center position, $p_\parallel\equiv mu$ denotes its parallel kinetic momentum, and (W,t) denotes the guiding-center energy-time canonical coordinates. Guiding-center dynamics has been shown to possess a noncanonical Hamiltonian structure (Littlejohn, 1981, 1983), i.e., $\dot{\mathbf{X}}\equiv\{\mathbf{X},H_{\rm gc}\}_{\rm gc}$ (with $u\equiv\hat{\mathsf{b}}\cdot\dot{\mathbf{X}}$), $\dot{p}_\parallel\equiv\{p_\parallel,H_{\rm gc}\}_{\rm gc}$, and $\dot{W}\equiv\partial H_{\rm gc}/\partial t$ are expressed in terms of a guiding-center Hamiltonian function $H_{\rm gc}$ and a noncanonical guiding-center Poisson bracket $\{,\}_{\rm gc}$.

When the characteristic time scale τ is also much longer than the bounce period (i.e., when the particle has executed many bounce cycles during time τ), the fast bounce angle θ_b can be asymptotically removed from the particle's orbital dynamics and a corresponding adiabatic invariant (the longitudinal invariant or bounce action J_b) can be constructed. The resulting bounce-averaged guiding-center (or bounce-center) dynamics takes place in a reduced four-dimensional phase space with spatial coordinates (α,β) and the energy-time canonical coordinates (k,t). Bounce-center dynamics in static magnetic fields has also been shown to possess a canonical Hamiltonian structure (Littlejohn, 1982b).

Last, when the characteristic time scale τ is much longer than the drift period (i.e., when the particle has executed many bounce-averaged drift-precession cycles during time τ), the fast drift angle θ_d can be asymptotically removed from the particle's orbital dynamics and a corresponding adiabatic invariant (the drift action J_d) can be constructed. The resulting drift-averaged bounce-center (or drift-center) dynamics takes place in a reduced two-dimensional phase space with energy-time coordinates (K,t).

A. Second and third adiabatic invariants

1. Longitudinal adiabatic invariant J_b

In the derivation of Sec. III, it was assumed that the velocity and the electric drift were of the same order, and that the rate of change of the fields was $O(1)$. When the electric drift is of the same order as the magnetic drifts, and the fields change more slowly $[O(\epsilon)]$, particles execute a *bounce* oscillation, moving back and forth along a field line, with direction reversed by the effective potential (3.11), before drifting to a significantly different field line and before the dynamics on the field line changes significantly. In this case, the longitudinal adiabatic invariant for the motion along the field line is nearly a constant of the motion.

In this section, we show how to obtain the longitudinal adiabatic invariant, and we derive the reduced equations of motion. This modified ordering is often valid in plasmas. For example, in toroidal equilibria (Kovrizhnykh, 1984), the outflows of the ions and electrons are in balance only if there is an electrostatic potential of the order of the particle energy, and the plasma evolves slowly, on the diffusion time.

The mathematical description of trapped-particle orbits in magnetized plasmas is facilitated using the magnetic coordinates (α,β,s) introduced in Sec. IV.A. The modified ordering implies that the Lagrangian (4.16) be modified to

$$\mathcal{L}_b=\epsilon^{-1}\alpha\dot{\beta}+u(b_\alpha\dot{\alpha}+b_\beta\dot{\beta}+\dot{s})-H_b, \tag{5.2}$$

where the Hamiltonian is $H_b=u^2/2+\mu B(\alpha,\beta,s,\epsilon t)+\Phi(\alpha,\beta,s,\epsilon t)$ for the study of the longitudinal adiabatic invariant. The b subscript on the Lagrangian (5.2) and the Hamiltonian signifies that these quantities are appropriate for the ordering described above. The removal of the factor ϵ^{-1} for the potential Φ and the slow temporal variation follow from the above discussion; this is the original ordering of Littlejohn (1983). Relatively slow temporal variation is also required for gauge invariance; the electric field coming from the vector potential must be of the same order as the electric field coming from the electrostatic potential. In addition, this slow temporal variation is needed for adiabatic theory to apply. These orderings are summarized in Table II. We note that in this ordering the polarization drift appears only in order ϵ^2.

It was noted in Sec. IV.A that flux coordinates suffer from being multivalued, and that this is a problem when a field line revisits a region repeatedly. However, this is not a problem in the present case, where we will be analyzing particles reflected by the effective potential and, so, confined to a portion of the field line.

The goal in the present calculation is analogous to that of guiding-center theory. We seek a transformation order by order in ϵ such that the Lagrangian has an ignorable coordinate. In guiding-center theory, we began by transforming to variables (perpendicular velocity and gyrophase) appropriate to the gyromotion one would find in a constant magnetic field. Here we begin by transforming to variables appropriate to the bounce motion that one would find if the field did not vary with time (constant $\tau \equiv \epsilon t$) and the particle remained on its initial field line (constant α and β), i.e., the motion to lowest order in ϵ. These variables are the action-angle variables of the parallel motion.

For constant magnetic labels $y^k \equiv (\alpha,\beta)$, the Lagrangian (5.2) reduces to

$$\mathcal{L}_{\parallel} \equiv u\dot{s} - H_b. \tag{5.3}$$

Hence, the motion is given by

$$\dot{s} = \frac{\partial H_b}{\partial u} = u \quad \text{and} \quad \dot{u} = -\frac{\partial H_b}{\partial s}. \tag{5.4}$$

For time-independent fields, the energy $E = H_b$ is a constant of motion. Hence, to this order, the particle moves along a field line until it reflects (mirrors) due to its interaction with the effective potential $V_b = \mu B + \Phi$. This occurs at the reflection (or *mirroring*) points defined by

$$E = V_b(\mathbf{y}, s_{\pm}, \tau), \tag{5.5}$$

where s_+ and s_- are the upper and lower turning points. If either reflection point does not exist, then there is no adiabatic invariant. The maximum parallel velocity occurs at the point of minimum effective potential. To simplify the calculations, we define the variable s such that its origin occurs at the minimum,

$$\frac{\partial V_b}{\partial s}(\mathbf{y}, s=0, \tau) = 0. \tag{5.6}$$

The longitudinal adiabatic invariant is, to lowest order, the bounce action defined as the loop integral of the phase-space part of the action around a loop of constant energy and constant slowly varying variables. In preparation for calculating this loop integral, we introduce the function

$$\hat{u}(\mathbf{y}, s, E, \tau) = \pm\sqrt{2(E - \mu B - \Phi)}, \tag{5.7}$$

found by solving for the parallel velocity along a field line. The sign depends on the direction of the particle motion along a field line. Thus, the lowest-order action is given by

$$J_{b0} = \frac{1}{2\pi}\oint \hat{u}(\mathbf{y}, s'; E, \tau)\, ds'. \tag{5.8}$$

This loop integral equals twice the integral between the turning points of $|\hat{u}|$. Equation (5.8) can be inverted to obtain the Hamiltonian

$$H_b = E = H_0(\mathbf{y}, J_{b0}, \tau). \tag{5.9}$$

To complete the transformation to (lowest-order) longitudinal action-angle coordinates (J_{b0}, Θ_0), we must find an appropriate angle coordinate Θ_0. To do this, we introduce the gauge function

$$F = \int_0^s \hat{u}\, ds' - J_{b0}\Theta_0, \tag{5.10}$$

whose derivative is to be subtracted from the Lagrangian (5.2). Equation (5.10) was chosen so that the first term of its derivative,

$$\begin{aligned}\dot{F} = {} & u\dot{s} - J_{b0}\dot{\Theta}_0 + \dot{y}^k \int_0^s ds' \left(\frac{\partial \hat{u}}{\partial y^k} + \frac{\partial \hat{u}}{\partial E}\frac{\partial H_0}{\partial y^k} \right) \\ & + \epsilon \int_0^s ds' \left(\frac{\partial \hat{u}}{\partial \tau} + \frac{\partial \hat{u}}{\partial E}\frac{\partial H_0}{\partial \tau} \right) \\ & + \dot{J}_{b0}\left(\frac{\partial H_0}{\partial J_{b0}} \int_0^s ds' \frac{\partial \hat{u}}{\partial E} - \Theta_0 \right),\end{aligned} \tag{5.11}$$

cancels the phase-space part of the Lagrangian (5.2) corresponding to parallel motion. The last term of Eq. (5.11) vanishes with the choice of angle

$$\Theta_0 = \frac{\partial H_0}{\partial J_{b0}} \int_0^s ds' \frac{\partial \hat{u}}{\partial E} = \frac{2\pi}{\tau_b}\int_0^s \frac{ds'}{\hat{u}}, \tag{5.12}$$

where $\partial H_0/\partial I_0 = \omega_b \equiv 2\pi/\tau_b$ defines the bounce frequency, with the bounce period defined as $\tau_b \equiv \oint \hat{u}^{-1} ds$. This equation shows that the angle is proportional to the transit time to the point in question. Moreover, it follows from Eqs. (5.8) and (5.12) that Θ_0 increases by 2π for one circuit of the constant-H_b contour in phase space. Thus, upon subtracting the derivative (5.11) from the Lagrangian (5.2), we obtain the Lagrangian

$$\begin{aligned}\mathcal{L}_b = {} & \epsilon^{-1}\alpha\dot{\beta} + J_{b0}\dot{\Theta}_0 - H_b \\ & + \dot{y}^k \left[u b_k - \int_0^s ds' \left(\frac{\partial \hat{u}}{\partial y^k} + \frac{\partial \hat{u}}{\partial E}\frac{\partial H_0}{\partial y^k} \right) \right],\end{aligned} \tag{5.13}$$

where

$$H_b \equiv H_0(\mathbf{y}, J_{b0}, \tau) + \epsilon \int_0^s ds' \left(\frac{\partial \hat{u}}{\partial \tau} + \frac{\partial \hat{u}}{\partial E}\frac{\partial H_0}{\partial \tau} \right). \tag{5.14}$$

Additional corrections are derived by Littlejohn (1982b).

Bounce-angle dependence remains in the Lagrangian (5.13) through explicit dependence on u and s. However, these terms can be removed by transforming to the coordinates

$$y_0^k \equiv y^k - \epsilon\eta^{k\ell}\left[u b_\ell - \int_0^s ds' \left(\frac{\partial \hat{u}}{\partial y^\ell} + \frac{\partial \hat{u}}{\partial E}\frac{\partial H_0}{\partial y^\ell} \right) \right], \tag{5.15}$$

where $\eta^{k\ell}$ is antisymmetric (with $\eta^{12} = -1$). These coordinates are single valued since the integrals in Eq. (5.15) vanish upon making a complete circuit in phase space, as follows from the fact that the action and the flux variables are independent, so that derivatives of Eq. (5.8) with respect to α and β vanish. Inserting these coordinates into the Lagrangian (5.13) and keeping terms only through zeroth order in the expansion parameter gives the bounce-center Lagrangian

$$\mathcal{L}_b = \epsilon^{-1}\alpha_0\dot{\beta}_0 + J_{b0}\dot{\Theta}_0 - H_b + O(\epsilon), \tag{5.16}$$

where

$$H_b \equiv H_0(\mathbf{y}_0, J_{b0}, \tau) + O(\epsilon). \tag{5.17}$$

The lowest-order equations for the bounce action-angle coordinates (J_{b0}, Θ_0) are

$$\dot{J}_{b0} = -\frac{\partial H_0}{\partial \Theta_0} \equiv 0 \quad \text{and} \quad \dot{\Theta}_0 = \frac{\partial H_0}{\partial J_{b0}} = \omega_b, \tag{5.18}$$

which implies that the lowest-order longitudinal (bounce) action J_{b0} is an (adiabatic) invariant.

At this point we restore the units in our equations so that they can be more easily used. The velocity function becomes

$$\hat{u}(\alpha,\beta,s,E,\tau) = \pm\sqrt{\frac{2}{m}(E - \mu B - e\Phi)}, \tag{5.19}$$

while the action becomes

$$J_{b0} = \frac{m}{2\pi}\oint \hat{u}(\mathbf{y},s',E,\tau)ds'. \tag{5.20}$$

This equation is inverted, as before, to obtain the Hamiltonian (5.9) as a function of the new variables. The modified flux variables are

$$y_0^k \equiv y^k - \frac{mc}{e}\eta^{k\ell}\left[ub_\ell - \int_0^s ds'\left(\frac{\partial \hat{u}}{\partial y^\ell} + \frac{\partial \hat{u}}{\partial E}\frac{\partial H_0}{\partial y^\ell}\right)\right]. \tag{5.21}$$

Finally, the new Lagrangian becomes

$$\mathcal{L}_b = \frac{e}{c}\alpha_0\dot{\beta}_0 + J_{b0}\dot{\Theta}_0 - H_0. \tag{5.22}$$

The adiabatic longitudinal theory is used to find the particle motion as follows. First, the function H_0 is found by integrating Eq. (5.8) and inverting it. This gives the functional form for the new variables $(\alpha_0,\beta_0,\Theta_0,J_{b0})$, which obey the equations of motion

$$\dot{y}_0^k = \frac{e}{c}\eta^{k\ell}\frac{\partial H_0}{\partial y^\ell}. \tag{5.23}$$

These equations are then integrated for constant J_{b0}, and Eqs. (5.18) are integrated by a direct integration, as the time dependence is then known explicitly. At any point in time one may find the values of the original variables by inverting Eq. (5.15) to find the original flux variables, and by inverting Eq. (5.12) to find the position along the field line as a function of the adiabatic phase. Indeed, the first step is facilitated using the fact that the differences between α_0 and α and between β_0 and β are small. Hence, Eq. (5.15) can be inverted as follows:

$$y^k \equiv y_0^k + \frac{mc}{e}\eta^{k\ell}\left[ub_\ell - \int_0^s ds'\left(\frac{\partial \hat{u}}{\partial y^\ell} + \frac{\partial \hat{u}}{\partial E}\frac{\partial H_0}{\partial y^\ell}\right)\right]_{\mathbf{y}_0}, \tag{5.24}$$

where by this notation we mean that the integrals are evaluated with the values of α_0 and β_0.

2. Drift adiabatic invariant J_d

Bounce-averaged drift-center motion (to lowest order) forms a closed curve on the space of bounce-averaged magnetic labels (α,β) parametrized by the drift angle ϕ. Note that the time-scale ordering consistent with the drift-center Hamiltonian dynamics involves time dependence at order ϵ^2, where the lowest-order drift-center Lagrangian is

$$\mathcal{L}_d = \epsilon^{-2}\left(\epsilon\frac{e}{c}\alpha\dot{\beta} - \bar{K}\right) \equiv \epsilon^{-1}J_d\dot{\phi} - \epsilon^{-2}\bar{K},$$

where a dot now represents a derivative with respect to the drift time scale $\tau \equiv \epsilon^2 t$. Here the third (drift) adiabatic invariant $J_d = (e/c)\Psi \equiv \bar{\Psi}$ is defined in terms of the magnetic flux Ψ enclosed by the bounce-averaged drifting guiding-center orbit

$$\Psi(\bar{K},t) \equiv \frac{1}{2\pi}\oint\left(\alpha\frac{\partial \beta}{\partial \phi} - \beta\frac{\partial \alpha}{\partial \phi}\right)d\phi, \tag{5.25}$$

where the magnetic labels $\alpha(\phi;\bar{K},t)$ and $\beta(\phi;\bar{K},t)$ are also functions of the lowest-order drift Hamiltonian $\bar{K}$ and time. In the case of an axisymmetric magnetic field $\mathbf{B} \equiv \nabla\psi \times \nabla\varphi$, for example, where the azimuthal angle $\beta \equiv \varphi$ is an exact ignorable angle, the magnetic-flux invariant is simply $\Psi \equiv \psi$. The drift-precession frequency $\omega_d \equiv \langle d\varphi/dt\rangle_b$ for this case is expressed in terms of the toroidal guiding-center angular frequency

$$\frac{d\varphi}{dt} \equiv \nabla\varphi\cdot\dot{\mathbf{X}} = \frac{c\mu}{e}\left(\frac{\partial B}{\partial \psi} - a\frac{\partial B}{\partial s}\right) + mu^2\frac{\partial a}{\partial s},$$

where $\hat{\mathsf{b}} = \nabla s + a\nabla\psi$. In general, the operation of drift-angle averaging is defined as $\langle\cdots\rangle_d \equiv \tau_d^{-1}\oint(\cdots)d\phi/\dot{\phi}$, where $\tau_d \equiv 2\pi/\omega_d$ defines the drift-precession period.

The time derivative of the lowest-order drift invariant (5.25) is

$$\frac{d\bar{\Psi}}{dt} \equiv \frac{\partial \bar{\Psi}}{\partial t} + \dot{\bar{K}}\frac{\partial \bar{\Psi}}{\partial \bar{K}} = \frac{1}{\omega_d}(\dot{\bar{K}} - \langle\dot{\bar{K}}\rangle_d),$$

where $\partial\bar{\Psi}/\partial\bar{K} \equiv \omega_d^{-1}$ defines the drift-precession frequency and $\partial\bar{\Psi}/\partial t \equiv -\omega_d^{-1}\langle\dot{\bar{K}}\rangle_d$ is expressed in terms of the drift-angle average $\langle\dot{\bar{K}}\rangle_d$. It turns out that, while $d\bar{\Psi}/dt \neq 0$, its drift-angle average $\langle d\bar{\Psi}/dt\rangle_d \equiv 0$ as is required by the general formulation of adiabatic invariance (Northrop and Teller, 1960; Northrop, 1963; Tao *et al.*, 2007) discussed next in Sec. V.B.

B. Higher-order adiabatic invariance

The three adiabatic invariants $J_a=(J_g,J_b,J_d)$ have so far been derived only to zeroth order in magnetic-field nonuniformity. The formal definitions of these approximate invariants are given in terms of asymptotic expansions in powers of $\epsilon_a \equiv \rho_a/L_B \ll 1$, where L_B represents the magnetic-field nonuniformity length scale and ρ_a denotes the cross-field displacement associated with the fast orbital motion involving the fast angle ζ_a. Hence, the asymptotic expansions for the first two adiabatic invariants are

$$J_a \equiv J_{a0} + \epsilon_a J_{a1} + \cdots, \tag{5.26}$$

where J_{a1} represents the first-order correction that explicitly involves the nonuniformity of the magnetic field. The first-order correction can be explicitly constructed from the invariance condition

$$\frac{dJ_a}{dt} = \frac{dJ_{a0}}{dt} + \epsilon_a \frac{dJ_{a1}}{dt} + \cdots \equiv 0, \tag{5.27}$$

where $dJ_{a0}/dt = O(\epsilon_a)$ and $dJ_{a1}/dt = \omega^a \partial J_{a1}/\partial\zeta_a + O(\epsilon_a)$. Note that the fast-angle average of the lowest-order action dynamics

$$\left\langle \frac{dJ_{a0}}{dt} \right\rangle_a \equiv 0 \tag{5.28}$$

vanishes identically (for time-independent fields), where the fast-angle average is

$$\langle \cdots \rangle_a \equiv \frac{1}{2\pi} \oint (\cdots) d\zeta_a.$$

From Eqs. (5.27) and (5.28), we easily solve for the fast-angle-dependent part

$$J_{a1} = -\omega_a^{-1} \int \frac{dJ_{a0}}{dt} d\zeta_a + \langle J_{a1} \rangle_a, \tag{5.29}$$

where the fast-angle-independent part $\langle J_{a1}\rangle_a$ is determined at higher order and

$$\omega_a \equiv \left\langle \frac{d\zeta_a}{dt} \right\rangle_a$$

is the fast-angle-averaged frequency. Since the lowest-order fast dynamics is represented by $d/dt = \omega_a \partial/\partial\zeta_a + O(\epsilon_a)$, we can easily verify that

$$\frac{d}{dt}(J_{a0} + \epsilon_a J_{a1} + \cdots) = \frac{dJ_{a0}}{dt} + \omega_a \frac{\partial J_{a1}}{\partial \zeta_a} + \cdots = O(\epsilon_a^2),$$

which vanishes (to order ϵ_a) when Eq. (5.29) is inserted. Hence, we have constructed an adiabatic invariant that is conserved up to order ϵ_a. We now proceed with derivations of these first-order corrections. More details can be found in Tao *et al.* (2007), where the derivation of first-order corrections to the three relativistic adiabatic invariants was performed.

1. First-order magnetic moment

It has long been known that, while the lowest-order magnetic moment $\mu_0 = m|\mathbf{v}_\perp|^2/2B$ is a suitable adiabatic invariant for most of a particle's orbit, corrections are needed to correctly describe the gyrophase-averaged motion of magnetically trapped particles (Belova *et al.*, 2003).

We first consider the derivation of the first-order correction μ_1 to the magnetic moment $\mu = \mu_0 + \epsilon\mu_1 + \cdots$, where the expression for μ_1 explicitly involves the nonuniformity of the background magnetic field. The exact time derivative of the lowest-order magnetic moment $\mu_0 = m|\mathbf{v}_\perp|^2/2B \equiv (e/mc)J_0$ yields

$$\frac{d\mu_0}{dt} = -\frac{\mu_0}{B}\frac{dB}{dt} + \frac{m}{B}\frac{d\mathbf{v}_\perp}{dt}\cdot \mathbf{v}_\perp$$
$$= -\mu_0 \mathbf{v}\cdot\nabla \ln B + (e\mathbf{E} - mv_\parallel \mathbf{v}\cdot\nabla\hat{\mathsf{b}})\cdot\frac{\mathbf{v}_\perp}{B}, \tag{5.30}$$

where we assume time-independent fields; for time-dependent fields, see Qin and Davidson (2006). Note that $\dot{\mu}_0$ is explicitly gyrophase dependent since the gyrophase average of Eq. (5.31) yields

$$\langle \dot{\mu}_0 \rangle = -\mu_0 v_\parallel (\hat{\mathsf{b}}\cdot\nabla \ln B + \nabla\cdot\hat{\mathsf{b}}) \equiv 0,$$

which vanishes as a result of $\nabla\cdot\mathbf{B} = \mathbf{0}$. Hence, according to Eqs. (5.28) and (5.29), the first-order gyrophase-dependent correction to the magnetic moment is

$$\tilde{\mu}_1 = -\frac{m\mathbf{v}_\perp}{B}\cdot\mathbf{v}_D + \frac{\mu_0 v_\parallel}{2\Omega}(\hat{\mathsf{a}}\hat{\mathsf{c}} + \hat{\mathsf{c}}\hat{\mathsf{a}}):\nabla\hat{\mathsf{b}}, \tag{5.31}$$

where the drift velocity is

$$\mathbf{v}_D \equiv \frac{c\hat{\mathsf{b}}}{eB}\times(e\nabla\Phi + \mu_0\nabla B + mv_\parallel^2 \hat{\mathsf{b}}\cdot\nabla\hat{\mathsf{b}}).$$

The gyroangle-independent part $\langle\mu_1\rangle$ is found to be (Kruskal, 1965)

$$\langle \mu_1 \rangle = -\mu_0 \frac{v_\parallel}{\Omega}\hat{\mathsf{b}}\cdot\nabla\times\hat{\mathsf{b}}. \tag{5.32}$$

This gyrophase-independent first-order correction is intimately connected to the first-order correction to the parallel guiding-center velocity $u \equiv \hat{\mathsf{b}}\cdot\dot{\mathbf{X}} = u_0 + \epsilon u_1 + \cdots$, where $u_0 \equiv v_\parallel$ and

$$u_1 \equiv -\left\langle \hat{\mathsf{b}}\cdot\frac{d\boldsymbol{\rho}}{dt}\right\rangle = \langle \mathbf{v}\cdot\nabla\hat{\mathsf{b}}\cdot\boldsymbol{\rho}\rangle = \frac{\mu B}{m\Omega}(\hat{\mathsf{b}}\cdot\nabla\times\hat{\mathsf{b}})$$

represents the so-called Baños drift (Baños, 1967; Northrop and Rome, 1978). These results have been recovered using Lie-transform perturbation methods discussed later.

2. First-order longitudinal invariant

The derivation of the first-order correction J_{b1} to the bounce action $J_b = J_{b0} + \epsilon J_{b1} + \cdots$ is, first, performed for the case of a time-independent magnetic field $\mathbf{B}=\nabla\alpha\times\nabla\beta$ (Northrop *et al.*, 1966) and then generalized to time-dependent electromagnetic fields.

We begin with the guiding-center time derivative of the lowest-order bounce action,

$$J_{b0}(\alpha,\beta,\mu,W) = \oint \sqrt{2m[W-e\Phi-\mu B]}\frac{ds}{2\pi}, \tag{5.33}$$

where $\mu=\mu_0+\epsilon\mu_1+\cdots$ denotes the exact magnetic moment, $\Phi(\alpha,\beta,s)$ and $B(\alpha,\beta,s)$ denote the scalar potential and magnetic strength, respectively, and W denotes the total guiding-center energy. For time-independent fields, $W=E$ is a constant of the guiding-center motion and the bounce frequency ω_b is defined through

$$\omega_b^{-1} \equiv \frac{\partial J_{b0}}{\partial E} = \frac{1}{2\pi}\oint \frac{ds}{\hat{u}}.$$

Next, we find

$$\frac{dJ_{b0}}{dt} = \frac{d\alpha}{dt}\frac{\partial J_{b0}}{\partial\alpha} + \frac{d\beta}{dt}\frac{\partial J_{b0}}{\partial\beta}, \tag{5.34}$$

where the partial derivation of J_{b0} with respect to $y^k=(\alpha,\beta)$,

$$\frac{\partial J_{b0}}{\partial y^k} = -\frac{1}{2\pi}\oint\frac{ds}{\hat{u}}\left(e\frac{\partial\Phi}{\partial y^k}+\mu\frac{\partial B}{\partial y^k}\right) \equiv -\omega_b^{-1}\left\langle e\frac{\partial\Phi}{\partial y^k}+\mu\frac{\partial B}{\partial y^k}\right\rangle_b, \tag{5.35}$$

is expressed in terms of the bounce-angle average

$$\langle\cdots\rangle_b \equiv \frac{1}{\tau_b}\oint(\cdots)\frac{ds}{\hat{u}} \tag{5.36}$$

and τ_b is the bounce period. The guiding-center equation $dy^k/dt=\dot{\mathbf{X}}\cdot\nabla y^k$ can be written as

$$\frac{dy^k}{dt} = \frac{c}{e}\eta^{k\ell}\left[e\frac{\partial\Phi}{\partial y^\ell}+\mu\frac{\partial B}{\partial y^\ell}+\frac{d}{dt}(mub_\ell)\right]. \tag{5.37}$$

Here the last term is obtained as follows. First, we note that $\nabla y^k=\eta^{k\ell}\mathbf{B}\times\partial\mathbf{X}/\partial y^\ell$ so that

$$\frac{dy^k}{dt} = \frac{c\hat{\mathsf{b}}}{eB}\times\left(e\nabla\Phi+\mu\nabla B+mu^2\frac{\partial\hat{\mathsf{b}}}{\partial s}\right)\cdot\nabla y^k = \frac{c}{e}\eta^{k\ell}\left[\left(e\frac{\partial\Phi}{\partial y^\ell}+\mu\frac{\partial B}{\partial y^\ell}+mu^2\frac{\partial\hat{\mathsf{b}}}{\partial s}\cdot\frac{\partial\mathbf{X}}{\partial y^\ell}\right) - \left(e\frac{\partial\Phi}{\partial s}+\mu\frac{\partial B}{\partial s}\right)b_\ell\right].$$

Next, we note that to lowest order in bounce dynamics,

$$u^2\frac{\partial\hat{\mathsf{b}}}{\partial s}\cdot\frac{\partial\mathbf{X}}{\partial y^\ell} = u^2\frac{\partial b_\ell}{\partial s} = u\frac{db_\ell}{dt},$$

and using

$$m\frac{du}{dt} = -\left(e\frac{\partial\Phi}{\partial s}+\mu\frac{\partial B}{\partial s}\right),$$

we combine these expressions to obtain

$$mu^2\frac{\partial\hat{\mathsf{b}}}{\partial s}\cdot\frac{\partial\mathbf{X}}{\partial y^\ell} - \left(e\frac{\partial\Phi}{\partial s}+\mu\frac{\partial B}{\partial s}\right)b_\ell = \frac{d}{dt}(mub_\ell).$$

Using Eq. (5.37) and the identity $\langle d(\cdots)/dt\rangle_b=0$ (to lowest order in the bounce dynamics), the final expression for Eq. (5.34) is

$$\frac{dJ_{b0}}{dt} = \frac{e}{c\omega_b}(\langle\dot{\alpha}\rangle_b\dot{\beta}-\dot{\alpha}\langle\dot{\beta}\rangle_b), \tag{5.38}$$

which explicitly satisfies the condition $\langle dJ_{b0}/dt\rangle_b\equiv 0$. Hence, the bounce-center phase-dependent first-order correction is

$$\tilde{J}_{b1} \equiv -\frac{e}{c\omega_b^2}\int(\langle\dot{\alpha}\rangle_b\dot{\beta}-\dot{\alpha}\langle\dot{\beta}\rangle_b)d\Theta'.$$

The general case of time-dependent electromagnetic fields is treated with the electric field written as

$$\mathbf{E} = -\nabla(\Phi+\Psi)-\frac{1}{c}\left(\frac{\partial\alpha}{\partial t}\nabla\beta-\frac{\partial\beta}{\partial t}\nabla\alpha\right),$$

where $\Psi\equiv(\alpha/c)\partial\beta/\partial t$ is defined so that $E_\parallel\equiv-\partial(\Phi+\Psi)/\partial s$. We replace Eq. (5.34) with

$$\frac{dJ_{b0}}{dt} = \frac{d\alpha}{dt}\frac{\partial J_{b0}}{\partial\alpha}+\frac{d\beta}{dt}\frac{\partial J_{b0}}{\partial\beta}+\frac{dW}{dt}\frac{\partial J_{b0}}{\partial W}+\frac{\partial J_{b0}}{\partial t},$$

where

$$\frac{\partial J_{b0}}{\partial W} = \omega_b^{-1} \quad \text{and} \quad \frac{\partial J_{b0}}{\partial t} = -\omega_b^{-1}\langle\dot{W}\rangle_b,$$

and we replace Φ with $\Phi+\Psi$ in Eqs. (5.35) and (5.37). Last, we obtain

$$\frac{dJ_{b0}}{dt} = \frac{e}{c\omega_b}(\langle\dot{\alpha}\rangle_b\dot{\beta}-\dot{\alpha}\langle\dot{\beta}\rangle_b)+\frac{1}{\omega_b}(\dot{W}-\langle\dot{W}\rangle_b), \tag{5.39}$$

so that the first-order correction to the longitudinal adiabatic invariant is

$$J_{b1} = -\frac{1}{\omega_b}\int\frac{dJ_{b0}}{dt}d\Theta+\langle J_{b1}\rangle_b, \tag{5.40}$$

where the bounce-averaged contribution is discussed later. Similar expressions were derived by Dubin and Krommes (1982); Littlejohn (1982b); and Brizard (1990) using Lie-transform perturbation methods.

C. Adiabatic invariance to arbitrary order

The derivation of adiabatic invariants expressed as asymptotic power series in terms of a small ordering parameter ϵ must be placed in the wider context of a near-identity transformation $\mathcal{T}_\epsilon:Z_0\to Z\equiv\mathcal{T}_\epsilon Z_0$ between the lowest-order phase-space coordinates $Z_0^\alpha=(\mathbf{x},u_0,\mu_0,\zeta_0)$

and the guiding-center phase-space coordinates $Z^\alpha = (\mathbf{X}, u, \mu, \zeta)$, where the guiding-center phase-space coordinates

$$Z^\alpha = Z_0^\alpha + \epsilon G_1^\alpha + \epsilon^2 \left(G_2^\alpha + G_1^\beta \frac{\partial G_1^\alpha}{\partial Z_0^\beta} \right) + \cdots \tag{5.41}$$

are expressed in terms of the generating vector fields $(\mathsf{G}_1, \mathsf{G}_2, \ldots)$. Since the near-identity transformation is invertible, we define the inverse near-identity transformation $\mathcal{T}_\epsilon^{-1}: Z \to Z_0 \equiv \mathcal{T}^{-1} Z$ in terms of the generating vector fields $(\mathsf{G}_1, \mathsf{G}_2, \ldots)$ as

$$Z_0^\alpha = Z^\alpha - \epsilon G_1^\alpha - \epsilon^2 \left(G_2^\alpha - \frac{1}{2} G_1^\beta \frac{\partial G_1^\alpha}{\partial Z_0^\beta} \right) + \cdots . \tag{5.42}$$

Note that the Jacobian $\mathcal{J}_0$ associated with the lowest-order coordinates Z_0 transforms to the new Jacobian $\mathcal{J}$ associated with the guiding-center coordinates Z according to

$$\mathcal{J} = \mathcal{J}_0 - \epsilon \frac{\partial}{\partial Z^\alpha} (\mathcal{J}_0 G_1^\alpha) + \cdots , \tag{5.43}$$

which guarantees that $\mathcal{J}_0 d^6 Z_0 \equiv \mathcal{J} d^6 Z$.

In the remainder of this section, we present the explicit expressions for the guiding-center and bounce-center phase-space transformations and refer the reader to Tao *et al.* (2007) for the details of the drift-center phase-space transformation.

1. Guiding-center transformation

Using Lie-transform perturbation methods, the guiding-center transformation has been carried out to first-order in magnetic-field nonuniformity and is expressed in terms of the relation between the particle position $\mathbf{x}$ and the guiding-center position $\mathbf{X}$,

$$\mathbf{x} \equiv \mathbf{X} + \boldsymbol{\rho}_\epsilon, \tag{5.44}$$

where the generalized gyroradius vector is defined as

$$\boldsymbol{\rho}_\epsilon \equiv -\epsilon G_1^{\mathbf{x}} - \epsilon^2 \left(G_2^{\mathbf{x}} - \frac{1}{2} G_1^\alpha \frac{\partial G_1^{\mathbf{x}}}{\partial Z^\alpha} \right) + \cdots , \tag{5.45}$$

where $\boldsymbol{\rho}_0 \equiv -G_1^{\mathbf{x}}$ denotes the lowest-order gyroradius vector, while the velocity-space components of the first-order generating field are

$$G_1^u = -u \boldsymbol{\rho}_0 \cdot (\hat{\mathsf{b}} \cdot \nabla \hat{\mathsf{b}}) + \frac{\mu B}{m\Omega} (\hat{\mathsf{b}} \cdot \nabla \times \hat{\mathsf{b}} + \mathsf{a}_1 : \nabla \hat{\mathsf{b}}),$$

$$G_1^\mu = -\frac{m\mathbf{v}_\perp}{B} \cdot \mathbf{v}_D - \mu \frac{u}{\Omega} (\hat{\mathsf{b}} \cdot \nabla \times \hat{\mathsf{b}} + \mathsf{a}_1 : \nabla \hat{\mathsf{b}})$$
$$= \frac{\Omega}{B} \frac{\partial S_3}{\partial \zeta} + \mu \left(\frac{\boldsymbol{\rho}_0}{3} \cdot \nabla \ln B - \frac{u}{\Omega} \hat{\mathsf{b}} \cdot \nabla \times \hat{\mathsf{b}} \right),$$

$$G_1^\zeta = -\boldsymbol{\rho}_0 \cdot \mathbf{R} - \frac{\Omega}{B} \frac{\partial S_3}{\partial \mu},$$

where the gyrophase-dependent scalar field

$$S_3 \equiv -\boldsymbol{\rho}_0 \cdot \frac{\hat{\mathsf{b}}}{\Omega} \times \left(\frac{2\mu}{3} \nabla B + m u^2 \hat{\mathsf{b}} \cdot \nabla \hat{\mathsf{b}} \right)$$
$$- \frac{\mu B}{\Omega} \left(\frac{u}{\Omega} \mathsf{a}_2 : \nabla \hat{\mathsf{b}} \right)$$

is derived at third order in the perturbation analysis (Littlejohn, 1983) and the dyadic tensors $\mathsf{a}_1 \equiv \partial \mathsf{a}_2 / \partial \zeta$ and a_2 are

$$\mathsf{a}_1 = -\tfrac{1}{2}(\hat{\mathsf{a}}\hat{\mathsf{c}} + \hat{\mathsf{c}}\hat{\mathsf{a}}) \quad \text{and} \quad \mathsf{a}_2 = \tfrac{1}{4}(\hat{\mathsf{c}}\hat{\mathsf{c}} - \hat{\mathsf{a}}\hat{\mathsf{a}}).$$

The spatial component $G_2^{\mathbf{x}}$ of the second-order vector field is expressed as

$$G_2^{\mathbf{x}} = \boldsymbol{\rho}_0 \left(\frac{u}{\Omega} \hat{\mathsf{b}} \cdot \nabla \times \hat{\mathsf{b}} \right) - \frac{\partial S_3}{\partial u} \frac{\hat{\mathsf{b}}}{m} + \frac{1}{2} \left(g_1^\mu \frac{\partial \boldsymbol{\rho}_0}{\partial \mu} + g_1^\zeta \frac{\partial \boldsymbol{\rho}_0}{\partial \zeta} \right),$$

where $g_1^\mu \equiv G_1^\mu - \mu(\boldsymbol{\rho}_0 \cdot \nabla \ln B)$ and $g_1^\zeta \equiv G_1^\zeta + \boldsymbol{\rho}_0 \cdot \mathbf{R}$, and the gyrogauge vector field is defined as $\mathbf{R} \equiv \nabla \hat{\mathsf{c}} \cdot \hat{\mathsf{a}} = \nabla \hat{1} \cdot \hat{2}$.

Note that the guiding-center Jacobian is constructed from the first-order generating field G_1 as

$$B_\|^* \equiv B - \frac{\partial}{\partial Z^\alpha} (B G_1^\alpha) + \cdots$$
$$= \nabla \cdot (B \boldsymbol{\rho}_0) + B \left(1 - \frac{\partial G_1^u}{\partial u} - \frac{\partial G_1^\mu}{\partial \mu} - \frac{\partial G_1^\zeta}{\partial \zeta} \right)$$
$$= B \left(1 + \frac{u}{\Omega} \hat{\mathsf{b}} \cdot \nabla \times \hat{\mathsf{b}} \right).$$

Note also that the kinetic energy is invariant to first order under the guiding—center transformation since

$$G_1^E \equiv B G_1^\mu + m u G_1^u - \mu \boldsymbol{\rho}_0 \cdot \nabla B \equiv 0.$$

Hence, the guiding-center and particle kinetic energies are identical to first order.

2. Bounce-center transformation

We first introduce the symplectic part of the parallel guiding-center phase-space Lagrangian

$$u \hat{\mathsf{b}} \cdot \dot{\mathbf{X}} = u \left[b_k \dot{y}^k + \frac{\partial s}{\partial J_b} \dot{J}_b + \frac{\partial s}{\partial \Theta} \dot{\Theta} + b_\tau \dot{\tau} \right] \equiv \Lambda_{\|\alpha} \dot{\mathcal{Z}}^\alpha,$$

expressed in terms of the coordinates $\mathcal{Z}^\alpha = (y^1, y^2, J_b, \Theta, w, \tau)$. Next, we construct the Lagrange matrix components

$$\Omega_{\|\alpha\beta} \equiv \frac{\partial \Lambda_{\|\beta}}{\partial \mathcal{Z}^\alpha} - \frac{\partial \Lambda_{\|\alpha}}{\partial \mathcal{Z}^\beta},$$

and express the guiding-center equations as

$$\dot{y}^k = \epsilon \eta^{k\ell} \left(\frac{\partial H}{\partial y^\ell} - \omega_b \Omega_{\|\ell\Theta} \right),$$

$$\dot{J}_b = \epsilon (\Omega_{\|\Theta\tau} + \Omega_{\|\Theta k} \epsilon^{-1} \dot{y}^k),$$

$$\dot{\Theta} = \omega_b + \epsilon (\Omega_{\|\tau J} + \epsilon^{-1} \dot{y}^k \Omega_{\|kJ}),$$

$$\dot{w} = \epsilon\left(\frac{\partial H}{\partial \tau} + \omega_b \Omega_{\|\Theta\tau}\right),$$

where first-order corrections are retained and the components $\Omega_{\|\alpha\Theta}$ depend explicitly on the bounce angle Θ. By defining the bounce-center phase-dependent functions

$$F_\alpha \equiv \int_0^\Theta \Omega_{\|\Theta\alpha} d\Theta',$$

we express the components of the first-order generating vector field for the bounce-center transformation as

$$G_1^k = -\eta^{k\ell} F_\ell,$$

$$G_1^J = \frac{\partial S_2}{\partial \Theta} + \frac{1}{2}\eta^{k\ell} F_k \frac{\partial F_\ell}{\partial \Theta},$$

$$G_1^\Theta = -\frac{\partial S_2}{\partial J_b} - \frac{1}{2}\eta^{k\ell} F_k \Omega_{\|I\ell},$$

$$G_1^w = F_\tau,$$

with $G_1^\tau \equiv 0$ and the scalar field S_2 is a solution to the differential equation

$$\omega_b \frac{\partial S_2}{\partial \Theta} = -\frac{\omega_b}{2}\eta^{k\ell}\left(F_k \frac{\partial F_\ell}{\partial \Theta} - \left\langle F_k \frac{\partial F_\ell}{\partial \Theta}\right\rangle_b\right) + \eta^{k\ell}\frac{\partial H}{\partial y^k} F_\ell - F_\tau.$$

Note that since

$$\langle \dot{y}^k \rangle_b = \eta^{k\ell}\frac{\partial H_b}{\partial y^\ell} \quad \text{and} \quad \langle \dot{w} \rangle_b = \frac{\partial H_b}{\partial \tau},$$

then

$$\begin{pmatrix} \eta^{k\ell}\Omega_{\|\Theta\ell} \\ \Omega_{\|\Theta\tau} \end{pmatrix} = \omega_b^{-1}\begin{pmatrix} \dot{y}^k - \langle \dot{y}^k \rangle_b \\ \dot{w} - \langle \dot{w} \rangle_b \end{pmatrix}$$

and

$$G_1^J = \frac{1}{\omega_b}\left(\eta^{k\ell}\frac{\partial H}{\partial y^k} F_\ell - F_\tau\right) + \frac{1}{2}\eta^{k\ell}\left\langle F_k \frac{\partial F_\ell}{\partial \Theta}\right\rangle_b = -\omega_b^{-1}\frac{dJ_{b0}}{dt},$$

where $\langle F_k \partial F_\ell / \partial \Theta \rangle \equiv 0$ (because the functions F_k are odd in Θ) and

$$\begin{pmatrix} \eta^{k\ell} F_\ell \\ F_\tau \end{pmatrix} = \frac{1}{\omega_b}\int \begin{pmatrix} \dot{y}^k - \langle \dot{y}^k \rangle_b \\ \dot{w} - \langle \dot{w} \rangle_b \end{pmatrix} d\Theta.$$

The bounce-center transformation presented here was also presented by Tao *et al.* (2007) in the relativistic limit. In the next section, we present the details of the derivation of relativistic Hamiltonian guiding-center equations of motion.

VI. RELATIVISTIC GUIDING-CENTER THEORY

The relativistic motion of a particle of rest mass m and charge e is described in eight-dimensional phase space in terms of the space-time coordinates $x^\mu = (x^0 = ct, \mathbf{x})$ and the four-momentum $p^\mu = m u^\mu$, with the four-velocity defined as

$$u^\mu = \frac{dx^\mu}{d\tau} = (u^0 = \gamma c, \mathbf{u} = \gamma \mathbf{v}), \tag{6.1}$$

where $\gamma = (1 - |\mathbf{v}|^2/c^2)^{-1/2} = (1 + |\mathbf{u}|^2/c^2)^{1/2}$ is the relativistic factor and $dx^\mu / d\tau = \gamma \dot{x}^\mu$ is the derivative with respect to proper time τ. Once again, we use the Minkowski space-time metric $g^{\mu\nu} = \mathrm{diag}(-1, +1, +1, +1)$ so that $u_\mu u^\mu = -c^2$.

The equation of motion for the four-momentum p^μ is

$$\frac{dp^\mu}{d\tau} = \frac{e}{c} F^{\mu\nu} u_\nu, \tag{6.2}$$

where

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \tag{6.3}$$

denotes the Faraday tensor. Here the space-time contravariant derivative is

$$\partial^\mu = g^{\mu\nu}\partial_\nu = (-\partial/\partial x^0, \nabla),$$

where $A^\mu = (A^0 = \Phi, \mathbf{A})$ is the electromagnetic four-potential.

A. Relativistic Hamiltonian formulations

We begin by introducing two Hamiltonian formulations for the relativistic equations of motion for a charged particle moving in an electromagnetic field. Each formulation is defined in terms of a Hamiltonian function H and a Poisson bracket $\{,\}$ derived from a phase-space Lagrangian.

The first formulation is based on a covariant description expressed in terms of the phase-space coordinates $\mathcal{Z}^{\mathsf{a}} = (x^\mu, p^\mu)$. The covariant formulation treats space and time as well as momentum and energy on equal footings. The second formulation, on the other hand, treats time and space separately and makes use of the extended phase-space coordinates $z^{\mathsf{a}} = (\mathbf{x}, \mathbf{p}; t, w)$, where the energy coordinate w is canonically conjugate to time t.

1. Covariant formulation

We begin our task of finding a suitable covariant (c) Hamiltonian formulation for Eqs. (6.1) and (6.2) in terms of a covariant Hamiltonian H_c and a covariant Poisson bracket $\{,\}_c$,

$$\frac{d\mathcal{Z}^{\mathsf{a}}}{d\tau} \equiv \{\mathcal{Z}^{\mathsf{a}}, H_c\}_c.$$

First, we start with the covariant relativistic phase-space Lagrangian

$$\mathcal{L}_c \equiv \left(p_\mu + \frac{e}{c}A_\mu\right)\dot{x}^\mu, \tag{6.4}$$

where $\dot{x}^\mu \equiv dx^\mu/d\tau$. From this Lagrangian, we obtain the covariant (8×8) Lagrange matrix

$$\boldsymbol{\omega}_c = \begin{pmatrix} (e/c)F_{\mu\nu} & -g_{\mu\nu} \\ g_{\mu\nu} & 0 \end{pmatrix}, \tag{6.5}$$

whose inverse yields the covariant Poisson matrix

$$\boldsymbol{\Pi}_c = \begin{pmatrix} 0 & g^{\mu\nu} \\ -g^{\mu\nu} & (e/c)F^{\mu\nu} \end{pmatrix}. \tag{6.6}$$

The Poisson bracket $\{\mathcal{A},\mathcal{B}\}_c$ of two functions $\mathcal{A}$ and $\mathcal{B}$ on eight-dimensional phase space is thus

$$\{\mathcal{A},\mathcal{B}\}_c \equiv g^{\mu\nu}\left(\frac{\partial\mathcal{A}}{\partial x^\mu}\frac{\partial\mathcal{B}}{\partial p^\nu} - \frac{\partial\mathcal{A}}{\partial p^\nu}\frac{\partial\mathcal{B}}{\partial x^\mu}\right) + \frac{e}{c}F^{\mu\nu}\frac{\partial\mathcal{A}}{\partial p^\mu}\frac{\partial\mathcal{B}}{\partial p^\nu}. \tag{6.7}$$

Substituting $\mathcal{A}=x^\mu$ or p^μ and $\mathcal{B}=H_c$ into Eq. (6.7), we find, respectively,

$$\{x^\mu,H_c\}_c = g^{\mu\nu}\frac{\partial H_c}{\partial p^\nu}, \tag{6.8}$$

$$\{p^\mu,H_c\}_c = -g^{\mu\nu}\frac{\partial H_c}{\partial x^\nu} + \frac{eF^{\mu\nu}}{c}\frac{\partial H_c}{\partial p^\nu}. \tag{6.9}$$

We recover the equations of motion relativistic particle dynamics (6.1) and (6.2) from Eqs. (6.8) and (6.9) if

$$\frac{\partial H_c}{\partial x^\mu} \equiv 0 \quad \text{and} \quad \frac{\partial H_c}{\partial p^\nu} \equiv u_\nu,$$

which implies that the covariant relativistic Hamiltonian must be of the form

$$H_c \equiv g_{\mu\nu}\frac{p^\mu p^\nu}{2m}. \tag{6.10}$$

Note that covariant relativistic particle motion takes place on the surface $H_c(\mathcal{Z}) \equiv -mc^2/2$ and, hence, H_c is a Lorentz scalar (i.e., it is not energylike). Furthermore, the covariant relativistic Hamiltonian (6.10) does not have a well-defined nonrelativistic limit (Jackson, 1975), which can make it impractical for some applications.

2. Noncovariant formulation

Because of the problems associated with a covariant Hamiltonian formulation, we turn our attention to a noncovariant Hamiltonian formulation of relativistic particle dynamics (Brizard and Chan, 1999). Here the time variable is to be treated differently from the spatial coordinates and we look for a noncovariant (energylike) Hamiltonian H and a noncovariant Poisson bracket $\{,\}$: $dz^{\mathsf{a}}/dt \equiv \{z^{\mathsf{a}},H\}$, where $z^{\mathsf{a}} \equiv (\mathbf{x},\mathbf{p};t,\mathcal{E})$ are eight-dimensional extended phase-space coordinates. The extended relativistic Hamiltonian is

$$H \equiv \gamma mc^2 + e\Phi - \mathcal{E}, \tag{6.11}$$

where

$$\gamma = \sqrt{1 + |\mathbf{p}/mc|^2}$$

is the relativistic factor expressed in terms of the relativistic kinetic momentum. Note that the Hamiltonian (6.11) has a well-defined classical limit and is an energy-like quantity. To complete this Hamiltonian formulation, we turn our attention to deriving a suitable expression for the extended phase-space Poisson bracket $\{,\}$.

The extended phase-space Lagrangian is

$$\mathcal{L} = \left(\mathbf{p} + \frac{q}{c}\mathbf{A}\right)\cdot\frac{d\mathbf{x}}{d\sigma} - \mathcal{E}\frac{dt}{d\sigma} - H \equiv \Gamma_{\mathsf{a}}\frac{dz^{\mathsf{a}}}{d\sigma} - H, \tag{6.12}$$

where σ represents a Hamiltonian orbit parameter in extended phase space and the physical particle motion in eight-dimensional extended phase space takes place on the surface $H=0$, or $\mathcal{E}=\gamma mc^2 + e\Phi$. By inverting the extended phase-space Lagrange matrix obtained from the symplectic part (Γ_{a}), we construct the extended phase-space Poisson bracket

$$\{F,G\} \equiv \left(\frac{\partial F}{\partial\mathcal{E}}\frac{\partial G}{\partial t} - \frac{\partial F}{\partial t}\frac{\partial G}{\partial\mathcal{E}}\right) + \left(\nabla F\cdot\frac{\partial G}{\partial\mathbf{p}} - \frac{\partial F}{\partial\mathbf{p}}\cdot\nabla G\right) + \frac{e}{c}\frac{\partial\mathbf{A}}{\partial t}\cdot\left(\frac{\partial F}{\partial\mathbf{p}}\frac{\partial G}{\partial\mathcal{E}} - \frac{\partial F}{\partial\mathcal{E}}\frac{\partial G}{\partial\mathbf{p}}\right) + \frac{e\mathbf{B}}{c}\cdot\frac{\partial F}{\partial\mathbf{p}}\times\frac{\partial G}{\partial\mathbf{p}}. \tag{6.13}$$

Hence, using Eqs. (6.11) and (6.13), we find

$$\frac{d\mathbf{x}}{d\sigma} = \{\mathbf{x},H\} = \frac{\mathbf{p}}{\gamma m},$$

$$\frac{d\mathbf{p}}{d\sigma} = \{\mathbf{p},H\} = e\mathbf{E} + \frac{\mathbf{p}}{\gamma m}\times\frac{e\mathbf{B}}{c},$$

$$\frac{d\mathcal{E}}{d\sigma} = \{\mathcal{E},H\} = e\frac{\partial\Phi}{\partial t} - \frac{e}{c}\dot{\mathbf{x}}\cdot\frac{\partial\mathbf{A}}{\partial t},$$

$$\frac{dt}{d\sigma} = \{t,H\} = +1.$$

Note that this noncovariant formulation separates the components of the electromagnetic four-potential: the scalar potential Φ appears explicitly in the Hamiltonian (6.11) while the vector potential $\mathbf{A}$ appears explicitly in the Poisson bracket (6.13).

B. Relativistic Hamiltonian guiding-center theory

Derivation of a relativistic Hamiltonian guiding-center theory follows steps similar to derivation of the nonrelativistic guiding-center theory. In the present section, we only present results of derivations presented elsewhere. Note that these Hamiltonian guiding-center formulations possess the same advantages (e.g., energy conservation) over the relativistic guiding-center equations presented by Northrop (1963).

1. Covariant formulation

The covariant formulation of relativistic Hamiltonian guiding-center theory was presented by Boghosian (1987) based on earlier work by Fradkin (1978). The fast gyromotion time scale is explicitly introduced by introduction of eigenvalues $\pm\lambda_E$ and $\pm i\lambda_B$ of the Faraday tensor (6.13), where

$$\begin{pmatrix}\lambda_E^2\\ \lambda_B^2\end{pmatrix} \equiv \mp\Lambda_1 + \sqrt{\Lambda_1^2+\Lambda_2^2}$$

are expressed in terms of the Lorentz invariants $\Lambda_1 \equiv \frac{1}{2}(|\mathbf{B}|^2-|\mathbf{E}|^2)$ and $\Lambda_2 \equiv \mathbf{E}\cdot\mathbf{B}$. We then introduce two projection operators $\mathcal{P}_\parallel$ and $\mathcal{P}_\perp$ defined as

$$\begin{pmatrix}P_\parallel^{\mu\nu}\\ P_\perp^{\mu\nu}\end{pmatrix} \equiv (\lambda_B^2+\lambda_E^2)^{-1}\left[\pm F^\mu_{\ \sigma}F^{\sigma\nu} + \begin{pmatrix}\lambda_B^2\\ \lambda_E^2\end{pmatrix} g^{\mu\nu}\right],$$

which satisfy the properties $\mathcal{P}_\parallel\cdot\mathcal{P}_\parallel=\mathcal{P}_\parallel$, $\mathcal{P}_\perp\cdot\mathcal{P}_\perp=\mathcal{P}_\perp$, $\mathcal{P}_\perp\cdot\mathcal{P}_\parallel=0=\mathcal{P}_\parallel\cdot\mathcal{P}_\perp$, and $P_\parallel^{\mu\nu}+P_\perp^{\mu\nu}=g^{\mu\nu}$. The operators $\mathcal{P}_\parallel$ and $\mathcal{P}_\perp$ project an arbitrary four-vector $V^\mu\equiv(V^0,\mathbf{V})$ onto the parallel two-flat and the perpendicular two-flat, respectively. For example, when these operators are expressed in the preferred frame where $\mathbf{E}=\mathbf{0}$, we find

$$P_\parallel^{\mu\nu}V_\nu = (V^0, V_\parallel\hat{\mathsf{b}}),$$

$$P_\perp^{\mu\nu}V_\nu = (0,\hat{\mathsf{b}}\times(\mathbf{V}\times\hat{\mathsf{b}})), \tag{6.14}$$

and, hence, the parallel two-flat combines the time coordinate with the spatial coordinate along the magnetic field while the perpendicular two-flat combines the two spatial coordinates that span the plane perpendicular to the local magnetic field.

These decompositions lead us to the covariant representation for the particle's relativistic four-vector velocity

$$u^\mu = q_\parallel(e_0^\mu\cosh\beta + e_1^\mu\sinh\beta) - q_\perp(e_2^\mu\sin\zeta + e_3^\mu\cos\zeta)$$
$$\equiv q_\parallel t^\mu + q_\perp c^\mu,$$

where $q_\parallel^2 \equiv -u_\mu\mathcal{P}_\parallel^{\mu\nu}u_\nu$ and $q_\perp^2\equiv u_\mu\mathcal{P}_\perp^{\mu\nu}u_\nu$, so that $q_\parallel^2-q_\perp^2 = c^2$, and the orthogonal basis four—vectors $(\mathsf{e}_0,\mathsf{e}_1,\mathsf{e}_2,\mathsf{e}_3)$ are used to define the parallel two-flat (spanned by e_0 and e_1) and the perpendicular two-flat (spanned by e_2 and e_3), with $\mathsf{e}_\mu\cdot\mathsf{e}_\nu=g_{\mu\nu}$. We also define $b^\mu\equiv\partial t^\mu/\partial\beta$ and $a^\mu\equiv-\partial c^\mu/\partial\zeta$.

We now write the covariant relativistic phase-space Lagrangian (6.4) as

$$\mathcal{L}_c \equiv \left(\frac{e}{c}A_\mu + mq_\parallel t_\mu + mq_\perp c_\mu\right)\dot{x}^\mu, \tag{6.15}$$

and the covariant Hamiltonian (6.10) as

$$H_c \equiv \frac{m}{2}(q_\perp^2 - q_\parallel^2). \tag{6.16}$$

The derivation of a covariant relativistic guiding-center Lagrangian from Eq. (6.15) proceeds similarly as in Sec. III.C, with the relativistic gyroradius four-vector defined as $\rho^\mu\equiv(q_\perp/\Omega_B)a^\mu$, where $\Omega_B\equiv e\lambda_B/mc$.

Working in the preferred frame where $\mathbf{E}=\mathbf{0}$, the covariant relativistic guiding-center (crgc) Lagrangian is expressed in terms the lowest-order guiding-center coordinates $Z^{\mathsf{a}}_{\rm crgc}=(X^\mu\equiv x^\mu-\rho^\mu, q_\parallel,\beta,\mu,\zeta)$ as

$$\mathcal{L}_{\rm crgc} = \left(\frac{e}{c}A_\mu + mq_\parallel t_\mu\right)\dot{X}^\mu + \mu\left(\frac{mc}{e}\right)\dot{\zeta}, \tag{6.17}$$

while the covariant relativistic guiding-center Hamiltonian is

$$H_{\rm crgc} = \mu\lambda_B - \frac{m}{2}q_\parallel^2. \tag{6.18}$$

We can now derive covariant relativistic guiding-center equations of motion expressed in terms of a covariant relativistic Poisson bracket $\{,\}_{\rm crgc}$ derived from Eq. (6.17) and the Hamiltonian (6.18) as $\dot{Z}^{\mathsf{a}}_{\rm crgc}=\{Z^{\mathsf{a}}_{\rm crgc},H_{\rm crgc}\}_{\rm crgc}$, with $\dot{\mu}=\{\mu,H_{\rm crgc}\}_{\rm crgc}\equiv0$ and the gyrophase angle ζ has become an ignorable angle. Instead of writing explicit expressions for these covariant relativistic guiding-center equations of motion, which are found in Boghosian (1987), we now present the noncovariant relativistic guiding-center equations of motion, which have greater applicability.

2. Noncovariant formulation

For weakly time-dependent fields, the relativistic guiding-center phase-space Lagrangian is expressed in terms of extended guiding-center phase-space coordinates $\mathcal{Z}^a\equiv(\mathbf{X},p_\parallel;\mu,\zeta;w,t)$ as

$$\mathcal{L}_{\rm rgc} = \left[\frac{e}{c}\mathbf{A}(\mathbf{X},t) + p_\parallel\hat{\mathsf{b}}(\mathbf{X},t)\right]\cdot\dot{\mathbf{X}} + \mu(mc/e)\dot{\zeta}$$
$$- w\dot{t} - \mathcal{H}_{\rm rgc}, \tag{6.19}$$

where $\dot{\mathcal{Z}}^a\equiv d\mathcal{Z}^a/d\sigma$. The relativistic guiding-center extended Hamiltonian is

$$H_{\rm rgc} \equiv \gamma mc^2 + e\Phi(\mathbf{X},t) - w, \tag{6.20}$$

where $\gamma=\sqrt{1+(2/mc^2)\mu B(\mathbf{X},t)+p_\parallel^2/(mc)^2}$ is the guiding-center relativistic factor and the relativistic guiding-center extended Poisson bracket is

$$\{F,G\}_{\rm rgc} \equiv \frac{e}{mc}\left(\frac{\partial F}{\partial\zeta}\frac{\partial G}{\partial\mu} - \frac{\partial F}{\partial\mu}\frac{\partial G}{\partial\zeta}\right)$$
$$+ \frac{\mathbf{B}^*}{B_\parallel^*}\cdot\left(\nabla^* F\frac{\partial G}{\partial p_\parallel} - \frac{\partial F}{\partial p_\parallel}\nabla^* G\right)$$
$$- \frac{c\hat{\mathsf{b}}}{eB_\parallel^*}\cdot\nabla^* F\times\nabla^* G$$
$$+ \left(\frac{\partial F}{\partial w}\frac{\partial G}{\partial t} - \frac{\partial F}{\partial t}\frac{\partial G}{\partial w}\right), \tag{6.21}$$

where the effective gradient operator ∇^* is

$$\nabla^* \equiv \nabla - \frac{e}{c}\frac{\partial \mathbf{A}^*}{\partial t}\frac{\partial}{\partial w},$$

and we have introduced the effective magnetic field

$$\mathbf{B}^* \equiv \nabla \times \mathbf{A}^* = \nabla \times [\mathbf{A} + (cp_\parallel/e)\hat{\mathsf{b}}] = \mathbf{B} + (cp_\parallel/e)\,\nabla \times \hat{\mathsf{b}}, \tag{6.22}$$

from which we define

$$B_\parallel^* \equiv \hat{\mathsf{b}}\cdot\mathbf{B}^* = B + (cp_\parallel/e)\hat{\mathsf{b}}\cdot\nabla\times\hat{\mathsf{b}}. \tag{6.23}$$

The Jacobian for the guiding-center transformation is $mB_\parallel^*$, i.e., $d^8z \equiv mB_\parallel^* d^8\mathcal{Z}$. We make an important remark here that the noncovariant relativistic guiding-center Poisson bracket (6.21) has the same form as the nonrelativistic Poisson bracket.

Using the relativistic guiding-center Hamiltonian (6.20) and Poisson bracket (6.21), the relativistic guiding-center Hamilton equations are expressed as $d\mathcal{Z}^a/d\sigma = \{\mathcal{Z}^a, \mathcal{H}_{\rm rgc}\}_{\rm rgc}$, with $dt/d\sigma=1$. The relativistic guiding-center velocity is

$$\frac{d\mathbf{X}}{dt} = \frac{p_\parallel}{m\gamma}\frac{\mathbf{B}^*}{B_\parallel^*} + \mathbf{E}^* \times \frac{c\hat{\mathsf{b}}}{B_\parallel^*}, \tag{6.24}$$

and the relativistic guiding-center parallel force equation is

$$\frac{dp_\parallel}{dt} = e\mathbf{E}^*\cdot\frac{\mathbf{B}^*}{B_\parallel^*}, \tag{6.25}$$

where we have introduced the effective electric field

$$\mathbf{E}^* \equiv -\nabla\Phi^* - \frac{1}{c}\frac{\partial \mathbf{A}^*}{\partial t} = \mathbf{E} - \frac{1}{e}\left(mc^2\,\nabla\,\gamma - p_\parallel\frac{\partial\hat{\mathsf{b}}}{\partial t}\right), \tag{6.26}$$

where the effective potentials are

$$\begin{pmatrix}\Phi^*\\ \mathbf{A}^*\end{pmatrix} \equiv \begin{pmatrix}\Phi\\ \mathbf{A}\end{pmatrix} + \frac{mc}{e}\begin{pmatrix}\gamma c\\ \gamma v_\parallel\hat{\mathsf{b}}\end{pmatrix}. \tag{6.27}$$

We note that the electromagnetic potentials in Eq. (6.27) are corrected by the parallel two-flat projection of the guiding-center four-velocity, defined in Eq. (6.14). Note that the relativistic guiding-center equations (6.24) and (6.25) are identical to those presented by Grebogi and Littlejohn (1984) if we substitute $\Phi^*\to\Phi$ and $\mathbf{E}^*\to\mathbf{E} - (\mu/e\gamma)\nabla B$. The relativistic guiding-center equations for the canonically conjugate coordinates (μ,ζ), on the other hand, are

$$\frac{d\mu}{dt} = -\frac{e}{mc}\frac{\partial H_{\rm rgc}}{\partial\zeta} = 0, \tag{6.28}$$

$$\frac{d\zeta}{dt} = \frac{e}{mc}\frac{\partial H_{\rm rgc}}{\partial\mu} \equiv \Omega\gamma^{-1}, \tag{6.29}$$

which completes the relativistic Hamiltonian formulation of guiding-center motion.

The relativistic guiding-center Hamiltonian equations (6.24) and (6.25) have the phase-space volume-preservation property

$$0 \equiv \frac{\partial B_\parallel^*}{\partial t} + \nabla\cdot(B_\parallel^*\dot{\mathbf{X}}) + \frac{\partial}{\partial p_\parallel}(B_\parallel^*\dot{p}_\parallel), \tag{6.30}$$

since

$$\frac{\partial B_\parallel^*}{\partial t} = \hat{\mathsf{b}}\cdot\frac{\partial\mathbf{B}^*}{\partial t} + \mathbf{B}^*\cdot\frac{\partial\hat{\mathsf{b}}}{\partial t} = -c\hat{\mathsf{b}}\cdot\nabla\times\mathbf{E}^* + \mathbf{B}^*\cdot\frac{\partial\hat{\mathsf{b}}}{\partial t},$$

$$\nabla\cdot(B_\parallel^*\dot{\mathbf{X}}) = c(\hat{\mathsf{b}}\cdot\nabla\times\mathbf{E}^* - \mathbf{E}^*\cdot\nabla\times\hat{\mathsf{b}}) + \frac{p_\parallel}{m}\mathbf{B}^*\cdot\nabla\gamma^{-1},$$

$$\frac{\partial}{\partial p_\parallel}(B_\parallel^*\dot{p}_\parallel) = e\left(\frac{\partial\mathbf{E}^*}{\partial p_\parallel}\cdot\mathbf{B}^* + \mathbf{E}^*\cdot\frac{\partial\mathbf{B}^*}{\partial p_\parallel}\right) = -\mathbf{B}^*\cdot\left(\frac{p_\parallel}{m}\nabla\,\gamma^{-1} + \frac{\partial\hat{\mathsf{b}}}{\partial t}\right) + c\mathbf{E}^*\cdot\nabla\times\hat{\mathsf{b}}.$$

One final note concerns the validity of the guiding-center approximation itself when considering applications of the relativistic guiding-center Hamiltonian equations (6.24), (6.25), (6.28), and (6.29). In standard guiding-center theory (Northrop, 1963), the small ordering parameter $\rho/L \equiv \epsilon \ll 1$, which scales with the mass of the guiding-center particle. Hence, since relativistic effects introduce the γm dependence on inertia, it would then appear that the relativistic guiding-center ordering parameter $\gamma\epsilon_0$ (where ϵ_0 denotes the characteristic rest-mass gyroradius) is no longer small only at very high kinetic energies $(\gamma \sim \epsilon_0^{-1} \gg 1)$. One can therefore confidently apply the relativistic guiding-center Hamiltonian equations (6.24), (6.25), (6.28), and (6.29) for relativistic charged particles with $\gamma \ll \epsilon_0^{-1}$.

VII. DYNAMICS IN TOROIDAL CONFINEMENT SYSTEMS

The confinement of pressure in magnetohydrodynamics requires a magnetic configuration in which the field lines lie on nested toroidal surfaces. But for collisionless plasmas, the fact that the orbits of particles can be large leads to large cross-field particle and energy transport (Hinton and Hazeltine, 1976; Kovrizhnykh, 1984), known as *neoclassical* transport. The large orbits arise due to the guiding-center drifts, which carry particles to different flux surfaces. This causes the resulting large transport—a particle moves far from its initial surface, then due to a collision its magnetic moment is changed, and the particle is on a new trajectory unrelated to the first. The diffusion coefficient for this random walk process is $D = \nu(\Delta\Psi)^2$, where ν is the collision frequency for the appropriate change of the magnetic moment, and $\Delta\Psi$ is either the width of the orbit in the flux variable (for the case of collision frequency small compared with the orbit frequency) or the typical change of the flux variable Ψ due to guiding-center drifts in one collision time (for the opposite case).

Here we consider the reduction of only the magnetic cross-flux-surface drift. [A recent review (Mynick, 2006) discussed the calculation and reduction of the cross-field collisional transport.] The self-consistently calculated electrostatic potential is found (Kovrizhnykh, 1984) to be of the same order as the magnetic potential, and so the ordering of Sec. V is appropriate, where the guiding-center Hamiltonian has the form of Eq. (5.2), in which the three terms are of the same order, and the term going as the square of the electric drift is dropped. In addition, the electrostatic potential is found to be dominantly a flux function. Hence, the electrostatic potential gives rise to drifts only within the flux surface. That we need consider only the magnetic cross-flux-surface drifts is convenient, as only the magnetic field comes from solving the scalar-pressure equilibrium equation.

In the case in which the magnetic drifts do not carry particles far from a flux surface, it is possible to develop a rigorous transport theory in which the distribution on each flux surface is nearly Maxwellian (Hinton and Hazeltine, 1976) even in the low-collisionality regime, where the collision time is long compared with the time needed for a trajectory to experience its maximum variation of flux variable Ψ. We call this theory *omnigeneous neoclassical theory*, as similar scalings of diffusion and thermal transport coefficients hold.

Palumbo (1968) suggested seeking *isodynamic* equilibria, those for which the flux surface crossing drift vanishes everywhere. This imposes the strict condition that the magnetic field be constant on a flux surface in the case in which there is net poloidal current. Hall and McNamara (1975), in examining open configurations, imposed the less stringent condition of *omnigeneity*, that the bounce averaged cross-flux-surface drift vanish.[2] Nührenberg and Zille (1988) proposed the condition of *quasihelicity*, one way in which to obtain three-dimensional equilibria that are integrable, with trajectories that remain close to the flux surfaces. Nuhrenberg showed numerically generated three-dimensional equilibria that approximately satisfied this condition. Cary and Shasharina (1997a, 1997b, 1997c) analyzed the less strict requirement of simple omnigeneity and showed a number of consequences and properties of the resulting systems. These properties could be sought through numerical means. The least stringent condition is that of Mynick *et al.* (1982), who proposed a class of stellarator configurations having reduced transport due to omnigeneous for either deeply or marginally trapped trajectories. We call such systems *specifically omnigeneous*. These systems were later discussed by Mynick (1983). Further classifications of the various systems are discussed by Mynick (2006).

For each of these types of systems, neoclassical transport is significantly reduced. The greatest reduction is for the isodynamic systems, which have neither neoclassical transport nor even the enhanced Pfirsch-Schlüter transport occurring in the collisional regime. Systems that are guiding-center integrable and those that are omnigeneous have the least troublesome neoclassical transport. For these systems, the orbit width is small in the guiding-center adiabatic parameter, and so the transport decreases with increasing magnetic field. With neither omnigeneity nor guiding-center integrability, there are trajectories whose width does not scale with magnetic field and is usually of the order of the machine radius in the absence of a strong electrostatic field. However, for transition omnigeneity, one is guaranteed the absence of transition trajectories, which are chaotic due to separatrix crossing and, so, cause transport even in the limit of zero collisionality. Finally, specific omnigeneity guarantees that at least some particles are omnigeneous and, hence, do not contribute to neoclassical transport.

Not surprisingly, the better the transport properties and, hence, the more specific the requirements, the more difficult such systems are to find. Bernardin *et al.* (1986), through expansion about the magnetic axis, showed that toroidal isodynamic configurations could not have closed flux surfaces without the magnetic field vanishing on axis. While Nührenberg and Zille (1988) found approximately guiding-center integrable configurations for large aspect ratio, Garren and Boozer (1991) showed that, in third order in an expansion away from the magnetic axis, guiding-center integrability cannot be satisfied. Some work by Meyer and Schmidt (1958) indicated that certain types of specific omnigeneity can be obtained, but little additional work has been done in this area.

In this section, we use noncanonical Hamiltonian guiding-center theory to review and add to the literature on improved confinement configurations. We begin by deriving the guiding-center equations of motion in flux variables. The condition for isodynamism follows from the guiding-center equations of motion. For isodynamic systems, both angles are ignorable in noncanonical Hamiltonian guiding-center theory. This implies constancy of the kinetic energy. We next consider the case of quasihelicity, which we obtain by requiring the phase-space Lagrangian coordinates to depend on the angles only through a single linear combination. A special case of this is when the angles are those of Boozer coordinates. Next we consider omnigeneity (zero bounce average drift off the flux surfaces). We first consider the conditions for specific omnigeneity, i.e., omnigeneity of various classes of particles. Finally, we show that the condition of omnigeneity is less restrictive than quasihelicity. We summarize by noting the hierarchy of conditions for improved confinement. Throughout our discussion, we use units such that $e=m=c=1$ and without the adiabatic ordering parameter.

[2]There is some confusion in the literature on these terms with isodynamic or isodynamism and omnigeneous or omnigeneity used interchangeably. Indeed, as noted by Catto and Hazeltine (1981), Hall and McNamara (1975) appear to state that by omnigeneity they mean that the instantaneous drift is within the surface. However, their later discussion indicates that by omnigeneity they mean that the bounce averaged drift is within the surface. We take omnigeneity to have the latter meaning.

A. Guiding-center equations of motion in flux coordinates

The guiding-center Lagrangian (3.2) for bounce ordering and in flux coordinates is

$$L_g = (ub_\Psi)\dot{\Psi} + (\Psi + ub_\theta)\dot{\theta} + [A_\varphi(\Psi) + ub_\varphi]\dot{\varphi} - H_g, \tag{7.1}$$

for which the Hamiltonian (3.49) becomes

$$H_g = \tfrac{1}{2}u^2 + \mu B(\Psi,\theta,\varphi) + \Phi(\Psi), \tag{7.2}$$

in keeping with our previous comments at the beginning of this section that the electrostatic potential is dominantly a flux function. (We have dropped the ordering parameter and the gyrophase term that does not affect the motion of the guiding center. We have also dropped the subscript F, as in this section we assume flux coordinates.)

The Lagrange tensor for these coordinates (Ψ,θ,φ,u) is found from the exterior derivative of the phase-space part of the differential action for this Lagrangian as in Eq. (2.25). The result is

$$\boldsymbol{\omega} = \begin{bmatrix} 0 & 1+uc_{\Psi\theta} & -(\iota + uc_{\varphi\Psi}) & -b_\Psi \\ -(1+uc_{\Psi\theta}) & 0 & uc_{\theta\varphi} & -b_\theta \\ \iota + uc_{\varphi\Psi} & -uc_{\theta\varphi} & 0 & -b_\varphi \\ b_\Psi & b_\theta & b_\varphi & 0 \end{bmatrix}, \tag{7.3}$$

where

$$c_{ji} \equiv \frac{\partial b_i}{\partial \psi^j} - \frac{\partial b_j}{\partial \psi^i}, \tag{7.4}$$

with $\psi^j = (\Psi,\theta,\varphi)$. The relation (2.43) implies that the conserved phase-space Jacobian in these variables is

$$\mathcal{J} = \mathcal{V}B_\parallel^* = b_\Psi uc_{\theta\varphi} + b_\theta(\iota + uc_{\varphi\Psi}) + b_\varphi(1 + uc_{\Psi\theta}), \tag{7.5}$$

where the spatial volume element $\mathcal{V}$ is given by Eqs. (4.7) and (4.8). The inverse of the Lagrange tensor, as noted in Sec. II, is the Poisson tensor,

$$\Pi = \frac{1}{\mathcal{J}}\begin{bmatrix} 0 & -b_\varphi & b_\theta & uc_{\theta\varphi} \\ b_\varphi & 0 & -b_\Psi & \iota + uc_{\varphi\Psi} \\ -b_\theta & b_\Psi & 0 & 1+uc_{\Psi\theta} \\ -uc_{\theta\varphi} & -(\iota + uc_{\varphi\Psi}) & -(1+uc_{\Psi\theta}) & 0 \end{bmatrix}. \tag{7.6}$$

The Poisson tensor acting on the gradient of the Hamiltonian gives the rate of change of the coordinates, according to Eq. (2.30). Thus, we obtain the equations of motion

$$\dot{\Psi} = \frac{1}{\mathcal{J}}\left(-b_\varphi\frac{\partial V_g}{\partial\theta} + b_\theta\frac{\partial V_g}{\partial\varphi} + u^2 c_{\theta\varphi}\right), \tag{7.7}$$

$$\dot{\theta} = \frac{1}{\mathcal{J}}\left(b_\varphi\frac{\partial V_g}{\partial\Psi} - b_\Psi\frac{\partial V_g}{\partial\varphi} + \iota u + u^2 c_{\varphi\Psi}\right), \tag{7.8}$$

$$\dot{\varphi} = \frac{1}{\mathcal{J}}\left(-b_\theta\frac{\partial V_g}{\partial\Psi} + b_\Psi\frac{\partial V_g}{\partial\theta} + u + u^2 c_{\Psi\theta}\right), \tag{7.9}$$

and

$$\dot{u} = -\frac{1}{\mathcal{J}}\left[uc_{\theta\varphi}\frac{\partial V_g}{\partial\Psi} + (\iota + uc_{\varphi\Psi})\frac{\partial V_g}{\partial\theta} + (1 + uc_{\Psi\theta})\frac{\partial V_g}{\partial\varphi}\right], \tag{7.10}$$

where

$$V_g = \mu B + \Phi \tag{7.11}$$

is the effective guiding-center potential.

B. Isodynamism

An equilibrium is isodynamic if the particles do not drift across flux surfaces, i.e., $\dot{\Psi}=0$. From the form of the covariant components of the magnetic field in Boozer coordinates, we can reduce the cross flux surface drift to the form

$$\dot{\Psi} = (\mu B + u^2)\left(B_\varphi\frac{\partial B}{\partial\theta} - B_\theta\frac{\partial B}{\partial\varphi}\right). \tag{7.12}$$

Thus, isodynamic equilibria satisfy

$$B_\varphi\frac{\partial B}{\partial\theta} - B_\theta\frac{\partial B}{\partial\varphi} = 0. \tag{7.13}$$

Hence, the magnetic field is constant along curves on a surface having parametric form

$$\theta = \theta_0 + B_\varphi\tau \tag{7.14}$$

and

$$\varphi = \varphi_0 + B_\theta\tau. \tag{7.15}$$

Typically, such as in a tokamak, both the net toroidal current $B_\theta(\Psi)$ and the net poloidal current $-B_\varphi(\Psi)$ flowing inside a flux surface Ψ vary continuously with Ψ. For values of Ψ such that the ratio B_θ/B_φ is irrational, a curve defined by Eqs. (7.14) and (7.15) covers a magnetic surface, and so the magnetic field is constant on a magnetic surface. By continuity one can then extend the constancy of the magnetic-field strength to surfaces for which B_θ/B_φ is rational. This is the case discussed early in the literature (Palumbo, 1968).

The constancy of magnetic-field strength implies that B_Ψ vanishes if the magnetic field corresponds to scalar pressure equilibrium. Expansion of Eq. (B37) in Fourier series in the angles shows that the amplitude of the harmonic $\exp(il\theta - in\varphi)$ of B_Ψ is nonzero only if either $l\iota - m$ is nonzero, or if the Jacobian $\mathcal{J}$ also has a nonzero amplitude for this harmonic. But the Jacobian $\mathcal{J}$ has nonzero amplitudes for only the (0,0) harmonic. Hence, in the typical case in which the surfaces on which the rotational transform is irrational are dense, B_Ψ is a function of Ψ alone. Moreover, for our modified Boozer coordinates, the average part of B_Ψ was shown to vanish. Thus, B_Ψ vanishes for isodynamical systems in modified Boozer coordinates.

For this generic case of nonvanishing and varying net toroidal and poloidal currents, the Lagrangian for an isodynamical system is independent of both the poloidal and toroidal modified Boozer angles. Hence, the conjugate momenta,

$$p_\theta \equiv \frac{e}{c}A_\theta + mub_\theta \tag{7.16}$$

and

$$p_\varphi \equiv \frac{e}{c}A_\varphi + mub_\varphi, \tag{7.17}$$

are both conserved. As these are both functions of only Ψ and u, Ψ and u are constants of motion. [Had we not chosen the (modified) Boozer coordinates, we would have had a hidden symmetry, as Ψ and u, and, therefore, p_θ and p_φ would still have been constants of motion, but this would not have been apparent from the Lagrangian, as through B_Ψ the Lagrangian would have had a dependence on the angles.]

Given that the case of a constant, finite, rational ratio of B_θ/B_φ seems improbable, the two remaining cases to discuss are those for which either B_θ or B_φ vanishes. For the case of no net poloidal current, $B_\varphi=0$, then B is a function of only the poloidal angle. This corresponds to field reversed configurations. In agreement with Bernardin *et al.* (1986), the magnetic field vanishes on axis. The lack of φ dependence implies that $p_\varphi = eA_\varphi(\Psi)/c$ is a constant of motion, making evident that Ψ is a constant of motion. In this case, without other considerations, the quantity B_Ψ need not vanish. A similar discussion applies to the case of no net toroidal current, except that as noted by Cary and Shasharina (1997b) closure of the magnetic axis requires some poloidal variation of B near the axis, and so this case cannot occur.

Finally, we note that open systems, for which the flux coordinates (Sec. IV.A.2) are appropriate, have been considered by Catto and Hazeltine (1981) and by Bernardin *et al.* (1986). Using a "long-thin" approximation, Catto and Hazeltine (1981) were able to construct mirror equilibria, while Bernardin *et al.* (1986) have shown that the magnetic axis of such systems must be straight.

C. Quasisymmetry

Quasisymmetry (Boozer, 1983) is the condition of having the magnitude of the magnetic-field strength depend on only some linear combination of the angles in Boozer coordinates. It includes quasipoloidal, in which B is independent of θ, quasitoroidal, in which B is independent of ϕ, and quasihelical, in which the magnetic-field strength depends on only a single linear combination,

$$\eta \equiv l_0\theta - n_0\varphi, \tag{7.18}$$

of the angles. In each of these cases there is an ignorable angle, so that the conjugate momentum is an invariant. Of course, toroidally symmetric systems, like the tokamak, are well known. (No example of a member of the isodynamic subset of these, which would have no net toroidal current, has come forth.) In contrast, no precise example of poloidally symmetric systems is available. This leaves only the possibility of quasihelical systems, in which the only angular dependence is on the variable (7.18), upon imposing the condition of no net toroidal current. (This latter condition can be relaxed, as one can obtain systems in which the bootstrap current provides net toroidal current, if the goal is simply not to have to drive current inductively.)

One might imagine that it is possible that the symmetry is manifest in a set of coordinates that are not Boozer coordinates. In the next section, we show that this is possible, although in an approximate sense, but to very high accuracy.

In Boozer coordinates for scalar pressure equilibrium, one need only demand that B be a function of the angle combination, and then it follows that B_Ψ also has this property. The argument follows from Eq. (B37). As before, expansion of this equation in Fourier series in the angles shows that the amplitude of the harmonic $\exp(il\theta - in\varphi)$ in B_Ψ is nonzero only either if $l\iota - m$ is nonzero or if the Jacobian $\mathcal{V}$ also has a nonzero amplitude for this harmonic. The Jacobian $\mathcal{V}$ has nonzero amplitudes only for harmonics satisfying

$$(l,n) = k(l_0,n_0), \tag{7.19}$$

as $\mathcal{V}$, related to the magnetic-field strength via Eq. (4.61), is a function of the angles only through η. Hence, in the typical case in which the surfaces on which the rotational transform is irrational are dense, B_Ψ also has nonzero amplitudes only for harmonics satisfying the relation (7.21), and so B_Ψ and the guiding-center Lagrangian are functions of only Ψ and η.

The invariant associated with this symmetry is found using the new variable set (Ψ, η, φ) for the guiding-center Lagrangian. With this replacement, Eq. (7.1) becomes

$$L_g = mub_\Psi\dot{\Psi} + \frac{1}{l}\left(\frac{e}{c}\Psi + mub_\theta\right)\dot{\eta} + \left[\frac{e}{c}\left(A_\varphi + \frac{n}{\ell}\Psi\right) + mu\left(b_\varphi + \frac{n}{\ell}b_\varphi\right)\right]\dot{\varphi} - h_g \tag{7.20}$$

with

$$h_g = \tfrac{1}{2}mu^2 + \mu B(\Psi,\eta) + e\Phi(\Psi). \tag{7.21}$$

(We restore units in the remainder of this section.) As this Lagrangian is a function of only (Ψ, η, u) (and independent of φ), the momentum,

$$P_s \equiv \frac{e}{c}A_s + mub_s, \tag{7.22}$$

conjugate to the ignorable coordinate φ is conserved, where

$$\frac{e}{c}A_s \equiv \frac{e}{c}\left(A_\varphi + \frac{n}{l}\Psi\right) \tag{7.23}$$

and

$$B_s \equiv B_\varphi + \frac{n}{l}B_\theta. \tag{7.24}$$

The subscript s denotes that these are the components associated with the symmetry variable. In the case $n=0$, this invariant reduces to the toroidal angular momentum, which is conserved in cases of axisymmetry, as discussed in Sec. III.F.

This invariant contains one term, eA_s/c, depending on only the flux variable and of one order higher in the guiding-center ordering that the other term, muB_s/B, which varies as the particle moves through space. Because the dominant term is a function of only the flux variable, the existence of this invariant implies that the flux variable is to lowest order an invariant, and, hence, the variation of the flux variable is small—first order in the guiding-center ordering.

Because the flux variable is, to lowest order, a constant of motion, a particle sees a variation of potential that is periodic, as it is a function of only the variable η. Conservation of the Hamiltonian (7.21) relates the parallel velocity to the magnetic-field value for two different points,

$$\tfrac{1}{2}mu_1^2 + \mu B_1 = \tfrac{1}{2}mu_2^2 + \mu B_2, \tag{7.25}$$

along a trajectory. The variation of the flux variable is related to the variation of the parallel velocity through Eq. (7.22). The constancy of the momentum P_s implies

$$\frac{e}{c}\frac{\partial A_s}{\partial \Psi}(\bar{\Psi})(\Psi_1 - \Psi_2) = mB_s\left(\frac{u_2}{B_2} - \frac{u_1}{B_1}\right), \tag{7.26}$$

where all quantities are evaluated at a flux variable value of $\bar{\Psi}$, which corresponds to some value of the flux variable on the trajectory. For maximum accuracy, Ψ is taken to be the mean value of the flux variable. As the flux variable scales as $\Psi \sim \frac{1}{2}Br^2$, Eq. (7.26) shows that the variation of the flux variable is small in the guiding-center ordering—it vanishes in large magnetic-field limit.

For passing particles, the variation of the flux variable is found by inserting the extreme values $u_1 = u_{\min}$, $B_1 = B_{\max}$, $u_2 = u_{\max}$, and $B_2 = B_{\min}$ into Eq. (7.26). The largest variation is found for the separatrix trajectory, where $u_1 = 0$, and so

$$u_{\max} = \sqrt{\frac{2\mu\Delta B}{m}}, \tag{7.27}$$

where

$$\Delta B \equiv B_{\max} - B_{\min}. \tag{7.28}$$

Inserting this into Eq. (7.26) gives the flux variable variation,

$$\Delta\Psi = \sqrt{2m\mu\Delta b}\left(\frac{e}{c}\frac{\partial A_s}{\partial \Psi}B_{\min}\right)^{-1}. \tag{7.29}$$

For the barely trapped particles just inside the separatrix, the variation of the flux variable is twice this value, as the particle moves in and out from the point where the parallel velocity vanishes by this amount.

Nührenberg and Zille (1988) were able to obtain numerical scalar pressure equilibria for which the amplitudes of the harmonics not being of the desired helicity were less than 2% of the value of the (0,0) harmonic. These results were obtained for an $(l=1, m=6)$ stellarator with rotational transform varying from 1.4 to 1.5 and of aspect ratio roughly 13. One might hope to obtain lower-aspect-ratio results. However, Garren and Boozer (1991) showed, by expansion near the magnetic axis, that it is possibility to satisfy the condition of guiding-center integrability only through second order in the inverse aspect ratio. Nevertheless, the results of Nührenberg and Zille (1988) have been used in the helical advanced stellarator (HELIAS) design.

D. Omnigeneity

Omnigeneous equilibria are those for which the bounce-averaged cross-flux-surface guiding-center drift vanishes. The guiding-center integrable systems just discussed have this property, as the flux variation is bounded. However, such systems exist only in the large-aspect-ratio limit. Hence, it is of interest to see whether systems with equally good transport properties but fewer restrictions are available.

In the discussion of such systems, it is useful to consider how the magnetic-field strength varies within the flux surface. Typical is the two-helicity model, often used in early discussions, in which the magnetic-field strength within a surface is of the form

$$B = \varepsilon_t(\Psi)\cos(\theta) + \varepsilon_h(\Psi)\cos(l\theta - n\varphi). \tag{7.30}$$

As the field line wraps on the surface, it encounters local maxima, which then form a closed curve on the surface. Similarly, the local minima form a closed curve.

These considerations show that, in general, there can exist transitioning particles, particles that change state from trapped to passing. For example, a particle could be trapped between two local maxima on one field line, but then its drift motion could carry it to a new field line on which the magnetic maxima are smaller; it would then change to a locally passing particle. This sort of motion is described by separatrix crossing theory (Cary *et al.*, 1986; Cary and Skodje, 1988), which shows such motion to be chaotic. Cary and Shasharina (1997a, 1997b, 1997c) noted that good transport qualities would require the elimination of such particles. This leads to two conditions. The first is that the local maxima must all have the same value of magnetic-field strength, while the second is that the bounce action,

$$J = \frac{1}{2\pi}\oint ds u = \frac{2}{\pi}(\iota B_\theta + B_\varphi)\int_{\min(B)}^{\max(B)} d\varphi \frac{u}{B}, \tag{7.31}$$

for marginally trapped particles must be a constant for the field lines within a surface.[3] Cary and Shasharina (1997a, 1997b, 1997c) then proved that this implied that the curve of maxima had to be straight in Boozer coordinates.

Cary and Shasharina went on to examine the consequences of requiring all trapped particles to be omnigeneous. They showed that this led to the requirement that all magnetic minima have the same value [also considered by Mynick *et al.* (1982)] and to the condition of isometry. Cary and Shasharina further showed that the requirement that all trapped particles be omnigeneous implied that all passing particles were omnigeneous as well. Isometry, which had been introduced by Skovoroda and Shafranov (1995), is the condition that the length along a field line between two contours of the same value of magnetic-field strength be constant on a magnetic surface. Skovoroda and Shafranov noted that isometric systems had omnigeneous trajectories, while Cary and Shasharina showed that, in fact, omnigeneity implied isometry.

Thus, in summary, the three Cary-Shasharina *integral conditions* for omnigeneity are as follows: (i) The magnetic maxima must have the same value. (ii) The curve of magnetic maxima must be straight in Boozer coordinates. (iii) The magnetic field must be isometric. (Isometry then implies that the magnetic minima all have the same value on a surface.)

Cary and Shasharina then proved an additional result: (iv) If the magnetic-field strength is analytic in the flux variables, then the contours are in fact straight in the Boozer angles. This then proves that the only choice for complete, analytic omnigeneity is the existence of quasisymmetry, the magnetic-field strength being a function of a single linear combination of the flux angles in Boozer coordinates.

One might consider this argument conclusive. However, the difference between analytic and nonanalytic functions can be very small. One can construct a nonanalytic magnetic-field-strength function that satisfies the Cary-Shasharina integral conditions and then truncates its Fourier series at some high mode number. Because only a small term is dropped, the function is still far from having a symmetry in Boozer angles, yet it is now analytic. Because the function remains close to the original, nonanalytic, exactly omnigeneous form, the trajectories remain close to being omnigeneous. Such systems are said to be *approximately omnigeneous*. They can be arbitrarily close to omnigeneous yet very far from quasisymmetric.

[3]Mynick *et al.* (1982) had previously looked at the consequences of only the first condition of having the magnetic maxima constant within a surface.

TABLE III. Classification of toroidal confinement systems from minimal to maximal deviation of trajectory from a flux surface.

Isodynamism
Quasisymmetry
Approximate omnigeneity
Specific omnigeneity

E. Specific omnigeneity

While the developments of the previous section indicate that one can obtain equilibria far from quasisymmetric while retaining near-full omnigeneity, one could imagine relaxing this even further, such that one requires only specific trajectories to be omnigeneous. Mynick (1983) considered two cases, one in which the deeply trapped particles were omnigeneous and one in which particles were omnigeneous at the local maximum of the magnetic field. The former condition implies that the magnetic minima on a flux surface all have the same value, while the latter condition implies that the magnetic maxima on a flux surface all have the same value. Mynick found the former case to have better transport properties. However, having the particles be omnigeneous at the magnetic maximum might not be expected to help much, as such trajectories are unstable. It is reasonable to expect that improvement comes about only when particles on the separatrix trajectory are omnigeneous. This implies the additional condition that the action enclosed by the separatrix be constant on a flux surface. For this case, there are no trajectories that transition between the locally passing and locally trapped states. The consequences of imposing this condition remain to be explored.

F. Hierarchy of improved confinement systems

To summarize this section, toroidally nested magnetic-field configurations can have varying degrees of deviation of guiding-center trajectories from the flux surfaces. For the isodynamic systems, the deviation vanishes. For quasisymmetry (which include toroidal symmetry), the trajectories drift off the flux surface but then return cyclically. The same is true for omnigeneous systems. These latter are known to reduce to quasisymmetric systems when the fields are analytic. However, very nearly omnigeneous systems are very far from quasisymmetric. This opens up a new avenue in the search for toroidal confinement systems with good orbit properties. Greater deviations still of the particle trajectories occur when one demands that only a few specific trajectories be omnigeneous. A promising condition is that there exist no transitioning particles, which follows from the constancy of the bounce separatrix action on flux surfaces. This hierarchy is summarized in Table III. As one moves up the hierarchy, there is more symmetry, and the particles deviate less from flux surfaces, but such systems are difficult to obtain or unobtainable altogether. For small-aspect-ratio systems, it is likely that one can obtain at best near omnigeneity—having at best a large class of particles being omnigeneous.

VIII. SUMMARY AND FUTURE DIRECTIONS

Guiding-center theory has been a powerful theoretical tool for understanding strongly magnetized plasmas. In this review, we have summarized the development and use of the Hamiltonian theory of guiding-center motion. Hamiltonian theory brings value through applicability of Liouville's theorem, which prevents the existence of attractors, and Noether's theorem, which allows one to prove the existence of invariants from symmetries.

Our approach has been through the phase-space Lagrangian formulation of mechanics. In this formulation, there is no restriction on the transformations that one can use to find coordinates in which the motion is simpler, though this comes at the cost of the dynamics being determined by multiple phase-space functions rather than just one, the Hamiltonian, as is the case for canonical coordinates. While more general, the phase-space approach retains the Liouville and Noether theorems.

We then applied noncanonical, perturbative coordinate transformations for systems with strong magnetic fields. Imposing the requirement that the gyrophase be absent from the new Lagrangian leads to a phase-space Lagrangian with a Noether theorem, corresponding to which is the magnetic moment invariant. Further discussion showed how to relate guiding-center currents to physical currents. Finally, we showed how to reduce the guiding-center Lagrangian to the Lagrangian that describes magnetic-field line flow.

To relate this work to the more familiar canonical-variable Hamiltonian mechanics, we showed multiple ways to obtain canonical coordinates. Further simplifications were shown for the case of toroidal magnetic fields with nested flux surfaces. In this case, the canonical coordinates are closely related to the flux coordinates.

In Secs. V and VI, we discussed various refinements of the standard Hamiltonian guiding-center theory by introducing higher-order adiabatic invariants and/or relativistic effects. These refinements greatly extend the applicability of the guiding-center equations.

In Sec. VII, we reviewed the classification of toroidal magnetic fields with respect to the off-surface drifts. We further noted that there is a hierarchy, with isodynamic systems having no off-surface drifts, quasisymmetric systems having an explicit symmetry in the Lagrangian, omnigeneous systems having no bounce-averaged off-surface drifts, and specifically omnigeneous systems having specific classes of particles with no bounce-averaged off-surface drifts. We noted that bounce-averaged omnigeneity and analyticity of the fields implies quasisymmetry, but that one could also have approximately omnigeneous systems with analytic fields that are very far from quasisymmetric systems.

We end our summary by discussing extensions of guiding-center Hamiltonian theory that have found applications in the development of the theoretical foundations of turbulent transport in strongly magnetized plasmas. One important application involves the development of low-frequency nonlinear gyrokinetic theory, which was recently reviewed by Brizard and Hahm (2007). Low-frequency gyrokinetic theory was initially motivated by the need to describe complex plasma dynamics over time scales that are long compared to the short gyromotion time scale. Thus, gyrokinetic theory was constructed as a generalization of guiding-center theory (Northrop, 1963; Littlejohn, 1983). For example, Taylor (1967) showed that, while the guiding-center magnetic-moment invariant (denoted μ) can be destroyed by low-frequency, short-perpendicular-wavelength electrostatic fluctuations, a new magnetic-moment invariant (denoted $\bar{\mu}$) can be constructed as an asymptotic expansion in powers of the amplitude (denoted ϵ) of the perturbation field, i.e., $\bar{\mu}=\bar{\mu}_0+\epsilon\bar{\mu}_1+\cdots$, where $\bar{\mu}_0\equiv\mu$ and $\bar{\mu}_1\equiv-\Omega^{-1}\int\dot{\mu}d\bar{\zeta}$ as follows from the general formalism discussed in Sec. V.B. This early result indicated that gyrokinetic theory could be built upon an additional transformation beyond the guiding-center phase-space coordinates, thereby constructing new *gyrocenter* phase-space coordinates, which describe gyroangle-averaged perturbed guiding-center dynamics.

The linear electrostatic and electromagnetic gyrokinetic equations have been successfully applied to the low-frequency stability analysis of many magnetized plasmas in various geometries. The nonlinear electrostatic and electromagnetic gyrokinetic equations, on the other hand, have been used to study the transport properties of turbulent magnetized plasmas; see Dimits *et al.* (2000); Batchelor *et al.* (2007); Brizard and Hahm (2007) for details and references.

ACKNOWLEDGMENTS

The authors gratefully acknowledge many discussions over the years with colleagues in the plasma physics and dynamics communities, including A. N. Kaufman, R. G. Littlejohn, J. D. Meiss, P. J. Morrison, H. E. Mynick, and R. B. White. The impetus for this work was initially supplied by R. D. Hazeltine. This work was supported by the Office of Fusion Energy Science of the U.S. Department of Energy through Grant No. DE-FG02-04ER54735.

APPENDIX A: NORTHROP LAGRANGIAN FORMULATION

The guiding-center equations of motion presented by Northrop (1963) can be derived from a guiding-center Lagrangian different from Eq. (3.2), with a guiding-center Hamiltonian different from Eq. (3.3). These guiding-center expressions can be constructed by following a procedure similar to Sec. III.D, where the choice of the gyroangle-independent displacement vector $\bar{\boldsymbol{\rho}}$ that leads to Eqs. (A1) and (A2) is $\bar{\boldsymbol{\rho}}\equiv\mathbf{0}$ in Eq. (3.45). A similar set of guiding-center equations for time-independent electric and magnetic fields was derived by Brizard (1995).

The Northrop guiding-center (Ngc) Lagrangian is

$$\mathcal{L}_{\rm Ngc} = \left[\frac{e}{c}\mathbf{A} + m(u\hat{\mathsf{b}} + \mathbf{v}_E)\right]\cdot\dot{\mathbf{X}} + J\dot{\zeta} - H_{\rm Ngc}, \tag{A1}$$

where the Northrop guiding-center Hamiltonian is

$$H_{\rm Ngc} = \frac{m}{2}|u\hat{\mathsf{b}} + \mathbf{v}_E|^2 + \mu B + e\Phi. \tag{A2}$$

1. Northrop Hamiltonian guiding-center dynamics

The Euler-Lagrange equations obtained from the Lagrangian (A1) are $u \equiv \hat{\mathsf{b}}\cdot\dot{\mathbf{X}}$ and

$$m\dot{u}\hat{\mathsf{b}} = e\mathbf{E}^{**} + \frac{e}{c}\dot{\mathbf{X}}\times\mathbf{B}^{**}, \tag{A3}$$

where the new effective fields $\mathbf{E}^{**} \equiv -\nabla\Phi^{**} - c^{-1}\partial\mathbf{A}^{**}/\partial t$ and $\mathbf{B}^{**} \equiv \nabla\times\mathbf{A}^{**}$ are expressed in terms of the effective potentials

$$e\Phi^{**} \equiv e\Phi + \mu B + (m/2)|\mathbf{v}_E|^2,$$

$$\mathbf{A}^{**} \equiv \mathbf{A} + (mc/e)(u\hat{\mathsf{b}} + \mathbf{v}_E). \tag{A4}$$

Note that, in a static magnetic field (where $\mathbf{E} = -\nabla\Phi$), the effective potential Φ^{**} is related to the effective potential Φ^* defined in Eq. (3.11) as follows:

$$\begin{aligned} e\Phi^{**}(\mathbf{X}+\bar{\boldsymbol{\rho}}) &= e\Phi(\mathbf{X}+\bar{\boldsymbol{\rho}}) + \mu B + \frac{m}{2}|\mathbf{v}_E|^2 \\ &= e(\Phi + \bar{\boldsymbol{\rho}}\cdot\nabla\Phi) + \mu B + \frac{m}{2}|\mathbf{v}_E|^2 \\ &= e\Phi + \mu B - \frac{m}{2}|\mathbf{v}_E|^2 \equiv e\Phi^*. \end{aligned}$$

The guiding-center equations of motion for $\dot{\mathbf{X}}$ and $\dot{u}$ are, respectively, expressed as

$$\dot{\mathbf{X}} = u\frac{\mathbf{B}^{**}}{B_\parallel^{**}} + \mathbf{E}^{**}\times\frac{c\hat{\mathsf{b}}}{B_\parallel^{**}} \tag{A5}$$

and

$$\dot{u} = \frac{e\mathbf{B}^{**}}{mB_\parallel^{**}}\cdot\mathbf{E}^{**}, \tag{A6}$$

where

$$B_\parallel^{**} \equiv \hat{\mathsf{b}}\cdot\mathbf{B}^{**} = B + \frac{mc}{e}\hat{\mathsf{b}}\cdot\nabla\times(u\hat{\mathsf{b}} + \mathbf{v}_E).$$

The presence of the $E\times B$ velocity in $\mathbf{A}^{**}$ implies that the polarization drift velocity now appears explicitly in the guiding-center velocity (A5).

2. Guiding-center conservation laws

The Northrop-Lagrangian guiding-center equations of motion (A5) and (A6) possess an important Hamiltonian property not shared by the original non-Hamiltonian equations (A7) and (A8) derived by Northrop (1963). This property involves the conservation of guiding-center phase-space volume, i.e., the guiding-center Liouville theorem

$$\frac{\partial B_\parallel^{**}}{\partial t} + \nabla\cdot(B_\parallel^{**}\dot{\mathbf{X}}) + \frac{\partial}{\partial u}(B_\parallel^{**}\dot{u}) = 0. \tag{A7}$$

Here, using Faraday's law ($\partial\mathbf{B}^{**}/\partial t = -c\nabla\times\mathbf{E}^{**}$), we find

$$\frac{\partial B_\parallel^{**}}{\partial t} = \hat{\mathsf{b}}\cdot\frac{\partial\mathbf{B}^{**}}{\partial t} + \mathbf{B}^{**}\cdot\frac{\partial\hat{\mathsf{b}}}{\partial t} = -c\hat{\mathsf{b}}\cdot\nabla\times\mathbf{E}^{**} + \mathbf{B}^{**}\cdot\frac{\partial\hat{\mathsf{b}}}{\partial t}. \tag{A8}$$

Next, we find

$$\begin{aligned} \nabla\cdot(B_\parallel^{**}\dot{\mathbf{X}}) &= u\,\nabla\cdot\mathbf{B}^{**} + \nabla\cdot(\mathbf{E}^{**}\times c\hat{\mathsf{b}}) \\ &= c(\hat{\mathsf{b}}\cdot\nabla\times\mathbf{E}^{**} - \mathbf{E}^{**}\cdot\nabla\times\hat{\mathsf{b}}), \end{aligned} \tag{A9}$$

where we used $\nabla\cdot\mathbf{B}^{**} \equiv 0$. Last, we find

$$\begin{aligned} \frac{\partial}{\partial u}(B_\parallel^{**}\dot{u}) &= \frac{e}{m}\left(\frac{\partial\mathbf{B}^{**}}{\partial u}\cdot\mathbf{E}^{**} + \mathbf{B}^{**}\cdot\frac{\partial\mathbf{E}^{**}}{\partial u}\right) \\ &= c\mathbf{E}^{**}\cdot\nabla\times\hat{\mathsf{b}} - \frac{\partial\hat{\mathsf{b}}}{\partial t}\cdot\mathbf{B}^{**}. \end{aligned} \tag{A10}$$

By combining Eqs. (A8)–(A10), we easily recover Eq. (A7). The conservation of phase-space volume by the Northrop-Lagrangian guiding-center equations of motion (A5) and (A6) plays a fundamental role in their numerical integration over long-time scales. The guiding-center equations (3.12) and (3.13) presented in Sec. III obey a similar phase-space volume conservation law, with $(\mathbf{E}^*, \mathbf{B}^*)$ replacing $(\mathbf{E}^{**}, \mathbf{B}^{**})$ in Eqs. (A8)–(A10).

The guiding-center equations of motion (A5) and (A6) satisfy other conservation laws when space-time symmetries exist. First, the time derivative of the Northrop guiding-center Hamiltonian (A2) is expressed as

$$\frac{dH_{\rm Ngc}}{dt} = e\frac{\partial\Phi^{**}}{\partial t} - \frac{e}{c}\frac{\partial\mathbf{A}^{**}}{\partial t}\cdot\dot{\mathbf{X}}, \tag{A11}$$

and, hence, the total guiding-center energy $E = H_{\rm Ngc}$ is conserved in the case of time-independent fields. Second, the time derivative of the total guiding-center canonical momentum $\mathbf{P} \equiv (e/c)\mathbf{A}^{**}$ is expressed as

$$\frac{\partial\mathbf{P}}{\partial t} = -e\,\nabla\,\Phi^{**} + \frac{e}{c}\,\nabla\,\mathbf{A}^{**}\cdot\dot{\mathbf{X}}, \tag{A12}$$

and, hence, the canonical momentum component $P_\alpha \equiv \mathbf{P}\cdot\partial\mathbf{X}/\partial\psi^\alpha$ is a constant of the guiding-center motion if the magnetic variable ψ^α is an ignorable coordinate, which follows from Noether's theorem.

3. Original Northrop equations

To the same order kept by Northrop (1963), the guiding-center equations of motion (A5) and (A6) become

$$\dot{\mathbf{X}}_N = \mathbf{U} + \frac{c\hat{\mathsf{b}}}{eB} \times \left(\mu \nabla B + m \frac{d_0 \mathbf{U}}{dt} \right), \tag{A13}$$

$$\dot{u}_N = \frac{\hat{\mathsf{b}}}{m} \cdot (e\mathbf{E} - \mu \nabla B) + \mathbf{v}_E \cdot \frac{d_0 \hat{\mathsf{b}}}{dt}, \tag{A14}$$

where the total time derivative is defined as $d_0/dt \equiv \partial/\partial t + \mathbf{U}\cdot\nabla$ to lowest order with $\mathbf{U} \equiv u\hat{\mathsf{b}} + \mathbf{v}_E$.

The original Northrop equations (A13) and (A14) satisfy the phase-space volume equation

$$\frac{\partial B}{\partial t} + \nabla \cdot (B\dot{\mathbf{X}}_N) + \frac{\partial}{\partial u}(B\dot{u}_N) = \nabla \cdot [B(\dot{\mathbf{X}}_N - \mathbf{U})] - cE_\parallel (\hat{\mathsf{b}} \cdot \nabla \times \hat{\mathsf{b}}). \tag{A15}$$

Hence, the original Northrop equations do not conserve phase-space volume since $\dot{\mathbf{X}}_N \neq \mathbf{U}$ and $\hat{\mathsf{b}}\cdot\nabla\times\hat{\mathsf{b}} \neq 0$ (in general). Last, the original Northrop equations of motion (A13) and (A14) satisfy the energy equation

$$\frac{d_N E}{dt} = e\frac{\partial \Phi}{\partial t} - \frac{e}{c}\frac{\partial \mathbf{A}}{\partial t} \cdot \dot{\mathbf{X}}_N + \mu \frac{\partial B}{\partial t} - \frac{\mu \hat{\mathsf{b}}}{\Omega} \times \nabla B \cdot \left(u \frac{d\hat{\mathsf{b}}}{dt} + \frac{d\mathbf{v}_E}{dt} \right), \tag{A16}$$

which does not vanish even for time-independent fields. While energy nonconservation appears at a higher order (i.e., at ϵ^2) than kept in the energy itself, its explicit nonconservation for time-independent fields presents difficulties when integration over long time scales is contemplated, which may result in unphysical results.

APPENDIX B: OTHER COORDINATE SYSTEMS FOR TOROIDAL MAGNETIC FIELDS WITH NESTED FLUX SURFACES

Section IV discussed canonical guiding-center theory starting from flux coordinates. Reviewed were two methods for obtaining flux coordinates, the second of which mixed the parallel velocity with the physical coordinates. In this appendix, we note that there are special flux coordinates for toroidal magnetic fields having nested flux surfaces, as occurs in MHD equilibria (Kruskal and Kulsrud, 1958) or can be obtained for vacuum fields by the Cary-Hanson technique (Cary, 1982, 1984a, 1984b; Hanson and Cary, 1984; Cary and Hanson, 1986). These are obtained by imposed additional restrictions allowed by the freedom of transformations within flux coordinates. This analysis shows that previously introduced canonical guiding-center coordinates (White and Chance, 1984) are special cases of what we found in Sec. IV.A.4.

We begin by reviewing the special toroidal magnetic coordinates. We start with the Hamada coordinates, in which the Jacobian is unity. We then discuss Boozer coordinates, in which the covariant angular components of the magnetic field are constant on flux surfaces.

1. Hamada coordinates

When the magnetic field is one of zero-flow scalar pressure MHD equilibrium

$$\mathbf{J} \times \mathbf{B} = c \nabla P, \tag{B1}$$

the current lines as well lie on magnetic surfaces, and so in the Clebsch representation only the same two components are nonzero as for the magnetic field. In this case, it is natural to seek coordinates such that in its Clebsch representation the current

$$\mathbf{J} = J_{\Psi\theta_h}(\Psi) \nabla \Psi \times \nabla \theta_h + J_{\varphi_h \Psi}(\Psi) \nabla \varphi_h \times \nabla \Psi \tag{B2}$$

also has components constant on flux surfaces as were found for the magnetic field. This defines the Hamada (1959) coordinates, denoted by the subscript h. In this section, we show how such coordinates can be obtained. Such coordinates have been known for a longer time than the Boozer coordinates. Discussing Hamada coordinates here allows for a comparison with the Boozer coordinates.

In this case of force-free equilibria ($\nabla P = 0$), flux coordinates already have this property, provided the rotational transform is irrational. The force-free condition implies

$$\mathbf{J} = \lambda \mathbf{B}, \tag{B3}$$

and the vanishing of the divergence of the current then implies ($\mathcal{V}^{-1} = \nabla\Psi \times \nabla\theta_F \cdot \nabla\varphi_F$)

$$\mathbf{B} \cdot \nabla \lambda = 0 = \frac{1}{\mathcal{V}} \left(\frac{\partial \lambda}{\partial \varphi_F} + \iota \frac{\partial \lambda}{\partial \theta_F} \right), \tag{B4}$$

from which it follows that λ is constant on a flux surface and so too, according to Eq. (B3), are the Clebsch representation components of the current for surfaces with irrational values of rotational transform. If the rotational transform varies from surface to surface, then continuity implies that the current has this property on all surfaces.

For the cases of nonzero pressure gradient, we introduce a transformation defined by

$$\theta_h = \theta_F + \iota g_h(\Psi, \theta_F, \varphi_F) \tag{B5}$$

and

$$\varphi_h = \varphi_F + g_h(\Psi, \theta_F, \varphi_F). \tag{B6}$$

As noted in Sec. IV.B.1, after any transformation of this type, one still has flux coordinates. Inserting this transformation into the representation (B2) gives

$$J_{\Psi\theta_h} + (\iota J_{\Psi\theta_h} - J_{\varphi_h\Psi}) \frac{\partial g_h}{\partial \theta_F} = J_{\Psi\theta_F} \tag{B7}$$

and

$$J_{\varphi_h\Psi} - (\iota J_{\Psi\theta_h} - J_{\varphi_h\Psi})\frac{\partial g_h}{\partial \varphi_F} = J_{\varphi_F\Psi}. \tag{B8}$$

Provided the quantity in parentheses does not vanish, these equations can be solved as were Eqs. (4.46) and (4.47). Solvability, as before, is guaranteed by the vanishing of the divergence of the current. Averaging over the flux coordinates gives

$$J_{\Psi\theta_h} = \bar{J}_{\Psi\theta_F} \tag{B9}$$

and

$$J_{\varphi_h\Psi} = \bar{J}_{\varphi_F\Psi}. \tag{B10}$$

Hence, all that remains is to determine the quantity in parentheses. This follows from writing out Eq. (B1) in flux variables. We find

$$\mathbf{J}\times\mathbf{B} = \mathcal{V}(\iota J_{\Psi\theta_h} - J_{\varphi_h\Psi})\,\nabla\Psi = \frac{dP}{d\Psi}\nabla\Psi. \tag{B11}$$

From this equation we find first that the quantity in parentheses does not vanish provided the pressure gradient does not vanish. Hence, Hamada coordinates always exist for MHD equilibria with nonzero pressure gradient. Second we see that the Jacobian in these coordinates is constant on flux surfaces, i.e., it is a function of the surface label Ψ only.

In the literature one can find Hamada coordinates defined by the property that the Jacobian is constant on flux surfaces. In this case, one cannot prove that Hamada coordinates exist for vacuum fields or for force-free equilibria, as then Eq. (B11) is identically satisfied regardless of the Jacobian. If instead, as here, Hamada coordinates are defined as those in which the Clebsch coordinates of the current are constant, then they do exist for vacuum and force-free fields, but they are not unique, as a transformation from any one set of flux variables to another does not affect this.

In Hamada coordinates, the Clebsch representation for the current can be integrated once to obtain the covariant representation of the magnetic field, up to the gradient of a scalar. We obtain

$$\mathbf{B} = \bar{B}_{\theta_h}\nabla\theta_h + \bar{B}_{\varphi_h}\nabla\varphi_h + \nabla\Phi_M, \tag{B12}$$

where

$$\frac{\partial \bar{B}_{\theta_h}}{\partial\Psi} = J_{\Psi\theta_h} \tag{B13}$$

and

$$\frac{\partial \bar{B}_{\varphi_h}}{\partial\Psi} = -J_{\varphi_h\Psi}. \tag{B14}$$

Analogous to the magnetic flux discussed earlier, B_{θ_h} gives the toroidal current flux between surfaces of flux variable values Ψ_1 and Ψ_2, while B_{φ_h}, gives minus the poloidal current flux between surfaces of flux variable values Ψ_1 and Ψ_2.

In fact, Eqs. (B13) and (B14) are uniquely specified by requiring the potential Ψ to be a single-valued function in the toroidal domain, which implies that all loop integrals of the form

$$\oint d\mathbf{x}\cdot\nabla\Psi \tag{B15}$$

vanish for loops regardless of whether they encircle the hole of the torus. [This specification, which we discussed presently, ensures that the first two terms on the right-hand side of Eq. (B12) contain the average angular covariant components. Hence, we have used the overbars on these terms.] As the integral of the magnetic field around a θ_h loop at constant Ψ and φ_h is the toroidal plasma current through the torus, and this must vanish at $\Psi=0$, we have

$$\bar{B}_{\theta_h}(\Psi=0) = 0. \tag{B16}$$

Similarly, the φ_h loop integral of the magnetic field at constant Ψ and θ_h gives $4\pi/c$ times the nonplasma or coil current I passing through the hole in the torus. Hence,

$$\bar{B}_{\varphi_h}(\Psi_{\rm edge}) = \frac{2I}{c}. \tag{B17}$$

Because Hamada coordinates are flux coordinates, we can apply the theory of Sec. IV.A.4 to obtain the canonical coordinates. As before, the canonical poloidal angle differs from the flux poloidal angle by a term proportional to the parallel velocity.

Angular dependence within the guiding-center Lagrangian is important for magnetic confinement, as derivatives with respect to the angles lead to nonzero $\dot{\Psi}$ and off-flux-surface dynamics, which leads to increased transport, as discussed in Sec. VII. For Hamada coordinates for scalar pressure equilibria, both the magnetic strength B and the magnetic scalar potential Φ_M are potentially functions of the angles. Consequently, in Hamada coordinates for a symmetry, such as a dependence on only a single linear combination of the angles, to exist, it must be present in both of these functions.

2. Boozer coordinates

Boozer coordinates $(\Psi,\theta_b,\varphi_b)$ are defined such that the angular covariant components of the magnetic field,

$$\mathbf{B} = B_{\theta_b}(\Psi)\,\nabla\theta_b + B_{\varphi_b}(\Psi)\,\nabla\varphi_b + B_\Psi(\Psi,\theta_b,\varphi_b)\,\nabla\Psi, \tag{B18}$$

are constant on a flux surface, while the remaining covariant component may have arbitrary dependence. This representation looks similar to the representation (B12), but it is significantly different. For Boozer coordinates, the relation

$$\mathcal{V} = \frac{1}{B^2} \tag{B19}$$

between the Jacobian $\mathcal{V}$ and the magnetic-field strength B can be found by taking the dot product of the Clebsch and covariant representations of the magnetic field. This relation shows that, on a flux surface, the variations of the Jacobian and inverse of the square of the magnetic field are related by an overall factor.

We first consider vacuum fields, for which we have the Hamada representation (B12). From the relations, (B12), (B16), and (B17), the covariant representation of the magnetic field has the form

$$\mathbf{B} = \left(\frac{2I}{c}\right) \nabla \varphi_h + \nabla \Phi_M, \tag{B20}$$

where the potential Φ_M is single valued, and the first term ensures that the loop integral of the magnetic field around the toroidal direction gives the total current through the center of the torus. Hence, the new toroidal angle

$$\varphi_b = \varphi_h + \frac{c}{2I}\Phi_M \tag{B21}$$

puts the covariant representation of the magnetic field in the correct form,

$$\mathbf{B} = \left(\frac{2I}{c}\right) \nabla \varphi_b. \tag{B22}$$

To ensure that our new variables are flux surfaces, the differences between the new and old poloidal and new and old toroidal coordinates must satisfy Eqs. (4.53) and (4.54). Hence, the relation between the new and old poloidal coordinates is

$$\theta_b = \theta_F + \frac{\iota c}{2I}\Phi_M. \tag{B23}$$

For vacuum magnetic fields, the factor in Eq. (B19) relating the Jacobian and the magnetic-field strength is a constant.

To analyze MHD equilibria, we introduce a transformation of the form of Eqs. (4.53) and (4.54),

$$\theta_h = \theta_b + \iota f_{hb}(\Psi, \theta_b, \varphi_b) \tag{B24}$$

and

$$\varphi_h = \varphi_b + f_{hb}(\Psi, \theta_b, \varphi_b), \tag{B25}$$

that ensures that both sets of coordinates are flux coordinates. Inserting this transformation into the covariant form (B12) and comparing with Eq. (B18) shows that the transformation function must satisfy the following two equations:

$$B_{\theta_b} = \bar{B}_{\theta_h} + \bar{B}_{\theta_h} \iota \frac{\partial f_{hb}}{\partial \theta_b} + \bar{B}_{\varphi_h} \frac{\partial f_{hb}}{\partial \theta_b} + \frac{\partial \Phi_M}{\partial \theta_b} \tag{B26}$$

and

$$B_{\varphi_b} = \bar{B}_{\varphi_h} + \bar{B}_{\theta_h} \iota \frac{\partial f_{hb}}{\partial \varphi_b} + \bar{B}_{\varphi_h} \frac{\partial f_{hb}}{\partial \varphi_b} + \frac{\partial \Phi_M}{\partial \varphi_b}. \tag{B27}$$

Thus, any solution of the form

$$f_{hb} = -\frac{\Phi_M}{\bar{B}_{\theta_h}\iota + \bar{B}_{\varphi_h}} + \bar{f}_{hb}(\Psi) \tag{B28}$$

guarantees that the covariant angular components

$$B_{\theta_b} = \bar{B}_{\theta_h} \tag{B29}$$

and

$$B_{\varphi_b} = \bar{B}_{\varphi_h} \tag{B30}$$

of the magnetic field in Boozer coordinates are functions of only the flux variable Ψ.

One additional convenient condition can be placed on these coordinates, namely, that the covariant component B_φ have vanishing flux-surface average. This condition specifies the function $\bar{f}_{hb}(\Psi)$. The transformation of Eqs. (B24) and (B25) applied to the magnetic field (B12) gives

$$B_\Psi = (\iota B_{\theta_b} + B_{\varphi_b})\frac{\partial \bar{f}_{hb}}{\partial \Psi} + B_{\theta_b}\frac{\partial \iota}{\partial \Psi}\bar{f}_{hb} + \Phi_M\left(\frac{\iota B'_{\theta_b} + B'_{\varphi_b}}{\iota B_{\theta_b} + B_{\varphi_b}}\right). \tag{B31}$$

Hence, the surface average value of the covariant Ψ component vanishes provided one chooses

$$(\iota B_{\theta_b} + B_{\varphi_b})\frac{\partial \bar{f}_{hb}}{\partial \Psi} + B_{\theta_b}\frac{\partial \iota}{\partial \Psi}\bar{f}_{hb} = -\langle \Phi_M \rangle \left(\frac{\iota B'_{\theta_b} + B'_{\varphi_b}}{\iota B_{\theta_b} + B_{\varphi_b}}\right). \tag{B32}$$

These modified Boozer coordinates will be useful in our discussion of isodynamism, where we show that B_Ψ vanishes for these coordinates.

Canonical coordinates that apply here are exactly like before, as Boozer coordinates are flux coordinates. The canonical toroidal angle is simply the Boozer toroidal angle, and its conjugate momentum is

$$p_\varphi = \frac{e}{c}A_\varphi + mub_{\varphi_b}. \tag{B33}$$

The canonical poloidal angle is

$$\theta_c \equiv \theta_b - \frac{u}{\Omega}B_\Psi, \tag{B34}$$

and its canonical momentum is

$$p_\theta = \frac{e}{c}\Psi + mub_{\varphi_b}. \tag{B35}$$

The guiding-center Lagrangian is Eq. (4.64), exactly as before, but with these new variables. These canonical coordinates were introduced by White and Chance (1984) and are accurate to through first order in the guiding-center equations. Boozer (1984) proposed using the regular poloidal angle without the correction of Eq. (B34). This is accurate when the curvature drift arising from B_Ψ is small. All of these coordinates are special cases of the canonical coordinates introduced in Sec. IV.B.2, as there it was shown that for any flux coordinates one can find the associated canonical guiding-center variables.

For vacuum fields, many of the complications disappear. In this case only the φ covariant component is nonzero. Hence, the canonical poloidal angle is the usual poloidal angle, its conjugate is Ψ up to a factor, and Eq. (B33) is easily solved to obtain the parallel velocity. Furthermore, there is only one angle-dependent quantity B in the guiding-center Lagrangian. The symmetries of this quantity as a function of the flux variables can imply invariants. For example, the angular dependence on only some linear combination of angles, as in the case of a single helicity, guarantees the invariance of a conserved momentum that is a linear combination of the two canonical momenta.

Without further analysis, it appears that for Boozer coordinates there are two separate functions B and B_Ψ in the Lagrangian having angular dependence. However, for magnetic fields of scalar pressure equilibria, one can show that these two quantities are related. To find this relationship, we first calculate the current from Eq. (B18),

$$\frac{4\pi}{c}\mathbf{J} = \nabla\times\mathbf{B} = \left(\frac{\partial B_\Psi}{\partial\theta} - \frac{\partial B_\theta}{\partial\Psi}\right)\nabla\theta\times\nabla\Psi + \left(\frac{\partial B_\Psi}{\partial\varphi} - \frac{\partial B_\varphi}{\partial\Psi}\right)\nabla\varphi\times\nabla\Psi. \tag{B36}$$

From this, the representation (4.40), the scalar pressure condition (B2), and Eq. (B19), it follows that

$$\iota\frac{\partial B_\Psi}{\partial\theta} + \frac{\partial B_\Psi}{\partial\varphi} = \frac{4\pi}{B^2}\frac{\partial P}{\partial\Psi} + \iota\frac{\partial B_\theta}{\partial\Psi} + \frac{\partial B_\varphi}{\partial\Psi}. \tag{B37}$$

This equation shows that B and B_Ψ are related. In particular, if B depends on only a particular linear combination of the angles, then the same is true for B_Ψ. Indeed, Fourier expansion of Eq. (B36) can be used to explicitly relate B and B_Ψ.

APPENDIX C: GUIDING-CENTER FOKKER-PLANCK FORMALISM

Magnetically confined plasmas found in nature and in laboratory devices are influenced by turbulent and collisional transport processes that play a major role in determining their particle and energy confinement properties (Balescu, 1988; Yoshikawa *et al.*, 2001). The study of the long-time confinement of magnetized plasmas involves the small dimensionless parameter $\epsilon_B \equiv \rho/L_B \ll 1$ defined as the ratio of the gyroradius ρ to the magnetic length scale L_B. Within the context of collisional transport theory in magnetized plasmas (Hinton and Hazeltine, 1976), a second small dimensionless parameter $\epsilon_\nu \equiv \nu/\Omega \ll 1$ is defined as the ratio of the characteristic collision frequency ν to the gyrofrequency Ω. These two ordering parameters, which appear in asymptotic expansions associated with the iterative solution of the collisional (Fokker-Planck) kinetic equation (Hinton *et al.*, 2003), also guarantee the existence of the first adiabatic invariant (i.e., magnetic moment) in a hierarchy of adiabatic invariants that underlies the long-time confinement of magnetized plasmas (Northrop, 1963). The existence of this first adiabatic invariant implies that the rapid gyromotion of a charged particle about a magnetic-field line is unaffected (to lowest order) by drift motion (associated with ϵ_B) and collisions (associated with ϵ_ν).

While both parameters ϵ_B and ϵ_ν are small in practice, it is useful to introduce the collisional parameter

$$\Delta \equiv \frac{\epsilon_B}{\epsilon_\nu} = \frac{\lambda_\nu}{L_B}, \tag{C1}$$

defined as the ratio of the collisional mean free path $\lambda_\nu \equiv \rho(\Omega/\nu) \gg \rho$ to the magnetic length scale L_B, in order to study collisional transport processes in complex magnetic geometries (Hinton and Hazeltine, 1976). The collisional parameter (C1) can be used to identify three distinct collisional regimes. In the collisional regime $\Delta \ll 1$, the collisional mean free path is much shorter than the magnetic length scale, so that the magnetic field may be treated in the *uniform* limit ($L_B \to \infty$). Hence, collisions are frequent enough to randomize the guiding-center drift motion and yield an isotropic pressure tensor $\mathsf{P} = p\mathbf{I}$ for each particle species. While magnetic spatial-gradient and curvature effects are ignored in this collisional regime, magnetic topology, however, may enter in a nontrivial way through magnetic-surface averaging of the Fokker-Planck collision operator (see, e.g., Pfirsch-Schlüter transport).

In the intermediate (drift) regime $\Delta \simeq 1$, collisions are infrequent enough to allow confined particles to sample the magnetic-field nonuniformity through their magnetic-drift motion between collisions ($L_B \simeq \lambda_\nu$). In the long-mean-free-path (or "collisionless") regime $\Delta \gg 1$, the collisional mean free path is much longer than the magnetic length scale ($\lambda_\nu \gg L_B$), and thus particles can sample the fully nonuniform magnetic field between collisions. Hence, although collisions are rare, they are not inconsequential, e.g., this low-collisionality regime yields an anisotropic Chew-Goldberger-Low pressure tensor $\mathsf{P} = p_\parallel\hat{\mathsf{b}}\hat{\mathsf{b}} + p_\perp(\mathbf{I} - \hat{\mathsf{b}}\hat{\mathsf{b}})$, and the finite magnetic length scale $L_B \ll \lambda_\nu$ cannot be ignored. Furthermore, the low-frequency ordering $\nu \ll \epsilon_B\Omega$ allows for the construction of a second adiabatic invariant, the bounce (longitudinal) action for magnetically trapped particles, which underlies Hamiltonian bounce-averaged guiding-center (or bounce-center) dynamics in nonuniform magnetic fields (Littlejohn, 1982a; Brizard, 2000). In this regime, collisions are thus insufficient to randomize the guiding-center drift motion and the resulting neoclassical transport processes can be dominated by large excursions from magnetic surfaces associated with complex (e.g., trapped-particle) guiding-center drift orbits (Hinton and Hazeltine, 1976).

The investigation of classical and neoclassical transport processes in complex magnetized plasmas is traditionally based on an iterative solution of the collisional Fokker-Planck kinetic equation for test-particle species a (Hinton and Hazeltine, 1976),

$$\frac{df_a(\mathbf{z},t)}{dt} \equiv \frac{\partial f_a(\mathbf{z},t)}{\partial t} + \frac{d\mathbf{z}}{dt}\cdot\frac{\partial f_a(\mathbf{z},t)}{\partial \mathbf{z}} = \sum_b \mathcal{C}_{ab}[f_a;f_b'](\mathbf{z},t), \tag{C2}$$

which describes the evolution of the distribution $f_a(\mathbf{z},t)$ in particle phase space $\mathbf{z}$ as a result of Hamiltonian (orbital) dynamics in phase space, represented by $dz^\alpha/dt = \{z^\alpha,h\}$ (where h and $\{,\}$ denote the particle Hamiltonian and Poisson bracket, respectively), and particle collisions in velocity space (between test-particle species a and field-particle species b) represented by the Fokker-Planck operator (Hinton and Hazeltine, 1976)

$$\mathcal{C}_{ab}[f_a;f_b'](\mathbf{x},\mathbf{p}) = -\frac{\partial}{\partial \mathbf{p}}\cdot\left(\mathbf{K}_{ab}\,f_a - \mathsf{D}_{ab}\cdot\frac{\partial f_a}{\partial \mathbf{p}}\right). \tag{C3}$$

Here the particle's kinetic momentum $\mathbf{p}=m\mathbf{v}$ is used and the Fokker-Planck collisional drag vector $\mathbf{K}_{ab}[f_b']$ and diffusion tensor $\mathsf{D}_{ab}[f_b']$ are

$$\mathbf{K}_{ab}[f_b'](\mathbf{z}) \equiv \Gamma_{ab}^K \frac{\partial H_b(\mathbf{z})}{\partial \mathbf{p}},$$

$$\mathsf{D}_{ab}[f_b'](\mathbf{z}) \equiv \Gamma_{ab}^D \frac{\partial^2 G_b(\mathbf{z})}{\partial \mathbf{p}\,\partial \mathbf{p}}, \tag{C4}$$

where $(\Gamma_{ab}^K,\Gamma_{ab}^D)=(m_a/m_b,\frac{1}{2}m_a^2)\Gamma_{ab}$, with $\Gamma_{ab} = 4\pi e_a^2 e_b^2 \ln\Lambda$, and the Rosenbluth potentials

$$\begin{pmatrix} H_b(\mathbf{z}) \\ G_b(\mathbf{z}) \end{pmatrix} \equiv \int d^6z'\,\delta^3(\mathbf{x}'-\mathbf{x})\begin{pmatrix} |\mathbf{v}'-\mathbf{v}|^{-1} \\ |\mathbf{v}'-\mathbf{v}| \end{pmatrix} f_b'(\mathbf{z}') \tag{C5}$$

are expressed as functionals over the entire field-particle phase space $\mathbf{z}'=(\mathbf{x}',\mathbf{p}'=m_b\mathbf{v}')$, which greatly facilitates our discussion of the transformation properties induced by phase-space transformations adopted for the test-particle and field-particle species. The presence of the delta function $\delta^3(\mathbf{x}'-\mathbf{x})$ ensures that collisions take place locally in physical space.

Classical transport coefficients (in the regime $\Delta \ll 1$) can appear explicitly in the Fokker-Planck collision operator (C3) if we formally introduce the transformation from particle phase-space coordinates $\mathbf{z}=(\mathbf{x},\mathbf{p})$ to the guiding-center phase-space coordinates $\mathbf{Z}\equiv(\mathbf{X},\mathcal{E},\mu,\zeta)$, where $\mathbf{X}$ denotes the guiding-center position, $\mathcal{E}$ denotes the guiding-center kinetic energy, μ denotes the guiding-center magnetic moment, and ζ denotes the guiding-center gyrophase. This transformation is expressed in terms of asymptotic expansions in powers of ϵ_B. In the uniform limit ($\epsilon_B=0$), however, it simplifies to $\mathbf{X}=\mathbf{x}-\rho_0$, $\mathcal{E}=|\mathbf{p}|^2/2m$, $\mu=|\mathbf{p}_\perp|^2/(2mB)$, and $\mathbf{p}_\perp = m\Omega\,\partial\rho_0/\partial\zeta$, where $\rho_0=\hat{\mathsf{b}}\times\mathbf{p}_\perp/m\Omega$ denotes the gyrophase-dependent gyroradius vector. Using this simplest guiding-center transformation, one obtains the guiding-center Fokker-Planck collision operator (Catto and Tsang, 1977; Brizard, 2004)

$$\mathcal{C}_{\rm gc}[\bar{F}](\mathbf{X},\mathcal{E},\mu) \equiv \langle e^{\boldsymbol{\rho}_0\cdot\nabla}\mathcal{C}[e^{-\boldsymbol{\rho}_0\cdot\nabla}\bar{F}]\rangle, \tag{C6}$$

where $\bar{F}$ denotes the reduced (gyrophase-independent) distribution of test-particle guiding centers (gyrophase averaging is denoted by an overbar) and the collision operator $\mathcal{C}$ denotes the original Fokker-Planck operator (C3) expressed in terms of $\mathbf{p}(\mathcal{E},\mu,\zeta)$. The reduced Fokker-Planck collision operator (C6) describes collisional drag and diffusion in five-dimensional guiding-center phase space $(\mathbf{X},\mathcal{E},\mu)$ and is, therefore, well suited to describe classical transport processes in the collisional regime ($\Delta \ll 1$), for which the magnetic field may be treated as spatially uniform (Xu and Rosenbluth, 1991; Dimits and Cohen, 1994).

The general rules for the transformation of an arbitrary bilinear collision operator were presented by Brizard (2004). Using Lie-transform methods, we obtained simpler and more compact expressions for transformed collision operators when compared to those obtained by the standard approach (Catto and Tsang, 1977; Xu and Rosenbluth, 1991; Dimits and Cohen, 1994), which could be appropriate for applications in gyrokinetic theory and gyrokinetic particle simulations.

The guiding-center Fokker-Planck collision operator presented by Brizard (2004) is written as

$$\mathcal{C}_{\rm gc}[\bar{F}] = -\frac{1}{\mathcal{J}_{\rm gc}}\frac{\partial}{\partial Z^\alpha}\left[\mathcal{J}_{\rm gc}\left(K_{\rm gc}^\alpha \bar{F} - D_{\rm gc}^{\alpha\beta}\frac{\partial \bar{F}}{\partial Z^\beta}\right)\right], \tag{C7}$$

where the guiding-center Fokker-Planck coefficients in guiding-center phase space

$$K_{\rm gc}^\alpha \equiv \langle \mathbf{K}_\epsilon\cdot\boldsymbol{\Delta}_\epsilon^\alpha\rangle,$$

$$D_{\rm gc}^{\alpha\beta} \equiv \langle \boldsymbol{\Delta}_\epsilon^\alpha\cdot\mathsf{D}_\epsilon\cdot\boldsymbol{\Delta}_\epsilon^\beta\rangle, \tag{C8}$$

are expressed in terms of $\boldsymbol{\Delta}_\epsilon^\alpha \equiv \{\mathbf{X}+\rho_\epsilon,Z^\alpha\}_\epsilon$, and $\mathcal{J}_{\rm gc} \equiv mB_\parallel^*/|v_\parallel|$ is the Jacobian for the guiding-center transformation, where $B_\parallel^*/B \equiv 1+\lambda_{\rm gc}$.

We may simplify our presentation (Brizard, 2004) by considering an isotropic field-particle distribution, so that the Rosenbluth potentials (C5) are functions of the normalized coordinate $\eta \equiv |\mathbf{p}|/m_a v_{{\rm T}b}$ (where $v_{{\rm T}b} = \sqrt{T_b/m_b}$), so that

$$\mathbf{K}_{ab} = \left(\frac{\Gamma_{ab}\,\eta H_b'}{2m_b\mathcal{E}}\right)\mathbf{p} \equiv -\nu\mathbf{p}, \tag{C9}$$

$$\begin{aligned}\mathsf{D}_{ab} &= \frac{\Gamma_{ab}m_a}{4\mathcal{E}}[\eta G_b'(\mathbf{I}-\hat{\mathsf{p}}\hat{\mathsf{p}}) + \eta^2 G_b''\hat{\mathsf{p}}\hat{\mathsf{p}}] \\ &\equiv \mathcal{D}_\perp(\mathbf{I}-\hat{\mathsf{p}}\hat{\mathsf{p}}) + \mathcal{D}_\parallel\hat{\mathsf{p}}\hat{\mathsf{p}},\end{aligned} \tag{C10}$$

where $\hat{\mathsf{p}} \equiv \mathbf{p}/|\mathbf{p}|$ and we, henceforth, omit species labels a and b.

In this isotropic case, the guiding-center Fokker-Planck drag and diffusion coefficients in Eq. (C7) are thus

$$(K_{\rm gc}^{\mathbf{X}}, D_{\rm gc}^{\mathbf{X}\mathcal{E}}, D_{\rm gc}^{\mathbf{X}\mu}) = (\nu_{\rm gc}, -\mathcal{D}_{\rm gc}^{\mathcal{E}}, -\mathcal{D}_{\rm gc}^{\mu})\boldsymbol{\rho}_{\rm gc}, \tag{C11}$$

where

$$\boldsymbol{\rho}_{\rm gc} \equiv \frac{\hat{\mathsf{b}}}{\Omega_{\|}^{*}} \times \mathbf{v}_{\rm gc} \tag{C12}$$

denotes the averaged guiding-center displacement, $\mathcal{D}^{\mathcal{E}} \equiv \mathcal{D}_{\|}/m$ and $\mathcal{D}^{\mu} \equiv \mu[(\mathcal{D}_{\|} - \mathcal{D}_{\perp})/|\mathbf{p}|^2]$,

$$\begin{pmatrix} K_{\rm gc}^{\mathcal{E}} & K_{\rm gc}^{\mu} \\ D_{\rm gc}^{\mathcal{E}\mathcal{E}} & D_{\rm gc}^{\mathcal{E}\mu} \end{pmatrix} = \begin{pmatrix} -2\mathcal{E}\nu_{\rm gc} & -(2-\lambda_{\rm gc})\nu_{\rm gc} \\ 2\mathcal{E}\mathcal{D}_{\rm gc}^{\mathcal{E}} & (2-\lambda_{\rm gc})\mathcal{D}_{\rm gc}^{\mathcal{E}} \end{pmatrix}$$

and

$$D_{\rm gc}^{\mu\mu} = (1-\lambda_{\rm gc})2\mu\left[2\left(1 - \frac{\mathcal{E}}{\mu B}\right)\mathcal{D}_{\rm gc}^{\mu} + \frac{\mathcal{D}_{\rm gc}^{\mathcal{E}}}{B}\right],$$

$$D_{\rm gc}^{\mathbf{XX}} = \frac{\mathbf{I}_{\perp}}{m\Omega^2}\left[\mathcal{D}_{\rm gc}^{\mathcal{E}} + \left(1 - \frac{2\mathcal{E}}{\mu B}\right)\mathcal{D}_{\rm gc}^{\mu} B\right] \equiv \mathcal{D}_{\rm gc}^{X}\mathbf{I}_{\perp}.$$

Here magnetic-field nonuniformity is represented by the terms $\lambda_{\rm gc}$ and $\rho_{\rm gc}$. While the magnetic-nonuniformity corrections associated with $\lambda_{\rm gc} \ll 1$ can be ignored for practical applications, the drag and diffusion coefficients (C11) depend explicitly on magnetic-field nonuniformity since the averaged guiding-center displacement (C12) directly involves the magnetic-field gradient and curvature.

REFERENCES

Alfvén, H., 1940, Ark. Mat., Astron. Fys. **27A**, 1.
Arnold, V. I., 1989, *Mathematical Methods of Classical Mechanics*, 2nd ed. (Springer, New York).
Balescu, R., 1988, *Transport Processes in Plasmas: 2. Neoclassical Transport Theory* (North-Holland, Amsterdam).
Baños, A., 1967, Plasma Phys. **1**, 305.
Batchelor, D. A., M. Beck, A. Becoulet, R. V. Budny, C. S. Chang, P. H. Diamond, J. Q. Dong, G. Y. Fu, A. Fukuyama, T. S. Hahm, D. E. Keyes, Y. Kishimoto, S. Klasky, L. L. Lao, K. Li, Z. Lin, B. Ludaescher, J. Manickam, N. Nakajima, T. Ozeki, N. Podhorszki, W. M. Tang, M. A. Vouk, R. E. Waltz, S. J. Wang, H. R. Wilson, X. Q. Xu, M. Yagi, and F. Zonca, 2007, Plasma Sci. Technol. **9**, 312.
Belova, E. V., N. N. Gorelenkov, and C. Z. Cheng, 2003, Phys. Plasmas **10**, 3240.
Bernardin, M. P., R. W. Moses, and J. A. Tataronis, 1986, Phys. Fluids **29**, 2605.
Birdsall, C. K., and A. B. Langdon, 1985, *Plasma Physics via Computer* (McGraw-Hill, New York).
Boghosian, B. M., 1987, Ph.D. thesis, University of California, Davis; see physics/0307148.
Boozer, A. H., 1980, Phys. Fluids **23**, 904.
Boozer, A. H., 1983, Phys. Fluids **26**, 1288.
Boozer, A. H., 1984, Phys. Fluids **27**, 2441.
Brizard, A. J., 1989, J. Plasma Phys. **41**, 541.
Brizard, A. J., 1990, Ph.D. Thesis, Princeton University, Princeton, NJ.
Brizard, A. J., 1992, Phys. Fluids B **4**, 1213.
Brizard, A. J., 1995, Phys. Plasmas **2**, 459.
Brizard, A. J., 2000, Phys. Plasmas **7**, 3238.
Brizard, A. J., 2004, Phys. Plasmas **11**, 4429.
Brizard, A. J., 2008, Commun. Nonlinear Sci. Numer. Simul. **13**, 24.
Brizard, A. J., and A. A. Chan, 1999, Phys. Plasmas **6**, 4548.
Brizard, A. J., and T. S. Hahm, 2007, Rev. Mod. Phys. **79**, 421.
Candy, J., and W. Rozmus, 1991, J. Comput. Phys. **92**, 230.
Cary, J. R., 1977, J. Math. Phys. **18**, 2432.
Cary, J. R., 1981a, Phys. Rep. **79**, 129.
Cary, J. R., 1981b, Phys. Fluids **24**, 2239.
Cary, J. R., 1982, Phys. Rev. Lett. **49**, 276.
Cary, J. R., 1984a, J. Math. Phys. **18**, 2432.
Cary, J. R., 1984b, Phys. Fluids **27**, 119.
Cary, J. R., and I. Doxas, 1993, J. Comput. Phys. **107**, 98.
Cary, J. R., D. F. Escande, and J. L. Tennyson, 1986, Phys. Rev. A **34**, 4256.
Cary, J. R., and J. D. Hanson, 1986, Phys. Fluids **29**, 2464.
Cary, J. R., C. L. Hedrick, and J. S. Tolliver, 1988, Phys. Fluids **31**, 1586.
Cary, J. R., and R. G. Littlejohn, 1983, Ann. Phys. (N.Y.) **151**, 1.
Cary, J. R., and S. G. Shasharina, 1997a, Phys. Rev. Lett. **78**, 674.
Cary, J. R., and S. G. Shasharina, 1997b, Phys. Plasmas **4**, 3323.
Cary, J. R., and S. G. Shasharina, 1997c, Plasma Phys. Rep. **23**, 509.
Cary, J. R., and R. T. Skodje, 1988, Phys. Rev. Lett. **61**, 1795.
Catto, P. J., and R. D. Hazeltine, 1981, Phys. Fluids **24**, 1663.
Catto, P. J., and K. T. Tsang, 1977, Phys. Fluids **20**, 396.
Dimits, A. M., G. Bateman, M. A. Beer, B. I. Cohen, W. Dorland, G. W. Hammett, C. Kim, J. E. Kinsey, M. Kotschenreuther, A. H. Kritz, L. L. Lao, J. Mandrekas, W. M. Nevins, S. E. Parker, A. J. Redd, D. E. Shumaker, R. Sydora, and J. Weiland, 2000, Phys. Plasmas **7**, 969.
Dimits, A. M., and B. I. Cohen, 1994, Phys. Rev. E **49**, 709.
Dobrott, D., and E. A. Frieman, 1971, Phys. Fluids **14**, 349.
Dubin, D. H. E., and J. A. Krommes, 1982, *Long Time Prediction in Dynamics* (Wiley, New York), pp. 251–280.
Forest, E., and R. D. Ruth, 1990, Physica D **43**, 105.
Fradkin, D. M., 1978, J. Phys. A **11**, 1069.
Gardner, C. S., 1959, Phys. Rev. **115**, 791.
Garren, D. A., and A. H. Boozer, 1991, Phys. Fluids B **3**, 2822.
Grebogi, C., and R. G. Littlejohn, 1984, Phys. Fluids **27**, 1996.
Goldstein, H., C. Poole, and J. Safko, 2002, *Classical Mechanics*, 3rd ed. (Addison-Wesley, San Francisco).
Hall, L. S., and B. McNamara, 1975, Phys. Fluids **18**, 552.
Hamada, S., 1959, Prog. Theor. Phys. **22**, 145.
Hanson, J. D., and J. R. Cary, 1984, Phys. Fluids **27**, 767.
Hazeltine, R. D., 1973, Plasma Phys. **15**, 77.
Hinton, F. L., and R. D. Hazeltine, 1976, Rev. Mod. Phys. **48**, 239.
Hinton, F. L., M. N. Rosenbluth, and R. E. Waltz, 2003, Phys. Plasmas **10**, 168.
Jackson, J. D., 1975, *Classical Electrodynamics*, 2nd ed. (Wiley, New York).
Kang, F., 1986, J. Comput. Math. **4**, 279.
Kaufman, A. N., 1986, Phys. Fluids **29**, 1736.
Kovrizhnykh, L. M., 1984, Nucl. Fusion **24**, 851.
Kruskal, M., 1962, J. Math. Phys. **3**, 806.

Kruskal, M., 1965, *Plasma Physics* (International Atomic Energy Agency, Vienna), p. 91.
Kruskal, M., and R. M. Kulsrud, 1958, Phys. Fluids **1**, 265.
Kulsrud, R. M., 1957, Phys. Rev. **106**, 205.
Landau, L. D., and E. M. Lifshitz, 1976, *Mechanics* (Pergamon, New York).
Littlejohn, R. G., 1979, J. Math. Phys. **20**, 2445.
Littlejohn, R. G., 1981, Phys. Fluids **24**, 1730.
Littlejohn, R. G., 1982a, J. Math. Phys. **23**, 742.
Littlejohn, R. G., 1982b, Phys. Scr., T **T2/I**, 119.
Littlejohn, R. G., 1983, J. Plasma Phys. **29**, 111.
Littlejohn, R. G., 1984, in *Contemporary Mathematics*, edited by J. E. Marsden (AMS, Providence, RI).
Littlejohn, R. G., 1985, Phys. Fluids **28**, 2015.
Littlejohn, R. G., 1988, Phys. Rev. A **38**, 6034.
Meiss, J. D., and R. D. Hazeltine, 1990, Phys. Fluids B **2**, 2563.
Meyer, F., and H. U. Schmidt, 1958, Z. Naturforsch. A **13**, 10005.
Morozov, A. I., and L. S. Solov'ev, 1966, in *Reviews of Plasma Physics*, edited by M. A. Leontovich (Consultants Bureau, New York), Vol. 2, p. 201.
Mynick, H. E., 1983, Phys. Fluids **26**, 1008.
Mynick, H. E., 2006, Phys. Plasmas **13**, 058102.
Mynick, H. E., T. K. Chu, and A. H. Boozer, 1982, Phys. Rev. Lett. **48**, 322.
Northrop, T. G., 1963, *Adiabatic Motion of Charged Particles* (Wiley, New York).
Northrop, T. G., C. S. Liu, and M. D. Kruskal, 1966, Phys. Fluids **9**, 1503.
Northrop, T. G., and J. A. Rome, 1978, Phys. Fluids **21**, 384.
Northrop, T. G., and E. Teller, 1960, Phys. Rev. **117**, 215.
Nührenberg, J., and R. Zille, 1988, Phys. Lett. A **129**, 113.
Palumbo, D., 1968, Nuovo Cimento B **53**, 507.
Qin, H., and R. C. Davidson, 2006, Phys. Rev. Lett. **96**, 085003.
Qin, H., and X. Guan, 2008, Phys. Rev. Lett. **100**, 035006.
Rutherford, P. H., 1970, Phys. Fluids **13**, 482.
Saletan, E. J., and A. H. Cromer, 1971, *Theoretical Mechanics* (Wiley, New York).
Skovoroda, A. A., and V. D. Shafranov, 1995, Plasma Phys. Rep. **21**, 886.
Sosenko, P. P., P. Bertrand, and V. K. Decyk, 2001, Phys. Scr. **64**, 264.
Stern, D. P., 1970, Am. J. Phys. **38**, 494.
Tao, X., A. A. Chan, and A. J. Brizard, 2007, Phys. Plasmas **14**, 092107.
Taylor, J. B., 1964, Phys. Fluids **7**, 767.
Taylor, J. B., 1967, Phys. Fluids **10**, 1357.
White, R. B., 1990, Phys. Fluids B **2**, 845.
White, R. B., 2008, *Turbulent Transport in Fusion Plasmas, First ITER International Summer School*, AIP Conf. Proc. No. 1013 (AIP, Melville, NY), pp. 59 and 127.
White, R. B., A. H. Boozer, and R. Hay, 1982, Phys. Fluids **25**, 575.
White, R. B., and M. S. Chance, 1984, Phys. Fluids **27**, 2455.
White, R. B., and L. E. Zakharov, 2003, Phys. Plasmas **10**, 573.
Wong, H. V., 1982, Phys. Fluids **25**, 1811.
Xu, X. Q., and M. N. Rosenbluth, 1991, Phys. Fluids B **3**, 627.
Yoshida, H., 1990, Phys. Lett. A **150**, 262.
Yoshikawa, A., S. I. Itoh, K. Itoh, and N. Yokoi, 2001, Plasma Phys. Controlled Fusion **43**, R1.