

Foundations of nonlinear gyrokinetic theory

A. J. Brizard*

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

T. S. Hahm

Plasma Physics Laboratory, Princeton University, Princeton, New Jersey 08543, USA

(Published 2 April 2007)

Nonlinear gyrokinetic equations play a fundamental role in our understanding of the long-time behavior of strongly magnetized plasmas. The foundations of modern nonlinear gyrokinetic theory are based on three pillars: (i) a gyrokinetic Vlasov equation written in terms of a gyrocenter Hamiltonian with quadratic low-frequency ponderomotivelike terms, (ii) a set of gyrokinetic Maxwell (Poisson-Ampère) equations written in terms of the gyrocenter Vlasov distribution that contain low-frequency polarization (Poisson) and magnetization (Ampère) terms, and (iii) an exact energy conservation law for the gyrokinetic Vlasov-Maxwell equations that includes all the relevant linear and nonlinear coupling terms. The foundations of nonlinear gyrokinetic theory are reviewed with an emphasis on rigorous application of Lagrangian and Hamiltonian Lie-transform perturbation methods in the variational derivation of nonlinear gyrokinetic Vlasov-Maxwell equations. The physical motivations and applications of the nonlinear gyrokinetic equations that describe the turbulent evolution of low-frequency electromagnetic fluctuations in a nonuniform magnetized plasmas with arbitrary magnetic geometry are discussed.

DOI: 10.1103/RevModPhys.79.421

PACS number(s): 52.30.Gz, 52.65.Tt

CONTENTS

I. Introduction	421	C. Nonlinear gyrocenter Hamiltonian dynamics	450
II. Basic Properties of Nonlinear Gyrokinetic Equations	426	D. Nonlinear gyrokinetic Vlasov equation	450
A. Physical motivations and nonlinear gyrokinetic orderings	426	E. Pull-back representation of the perturbed Vlasov distribution	451
B. Frieman-Chen nonlinear gyrokinetic equation	428	VI. Gyrokinetic Variational Formulation	452
C. Modern nonlinear gyrokinetic equations	431	A. Nonlinear gyrokinetic Vlasov-Maxwell equations	452
III. Simple Forms of Nonlinear Gyrokinetic Equations	432	B. Gyrokinetic energy conservation law	453
A. General gyrokinetic Vlasov-Maxwell equations	432	VII. Summary	454
B. Electrostatic fluctuations	433	Acknowledgments	454
C. Shear-Alfvénic magnetic fluctuations	435	Appendix A: Mathematical Primer	455
1. Hamiltonian (p_{\parallel}) formulation	436	Appendix B: Unperturbed Guiding-Center Hamiltonian Dynamics	456
2. Symplectic (v_{\parallel}) formulation	437	1. Guiding-center phase-space transformation	456
D. Compressional magnetic fluctuations	437	2. Guiding-center Hamiltonian dynamics	457
E. Reduced fluid equations from moments of the nonlinear gyrokinetic Vlasov equation	438	3. Guiding-center pull-back transformation	457
IV. Lie-Transform Perturbation Theory	440	4. Bounce-center Hamiltonian dynamics	458
A. Single-particle extended Lagrangian dynamics	441	Appendix C: Push-Forward Representation of Fluid Moments	459
B. Perturbation theory in extended phase space	442	1. Push-forward representation of fluid moments	459
C. Near-identity phase-space transformations	443	2. Push-forward representation of gyrocenter fluid moments	460
D. Lie-transform methods	444	Appendix D: Extensions of Nonlinear Gyrokinetic Equations	460
1. Transformed extended Poisson-bracket structure	445	1. Strong $E \times B$ flow shear	461
2. Transformed extended Hamiltonian	445	2. Bounce-center-kinetic Vlasov equation	463
E. Reduced Vlasov-Maxwell equations	446	References	465
V. Nonlinear Gyrokinetic Vlasov Equation	447		
A. Unperturbed guiding-center Hamiltonian dynamics	448		
B. Perturbed guiding-center Hamiltonian dynamics	448		

I. INTRODUCTION

Magnetically confined plasmas, found either in laboratory devices (e.g., fusion-related plasmas) or in nature (e.g., astrophysical plasmas), exhibit a wide range of spatial and temporal scales. An important class of plasma space-time scales, which includes some of the shortest and longest plasma scales, is defined by the orbits of

*Electronic address: abrizard@smcvt.edu

magnetically confined charged particles. It has long been known that the long-time magnetic confinement of charged particles (Northrop, 1963) depends on the existence of adiabatic invariants (based on a separation of orbital time scales) and/or exact invariants (associated with exact symmetries of the confining magnetic field). This review begins a brief survey of orbital time scales by considering the motion of single charged particles in a plasma confined by a strong magnetic field, represented in divergenceless form as

$$\mathbf{B} = \nabla\alpha \times \nabla\beta \equiv B(\alpha, \beta, s)\hat{\mathbf{b}}(\alpha, \beta, s). \quad (1)$$

Here the Euler potentials (α, β) are time-dependent magnetic-line labels (Stern, 1970), s denotes the parallel coordinate along a field line, and $\hat{\mathbf{b}} \equiv \partial\mathbf{x}/\partial s$ denotes the unit vector along the magnetic-field line. Charged-particle motion in a uniform magnetic field $\mathbf{B}_0 = B_0\hat{\mathbf{b}}_0$ (in the absence of an electric field) is characterized by (a) a gyromotion perpendicular to a magnetic-field line with constant gyrofrequency $\Omega = eB_0/mc$ (for a particle of mass m and charge e) and constant gyroradius $r_\perp = v_\perp/\Omega$, where $v_\perp = |\mathbf{v} \times \hat{\mathbf{b}}_0|$ denotes the magnitude of the particle's velocity perpendicular to the magnetic field, and (b) a parallel motion along a magnetic-field line with constant parallel velocity $v_\parallel = \mathbf{v} \cdot \hat{\mathbf{b}}_0$.

When the magnetic field (1) is spatially inhomogeneous the magnitude $B = |\nabla\alpha \times \nabla\beta|$ and direction $\hat{\mathbf{b}} = \mathbf{B}/B$ are not constant and the velocity components v_\perp and v_\parallel are no longer conserved. The primary terms associated with magnetic-field inhomogeneity are represented by the perpendicular gradient $\hat{\mathbf{b}} \times \nabla \ln B$ as well as the magnetic curvature $\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}}$ and the parallel gradient $\hat{\mathbf{b}} \cdot \nabla \ln B$.¹ Charged particles confined by a strong magnetic field with weak inhomogeneity execute three types of quasiperiodic motion (Kruskal, 1962; Northrop, 1963): (i) a rapid gyromotion about a single magnetic-field line, with a gyroperiod $\tau_g(\mathbf{x})$ that depends on the particle's spatial position $\mathbf{x} = (\alpha, \beta, s)$; (ii) an intermediate bounce (or transit) motion along a magnetic-field line (affected by the parallel gradient), with an intermediate characteristic time scale $\tau_b(\mathbf{y}; \mathcal{E}, J_g)$ that depends on the particle's energy \mathcal{E} and its gyroaction $J_g = (mc/e)\mu$ (where μ denotes the magnetic moment) as well as the field-line labels $\mathbf{y} = (\alpha, \beta)$; and (iii) a slow drift (bounce-averaged precession) motion across magnetic-field lines (driven by the perpendicular gradient and the magnetic curvature term), with a long characteristic time scale $\tau_d(\mathcal{E}, J_g, J_b)$ that depends on the bounce action J_b as well

as the energy \mathcal{E} and the gyroaction J_g . The use of magnetic coordinates (α, β) in the magnetic-field representation (1) facilitates the general description of bounce-averaged guiding-center motion, where the fast gyromotion and bounce-motion time scales are decoupled from the drift motion of magnetically confined charged particles; see Appendixes B.4 and D.2. For the remainder of the paper, the particular magnetic geometry is ignored, the work presented here is suitable for applications in gyrokinetic studies of laboratory and space magnetized plasmas.

The orbital time scales $\tau_g \ll \tau_b \ll \tau_d$ are typically well separated for charged particles magnetically confined in a strong magnetic field with weak inhomogeneity (i.e., $\rho|\nabla \ln B| \ll 1$). For example, the orbital time scales of a 10-keV proton equatorially mirroring at geosynchronous orbit (Schulz and Lanzerotti, 1974) are $\tau_g \sim 0.33$ s $\ll \tau_b \sim 33$ s $\ll \tau_d \sim 10^5$ s. It has long been understood (Northrop and Teller, 1960) that the stability and longevity of Earth's radiation belts was due to the adiabatic invariance of the three actions (J_g, J_b, J_d) , where the drift action $J_d \equiv (e/c)\Phi_B$ is defined in terms of the magnetic flux Φ_B enclosed by the bounce-averaged precession motion of magnetically trapped charged particles. In a high-temperature magnetized plasma the typical energy-confinement time τ_E (great interest in the development of fusion energy) generally satisfies the condition $\tau_E \gg \tau_b \gg \tau_g$ so the time scales associated with the charge particle's gyromotion and bounce or transit motion are much smaller than the transport time scale of interest.

Understanding the nonlinear dynamics of magnetically confined plasmas is a formidable task. There exist a wide variety of instabilities in inhomogeneous magnetically confined plasmas whose nonlinear behavior is, in general, different from their linear behavior. In plasma turbulence, the “inertial” range is relatively narrow in wave-vector \mathbf{k} space due to Landau damping from wave-particle resonant interaction, in contrast to the inertial range in fluid turbulence (Frisch, 1995) that exists over several decades in \mathbf{k} space and for which the Reynolds number \mathcal{R} (a dimensionless number characterizing the ratio between nonlinear coupling and classical dissipation) satisfies the condition $\mathcal{R} \gg 1$ [see, for example, Dimotakis (2000)]. Plasma turbulence involves a plethora of additional dimensionless parameters associated with the orbital dynamics of magnetically confined charged particles not present in fluid turbulence. Many aspects of the nonlinear dynamics involved in the evolution toward such a saturated state that often exhibits self-organized large-scale motion and are not yet well understood. It is important to note that many plasmas of interest in magnetic fusion and in astrophysics are “collisionless” on the particle dynamics and turbulence time scales, since typical charged particles can execute many gyrations and bounces or transits, and collective waves can undergo many cycles before particles suffer a 90° Coulomb collision. A collisionless kinetic description is desirable for such plasmas.

¹Additional terms associated with magnetic-field inhomogeneity include $\hat{\mathbf{b}} \cdot \nabla \times \hat{\mathbf{b}} = (4\pi/c)J_\parallel/B$, which is related to the plasma current J_\parallel flowing along magnetic-field lines, and $\mathbf{R} \equiv \nabla\hat{\mathbf{1}} \cdot \hat{\mathbf{2}}$, where $\hat{\mathbf{1}} \equiv \nabla\alpha/|\nabla\alpha|$ and $\hat{\mathbf{2}} \equiv \hat{\mathbf{b}} \times \hat{\mathbf{1}}$ denote local unit vectors perpendicular to $\hat{\mathbf{b}}$ (Littlejohn, 1988); these secondary terms appear in the Hamiltonian formulation of guiding-center theory (Littlejohn, 1983).

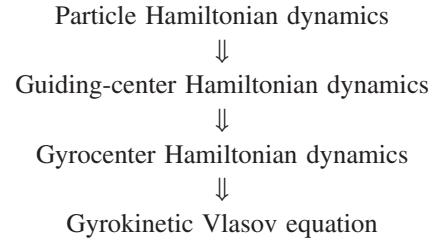
The development of low-frequency gyrokinetic theory was motivated by the need to describe complex plasma dynamics over time scales that are long compared to the gyromotion time scale. Gyrokinetic theory was developed as a generalization of guiding-center theory (Northrop, 1963; Littlejohn, 1983). Taylor (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 of the perturbation field. This result indicated that gyrokinetic theory could be built upon an additional transformation beyond the guiding-center phase-space coordinates, constructing new gyrocenter phase-space coordinates that describe gyroangle-averaged perturbed guiding-center dynamics. This step was not considered the highest priority at the time, and Rutherford and Frieman (1968) followed a more conventional approach by developing the linear gyrokinetic theory of low-frequency drift-wave (universal) instabilities in general magnetic geometry.²

Nonlinear gyrokinetic theory focuses on low-frequency electromagnetic fluctuations that are observed in inhomogeneous magnetized plasmas found in laboratory devices and nature; see, for example, Frieman and Chen (1982); Dubin *et al.* (1983); Hahm (1988, 1996); Hahm *et al.* (1988); and Brizard (1989a). Plasma microturbulence and its associated anomalous transport have been subjects of active research and wide interest over many years. The following review papers have addressed this subject with different emphases: Tang (1978) on linear instabilities in magnetized plasmas, Horton (1999) on further developments in linear and nonlinear theories and simulations, Krommes (2002) on analytical aspects of statistical closure, and Diamond *et al.* (2005) on the self-regulation of turbulence and transport by zonal flows. Gyrokinetic simulations now play a major role in the investigation of low-frequency plasma turbulence and its associated transport in magnetized plasmas; see Table I for a survey of applications of nonlinear gyrokinetic equations.

The foundations of modern nonlinear gyrokinetic theory presented in this review paper are three important mutually dependent pillars: (I) a gyrokinetic Vlasov equation written in terms of a gyrocenter Hamiltonian that contains quadratic low-frequency ponderomotive-like terms, (II) a set of gyrokinetic Maxwell (Poisson-Ampère) equations that contain low-frequency polarization (Poisson) and magnetization (Ampère) terms derived from the quadratic nonlinearities in the gyro-

center Hamiltonian, and (III) an exact energy conservation law for the self-consistent gyrokinetic Vlasov-Maxwell equations that includes the relevant linear and nonlinear coupling terms.

I. Nonlinear gyrokinetic Vlasov equation



Our derivation of the nonlinear gyrokinetic Vlasov equation (I) proceeds in two steps; each step involves the asymptotic decoupling of the fast gyromotion time scale from a set of Hamilton equations by Lie-transform methods (Littlejohn, 1982a; Brizard, 1990; Qin and Tang, 2004). The first step derives the guiding-center (gc) Hamilton equations through the elimination of the gyroangle associated with the gyromotion of charged particles about equilibrium magnetic-field lines. As a result of the guiding-center transformation, the gyroangle becomes an ignorable coordinate, and the guiding-center magnetic moment $\mu = \mu_0 + \dots$ (where $\mu_0 \equiv m|\mathbf{v}_\perp|^2/2B$ denotes the lowest-order term) is treated as a dynamical invariant within the guiding-center Hamiltonian dynamics. The introduction of low-frequency electromagnetic fluctuations (within the guiding-center Hamiltonian formalism) results in the destruction of the guiding-center magnetic moment due to the reintroduction of the gyroangle dependence in the perturbed guiding-center Hamiltonian system. In the second step, a new set of gyrocenter (gy) Hamiltonian equations is described through the elimination of the gyroangle from the perturbed guiding-center equations. As a result of the gyrocenter transformation, a new gyrocenter magnetic moment $\bar{\mu} = \mu + \dots$ is constructed as an adiabatic invariant and the gyrocenter gyroangle $\bar{\zeta}$ is an ignorable coordinate. Within the gyrocenter Hamiltonian formalism, the gyrokinetic Vlasov equation expresses the fact that the gyrocenter Vlasov distribution $\bar{F}(\bar{\mathbf{X}}, \bar{v}_\parallel, t; \bar{\mu})$ is constant along a gyrocenter orbit in gyrocenter phase space $(\bar{\mathbf{X}}, \bar{v}_\parallel; \bar{\mu}, \bar{\zeta})$,

$$\frac{\partial \bar{F}}{\partial t} + \frac{d\bar{\mathbf{X}}}{dt} \cdot \bar{\nabla} \bar{F} + \frac{d\bar{v}_\parallel}{dt} \frac{\partial \bar{F}}{\partial \bar{v}_\parallel} = 0, \quad (2)$$

where $d\bar{\mu}/dt \equiv 0$ and $\partial \bar{F} / \partial \bar{\zeta} \equiv 0$ (both by definition). Here $\bar{\mathbf{X}}$ denotes the gyrocenter position, $\bar{v}_\parallel \equiv \hat{\mathbf{b}} \cdot d\bar{\mathbf{X}}/dt$ denotes the gyrocenter parallel velocity ($\hat{\mathbf{b}} = \mathbf{B}/B$), and the gyrocenter equations of motion ($d\bar{\mathbf{X}}/dt, d\bar{v}_\parallel/dt$) are independent of the gyrocenter gyroangle $\bar{\zeta}$ (explicit expressions are given below). This two-step Lie-transform approach plays a fundamental role in the construction of modern gyrokinetic theory (Dubin *et al.*, 1983; Hahm,

²The conventional approach to deriving the gyrokinetic Vlasov equation is based on an iterative solution of the gyroangle-averaged Vlasov equation perturbatively expanded in powers of the finite-Larmor-radius dimensionless parameter ρ/L (Hastie *et al.*, 1967). The modern Lie-transform approach presented here is based on a two-step transformation procedure from particle to gyrocenter phase space.

TABLE I. Applications of nonlinear gyrokinetic equations.

Instability	Nonlinear theory	Nonlinear simulation
Drift (universal or dissipative) instability	Frieman and Chen (1982) Smith <i>et al.</i> (1985) Hahm (1992) Mattror and Diamond (1989) Hahm and Tang (1990) Mattror (1992) Rosenbluth and Hinton (1998) Chen <i>et al.</i> (2000)	Lee (1983) Lee <i>et al.</i> (1984) Parker and Lee (1993) Lee and Tang (1988) Sydora <i>et al.</i> (1990) Parker <i>et al.</i> (1993) Dimits <i>et al.</i> (1996) Lin <i>et al.</i> (1999) Refs. in Diamond <i>et al.</i> (2005) Hatzky <i>et al.</i> (2002) Villard, Allfrey, <i>et al.</i> (2004) Sydora (1990) Chen and Parker (2001) Ernst <i>et al.</i> (2004) Dannert and Jenko (2005) Depret <i>et al.</i> (2005) Jenko <i>et al.</i> (2000) Idomura <i>et al.</i> (2000) Dorland <i>et al.</i> (2000) Lin, Chen, <i>et al.</i> (2005) Sarazin <i>et al.</i> (2005) Lee <i>et al.</i> (2001) Parker <i>et al.</i> (2004) Briguglio <i>et al.</i> (1998) Jenko and Scott (1999) Chen <i>et al.</i> (2001) Naitou <i>et al.</i> (1995) Matsumoto <i>et al.</i> (2005, 2003) Sydora (2001) Parker <i>et al.</i> (2004) Park <i>et al.</i> (1992) Santoro and Chen (1996) Zonca <i>et al.</i> (2002) Todo <i>et al.</i> (2003)
Ion temperature gradient (ITG) mode		
ITG turbulence with zonal flows		
ITG turbulence with velocity space nonlinearity addressing energy conservation		
Trapped electron mode	Similon and Diamond (1984) Gang <i>et al.</i> (1991) Hahm and Tang (1991)	
Trapped-ion mode	Hahm and Tang (1996)	
Electron-temperature-gradient (ETG) mode	Kim <i>et al.</i> (2003) Chen <i>et al.</i> (2005) Idomura (2006)	
Interchange turbulence		
(Kinetic) shear-Alfvén wave	Frieman and Chen (1982) Hahm <i>et al.</i> (1988) Briguglio <i>et al.</i> (2000) Chen <i>et al.</i> (2001)	
Drift-Alfvén turbulence		
Tearing and internal kink instability		
Microtearing and drift-tearing mode		
Energetic particle driven MHD instabilities	Chen (1994) Vlad <i>et al.</i> (1999) Zonca <i>et al.</i> (2005)	
Geomagnetic pulsation	Chen and Hasegawa (1994)	
Whistler lower-hybrid instability		Lin, Wang, <i>et al.</i> (2005)

1988; Hahm *et al.*, 1988; Brizard, 1989a):

$$\begin{array}{c}
 \text{II. Gyrokinetic Maxwell equations} \\
 \hline
 \text{Maxwell equations} \\
 \Downarrow \\
 \text{Gyrokinetic Maxwell equations}
 \end{array}$$

A self-consistent description of low-frequency electromagnetic fluctuations, produced by charged-particle motion, is based on the derivation of gyrokinetic Maxwell equations (II) expressed in terms of moments of the gyrocenter Vlasov distribution. The transformation from particle moments to gyrocenter moments again involves two steps (associated with the guiding-center and gyro-

center phase-space transformations), and each step introduces a polarization density and a magnetization current in the gyrokinetic Maxwell (Poisson-Ampère) equations,

$$\nabla^2(\Phi + \delta\phi) = -4\pi(\bar{\rho} + \rho_{\text{pol}}), \quad (3)$$

$$\nabla \times (\mathbf{B} + \delta\mathbf{B}) = \frac{4\pi}{c}(\bar{\mathbf{J}} + \mathbf{J}_{\text{pol}} + \mathbf{J}_{\text{mag}}), \quad (4)$$

where $\bar{\rho}$ and $\bar{\mathbf{J}}$ denote charge and current densities evaluated as moments of the gyrocenter Vlasov distribution \bar{F} , while the polarization density $\rho_{\text{pol}} \equiv -\nabla \cdot \mathbf{P}_{\text{gy}}$, the polarization current $\mathbf{J}_{\text{pol}} \equiv \partial \mathbf{P}_{\text{gy}} / \partial t$, and the divergence-

less magnetization current $\mathbf{J}_{\text{mag}} \equiv c\nabla \times \mathbf{M}_{\text{gy}}$ are defined in terms the gyrocenter polarization vector \mathbf{P}_{gy} and the gyrocenter magnetization vector \mathbf{M}_{gy} .³ The guiding-center magnetization current $\mathbf{M}_{\text{gc}} \equiv -\|\mu\|_{\text{gc}}\hat{\mathbf{b}}$ (where $\|\cdots\|_{\text{gc}}$ denotes a moment with respect to the guiding-center Vlasov distribution) accounts for the difference between the particle current and guiding-center current. Gyrocenter polarization and magnetization effects, on the other hand, involve expressions for ρ_{pol} , \mathbf{J}_{pol} , and \mathbf{J}_{mag} in which the perturbed electromagnetic fields ($\delta\mathbf{E}$, $\delta\mathbf{B}$) appear explicitly. The presence of self-consistent gyrocenter polarization effects within the nonlinear electrostatic gyrokinetic formalism (Dubin *et al.*, 1983; Hahm, 1988), for example, yields important computational advantages in gyrokinetic electrostatic simulations (Lee, 1983). The self-consistency of the nonlinear gyrokinetic Vlasov-Maxwell equations is guaranteed by the Lie-transform perturbation approach.

Under standard gyrokinetic space-time orderings considered in this work (see Sec. II.A), the quasineutrality condition $\rho \equiv \bar{\rho} + \rho_{\text{pol}} = 0$, associated with the gyrokinetic Poisson equation (3), is consistent with the charge conservation law $\nabla \cdot (\mathbf{J} + \mathbf{J}_{\text{pol}}) = 0$, associated with the gyrokinetic Ampère equation (4). These relations ensure that the gyrocenter charge conservation law $\partial\bar{\rho}/\partial t + \nabla \cdot \bar{\mathbf{J}} = 0$ is satisfied, i.e.,

$$\frac{\partial\bar{\rho}}{\partial t} = -\nabla \cdot \bar{\mathbf{J}} = \nabla \cdot \mathbf{J}_{\text{pol}} = -\frac{\partial\rho_{\text{pol}}}{\partial t}.$$

This relation guarantees the proper treatment of polarization (and magnetization) effects in gyrokinetic theory:

III. Gyrokinetic energy conservation law

Gyrokinetic variational formulation

↓

Gyrokinetic Vlasov-Maxwell equations

↓

Gyrokinetic energy conservation law (Noether)

The self-consistent polarization and magnetization effects appearing in the gyrokinetic Maxwell equations (3) and (4) can be computed either directly by the *push-forward* method, transforming particle moments into guiding-center moments and then into gyrocenter moments, or by a variational method from a nonlinear low-frequency gyrokinetic action functional. While the direct approach has the advantage of being the simplest method to use (Dubin *et al.*, 1983; Hahm *et al.*, 1988; Brizard, 1989a, 1990), the variational approach (Brizard, 2000a, 2000b) has the advantage of allowing a direct derivation of an exact energy conservation law (III) for the nonlinear gyrokinetic Vlasov-Maxwell equations

through the Noether method (Brizard, 2005a).

This paper reviews the modern foundations of nonlinear gyrokinetic theory by presenting the Lagrangian and Hamiltonian Lie-transform perturbation methods used to derive self-consistent, energy-conserving gyrokinetic Vlasov-Maxwell equations describing the nonlinear turbulent evolution of low-frequency, short-perpendicular-wavelength electromagnetic fluctuations in nonuniform magnetized plasmas. While prototype modern gyrokinetic theories (Dubin *et al.*, 1983; Hagan and Frieman, 1985; Yang and Choi, 1985; Hahm *et al.*, 1988) relied on the Hamiltonian Lie-Darboux perturbation method (Littlejohn, 1979, 1981), this work is focused on applications of the phase-space-Lagrangian Lie perturbation method (Littlejohn, 1982a, 1983), which is superior in its clarity and efficiency.

The remainder of the paper is organized as follows. In Sec. II, the physical motivations for the nonlinear gyrokinetic analysis of turbulent magnetized plasmas are presented. While this discussion tends to focus on fusion-related magnetically confined plasmas, these nonlinear gyrokinetic equations are valid for arbitrary magnetic geometry and, will find applications in space plasma physics (Lin, Wang, *et al.*, 2005) and astrophysics (Howes *et al.*, 2006). A brief discussion of the nonlinear gyrokinetic equation derived by conventional (iterative) methods by Frieman and Chen (1982) is also presented as well as an outline of the modern (i.e., Lie-transform) derivation of the nonlinear gyrokinetic Vlasov equation. In Sec. III, simplified forms of the nonlinear gyrokinetic Vlasov-Maxwell equations that are recommended for simulations and analytic applications are presented. The nonlinear gyrokinetic equations suitable to describe electrostatic fluctuations as well as shear-Alfvénic and compressional magnetic fluctuations in a strongly magnetized plasma of arbitrary geometry are presented. These equations are supplemented by explicit expressions for the appropriate gyrokinetic Poisson-Ampère equations as well as an explicit form for the gyrokinetic energy-conservation law. Last, simplified nonlinear gyrofluid equations (derived as moments of the nonlinear gyrokinetic Vlasov equation) are presented to discuss the intimate connection between nonlinear gyrokinetic dynamics and reduced fluid dynamics such as reduced magnetohydrodynamics (RMHD). Section IV introduces the basic concepts of Lie-transform perturbation theory applied to the transformation of the Vlasov-Maxwell equations induced by an arbitrary near-identity phase-space transformation. These methods are then applied in Sec. V to derive the nonlinear gyrocenter Hamiltonian as well as the corresponding nonlinear gyrokinetic Vlasov equation. Section VI shows how the self-consistent gyrokinetic Poisson-Ampère equations are derived from a gyrokinetic variational principle. The existence of such a variational principle guarantees the existence of exact conservation laws for nonlinear gyrokinetic Vlasov-Maxwell equations. The gyrokinetic energy conservation law is derived explicitly since it plays a fundamental role in the development of energy-conserving gyrokinetic simulation techniques. Last, a

³See Appendix C for derivation of the reduced polarization density and reduced magnetization current by Lie-transform methods; see also Sec. 6.7 of Jackson (1975) for additional details.

summary of the paper is presented in Sec. VII. A number of appendixes are included that present either supporting material not central to the topic of this paper or extensions of the gyrokinetic Vlasov-Maxwell equations presented in the text. For readers who are mainly interested in basic concepts, fusion applications, and the simpler forms of the nonlinear gyrokinetic equations useful for simulations, we recommend Secs. II and III, and Appendix D.1.

Other developments in gyrokinetic theory not presented here include derivation of nonlinear relativistic gyrokinetic Vlasov-Maxwell equations (Brizard and Chan, 1999), investigation of the thermodynamic properties of the gyrokinetic equations (Krommes *et al.*, 1986; Krommes, 1993a, 1993b; Sugama *et al.*, 1996; Watanabe and Sugama, 2006), inclusion of a reduced (guiding-center) Fokker-Planck collision operator into the gyrokinetic formalism (Dimits and Cohen, 1994; Brizard, 2004), and derivation of high-frequency linear gyrokinetics (Lee *et al.*, 1983; Tsai *et al.*, 1984; Lashmore-Davies and Dendy, 1989; Qin *et al.*, 1999, 2000).

II. BASIC PROPERTIES OF NONLINEAR GYROKINETIC EQUATIONS

In many plasmas found in both laboratory devices and nature, the temporal scales of collective electrostatic and/or electromagnetic fluctuations of interest are much longer than a period of a charged particle's cyclotron motion (gyromotion) due to a strong background magnetic field, while the spatial scales of such fluctuations are much smaller than the scale length of the background magnetic-field inhomogeneity. In these circumstances, the details of the charged particle's gyration, such as the gyrophase, are not of dynamical significance, and it is possible to develop a reduced set of dynamical equations that captures the essential features of the low-frequency phenomena of interest.

By decoupling the nearly periodic gyromotion (Kruskal, 1962), one can derive the gyrokinetic equation (2) that describes the spatiotemporal evolution of the gyrocenter distribution function defined over a *reduced* (4+1)-dimensional gyrocenter phase space $(\tilde{\mathbf{X}}, \tilde{v}_\parallel; \bar{\mu})$, a key feature of the modern nonlinear gyrokinetic approach. In simulating strongly magnetized plasmas, an enormous amount of computing time is saved by using a time step greater than the gyroperiod and reducing the number of dynamical variables by 1.

An excellent example in which the nonlinear gyrokinetic formulations have applied well and have made a deep and long-lasting impact is the theoretical study of microturbulence in tokamak and stellarator devices. Experimental measurements over the past three decades have led to the belief that, in the absence of macroscopic magnetohydrodynamic (MHD) instabilities,⁴ tokamak

microturbulence is responsible for the anomalous transport of plasma particles, heat, and toroidal angular momentum commonly found to appear at higher levels than predictions from classical and neoclassical collisional transport theories (Rosenbluth *et al.*, 1972; Hinton and Hazeltine, 1976; Chang and Hinton, 1982; Hinton and Wong, 1985; Connor *et al.*, 1987).

A. Physical motivations and nonlinear gyrokinetic orderings

Experimental observations of magnetically confined high-temperature plasmas indicate that they represent strongly turbulent systems (Liewer, 1985; Wootton *et al.*, 1990). While there are various ways to excite instabilities in plasmas (such as the velocity-space gradient, current gradient, and an energetic-particle population), the observed turbulent fluctuations are believed to originate primarily from collective instabilities driven by the expansion free energy associated with radial gradients in temperature or density (Tang, 1978; Horton, 1999), and is thought to be related to fluctuation-induced transport processes due to saturated (finite-amplitude) plasma microturbulence, whose characteristic time scale is much longer than the gyroperiod. The fluctuation spectra are characterized by the following features: (i) the characteristic (mean) frequency $\omega_{\mathbf{k}}$ (for fixed wave vector \mathbf{k}) and perpendicular wavelength $\lambda_\perp (=2\pi/|\mathbf{k}_\perp|)$ of the fluctuation spectrum are typical of drift-wave turbulence theories (Horton, 1999); (ii) the frequency spectrum is broad ($\Delta\omega \sim \omega_{\mathbf{k}}$) at fixed wave vector \mathbf{k} ; (iii) there are fluctuations in density, temperature, electrostatic potential, and magnetic field, with each fluctuating quantity having its own spatial profile across the plasma discharge; and (iv) the fixed-frequency fluctuation spectrum is highly anisotropic in wave vector ($k_\parallel \ll k_\perp$), i.e., parallel wavelengths are much longer than perpendicular wavelengths.

From experimental observations [see references listed in Wootton *et al.* (1990)], the typical fluctuation frequency spectrum is found to be broadband ($\Delta\omega \sim \omega_{\mathbf{k}}$) at fixed wave vector \mathbf{k} . Its characteristic mean frequency (in the plasma frame rotating with $E \times B$ velocity) is on the order of the diamagnetic frequency $\omega_* \equiv \mathbf{k} \cdot \mathbf{v}_D \ll \Omega$, where the diamagnetic velocity \mathbf{v}_D [$\equiv (cT/eB)\hat{\mathbf{b}} \times \nabla \ln P$] is caused by a perpendicular gradient in plasma pressure P and $\Omega \equiv eB/m_i c$ is the ion gyrofrequency. Using some typical plasma parameters (temperature $T = 10$ keV and magnetic field $B = 50$ kG with a typical gradient length scale $L \sim 100$ cm), the thermal ion gyroradius is $\rho_i \sim 0.2$ cm and the frequency ratio is $\omega_*/\Omega = (k_\theta \rho_i) \rho_i / L \sim 10^{-3}$, where $k_\theta \sim 1 \text{ cm}^{-1}$ denotes the poloidal component of the wave vector \mathbf{k} (see Fig. 1). Its correlation lengths in both the radial and poloidal directions are on the order of several ion gyroradii, much shorter than the macroscopic gradient-scale length L (Mazzucato, 1982; Fonck *et al.*, 1993; McKee *et al.*, 2003). Its wavelength (or correlation length) along the equilibrium magnetic field is rarely measured, in particular inside the last closed magnetic surface. Some measurements at the scrape-off layer of tokamak plasmas

⁴MHD instabilities sometimes lead to a catastrophic termination of a plasma discharge called a disruption, or otherwise severely limit the performance of plasmas.

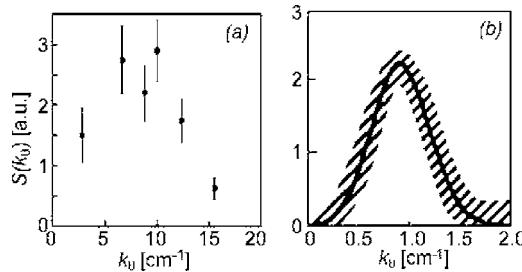


FIG. 1. Spatial wave-number spectra obtained from spatial correlation coefficients in the poloidal direction for (a) the Adiabatic Toroidal Compression tokamak (Mazzucato, 1982) and (b) the Tokamak Fusion Test Reactor (Fonck *et al.*, 1993).

indicate that it is much less than the connection length qR (Zweben and Medley, 1989; Endler *et al.*, 1995), where q denotes the safety factor and R denotes the major radius of the plasma torus. Last, the relative density fluctuation level $\delta n/n_0$ ranges typically from well under 1% at the core (near the magnetic axis) (Mazzucato *et al.*, 1996; Nazikian *et al.*, 2005) to the order of 10% at the edge (see Fig. 2). Although fluctuations in electric and magnetic fields in the interior of tokamaks are rarely measured, estimates indicate that $e\delta\phi/T_e \sim \delta n/n_0$ and $|\delta\mathbf{B}|/B_0 \sim 10^{-4}$.

From these spatiotemporal scales of tokamak microturbulence, one can make a very rough estimate of the transport coefficient D_{turb} using a dimensional analysis based on a random-walk argument,

$$D_{\text{turb}} \sim \frac{(\Delta r)^2}{\Delta t} \sim \frac{\Delta\omega}{k_r^2} \propto \frac{\omega_*}{k_r^2} \sim \left(\frac{k_\theta}{k_r^2 \rho_i} \right) \frac{\rho_i c T_i}{L e B}.$$

While this elementary estimate reveals neither dynamical insights on the detailed transport process by random $E \times B$ motion nor the significant role played by fluctuation amplitude in determining transport scalings, it illustrates the basic relation between the spatiotemporal scales of fluctuations and transport scaling. More detailed heuristic discussions have been given by Horton (1999) and Krommes (2006). If it is assumed that $k_\theta \sim k_r \propto \rho_i^{-1}$ as observed in gyrokinetic simulations with self-generated zonal flows (see Table I) and in some ex-

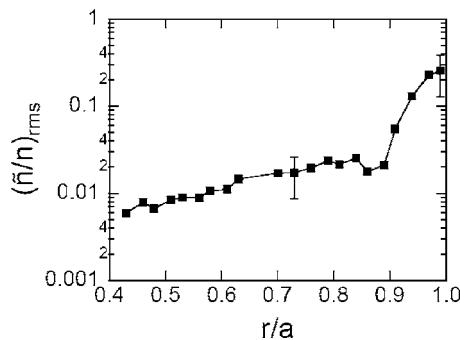


FIG. 2. Spatial profile of the total rms density-fluctuation amplitude obtained by beam emission spectroscopy on the Tokamak Fusion Test Reactor (Fonck *et al.*, 1993).

periments (Fonck *et al.*, 1993; McKee *et al.*, 2003), then

$$D_{\text{turb}} \sim \frac{\rho_i c T_i}{L e B}.$$

This is called gyro-Bohm scaling because the Bohm scaling ($\sim c T_i / e B$) is reduced by a factor $\rho_i/L \ll 1$, the ratio of gyroradius to a macroscopic length scale. This scaling is expected when local physics dominates. It can be modified due to a variety of mesoscale phenomena (Itoh and Itoh, 2001) such as turbulence spreading⁵ and avalanches.⁶

Some early simulations without self-generated zonal flows (see Table I) reported a radial step $\Delta r \propto \sqrt{L \rho_i} \gg \rho_i$, with $k_\theta \propto \rho_i^{-1}$, consistent with the Bohm scaling $D_{\text{turb}} \sim c T_i / e B$. While the distinction between Bohm scaling and gyro-Bohm scaling may seem obvious, it is complicated by many subtle issues (Lin *et al.*, 2002; Waltz *et al.*, 2002).

The nonlinear gyrokinetic formalism pursues a *dynamical reduction* of the original Vlasov-Maxwell equations for both computational and analytic feasibility while retaining the general description of the relevant physical phenomena. In this section, the standard nonlinear gyrokinetic ordering (Frieman and Chen, 1982) is described with an emphasis on physical motivations. The understanding of microturbulence based on experimental observations was incomplete when earlier versions of nonlinear gyrokinetic equations were being developed. The original motivation of the ordering may have been somewhat different than presented here. The adiabatic invariance of the new magnetic moment $\bar{\mu}$ is established on the fundamental fluctuation-based space-time orderings $\omega \ll \Omega$ and $|\mathbf{k}_\perp| \gg L^{-1}$, which have a broad experimental basis for plasma instabilities in strongly magnetized plasmas.

The nonlinear gyrokinetic Vlasov-Maxwell equations are traditionally derived through a multiple space-time-scale expansion that relies on the existence of one or more small (dimensionless) ordering parameters (Frieman and Chen, 1982). These ordering parameters are defined in terms of the following characteristic physical parameters associated with the background magnetized plasma (represented by the Vlasov distribution F and the magnetic field \mathbf{B}) and fluctuation fields (represented by the perturbed Vlasov distribution δF and the perturbed electric and magnetic fields $\delta \mathbf{E}$ and $\delta \mathbf{B}$): ω is the characteristic fluctuation frequency, k_\parallel is the characteristic fluctuation parallel wave number, $|\mathbf{k}_\perp|$ is the characteristic fluctuation perpendicular wave number, Ω is the

⁵See Garbet *et al.*, 1994; Kim *et al.*, 2003; Chen *et al.*, 2004; Hahm, Diamond, *et al.*, 2004; Lin and Hahm, 2004; Villard, Angelino, *et al.*, 2004b; Zonca *et al.*, 2004; Gurcan *et al.*, 2005; Hahm *et al.*, 2005; Itoh *et al.*, 2005; Naulin *et al.*, 2005; Waltz and Candy, 2005; Gurcan *et al.*, 2006a, 2006b; Wang *et al.*, 2006; Yagi *et al.*, 2006.

⁶See Diamond and Hahm, 1995; Newman *et al.*, 1996; Garbet and Waltz, 1998; Naulin *et al.*, 1998; Sarazin and Ghendrih, 1998; Politzer *et al.*, 2002.

ion cyclotron frequency, $v_{\text{th}} = \sqrt{T_i/m_i}$ is the ion thermal speed, $\rho_i = v_{\text{th}}/\Omega$ is the ion thermal gyroradius, L_B is the characteristic background magnetic-field nonuniformity length scale, and L_F is the characteristic background plasma density and temperature nonuniformity length scale.

The background plasma is described in terms of the (guiding-center) small parameter $\epsilon_B \equiv \rho_i/L_B$ as

$$|\rho_i \nabla \ln B| \sim \epsilon_B \quad \text{and} \quad \left| \frac{1}{\Omega} \frac{\partial \ln B}{\partial t} \right| \sim \epsilon_B^3, \quad (5)$$

where the background time-scale ordering (ϵ_B^3) is consistent with the transport time-scale ordering (Hinton and Hazeltine, 1976): The background Vlasov distribution F satisfies a similar space-time ordering⁷ with $\epsilon_F \equiv \rho_i/L_F$. The background magnetized plasma is treated as a static, nonuniform medium that is perturbed by low-frequency electromagnetic fluctuations characterized by short wavelengths perpendicular to the background magnetic field and long wavelengths parallel to it.

The fluctuating fields ($\delta f, \delta \mathbf{E}, \delta \mathbf{B}$) are described in terms of two space-time ordering parameters ($\epsilon_\perp, \epsilon_\omega$),

$$|\mathbf{k}_\perp| \rho_i \equiv \epsilon_\perp \sim 1 \quad \text{and} \quad \frac{\omega}{\Omega} \sim \epsilon_\omega \ll 1. \quad (6)$$

While microturbulence spectra of present-day high-temperature plasmas typically peak around $\epsilon_\perp \sim 0.1\text{--}0.2 < 1$ at nonlinear saturation, the linear growth rates are often highest at $\epsilon_\perp \sim 1$, in particular for trapped-electron-driven turbulence. Since shorter-wavelength fluctuations can affect longer-wavelength modes via nonlinear interactions, it is desirable to have an accurate description of the relatively short-wavelength fluctuations ($\epsilon_\perp \sim 1$). Since it is important to have an ordering in which a strong wave-particle interaction (e.g., Landau damping) is captured at the lowest order, we require that $\omega \sim |k_\parallel| v_{\text{th}}$ and have the ordering

$$\frac{|k_\parallel|}{|\mathbf{k}_\perp|} \sim \frac{\epsilon_\omega}{\epsilon_\perp}. \quad (7)$$

The most dangerous plasma instabilities in a strong magnetic field tend to satisfy the parallel ordering $k_\parallel \ll k_\perp$ (with $\epsilon_\omega \ll \epsilon_\perp$). The relative fluctuation levels are described in terms of the amplitude ordering parameter ϵ_δ ,

$$\left| \frac{\delta f}{F} \right| \sim \frac{c |\delta \mathbf{E}_\perp|}{B v_{\text{th}}} \sim \frac{|\delta \mathbf{B}|}{B} \sim \epsilon_\delta \ll 1. \quad (8)$$

The electric fluctuation ordering,

⁷The ratio $\epsilon_B/\epsilon_F < 1$ can be used to formally define an auxiliary ordering parameter known as the inverse-aspect-ratio parameter a/R in toroidal magnetized plasmas, where a and R denote the minor and major radii; we shall not make use of this auxiliary ordering parameter in the present work and, henceforth, assume that $\epsilon_F \sim \epsilon_B$.

$$\epsilon_\delta \sim \frac{c |\delta \mathbf{E}_\perp|}{B v_{\text{th}}} \sim \epsilon_\perp \frac{e \delta \phi}{T_i}, \quad (9)$$

implies that, for $\epsilon_\perp \sim 1$ and $T_e \sim T_i$, we have $e \delta \phi / T_e \sim \epsilon_\delta$. A covariant description of electromagnetic fluctuations requires that the parallel component $\delta A_\parallel \equiv \hat{\mathbf{b}} \cdot \delta \mathbf{A}$ of the perturbed vector potential satisfies the amplitude ordering $(v_\parallel/c) \delta A_\parallel \sim \delta \phi$, where the parallel particle velocity is $v_\parallel \sim v_{\text{th}}$, so that

$$\epsilon_\delta \sim \frac{v_\parallel e \delta A_\parallel}{c} \sim \frac{|\delta \mathbf{B}_\perp|}{\epsilon_\perp B}, \quad (10)$$

which implies that, for $\epsilon_\perp \sim 1$, we have $|\delta \mathbf{B}_\perp|/B \sim \epsilon_\delta$. Hence, the orderings (9) and (10) imply that the terms $|k_\parallel| \delta \phi$ and $(\omega/c) \delta A_\parallel$ in the parallel perturbed electric field δE_\parallel have similar orderings (for $\epsilon_\perp \sim 1$),

$$\frac{|k_\parallel| \delta \phi}{(\omega/c) \delta A_\parallel} \sim \frac{|k_\parallel|/|\mathbf{k}_\perp|}{\omega/\Omega} \sim \frac{1}{\epsilon_\perp},$$

so that $|\delta E_\parallel|/|\delta \mathbf{E}_\perp| \sim |k_\parallel|/|\mathbf{k}_\perp| \sim \epsilon_\omega/\epsilon_\perp \ll 1$. Last, for a fully electromagnetic gyrokinetic ordering (and high- β plasmas, where $\beta \equiv 8\pi P/B^2$), we require that $|\delta B_\parallel|/B \sim \beta \epsilon_\delta$ (Brizard, 1992);⁸ note that we use the electromagnetic gauge condition $\nabla_\perp \cdot \delta \mathbf{A}_\perp = 0$, which represents the low-frequency gyrokinetic limit of the Lorenz gauge $c^{-1} \partial_t \delta \phi + \nabla \cdot \delta \mathbf{A} = 0$. While ϵ_ω and ϵ_δ are comparable in practice (e.g., $\epsilon_\omega \sim \epsilon_\delta \sim 10^{-3}$), it is useful to keep these parameters separate for ordering purposes and greater flexibility. Note that, because of the perpendicular ordering $\epsilon_\perp \sim 1$, full finite-Larmor-radius (FLR) effects must be retained in the nonlinear gyrokinetic formalism.

The regimes of validity of various drift-kinetic and gyrokinetic theories are summarized in Fig. 3 in terms of the normalized electrostatic potential $(L/\rho_i) e \delta \phi / T_e \sim \epsilon_\delta / \epsilon_\perp \epsilon_B$ (Dimits *et al.*, 1992; Hazeltine and Hinton, 2005). Note that one needs $\epsilon_\delta \ll 1$ for any perturbative nonlinear kinetic equations: for drift-kinetic theories, we require $\epsilon_\perp \ll 1$, while for gyrokinetic theories ($\epsilon_\perp \sim 1$), we require $\epsilon_\delta \sim \epsilon_B$. The latter ordering implies that the linear drive term ($\sim \hat{\mathbf{b}} \times \nabla \delta \phi \cdot \nabla F$) is of the same order as the nonlinear $E \times B$ coupling term ($\sim \hat{\mathbf{b}} \times \nabla \delta \phi \cdot \nabla \delta f$). This is a generic situation for strong turbulence, which yields a nonlinear saturation roughly at a *mixing length* level $\delta n \sim (\rho_i/L)n$; nonetheless, with a subsidiary ordering, the nonlinear gyrokinetic equations can describe weak turbulence as well (Sagdeev and Galeev, 1969).

B. Frieman-Chen nonlinear gyrokinetic equation

The first significant work on nonlinear gyrokinetic equations in general magnetic geometry was by Frieman and Chen (1982), who used a conventional approach

⁸While $e \delta \phi / T_e \gg |\delta \mathbf{B}_\perp|/B \gg |\delta B_\parallel|/B$ for typical low-to-modest β tokamak plasmas, the fluctuation ordering (8) is retained for its generality (which makes it applicable to high- β devices such as spherical tori).

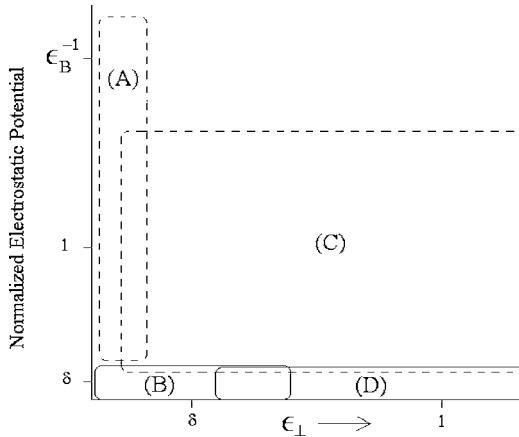


FIG. 3. Regimes of validity of nonlinear (A) and linear (B) drift-kinetic and nonlinear (C) and linear (D) gyrokinetic theories displayed on a plot of normalized electrostatic potential $(L/\rho_i)e\delta\phi/T_e \sim \epsilon_\delta/\epsilon_\perp\epsilon_B$ vs ϵ_\perp , where the ordering parameters $(\epsilon_\delta, \epsilon_\perp, \epsilon_B)$ are defined in the text and $\delta \ll 1$ denotes an ordering parameter that distinguishes linear from nonlinear theories or drift-kinetic from gyrokinetic theories.

based on a maximal multiple-scale-ordering expansion involving a single ordering parameter $\epsilon (\epsilon \sim \epsilon_B \sim \epsilon_\omega \sim \epsilon_\delta)$. Here the linear-physics-driven terms are ordered at $\epsilon_\omega\epsilon_\delta$ and $\epsilon_B\epsilon_\delta$ (which recognizes the crucial role played by the background magnetic-field nonuniformity with $\epsilon_B \sim \epsilon_\omega$), while the nonlinear coupling terms are ordered at ϵ_δ^2 . The main purpose of the Frieman-Chen (FC) gyrokinetic equations was for analytic applications and it has served its original motivation during the past two decades. For instance, many nonlinear kinetic theories of tokamak microturbulence [see Table I and references in Horton (1999)] have used the FC equations as the starting point. The number of assumptions on the general FC ordering was minimal at least in the context of nonlinear gyrokinetics (we elaborate on this point later on).

The material presented here only summarizes some aspects of the work of Frieman and Chen (1982) relevant to our discussion (and we use notation consistent with the remainder of our paper). We begin with the Vlasov equation

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \frac{d\mathbf{z}}{dt} \cdot \frac{\partial f}{\partial \mathbf{z}} = 0, \quad (11)$$

where $\mathbf{z}=(\mathbf{x}, \mathbf{p})$ denote the particle phase-space coordinates and $f(\mathbf{z}, t)$ denotes the Vlasov particle distribution. Here the Vlasov equation (11) states that the particle distribution $f(\mathbf{z}(t; \mathbf{z}_0), t) = f(\mathbf{z}_0; 0)$ is a constant along an exact particle orbit $\mathbf{z}(t; \mathbf{z}_0)$, where $\mathbf{z}_0 \equiv \mathbf{z}(0; \mathbf{z}_0)$ denotes the orbit's initial condition.

Following the conventional (iterative) approach initially used by Hastie *et al.* (1967), Frieman and Chen (1982) introduce decompositions of the Vlasov distribution $f = F + \epsilon_\delta \delta f$ and the particle's equations of motion $d\mathbf{z}/dt = d\mathbf{Z}/dt + \epsilon_\delta \delta \mathbf{z}/dt$, where $(F, d\mathbf{Z}/dt)$ represent the background plasma dynamics and $(\delta f, d\delta \mathbf{z}/dt)$ represent the perturbed plasma dynamics associated with the pres-

ence of short-wavelength fluctuating electromagnetic fields $\delta \mathbf{E} = -\nabla \delta \phi - c^{-1} \partial_t \delta \mathbf{A}$ and $\delta \mathbf{B} = \nabla \times \delta \mathbf{A}$.

Next, Frieman and Chen introduce a *short-space-scale* averaging (denoted by an overbar) with the definitions $\bar{f} \equiv F$ and $\overline{(d\mathbf{z}/dt)} \equiv d\mathbf{Z}/dt$. Hence, the short-space-scale average of the Vlasov equation (11) yields

$$\frac{\partial F}{\partial t} + \frac{d\mathbf{Z}}{dt} \cdot \frac{\partial F}{\partial \mathbf{z}} \equiv \frac{DF}{Dt} = -\epsilon_\delta^2 \overline{\left(\frac{d\delta \mathbf{z}}{dt} \cdot \frac{\partial \delta f}{\partial \mathbf{z}} \right)}, \quad (12)$$

which describes the long-time evolution of the background Vlasov distribution F as a result of the background plasma dynamics (represented by the averaged Vlasov operator D/Dt) and the nonlinear (ϵ_δ^2) short-wavelength-averaged (quasilinear) influence of the fluctuating fields. Here, Frieman and Chen (1982) expand the solution $F = F_0 + \epsilon_B F_1 + \dots$ for the background Vlasov equation (12) up to first order in ϵ_B and without the nonlinear fluctuation-driven term ($\epsilon_\delta = 0$), where the gyroangle-dependent part of the first-order correction

$$\epsilon_B \tilde{F}_1 = - \left(\int \frac{d\zeta}{\Omega} \dot{\mu}_0 \right) \frac{\partial F_0}{\partial \mu} \quad (13)$$

is expressed in terms of the lowest-order distribution $F_0(\mathbf{X}_\perp, \mathcal{E}, \mu)$, which is a function of the perpendicular components of the guiding-center position \mathbf{X}_\perp (i.e., $\hat{\mathbf{b}} \cdot \nabla F_0 = 0$), the (lowest-order) guiding-center kinetic energy \mathcal{E} (with $\mathcal{E}_0 = 0$), and the (lowest-order) guiding-center magnetic moment μ . In Eq. (13), $\dot{\mu}_0$ denotes the time derivative of the lowest-order magnetic moment, which is ordered at ϵ_B for a time-independent magnetic field and is explicitly gyroangle dependent.⁹

By subtracting the averaged Vlasov equation (12) from the Vlasov equation (11), we obtain the *fluctuating* Vlasov equation

$$\frac{D\delta f}{Dt} = - \frac{d\delta \mathbf{z}}{dt} \cdot \frac{\partial F}{\partial \mathbf{z}} - \frac{d\delta \mathbf{z}}{dt} \cdot \frac{\partial \delta f}{\partial \mathbf{z}} + \overline{\left(\frac{d\delta \mathbf{z}}{dt} \cdot \frac{\partial \delta f}{\partial \mathbf{z}} \right)}, \quad (14)$$

where the left side contains terms of order $\epsilon_\omega\epsilon_\delta$ and $\epsilon_B\epsilon_\delta$, while the first term on the right side provides the linear drive for δf (at order $\epsilon_B\epsilon_\delta$) and the remaining terms involve the short-spatial-scale nonlinear coupling (at order ϵ_δ^2). Note that a quasilinear formulation is obtained from Eq. (14) by retaining only the first term on the right side and substituting the (eikonal) solution for δf (as a functional of F) into the averaged Vlasov equation (12).

Next, Frieman and Chen (1982) introduce a decomposition of the perturbed Vlasov distribution δf in terms of

⁹Using (\mathcal{E}, μ) rather than $(v_{||}, \mu)$ as the velocity-space coordinates reduces the number of nonzero terms when either F_0 is isotropic in velocity space (i.e., $\partial F_0 / \partial \mu = 0$ at constant \mathcal{E}), or the electromagnetic fields are time independent such that $\dot{\mathcal{E}} = 0$. Thus, it can sometimes be advantageous to the $(v_{||}, \mu)$ formulation. However, for more complex realistic nonlinear applications, we find the $(v_{||}, \mu)$ formulation more straightforward in describing physics.

its *adiabatic* and *nonadiabatic* components (Antonsen and Lane, 1980; Catto *et al.*, 1981),

$$\delta f \equiv \left[e\delta\phi \frac{\partial}{\partial \mathcal{E}} + \frac{e}{B} \left(\delta\phi - \frac{v_{\parallel}}{c} \delta A_{\parallel} \right) \frac{\partial}{\partial \mu} \right] F_0 + e^{-\rho} \nabla \cdot \left(\delta g - \frac{e \langle \delta\psi_{gc} \rangle}{B} \frac{\partial F_0}{\partial \mu} \right). \quad (15)$$

Here δg denotes the gyroangle-independent nonadiabatic part of the perturbed Vlasov distribution, $\delta A_{\parallel} \equiv \delta \mathbf{A} \cdot \hat{\mathbf{b}}$ denotes the component of the perturbed vector potential parallel to the background magnetic field $\mathbf{B} = B\hat{\mathbf{b}}$ (v_{\parallel} denotes the parallel component of the guiding-center velocity), $\langle \dots \rangle$ denotes gyroangle averaging (ρ denotes the lowest-order gyroangle-dependent gyroradius vector), and the effective first-order gyroaveraged potential is

$$\langle \delta\psi_{gc} \rangle \equiv \left\langle e^{\rho} \nabla \left(\delta\phi - \frac{\mathbf{v}}{c} \cdot \delta \mathbf{A} \right) \right\rangle = \left\langle \delta\phi_{gc} - \frac{\mathbf{v}}{c} \cdot \delta \mathbf{A}_{gc} \right\rangle. \quad (16)$$

Note that all terms on the right side of Eq. (15) are ordered at ϵ_{δ} and an additional higher-order adiabatic term $\nabla F_0 \cdot \delta \mathbf{A} \times \hat{\mathbf{b}} / B$ (of order $\epsilon_B \epsilon_{\delta}$) has been omitted. For the sake of clarity in the discussion presented below, we refer to the adiabatic terms involving the perturbed potentials ($\delta\phi, \delta A_{\parallel}$) evaluated at the particle position as the *particle* adiabatic terms, while the adiabatic term involving the effective first-order Hamiltonian (16), where perturbed potentials are evaluated at the guiding-center position, as the *guiding-center* adiabatic term.

By substituting the nonadiabatic decomposition (15) into the fluctuating Vlasov equation (14), Frieman and Chen (1982) obtain the nonlinear gyrokinetic equation for the nonadiabatic part δg of the perturbed Vlasov distribution,

$$\begin{aligned} \frac{d_{gc}\delta g}{dt} = & - \left(e \frac{\partial \langle \delta\psi_{gc} \rangle}{\partial t} \frac{\partial}{\partial \mathcal{E}} + \frac{c\hat{\mathbf{b}}}{B} \times \nabla \langle \delta\psi_{gc} \rangle \cdot \nabla \right) F_0 \\ & - \frac{c\hat{\mathbf{b}}}{B} \times \nabla \langle \delta\psi_{gc} \rangle \cdot \nabla \delta g, \end{aligned} \quad (17)$$

where $D/Dt \equiv d_{gc}/dt$ denotes the unperturbed averaged Vlasov operator expressed in guiding-center coordinates $(\mathbf{X}, \mathcal{E}, \mu)$. Note that the time evolution of the nonadiabatic part δg depends solely on the effective first-order Hamiltonian (16). The terms appearing on the left side of Eq. (17), as well as the F_0 terms on the right side, are ordered at $\epsilon_{\omega} \epsilon_{\delta}$ and $\epsilon_B \epsilon_{\delta}$, while the last term on the right side is ordered at ϵ_{δ}^2 and thus represents the nonlinear coupling terms, which are absent from previous linear gyrokinetic models (Antonsen and Lane, 1980; Catto *et al.*, 1981). The nonlinear coupling terms include the (linear) perturbed $E \times B$ velocity $(c\hat{\mathbf{b}}/B) \times \nabla \langle \delta\phi_{gc} \rangle$, the magnetic flutter velocity $(v_{\parallel}/B) \langle \delta \mathbf{B}_{\perp gc} \rangle$, and the perturbed grad- B drift velocity $(-\hat{\mathbf{b}}/B) \times \nabla \langle \mathbf{v}_{\perp} \cdot \delta \mathbf{A}_{\perp gc} \rangle$. The nonlin-

ear gyrokinetic Vlasov equation (48) derived by Frieman and Chen (1982) contains additional terms, defined in their equation (45), that are subsequently omitted in their final equation (50) shown above as Eq. (17).

A self-consistent description of nonlinear gyrokinetic dynamics requires that the Maxwell equations for the perturbed electromagnetic fields $\delta \mathbf{E}$ and $\delta \mathbf{B}$ be expressed in terms of particle charge and current densities represented as fluid moments of the nonadiabatic part δg of the perturbed Vlasov distribution (15). For example, using the nonadiabatic decomposition (15), the perturbed particle fluid density $\delta n \equiv \int d^3 p \delta f$ is

$$\delta n = \int d^3 P \left\langle e^{-\rho} \nabla \left(\delta g - \frac{e \langle \delta\psi_{gc} \rangle}{B} \frac{\partial F_0}{\partial \mu} \right) \right\rangle, \quad (18)$$

where $d^3 P$ denotes the (\mathcal{E}, μ) integration in guiding-center coordinates of a gyroangle-averaged integrand. Note that the particle adiabatic terms in Eq. (15) have canceled out of Eq. (18) and only the guiding-center adiabatic and nonadiabatic terms contribute to δn . We show later that this guiding-center adiabatic contribution leads to the so-called polarization density (see Appendix C.2); a similar treatment for the perturbed particle moment $\int d^3 p \mathbf{v} \delta f$ leads to the cancellation of particle adiabatic terms and the definition of the magnetization current in terms of the guiding-center adiabatic and nonadiabatic terms.

The Frieman-Chen nonlinear gyrokinetic Vlasov equation (17) is contained in modern versions of the nonlinear gyrokinetic Vlasov equation (Brizard, 1989a), as shown in Sec. V.E. The Frieman-Chen formulation contains the polarization density (while there was no explicit mention about it in the FC paper). For most analytic applications (see Table I) a separate treatment of this term is not necessary. However, an explicit treatment of the polarization density as the dominant shielding term in the gyrokinetic Poisson equation (Lee, 1983) has provided computational advantage (Lee, 1987) in nonlinear gyrokinetic simulations. It is also a key quantity in relating the nonlinear gyrokinetic approach to reduced magnetohydrodynamics (Hahm *et al.* 1988; Brizard, 1992) (see Sec. III.E).

The major difficulties encountered in the conventional Frieman-Chen derivation of Eq. (17) involve (a) inserting the solution (13) for the first-order correction F_1 to the background Vlasov distribution F_0 into the first term on the right of Eq. (14) and (b) constructing a new magnetic moment $\bar{\mu}$ that is invariant at first order in ϵ_B and ϵ_{δ} . While the Frieman-Chen equations are valid up to order ϵ^2 and for practical purposes,¹⁰ including initial interactions of linear modes and the early phase following nonlinear saturation, their work did not consider preserving the conservation laws of the original Vlasov-Maxwell equations (e.g., total energy and momentum).

¹⁰The Frieman-Chen paper was published in 1982 at a time when computer power and plasma-diagnostic capabilities were far lower than the present-day equivalents.

For instance, the sum of the kinetic energy and field energy, as well as phase-space volume, are not conserved up to the nontrivial order. A lack of phase-space-volume conservation can introduce fictitious dissipation that can affect the long-term behavior of the (presumed) Hamiltonian system. Moreover, by ignoring the $\mathcal{O}(\epsilon^3)$ nonlinear wave-particle interactions due to parallel-velocity-space nonlinearity, one cannot describe phase-space structures such as clumps (Dupree, 1972) and holes (Dupree, 1982; Terry *et al.*, 1990). Ignoring the nonlinear wave-particle interactions can also artificially limit the energy exchange between particles and waves.

C. Modern nonlinear gyrokinetic equations

In contrast to conventional methods used for deriving nonlinear gyrokinetic equations (based on a regular perturbation expansion in terms of small parameters and a direct gyrophase average), the modern nonlinear gyrokinetic derivation pursues a reduction of dynamical dimensionality via phase-space coordinate transformations. The modern derivation of the nonlinear gyrokinetic Vlasov equation is based on the construction of a time-dependent phase-space transformation from (old) particle coordinates $\mathbf{z}=(\mathbf{x}, \mathbf{p})$ to (new) gyrocenter phase-space coordinates $\bar{\mathbf{Z}}$ (to be defined later) such that the new *gyrocenter* equations of motion $d\bar{\mathbf{Z}}/dt$ are independent of the fast gyromotion time scale at arbitrary orders in ϵ_B and ϵ_δ . This transformation effectively decouples the fast gyromotion time scale from the slow *reduced* time scales. In the course of this derivation, the important underlying symmetry and conservation laws of the original system are retained. In contrast to the conventional derivation (where different small parameters are lumped together via a particular ordering), various expansion parameters appear at different stages of the derivation. This feature makes modifications of ordering for specific applications more transparent, such as nonlinear gyrokinetic equations with strong $E \times B$ shear flows as described in Appendix D.1.

The phase-space transformation $\mathbf{z} \rightarrow \bar{\mathbf{Z}}$ is formally expressed in terms of an asymptotic expansion in powers of the perturbation-amplitude ordering parameter ϵ_δ ,

$$\mathbf{z} \equiv \sum_{n=0} \epsilon_\delta^n \bar{\mathbf{Z}}_n(\mathbf{z}), \quad (19)$$

where the lowest-order term $\bar{\mathbf{Z}}_0(\mathbf{z})$ is expressed in terms of an asymptotic expansion in powers of the background-plasma ordering parameter ϵ_B associated with the guiding-center transformation (see Fig. 4). The original particle dynamics $d\mathbf{z}/dt$ in Eq. (11) can be represented as a Hamiltonian system $d\mathbf{z}/dt \equiv \{\mathbf{z}, H\}_z$, where $H(\mathbf{z}, t)$ denotes the particle Hamiltonian and $\{\cdot, \cdot\}_z$ denotes the Poisson bracket in particle phase space with coordinates \mathbf{z} (which are, generically, noncanonical). Since Hamiltonian systems have important conservation properties, e.g., the Liouville theorem associated with the invariance of the phase-space volume under Hamil-

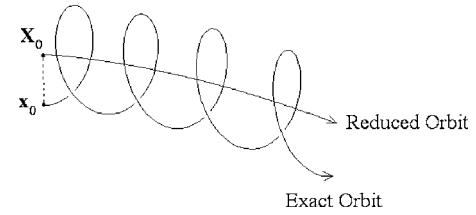


FIG. 4. Exact and reduced single-particle orbits in a magnetic field.

tonian evolution (Goldstein *et al.*, 2002), the new equations of motion $d\bar{\mathbf{Z}}/dt$ must also be expressed as a Hamiltonian system in terms of a new Hamiltonian $\bar{H}(\bar{\mathbf{Z}}, t)$ and a new Poisson bracket $\{\cdot, \cdot\}_{\bar{\mathbf{Z}}}$ such that $d\bar{\mathbf{Z}}/dt \equiv \{\bar{\mathbf{Z}}, \bar{H}\}_{\bar{\mathbf{Z}}}$.

The phase-space transformation $\mathbf{z} \rightarrow \bar{\mathbf{Z}}$ induces a transformation from the (old) particle Vlasov distribution f to a (new) reduced Vlasov distribution \bar{F} , subject to the scalar-invariance property $\bar{F}(\bar{\mathbf{Z}}) = f(\mathbf{z})$, such that the new Vlasov distribution \bar{F} is constant along a reduced orbit $\bar{\mathbf{Z}}(t)$. From the scalar-invariance property, the induced transformation $f \rightarrow \bar{F}$ is

$$f(\mathbf{z}) \equiv \bar{F}(\bar{\mathbf{Z}}) = \bar{F}\left(\sum_{n=0} \epsilon_\delta^n \bar{\mathbf{Z}}_n(\mathbf{z})\right), \quad (20)$$

that generates an asymptotic expansion in powers of ϵ_δ ,

$$f \equiv \sum_{n=0} \epsilon_\delta^n f_n(\bar{F}), \quad (21)$$

where each term $f_n(\bar{F})$ is expressed in terms of derivatives of the new Vlasov distribution \bar{F} . The infinitesimal constant translation $x \rightarrow X = x + \epsilon$ and the induced transformation $f \rightarrow F$ satisfies

$$f(x) \equiv F(X) = F(x + \epsilon) = \sum_{n=0} \frac{\epsilon^n}{n!} \frac{d^n F(x)}{dx^n} = \sum_{n=0} \epsilon^n f_n(F(x)),$$

where $f_n(F) \equiv (1/n!) d^n F / dx^n$ shows that the two functions f and F are formally different functions. The non-adiabatic decomposition (15) is a similar asymptotic expansion $f = f_0 + \epsilon_\delta \delta f + \dots$, where f_0 and δf are expressed in terms of a reduced Vlasov distribution \bar{F} .

For the new Vlasov kinetic theory to be dissipation-free, the phase-space transformation (19) must be invertible (i.e., entropy-conserving) so that no information on the fast-time-scale particle dynamics is lost. Hence, the following inverse relations are required:

$$\mathbf{z} \equiv \sum_{n=0} \epsilon_\delta^n \mathbf{z}_n(\bar{\mathbf{Z}}), \quad (22)$$

$$\bar{F}(\bar{\mathbf{Z}}) \equiv f(\mathbf{z}) = f\left(\sum_{n=0} \epsilon_\delta^n \mathbf{z}_n(\bar{\mathbf{Z}})\right), \quad (23)$$

$$\bar{F} \equiv \sum_{n=0} \epsilon_\delta^n \bar{F}_n(f). \quad (24)$$

These are justified by the smallness of the ordering parameter $\epsilon_\delta \ll 1$. Within canonical Hamiltonian perturbation theory (Goldstein *et al.*, 2002) the relation between the (old) particle Hamiltonian H and the (new) reduced Hamiltonian \bar{H} is expressed as

$$\bar{H}(\bar{\mathbf{Z}}, t) \equiv H(\mathbf{z}, t) - \frac{\partial S}{\partial t}(\mathbf{z}, t), \quad (25)$$

where S denotes the scalar field that generates the time-dependent canonical transformation $\mathbf{z} \rightarrow \bar{\mathbf{Z}}$ and each function (H, S, \bar{H}) is itself expressed as a power expansion in ϵ_δ (and ϵ_B).

The modern formulation of the nonlinear gyrokinetic Vlasov theory provides powerful algorithms to construct the phase-space transformations (19) and (22) and induced transformations (21) and (24). These algorithms are based on applications of differential geometric methods associated with Lie-transform operators (see Appendix A for a primer on these mathematical methods).

III. SIMPLE FORMS OF NONLINEAR GYROKINETIC EQUATIONS

Simplified forms of the nonlinear gyrokinetic equations recommended for simulations and analytic applications in laboratory (e.g., fusion) and space plasma physics as well as astrophysics are presented. This section provides a quick reference to readers who are mainly interested in applications of nonlinear gyrokinetic formulations, rather than theoretical derivations of mathematical structures thereof. The full nonlinear gyrokinetic equations will be derived in Secs. V and VI.

In this review, recent remarkable progress in nonlinear gyrokinetic simulations (Tang, 2002; Tang and Chan, 2005) is not covered. The physics issues that arise when nonlinear gyrokinetic equations are simplified and applied to specific collective waves and instabilities in magnetized plasmas are described. The basic physics underlying the magnetic confinement of plasmas has been described by Boozer (2004). A relatively complete survey of fusion-relevant instabilities can be found in Connor and Wilson (1994) and Horton (1999), while a partial summary of applications of nonlinear gyrokinetic formulations is provided in Table I. Some early simulations were performed as the modern nonlinear gyrokinetic formulations were being developed. Not all of the gyrokinetic theoretical knowledge discussed in this review article was available then. The overbar notation used to identify the gyrocenter coordinates is omitted for the remainder of this section (unless otherwise needed).

A. General gyrokinetic Vlasov-Maxwell equations

For nonlinear simulations, the nonlinear gyrokinetic Vlasov equation (2) for the gyrocenter distribution F is

written in terms of the Hamiltonian gyrocenter equations of motion,

$$\frac{d\mathbf{X}}{dt} = v_\parallel \frac{\mathbf{B}^*}{B_\parallel^*} + \frac{c\hat{\mathbf{b}}}{eB_\parallel^*} \times (\mu \nabla B + e \nabla \delta\Psi_{gy}) \quad (26)$$

and

$$\frac{dp_\parallel}{dt} = - \frac{\mathbf{B}^*}{B_\parallel^*} \cdot (\mu \nabla B + e \nabla \delta\Psi_{gy}) - \frac{e}{c} \frac{\partial \delta A_{\parallel gy}}{\partial t}, \quad (27)$$

where the effective gyrocenter perturbation potential $\delta\Psi_{gy}$ contains terms that are linear in the perturbed electromagnetic potentials ($\delta\phi, \delta\mathbf{A}$) [e.g., the effective linear potential $\langle \delta\psi_{gc} \rangle$ defined in Eq. (16)] and terms that are nonlinear (quadratic) in $(\delta\phi, \delta\mathbf{A})$. Explicit forms for this effective gyrocenter perturbation potential are given below for limiting cases: electrostatic fluctuations (Sec. III.B), shear-Alfvénic magnetic fluctuations (Sec. III.C), and compressional-Alfvénic magnetic fluctuations (Sec. III.D). The gyrocenter Poisson-bracket (symplectic) structure is represented by the modified magnetic field (with $B_\parallel^* \equiv \hat{\mathbf{b}} \cdot \mathbf{B}^*$),

$$\mathbf{B}^* \equiv \mathbf{B} + (c/e)p_\parallel \nabla \times \hat{\mathbf{b}} + \delta\mathbf{B}_{gy}, \quad (28)$$

where the first term denotes the background magnetic field $\mathbf{B} \equiv B\hat{\mathbf{b}}$, the second term is associated with the guiding-center curvature drift, and the third term represents the symplectic magnetic perturbation $\delta\mathbf{B}_{gy} \equiv \nabla \times \delta\mathbf{A}_{gy}$ that may or may not be present depending on the choice of gyrocenter model adopted (see below).

The perturbed linear gyrocenter dynamics contained in Eq. (26) includes the linear perturbed $E \times B$ velocity $\delta\mathbf{u}_E = (c\hat{\mathbf{b}}/B) \times \nabla \delta\phi$, the perturbed magnetic-flutter velocity $v_\parallel \delta\mathbf{B}_\perp/B$, and the perturbed grad- B velocity $(c\hat{\mathbf{b}}/eB) \times \mu \nabla \delta B_\parallel$. When magnetic perturbations are present, the gyrocenter parallel momentum p_\parallel appearing in Eq. (27) is either a *canonical* momentum if the magnetic perturbation $\delta\mathbf{A}_{gy}$ is chosen so that $\delta A_{\parallel gy} = \hat{\mathbf{b}} \cdot \delta\mathbf{A}_{gy} = 0$ or a *kinetic* momentum (i.e., $p_\parallel = mv_\parallel$) if $\delta A_{\parallel gy} \neq 0$. The magnetic-flutter velocity $v_\parallel \delta\mathbf{B}_\perp/B$ is either included in $v_\parallel \delta\mathbf{B}_{gy}/B$ (in the symplectic gyrocenter model) or in $(c\hat{\mathbf{b}}/B) \times \nabla \delta\Psi_{gy}$ (in the Hamiltonian gyrocenter model) (Hahm *et al.*, 1988; Brizard, 1989a), while the inductive part $(\partial_t \delta A_{\parallel gy})$ of the perturbed electric field appears on the right side of the gyrocenter parallel force equation (27) only in the symplectic gyrocenter model. For specific application of nonlinear gyrokinetics, not all terms in Eqs. (26) and (27) are used simultaneously; all of the terms are included here for easy reference within this section.

A closed self-consistent description of the interactions involving the perturbed electromagnetic field and a Vlasov distribution of gyrocenters implies that the gyrokinetic Maxwell equations are written with charge-current densities expressed in terms of the gyrocenter distribution function. The gyrokinetic Poisson equation (3) is

$$\nabla^2 \delta\phi = -4\pi \sum e \int d^3p \langle \mathbf{T}_\epsilon \bar{F} \rangle \\ \equiv -4\pi [e(\bar{N}_i - \bar{N}_e) + \rho_{\text{pol}}], \quad (29)$$

where $\mathbf{T}_\epsilon \equiv \mathbf{T}_{\text{gc}} \mathbf{T}_{\text{gy}}$ denotes an operator (defined in Sec. IV) that transforms functions on gyrocenter phase space into functions on guiding-center phase space and then into functions on particle phase space, Σ denotes a sum over charged-particle species (e.g., ions and electrons), $\int d^3p = 2\pi mB \int dp_\parallel d\mu$ denotes an integration over the gyrocenter coordinates p_\parallel and μ , and gyroangle averaging is denoted as $\langle \cdots \rangle \equiv (2\pi)^{-1} \oint d\zeta \langle \cdots \rangle$. On the right side of Eq. (29), $\bar{N}_j \equiv \int d^3p \bar{F}_j$ is the gyrocenter density for particle species j ($=i$ or e), and the gyrocenter polarization density is

$$\rho_{\text{pol}} \equiv \sum e \int d^3p (\langle \mathbf{T}_\epsilon \bar{F} \rangle - \bar{F}) \equiv -\nabla \cdot \mathbf{P}_{\text{gy}}, \quad (30)$$

that is then expressed in terms of the divergence of the gyrocenter polarization vector \mathbf{P}_{gy} (defined below). The gyrokinetic Ampère equation (4) is

$$\nabla \times (\mathbf{B} + \delta\mathbf{B}) = \frac{4\pi}{c} \sum e \int d^3p \langle \mathbf{v} \mathbf{T}_\epsilon \bar{F} \rangle \\ \equiv \frac{4\pi}{c} [(\bar{\mathbf{J}}_i + \bar{\mathbf{J}}_e) + (\mathbf{J}_{\text{pol}} + \mathbf{J}_{\text{mag}})], \quad (31)$$

where $\bar{\mathbf{J}}_j \equiv e_j \int d^3p \langle \mathbf{T}_\epsilon^{-1} \mathbf{v} \rangle \bar{F}_j$ denotes the gyrocenter current density associated with particle species j , $\mathbf{T}_\epsilon^{-1} \mathbf{v} \equiv \mathbf{T}_{\text{gy}}^{-1} (\mathbf{T}_{\text{gc}}^{-1} \mathbf{v})$ denotes the particle velocity transformed into gyrocenter phase space, and the gyrocenter polarization and magnetization currents are defined as

$$\mathbf{J}_{\text{pol}} + \mathbf{J}_{\text{mag}} = \sum e \int d^3p \langle \mathbf{v} \mathbf{T}_\epsilon \bar{F} - \mathbf{T}_\epsilon^{-1} \mathbf{v} \bar{F} \rangle \\ \equiv \frac{\partial \mathbf{P}_{\text{gy}}}{\partial t} + c \nabla \times \mathbf{M}_{\text{gy}}, \quad (32)$$

that are expressed, respectively, in terms of the time derivative of the gyrocenter polarization vector \mathbf{P}_{gy} and the curl of the gyrocenter magnetization vector \mathbf{M}_{gy} (defined below). By definition, the gyrocenter polarization charge conservation law

$$0 = \frac{\partial \rho_{\text{pol}}}{\partial t} + \nabla \cdot \mathbf{J}_{\text{pol}} \equiv \frac{\partial}{\partial t} (-\nabla \cdot \mathbf{P}_{\text{gy}}) + \nabla \cdot \left(\frac{\partial \mathbf{P}_{\text{gy}}}{\partial t} \right) \quad (33)$$

is explicitly satisfied. A related discussion on the polarization drift in a simple limit can be found in Appendix C of Krommes (2002). This relation has been extended to the neoclassical polarization associated with trapped particles (Fong and Hahm, 1999). The gyrocenter magnetization current $\mathbf{J}_{\text{mag}} \equiv c \nabla \times \mathbf{M}_{\text{gy}}$ is explicitly divergenceless.

The variational formulation for the nonlinear gyrokinetic Vlasov-Maxwell equations reveals that these gyrokinetic polarization and magnetization effects are also associated with derivatives of the nonlinear gyrocenter

Hamiltonian $e\delta\Psi_{\text{gy}}$ with respect to the perturbed electric and magnetic fields $\delta\mathbf{E}$ and $\delta\mathbf{B}$, respectively,

$$\mathbf{P}_{\text{gy}} \equiv -\sum e \int d^3p F \left(\frac{\partial \delta\Psi_{\text{gy}}}{\partial \delta\mathbf{E}} \right) = \sum \int d^3p F \boldsymbol{\pi}_{\text{gy}}, \quad (34)$$

$$\mathbf{M}_{\text{gy}} \equiv -\sum e \int d^3p F \left(\frac{\partial \delta\Psi_{\text{gy}}}{\partial \delta\mathbf{B}} \right) \\ = \sum \int d^3p F \left(\boldsymbol{\mu}_{\text{gy}} + \boldsymbol{\pi}_{\text{gy}} \times \frac{p_\parallel}{mc} \hat{\mathbf{b}} \right), \quad (35)$$

where $\boldsymbol{\pi}_{\text{gy}}$ is the gyrocenter electric-dipole moment, $\boldsymbol{\mu}_{\text{gy}}$ is the gyrocenter magnetic-dipole moment, and the gyrocenter magnetization vector (35) includes a moving-electric-dipole contribution $\boldsymbol{\pi}_{\text{gy}} \times (p_\parallel/mc) \hat{\mathbf{b}}$ (Jackson, 1975).

The nonlinear gyrokinetic Vlasov-Maxwell equations possess an exact energy conservation law $dE/dt = 0$, where the global gyrokinetic energy integral is (Brizard, 1989a, 1989b)

$$E = \int d^6Z F \left(\frac{p_\parallel^2}{2m} + \mu B + e\delta\Psi_{\text{gy}} - e\langle \mathbf{T}_{\text{gy}}^{-1} \delta\phi_{\text{gc}} \rangle \right) \\ + \int \frac{d^3x}{8\pi} (|\nabla \delta\phi|^2 + |\mathbf{B} + \delta\mathbf{B}|^2), \quad (36)$$

where $\mathbf{T}_{\text{gy}}^{-1} \delta\phi_{\text{gc}}$ is the gyrocenter *push-forward* of the perturbed scalar potential. This exact conservation law either is constructed directly from the nonlinear gyrokinetic equations or is derived by applying the Noether method within a gyrokinetic variational formulation.

The gyrocenter pull-back and push-forward operators $\mathbf{T}_\epsilon = \mathbf{T}_{\text{gc}} \mathbf{T}_{\text{gy}}$ and $\mathbf{T}_\epsilon^{-1} = \mathbf{T}_{\text{gy}}^{-1} \mathbf{T}_{\text{gc}}^{-1}$ appearing in Eqs. (29)–(32) are the fundamental tools used in the modern derivation of the nonlinear gyrokinetic Vlasov equation (2), with gyrocenter equations of motion given by Eqs. (26) and (27), the gyrokinetic Poisson-Ampère equations (29) and (31), and the global gyrokinetic energy invariant (36), defined in Sec. IV. The classical forms (30) and (32) for the polarization density and current and the magnetization current (Jackson, 1975) arise from the push-forward representation of fluid moments (see Appendix C). For the remainder of this section, various limiting cases of the nonlinear gyrokinetic equations are presented in general magnetic geometry (while some applications in Table I were made in simple geometry).

B. Electrostatic fluctuations

When only electrostatic fluctuations are present (i.e., $\delta\mathbf{A} = \mathbf{0}$) the electrostatic nonlinear gyrokinetic equations in general geometry (Hahm, 1988) can be used to study most drift-wave-type fluctuations driven by the expansion free energy associated with the gradients in density and temperature. The sound-wave dynamics as well as linear and nonlinear Landau damping (Sagdeev and Galeev, 1969) are contained within the nonlinear gyrokinetic formulations. The electrostatic nonlinear gyrokinetic equations can be used for ion dynamics associated

with the ion-temperature-gradient (ITG) instability, electron drift waves including the trapped-electron mode (TEM), collisionless trapped-ion modes (TIM), and universal and dissipative drift instabilities. These gyrokinetic equations can be used to study electron dynamics of the electron-temperature-gradient (ETG) instability. While an unmagnetized “adiabatic” ion response is commonly used for ETG studies, a more accurate treatment of ion dynamics associated with the ETG instability is possible with a gyrokinetic formulation with proper high- \mathbf{k} behavior. A recent simulation (Candy and Waltz, 2006) shows its importance. The nonlinear gyrokinetic formulations can also be used to study zonal flows (Diamond *et al.*, 2005) and geodesic acoustic modes (Winsor *et al.*, 1968) that are typically linearly stable. A partial summary of nonlinear gyrokinetic applications of these examples is listed in Table I.

In the electrostatic case, the modified magnetic field (28) has the form that appears in the guiding-center theory (with $\delta\mathbf{B}_{gy} \equiv \mathbf{0}$), and the effective gyrocenter perturbation potential $\delta\Psi_{gy}$ in the gyrocenter equations of motion (26) and (27) is expressed as

$$\delta\Psi_{gy} = \langle \delta\phi_{gc} \rangle - \frac{e}{2B} \frac{\partial}{\partial \mu} \langle \delta\tilde{\phi}_{gc}^2 \rangle, \quad (37)$$

which retains full FLR effects in both the linear and the nonlinear terms,¹¹ where $\delta\tilde{\phi}_{gc} \equiv \delta\phi_{gc} - \langle \delta\phi_{gc} \rangle$ denotes the gyroangle-dependent part of $\delta\phi_{gc} \equiv \exp(\boldsymbol{\rho} \cdot \nabla) \delta\phi$. Only the gyrokinetic Poisson equation (29) is relevant in the electrostatic limit, where the gyrocenter pull-back $T_{gy}F$ consistent with the simplified effective gyrocenter perturbation potential (37) is

$$T_{gy}F = F + \frac{e \delta\tilde{\phi}_{gc}}{B} \frac{\partial F}{\partial \mu}.$$

Thus, the integrand on the right side of the gyrokinetic Poisson equation (29) includes the polarization term

$$e^{-\boldsymbol{\rho} \cdot \nabla} (T_{gy}F - F) = -\frac{ef_0}{T_\perp} (\delta\phi - e^{-\boldsymbol{\rho} \cdot \nabla} \langle \delta\phi_{gc} \rangle), \quad (38)$$

where $f_0 \equiv e^{-\boldsymbol{\rho} \cdot \nabla} F_0$ denotes the background particle Vlasov distribution expressed in terms of a Maxwellian distribution F_0 in μ (with temperature T_\perp). The gyrokinetic energy invariant (36) includes the perturbation term

$$\delta\Psi_{gy} - \langle T_{gy}^{-1} \delta\phi_{gc} \rangle = \frac{e}{2B} \frac{\partial}{\partial \mu} \langle \delta\tilde{\phi}_{gc}^2 \rangle, \quad (39)$$

which is consistent with the effective gyrocenter perturbation potential (37) and the gyrokinetic Poisson equation (29) with the polarization density (38).

The energy-conserving gyrokinetic Vlasov-Poisson equations constructed with the effective gyrocenter perturbation potential (37) and the gyrokinetic polarization

density (38) can be written in a more definite form involving modified Bessel functions (Dubin *et al.*, 1983) by Fourier transforming the gyrokinetic Poisson equation (29) into \mathbf{k} space,

$$n_0(|\mathbf{k}|^2 \lambda_{Di}^2) \frac{e \delta\phi_{\mathbf{k}}}{T_{i\perp}} = \delta n_{i\mathbf{k}} - \delta n_{e\mathbf{k}}, \quad (40)$$

where $\lambda_{Di}^2 \equiv T_{i\perp}/4\pi n_0 e^2$ and the perturbed ion fluid density

$$\begin{aligned} \delta n_{i\mathbf{k}} &= \delta N_{i\mathbf{k}} - n_0(1 - \Gamma_0) \frac{e \delta\phi_{\mathbf{k}}}{T_{i\perp}} \\ &+ n_0 [\rho_i^2 (i\mathbf{k}_\perp \cdot \nabla \ln n_0) (\Gamma_1 - \Gamma_0)] \frac{e \delta\phi_{\mathbf{k}}}{T_{i\perp}} \end{aligned} \quad (41)$$

is expressed in terms of the perturbed ion gyrofluid density $\delta N_{i\mathbf{k}} \equiv \int d^3 p \langle e^{-i\boldsymbol{\rho} \cdot \mathbf{k}_\perp} \rangle \delta F_{i\mathbf{k}}$, and $\Gamma_n(b) \equiv I_n(b)e^{-b}$ is expressed in terms of modified Bessel functions I_n (of order n), with $b \equiv |\mathbf{k}_\perp|^2 \rho_i^2$. While the last term in Eq. (41) is smaller than the leading term, it is needed to preserve the polarization density as a divergence of a polarization vector (Dubin *et al.*, 1983; Hahm *et al.*, 1988). The invariant energy for these electrostatic gyrokinetic equations is

$$\begin{aligned} E &= \int d^6 Z \delta F_i \left(\mu B + \frac{m_i}{2} v_\parallel^2 \right) + \int d^6 z \delta f_e \left(\frac{m_e}{2} v^2 \right) \\ &+ \int \frac{d^3 x}{8\pi} |\delta\mathbf{E}|^2 + \frac{n_0 e^2}{2T_i} \int \frac{d^3 \mathbf{k}}{(2\pi)^3} (1 - \Gamma_0) |\delta\phi_{\mathbf{k}}|^2, \end{aligned} \quad (42)$$

which provides an accurate linear response for arbitrary $k_\perp \rho_i$ and the dominant $E \times B$ nonlinearity as well as the parallel velocity-space nonlinearity needed for most applications.

The nonlinear gyrokinetic Vlasov-Poisson equations (37)–(39) can be further simplified by taking the long-wavelength (drift-kinetic) limit ($k_\perp \rho_i \ll 1$) of the nonlinear correction in the effective gyrocenter perturbation potential (37),

$$e \delta\Psi_{gy} = e \langle \delta\phi_{gc} \rangle - \frac{m}{2} |\delta\mathbf{u}_E|^2, \quad (43)$$

where the second term is the normalized kinetic energy associated with the perturbed $E \times B$ drift (Scott, 2005). There is a one-to-one correspondence with this term, the polarization density term in the gyrokinetic Poisson equation, and the sloshing energy term in the energy invariant. In the same drift-kinetic limit (Dubin *et al.*, 1983), the linear gyrocenter polarization vector in the gyrokinetic Poisson equation (29) is expressed in terms of the gyrocenter electric-dipole moment,

$$\boldsymbol{\pi}_{gy} \equiv -e \frac{\partial \delta\Psi_{gy}}{\partial \delta\mathbf{E}_\perp} = -\frac{mc^2}{B^2} \nabla_\perp \delta\phi \equiv \frac{c\hat{\mathbf{b}}}{B} \times (m \delta\mathbf{u}_E) \quad (44)$$

that is directly related to the nonlinear terms in the effective gyrocenter perturbation potential (43); note that the linear term $\langle \delta\phi_{gc} \rangle$ contains the guiding-center polarization vector that is automatically included in the defi-

¹¹An additional nonlinear term involving the multidimensional expression $\hat{\mathbf{b}} \cdot (\nabla \delta\tilde{\Phi}_{gc} \times \nabla \delta\tilde{\phi}_{gc})$, where $\delta\tilde{\Phi}_{gc} = \int \delta\tilde{\phi}_{gc} d\zeta$, is omitted.

nition of the ion gyrofluid density. Because the gyrocenter electric-dipole moment (44) is proportional to the particle's mass, the dominant contribution to the polarization density comes from the ion species. Representing the polarization drift as a shielding term in the gyrokinetic Poisson equation provided one of the principal computational advantages of the gyrokinetic approach. The energy invariant consistent with the simplified effective gyrocenter perturbation potential (43) and the gyrokinetic Poisson equation (29), with gyrocenter polarization vector (44), includes the nonlinear term

$$e\delta\Psi_{\text{gy}} - e\langle T_{\text{gy}}^{-1}\delta\phi_{\text{gc}} \rangle = \frac{m}{2}|\delta\mathbf{u}_E|^2. \quad (45)$$

The simplified nonlinear gyrokinetic Vlasov-Poisson equations based on Eqs. (43)–(45) highlight the three pillars of nonlinear gyrokinetic theory: a gyrocenter Hamiltonian (43) that contains nonlinear (quadratic) terms, a gyrokinetic Poisson equation that contains a polarization density derived from the nonlinear gyrocenter Hamiltonian (44), and an energy invariant that includes all relevant nonlinear coupling terms (45).

While the nonlinear electrostatic gyrokinetic equations based on Eqs. (37)–(39) have a clear physical meaning, this set has not been utilized much for applications due to its complexity. For tokamak core turbulence, the relative density-fluctuation amplitude is typically less than 1%, and nonlinear corrections to the effective potential are small. However, these nonlinear corrections may play important roles in edge turbulence where the relative fluctuation amplitude is high, typically greater than 10% (see Fig. 2).

For both simulation and analytic applications, the distribution function $F=F_0+\delta F$ is often split into the equilibrium part F_0 and the perturbed part δF , with $\delta F/F_0 \sim \epsilon_\delta$. One can write the equilibrium part and the perturbed part of Eqs. (26) and (27) separately. Then, Eqs. (2)–(27) become

$$\frac{\partial\delta F}{\partial t} + \frac{d\mathbf{Z}}{dt} \cdot \frac{\partial\delta F}{\partial\mathbf{Z}} = -\frac{d\delta\mathbf{Z}}{dt} \cdot \frac{\partial F_0}{\partial\mathbf{Z}}, \quad (46)$$

where the perturbed equations of motion are

$$\frac{d\delta\mathbf{X}}{dt} = \frac{c\hat{\mathbf{b}}}{B_\parallel^*} \times \nabla \langle \delta\phi_{\text{gc}} \rangle \quad \text{and} \quad \frac{d\delta v_\parallel}{dt} = -\frac{\mathbf{B}^*}{mB_\parallel^*} \cdot \nabla \langle \delta\phi_{\text{gc}} \rangle,$$

and the full equations of motion are

$$\frac{d\mathbf{X}}{dt} = v_\parallel \frac{\mathbf{B}^*}{B_\parallel^*} + \frac{c\hat{\mathbf{b}}}{eB_\parallel^*} \times (e\nabla \langle \delta\phi_{\text{gc}} \rangle + \mu\nabla B) \quad (47)$$

and

$$\frac{dv_\parallel}{dt} = -\frac{\mathbf{B}^*}{mB_\parallel^*} \cdot (e\nabla \langle \delta\phi_{\text{gc}} \rangle + \mu\nabla B). \quad (48)$$

The second term on the left side of Eq. (46) contains the dominant $E \times B$ nonlinearity $[(d\delta\mathbf{X}/dt) \cdot \nabla \delta F]$ and the subdominant velocity-space parallel nonlinearity $[(d\delta v_\parallel/dt)\partial\delta F/\partial v_\parallel]$. If this parallel velocity-space nonlinearity is ignored, the physics contained in Eqs. (46)–(48)

is essentially the same as the electrostatic limit of the Frieman-Chen gyrokinetic equation (17). One consequence of omitting this term is that the energy invariant (42) is not conserved to the same order. While most turbulence simulations have not kept this small subdominant term for simplicity, some simulations (Sydora *et al.*, 1996; Hatzky *et al.*, 2002; Kniep *et al.*, 2003; Villard, Angelino, *et al.*, 2004) have kept it. In principle, simulations with this term should have better energy-conservation properties and, therefore, less time-accumulated error. This term is crucial in long-time simulations; this topic is an active area of current research.

C. Shear-Alfvénic magnetic fluctuations

It has been shown by Hahm *et al.* (1988) that the reduced magnetohydrodynamic (MHD) equations (whose derivation makes use of the ratio $k_\parallel/k_\perp \ll 1$ as an expansion parameter) can be recovered from the electromagnetic nonlinear gyrokinetic equations. For finite- β plasmas (with $m_e/m_i < \beta \ll 1$), the perpendicular magnetic fluctuations $\delta\mathbf{B}_\perp = \nabla_\perp \delta A_\parallel \times \hat{\mathbf{b}}$ become important as the magnetic flutter $v_\parallel \delta\mathbf{B}_\perp / B$ becomes comparable to the perturbed (linear) $E \times B$ velocity $(c\hat{\mathbf{b}}/B) \times \nabla_\perp \delta\phi$ (i.e., $\delta A_\parallel v_\parallel / c \sim \delta\phi$). The physics associated with shear-Alfvén waves and instabilities (which include a wide variety of MHD instabilities) can be studied using the gyrokinetic approach.

Early computational applications to MHD modes consisted of various hybrid approaches with nonlinear gyrokinetic description of energetic particle dynamics and MHD description of bulk plasmas (Fu and Park, 1995; Santoro and Chen, 1996; Briguglio *et al.*, 1998). The nonlinear gyrokinetic approach has been applied to the classical tearing and kink instabilities when the free energy comes from the radial gradient of the equilibrium plasma current. For these simulations, the electron dynamics should include the radial variation of the equilibrium current along the perturbed magnetic field to describe the release of the current free energy. The electromagnetic modifications of drift-wave turbulence, often referred to as “drift-Alfvén” turbulence (see Table I), is an outstanding topic in magnetic confinement physics, including edge turbulence (Scott, 1997). There have been nonlinear simulations of shear-Alfvén fluctuations based on nonlinear gyrokinetic formulations; some examples are listed in Table I.

Since the early days of modern nonlinear gyrokinetic theory (Hahm *et al.*, 1988) it has become apparent that there can be at least two different versions of electromagnetic nonlinear gyrokinetic equations. One version is the Hamiltonian formulation that uses the parallel canonical momentum p_\parallel as an independent variable; the other is the symplectic formulation, where the parallel velocity v_\parallel is an independent variable. Each approach has its own advantages. This work is confined to the case in which nonlinear modifications of the perturbed potential are expressed in the drift-kinetic limit that may turn

out to be important in the nonlinear gyrokinetic simulation of edge turbulence.

1. Hamiltonian (p_{\parallel}) formulation

In the Hamiltonian formulation, the magnetic perturbation $\delta A_{\parallel} \equiv \hat{\mathbf{b}} \cdot \delta \mathbf{A}$ is treated as part of the gyrocenter Hamiltonian, with all linear and nonlinear δA_{\parallel} terms included in the effective gyrocenter potential $\delta\Psi_{\text{gy}}$ in Eqs. (26) and (27) while the symplectic magnetic perturbation is $\delta \mathbf{A}_{\text{gy}} \equiv \mathbf{0}$. The linear perturbation potential $\delta\psi_{\text{gc}} = \delta\phi_{\text{gc}} - (p_{\parallel}/mc)\delta A_{\parallel\text{gc}}$ is manifestly covariant. The gyrocenter parallel velocity $v_{\parallel} \equiv \hat{\mathbf{b}} \cdot d\mathbf{X}/dt$ is expressed in terms of the gyrocenter parallel canonical momentum p_{\parallel} and the perturbed parallel vector potential $e\delta A_{\parallel}/c$,

$$v_{\parallel} \equiv \frac{1}{m} \left(p_{\parallel} - \epsilon \frac{e}{c} \langle \delta A_{\parallel\text{gc}} \rangle + \dots \right).$$

For this reason, the Hamiltonian formulation is sometimes referred to as the “canonical-momentum” formulation or the P_z formulation following the terminology from early work (Hahm *et al.*, 1988) in a straight magnetic field.

This formulation deserves two important remarks. First, the expression $\partial/\partial t$ associated with the parallel induction electric field is absent on the right side of Eq. (27). This computationally desirable feature (Hahm *et al.*, 1988) is one of the motivations for the canonical-momentum formulation. Second, in the perpendicular gyrocenter velocity (26), the second term $\hat{\mathbf{b}} \times \nabla \langle \delta\psi_{\text{gc}} \rangle$ contains both the perturbed $E \times B$ velocity and the magnetic-flutter motion along the perturbed magnetic field (which often becomes stochastic). The $E \times B$ drift turbulence is the most likely anomalous transport mechanism in magnetically confined plasmas with low to moderate values of β ; a detailed explanation of this mechanism can be found in Scott (2003). The test-particle transport in stochastic magnetic fields has been thoroughly studied (Rechester and Rosenbluth, 1978; Krommes *et al.*, 1983).

For the Hamiltonian gyrocenter formulation of shear-Alfvénic fluctuations, the effective gyrocenter perturbation potential is (Hahm *et al.*, 1988)

$$e\delta\Psi_{\text{gy}} = e\langle \delta\psi_{\text{gc}} \rangle + \frac{e^2\delta A_{\parallel}^2}{2mc^2} - \frac{m}{2} \left| \delta \mathbf{u}_E + \frac{p_{\parallel}}{m} \frac{\delta \mathbf{B}_{\perp}}{B} \right|^2, \quad (49)$$

where the linear term retains full FLR effects while the nonlinear terms are given in the drift-kinetic limit ($\epsilon_{\perp} \rightarrow 0$), with $\delta \mathbf{B}_{\perp} \equiv \nabla_{\perp} \delta A_{\parallel} \times \hat{\mathbf{b}}$ denoting the perturbed magnetic field.

These gyrokinetic Maxwell's equations consist of the gyrokinetic Poisson equation (29), with the linear gyrocenter polarization vector expressed in terms of the gyrocenter electric-dipole moment

$$\boldsymbol{\pi}_{\text{gy}} \equiv -e \frac{\partial \delta\Psi_{\text{gy}}}{\partial \mathbf{E}_{\perp}} = -\frac{mc^2}{B^2} \left(\nabla_{\perp} \delta\phi - \frac{p_{\parallel}}{mc} \nabla_{\perp} \delta A_{\parallel} \right)$$

$$= \frac{c\hat{\mathbf{b}}}{B} \times \left(m \delta \mathbf{u}_E + p_{\parallel} \frac{\delta \mathbf{B}_{\perp}}{B} \right), \quad (50)$$

where magnetic-flutter motion along perturbed magnetic-field lines contributes to the polarization density, and the gyrokinetic parallel Ampère equation

$$-\nabla_{\perp}^2 \delta A_{\parallel} = -\left(\frac{\omega_p^2}{c^2} \right) \delta A_{\parallel} + \frac{4\pi}{c} (J_{i\parallel} + J_{e\parallel}) + 4\pi \nabla \cdot (\mathbf{M}_{\text{gy}} \times \hat{\mathbf{b}}), \quad (51)$$

where the parallel current densities $J_{j\parallel}$ ($j=i$ or e) involve moments of p_{\parallel}/m_j , and the linear gyrocenter magnetization vector

$$\begin{aligned} \mathbf{M}_{\text{gy}} &\equiv -\sum e \int d^3 p f_0 \left(\frac{\partial \delta\Psi_{\text{gy}}}{\partial \delta \mathbf{B}_{\perp}} \right) \\ &= \sum \int d^3 p f_0 \left(\boldsymbol{\pi}_{\text{gy}} \times \frac{p_{\parallel}}{mc} \hat{\mathbf{b}} \right) \end{aligned} \quad (52)$$

only displays the moving-electric-dipole contribution. The explicit appearance of the collisionless skin depth (ω_p/c) on the right side of Eq. (51), whose dominant contribution comes from the electron species, might suggest that it can be a characteristic correlation length for electromagnetic turbulence in magnetized plasmas. While this alone is not sufficient theoretical evidence, turbulence at the scale of the collisionless skin depth has been simulated (Yagi *et al.*, 1995; Horton *et al.*, 2000) and measured in experiments (Wong *et al.*, 1997). This canonical-momentum formulation has been widely used for kink mode, classical tearing mode, and drift-tearing mode nonlinear gyrokinetic simulations, as listed in Table I.

The gyrokinetic energy invariant consistent with the effective gyrocenter perturbation potential (49), the gyrokinetic polarization density (50), and the gyrokinetic parallel magnetization current (52) includes the terms

$$\begin{aligned} e\delta\Psi_{\text{gy}} - e\langle \mathbf{T}_{\text{gy}}^{-1} \delta\phi_{\text{gc}} \rangle &= -\frac{ep_{\parallel}}{mc} \langle \delta A_{\parallel\text{gc}} \rangle + \frac{e^2\delta A_{\parallel}^2}{2mc^2} \\ &+ \frac{1}{2} \left(m |\delta \mathbf{u}_E|^2 - \frac{p_{\parallel}^2}{mB^2} |\delta \mathbf{B}_{\perp}|^2 \right). \end{aligned} \quad (53)$$

In the Hamiltonian gyrocenter model of shear-Alfvénic fluctuations, the magnetic-flutter perturbed motion changes the gyrocenter polarization density (50), while the perturbed $E \times B$ motion contributes to the gyrocenter magnetization current (52). If one drops contributions to the gyrocenter polarization and magnetization vectors that are subdominant compared to other terms in the gyrokinetic Poisson-Ampère equations, they must be dropped simultaneously in the effective gyrocenter perturbation potential and the gyrokinetic invariant.

2. Symplectic (v_{\parallel}) formulation

While the Hamiltonian formulation has some computational advantages (Hahm *et al.*, 1988) and is in an explicitly covariant form suitable for renormalization (Krommes and Kim, 1988), it is more straightforward to identify the physical meaning of various terms in an alternative “symplectic” (or v_z) formulation. In the Hamiltonian formulation it is often cumbersome to calculate the moment of p_{\parallel} numerically while using the analytic expression $(\omega_p/c)^2 \delta A_{\parallel}$ that appears explicitly in Ampère’s law (51), up to an accuracy that is sufficient for calculation of their difference, i.e., the moment of mv_{\parallel} (Chen and Parker, 2001; Hammett, 2001; Chen *et al.*, 2003; Mishchenko *et al.*, 2004; Lin, Wang, *et al.*, 2005). It is more computationally efficient and easier to understand the physics if a smaller term v_{\parallel} is used as an independent variable. In the symplectic formulation, the perturbed parallel vector potential δA_{\parallel} appears explicitly in the gyrocenter Poisson bracket (where $\delta \mathbf{A}_{\text{gy}} = \langle \delta A_{\parallel \text{gc}} \rangle \hat{\mathbf{b}}$ and $\delta \mathbf{B}_{\text{gy}} = \nabla \times \delta \mathbf{A}_{\text{gy}}$), rather than in the gyrocenter Hamiltonian. The resulting Euler-Lagrange equations contain the induction part of the electric field with $\partial \langle \delta A_{\parallel \text{gc}} \rangle / \partial t$, although this is not a computationally attractive feature. As discussed by Cummings (1995), the presence of the term $\partial \delta A_{\parallel \text{gy}} / \partial t$ in Eq. (27) makes the gyrokinetic Vlasov equation unsuitable for numerical integration by the method of characteristics. The perturbed parallel electric field δE_{\parallel} from the generalized parallel Ohm’s law should be calculated, which involves the calculation of a parallel-pressure moment.

This symplectic version of the electromagnetic nonlinear gyrokinetic equation is more suitable in showing its relation to various reduced fluid equations by taking moments of the nonlinear gyrokinetic Vlasov equation (Brizard, 1992). One of the key points in understanding the shear-Alfvén physics in the context of the electromagnetic nonlinear gyrokinetic formulation is that the “vorticity evolution” in the reduced MHD equation is equivalent to the evolution of the ion polarization density that is the difference between the ion and electron gyrocenter densities (Hahm *et al.*, 1988; Brizard, 1992). It is straightforward to extend a simple illustration of deriving the vorticity evolution equation of reduced MHD in Hahm *et al.* (1988) to include the gyrocenter drift due to magnetic-field inhomogeneity (the driving term for ballooning and interchange instability) and the variation of the equilibrium current along the perturbed magnetic field (the driving term for kink, tearing, and peeling instabilities).

For the symplectic formulation of shear-Alfvénic fluctuations, the effective gyrocenter perturbation potential is (Hahm *et al.*, 1988)

$$e \delta \Psi_{\text{gy}} = e \langle \delta \phi_{\text{gc}} \rangle + \frac{\mu}{2B} |\delta \mathbf{B}_{\perp}|^2 - \frac{m}{2} \left| \delta \mathbf{u}_E + v_{\parallel} \frac{\delta \mathbf{B}_{\perp}}{B} \right|^2, \quad (54)$$

where the linear term that retains full FLR effects includes only the perturbed electrostatic potential [the

symplectic perturbed magnetic potential $\delta \mathbf{A}_{\text{gy}} \equiv \langle \delta A_{\parallel \text{gc}} \rangle \hat{\mathbf{b}}$ appears in the modified magnetic field (28) and the inductive part $\partial_t \langle \delta A_{\parallel \text{gc}} \rangle$ of the parallel perturbed electric field] while the nonlinear terms in Eq. (54) are in the drift-kinetic limit. The gyrokinetic Maxwell’s equations consist of the gyrokinetic Poisson equation (29), where the gyrocenter electric-dipole moment $\boldsymbol{\pi}_{\text{gy}}$ is given by Eq. (50), and the gyrokinetic parallel Ampère equation

$$-\nabla_{\perp}^2 \delta A_{\parallel} = \frac{4\pi}{c} (J_{i\parallel} + J_{e\parallel}) + 4\pi \nabla \cdot (\mathbf{M}_{\text{gy}} \times \hat{\mathbf{b}}), \quad (55)$$

where the linear gyrocenter magnetization vector has the moving-electric-dipole contribution shown in Eq. (52) and an intrinsic gyrocenter magnetic-dipole moment contribution

$$\boldsymbol{\mu}_{\text{gy}} \equiv -\mu \frac{\delta \mathbf{B}_{\perp}}{B}. \quad (56)$$

In the symplectic formulation, the skin-depth term is absent from the parallel Ampère equation (55) while a new intrinsic magnetic-dipole moment (56) contributes to the gyrocenter magnetization vector.

The gyrokinetic energy invariant consistent with the effective gyrocenter perturbation potential (54), the gyrokinetic polarization density, and the gyrokinetic parallel magnetization current include the terms

$$e \delta \Psi_{\text{gy}} - e \langle \mathbf{T}_{\text{gy}}^{-1} \delta \phi_{\text{gc}} \rangle = \frac{\mu}{2B} |\delta \mathbf{B}_{\perp}|^2 + \frac{m}{2} \left(|\delta \mathbf{u}_E|^2 - \frac{v_{\parallel}^2}{B^2} |\delta \mathbf{B}_{\perp}|^2 \right). \quad (57)$$

An energy-conserving set of nonlinear drift-Alfvén fluid equations, with linear and nonlinear terms both expressed in the drift-kinetic limit, was derived by Brizard (2005b) using variational methods.

D. Compressional magnetic fluctuations

The treatment of the compressional Alfvén wave is beyond the scope of the low-frequency nonlinear gyrokinetic formulation. If $\omega \sim k_{\perp} v_A$, then $\omega/\Omega \sim k_{\perp} v_A/\Omega \sim k_{\perp} \rho_i / \beta^{1/2}$. Therefore, with the full gyrokinetic FLR description ($k_{\perp} \rho_i \sim 1$), it is impossible to satisfy the low-frequency gyrokinetic ordering $\omega/\Omega \ll 1$ for a typical value of $\beta < 1$ encountered in magnetically confined plasmas. To describe the compressional Alfvén wave, it is necessary to use a drift-kinetic description (i.e., $k_{\perp} \rho_{\text{th}} \ll 1$ for all particle species). It has been shown that it is possible to decouple the gyromotion from dynamics associated with high-frequency waves with $\omega/\Omega > 1$ and develop a high-frequency linear gyrokinetic equation (Tsai *et al.*, 1984; Lashmore-Davies and Dendy, 1989). It has also been shown that the phase-space Lagrangian and Lie-transform perturbation method can be useful in deriving the linear high-frequency gyrokinetic equation in a more transparent way. It is instructive to follow the

derivation of the compressional Alfvén wave linear dispersion relation from the high-frequency gyrokinetic approach (Qin *et al.*, 1999). A satisfactory nonlinear high-frequency gyrokinetic formulation has not been derived to date and progress in the linear high-frequency gyrokinetic formulation is not discussed in this review.

Although the compressional Alfvén wave does not exist within the nonlinear (low-frequency) gyrokinetic formulation, the compressional component ($\delta B_{\parallel} \equiv \hat{\mathbf{b}} \cdot \delta \mathbf{B}$) of the perturbed magnetic field $\delta \mathbf{B}$, it becomes important as the plasma β value is increased. The magnetic field δB_{\parallel} must be kept for a quantitatively accurate description of fluctuations in relatively high- β plasmas, for example, those encountered in spherical tori including the National Spherical Torus Experiment (Ono *et al.*, 2003) and the Mega-Amp Spherical Tokamak (Sykes *et al.*, 2001).

The fully electromagnetic nonlinear gyrokinetic Vlasov equation is presented in Sec. V. We use the Hamiltonian gyrocenter model (with $\delta \mathbf{A}_{gy} \equiv \mathbf{0}$) and express the effective gyrocenter perturbation potential in the drift-kinetic limit as

$$\begin{aligned} e\delta\Psi_{gy} = & e\left(\delta\phi - \frac{p_{\parallel}}{mc}\delta A_{\parallel}\right) + \mu\delta B_{\parallel} + \frac{e^2\delta A_{\parallel}^2}{2mc^2} - \frac{m}{2} \left| \delta \mathbf{u}_E \right. \\ & \left. + \frac{p_{\parallel}}{m}\frac{\delta \mathbf{B}_{\perp}}{B} \right|^2 - \frac{e}{c}\delta \mathbf{A}_{\perp} \cdot \left(\delta \mathbf{u}_E + \frac{p_{\parallel}}{m}\frac{\delta \mathbf{B}_{\perp}}{B} \right). \end{aligned} \quad (58)$$

The linear perpendicular gyrocenter dynamics is represented by the linear perturbed $E \times B$ velocity ($\hat{\mathbf{b}} \times \nabla \delta\phi$), the linear magnetic-flutter velocity ($v_{\parallel} \delta \mathbf{B}_{\perp}$), and the linear perturbed grad- B drift ($\mu \hat{\mathbf{b}} \times \nabla \delta B_{\parallel}$). While in most fusion plasmas (with $\beta < 1$) the radial transport due to this last term is subdominant to other transport mechanisms driven by $E \times B$ transport and magnetic-flutter transport, this drift can be important in geophysical applications (Chen, 1999) and the cross-field diffusion of cosmic rays due to turbulence (Otsuka and Hada, 2003); linear gyrokinetic simulations are currently extended to high- β astrophysical plasmas (Howes *et al.*, 2006). Although the nonlinear terms in the gyrocenter potential (58) are small compared to the linear terms, they play an important role in contributing to the gyrocenter polarization and (intrinsic) magnetization vectors

$$\boldsymbol{\pi}_{gy} = \frac{c\hat{\mathbf{b}}}{B} \times \left(\frac{e}{c}\delta \mathbf{A}_{\perp} + m\delta \mathbf{u}_E + p_{\parallel}\frac{\delta \mathbf{B}_{\perp}}{B} \right), \quad (59)$$

$$\boldsymbol{\mu}_{gy} = -\mu \hat{\mathbf{b}}. \quad (60)$$

The lowest-order contribution of the intrinsic magnetization current in the perpendicular gyrokinetic Ampère equation, derived from the gyrocenter magnetic-dipole moment (60), yields the perpendicular pressure balance condition $\delta B_{\parallel} + 4\pi\delta P_{\perp}/B = 0$, where δP_{\perp} is the perturbed perpendicular (total) pressure (Tang *et al.*, 1980; Brizard, 1992); a straightforward demonstration of this condition can be found in Roach *et al.* (2005). The corresponding energy invariant includes the terms

$$\begin{aligned} e\delta\Psi_{gy} - e\langle T_{gy}^{-1}\delta\phi_{gc} \rangle = & \mu\delta B_{\parallel} - \frac{ep_{\parallel}}{mc}\delta \mathbf{A} \cdot \left(\hat{\mathbf{b}} + \frac{\delta \mathbf{B}_{\perp}}{B} \right) \\ & + \frac{1}{2} \left(m|\delta \mathbf{u}_E|^2 - \frac{p_{\parallel}^2}{mB^2}|\delta \mathbf{B}_{\perp}|^2 \right). \end{aligned} \quad (61)$$

The polarization velocity $e^{-1}\partial \boldsymbol{\pi}_{gy}/\partial t$ based on Eq. (59) includes the inductive part $(\hat{\mathbf{b}}/B) \times \partial \delta \mathbf{A}_{\perp}/\partial t$ of the perturbed $E \times B$ velocity in addition to the polarization velocity $(\hat{\mathbf{b}}/\Omega) \times \partial \delta \mathbf{u}_E/\partial t$. The variational expressions (59) and (60) for the electric-dipole and magnetic-dipole moments are rederived in Appendix C by the push-forward method.

E. Reduced fluid equations from moments of the nonlinear gyrokinetic Vlasov equation

The need to carry out long-time simulations of turbulent magnetized plasmas with increasingly realistic geometries has also been addressed by the development of reduced nonlinear fluid models derived from perturbative expansions of exact fluid models based on space-time and amplitude orderings similar to the nonlinear gyrokinetic orderings (5)–(8). Examples of such nonlinear reduced fluid models include the nonlinear reduced MHD equations (Strauss, 1976, 1977), the Hasegawa-Mima nonlinear drift-wave equation (Hasegawa and Mima, 1978), and the Hasegawa-Wakatani FLR-corrected nonlinear reduced fluid equations (Hasegawa and Wakatani, 1983).

While reduced fluid models cannot capture all kinetic physics described by nonlinear gyrokinetic equations, their use is justified for many applications (such as the investigation of macroscopic MHD stability). They are also useful for turbulence studies due to their relative simplicity and their closer connection to nonlinear theories that are primarily based on the $E \times B$ nonlinearity. As a result, the knowledge gained in the fluid-turbulence community can serve as a useful guide for plasma turbulence studies.

An important class of reduced fluid models involves the so-called gyrofluid models that are derived by taking gyrocenter-velocity moments of the nonlinear gyrokinetic Vlasov equation while keeping finite-gyroradius effects (Brizard, 1992). One obtains a hierarchy of evolution equations for gyrocenter-fluid moments, i.e., for density, parallel velocity, pressure, etc. To obtain a closed set of these gyrofluid equations, one needs to invoke a *closure approximation* (i.e., expressions for higher-order fluid moments in terms of lower-order fluid moments). In the simulation community, the so-called Landau-closure approach that places emphasis on accurate linear Landau damping and the linear growth rate has been most widely adopted (Dorland and Hammett, 1993; Waltz *et al.*, 1994; Beer, 1995; Snyder and Hammett, 2001). In this approach, some kinetic effects such as linear Landau damping and a limited form of nonlinear Landau damping (e.g., elastic Compton scattering)

have been successfully incorporated in gyrofluid models. These models do not accurately treat the strongly nonlinear wave-particle interactions (e.g., inelastic Compton scattering and trapping) (Sagdeev and Galeev, 1969). While more accurate closures involving nonlinear kinetic effects (Mator, 1992) and the treatment of damped modes (Sugama *et al.*, 2001) have been developed, they have not been as widely used in simulations as the Landau-closure-based models. To date, gyrofluid models cannot describe the zonal-flow damping accurately and can overestimate the turbulence level (Rosenbluth and Hinton, 1998). More discussions on zonal flows can be found in Appendix D.1, and a hierarchy of reduced kinetic and fluid models has been summarized by Diamond *et al.* (2005).

While a complete survey of collisionless and collisional reduced fluid models is beyond the scope of the present review of nonlinear gyrokinetic theory, it is instructive to see how nonlinear gyrokinetic theory is connected to nonlinear reduced fluid models. For this purpose (and using standard normalizations used in reduced fluid models that are presented below), the nonlinear gyrocenter-fluid equations derived from moments of the nonlinear gyrokinetic Vlasov equation are compared with the normalized Hasegawa-Mima equation (Hasegawa and Mima, 1978),

$$\frac{d}{dt}(\delta\phi - \nabla_{\perp}^2 \delta\phi) = \hat{\omega}_* \delta\phi, \quad (62)$$

and the normalized (collisionless) Hasegawa-Wakatani equations (Hasegawa and Wakatani, 1983),

$$-\nabla_{\parallel} \nabla_{\perp}^2 \delta A_{\parallel} = \left(\frac{d}{dt} + \hat{\omega}_*^{p(i)} \right) \nabla_{\perp}^2 \delta\phi - \{\nabla_{\perp}^2 \delta\phi, \delta p_{\perp i}\}_{\mathbf{x}} - \sum (\hat{\omega}_{\nabla} \delta p_{\perp} + \hat{\omega}_{\kappa} \delta p_{\parallel}), \quad (63)$$

$$\frac{\partial \delta A_{\parallel}}{\partial t} = -\nabla_{\parallel}(\delta\phi - \delta p_{\parallel e}) + \hat{\omega}_*^{p(e)} \delta A_{\parallel}, \quad (64)$$

$$\frac{dp_{\perp i}}{dt} = \frac{dp_{\parallel i}}{dt} = \frac{dp_i}{dt} = 0, \quad (65)$$

$$\frac{dp_{\perp e}}{dt} = \frac{dp_{\parallel e}}{dt} = \frac{dp_e}{dt} = -\nabla_{\parallel} \nabla_{\perp}^2 \delta A_{\parallel}. \quad (66)$$

Here the normalizations are standard (i.e., space-time scales are ordered with $\rho_s = \Omega^{-1} \sqrt{T_{e0}/m_i}$ and Ω^{-1} , $\delta\phi \sim \epsilon_{\delta} T_{e0}/e$, $\delta A_{\parallel} \sim \epsilon_{\delta} B_0 \rho_s$, etc.), the normalized linear operators are the drift operator $\hat{\omega}_* = \hat{\mathbf{b}}_0 \times \nabla \ln n_0 \cdot \nabla_{\perp}$, the diamagnetic operator $\hat{\omega}_*^{p(j)} = \hat{\mathbf{b}}_0 \times \nabla \ln p_{j0} \cdot \nabla_{\perp}$, the ∇B operator $\hat{\omega}_{\nabla} = \hat{\mathbf{b}}_0 \times \nabla \ln B_0 \cdot \nabla_{\perp}$, and the curvature operator $\hat{\omega}_{\kappa} = \hat{\mathbf{b}}_0 \times (\hat{\mathbf{b}}_0 \cdot \nabla \hat{\mathbf{b}}_0) \cdot \nabla_{\perp}$, while the nonlinear operators d/dt and $\hat{\mathbf{b}} \cdot \nabla$ are defined as

$$\frac{d}{dt} \equiv \frac{\partial}{\partial t} + \{\delta\phi, \}_{\mathbf{x}} \text{ and } \nabla_{\parallel} \equiv \hat{\mathbf{b}}_0 \cdot \nabla - \{\delta A_{\parallel}, \}_{\mathbf{x}},$$

where $\{f, g\}_{\mathbf{x}} \equiv \hat{\mathbf{b}}_0 \cdot (\nabla f \times \nabla g)$ is the normalized spatial Poisson bracket. The Hasegawa-Mima equation (62) is the paradigm model equation describing nonlinear drift-wave dynamics of electrostatic fluctuations assuming adiabatic electrons ($\delta n_e = \delta\phi$) in a neutralizing cold-ion fluid background with uniform magnetic field. On the other hand, the Hasegawa-Wakatani equations (63)–(66) that contain the low-beta reduced MHD equations (Strauss, 1976) $d\nabla_{\perp}^2 \delta\phi/dt = -\nabla_{\parallel} \nabla_{\perp}^2 \delta A_{\parallel}$ and $\partial \delta A_{\parallel} / \partial t = -\nabla_{\parallel} \delta\phi$ include the vorticity equation (63), the parallel Ohm's law (64), the incompressible isotropic ion pressure equation (65), and the isothermal isotropic electron pressure equation (66). The vorticity equation (63) can be derived from the ion fluid equation of motion either by taking the parallel component of the ion fluid vorticity $\hat{\mathbf{b}}_0 \cdot \nabla \times \delta \mathbf{u} = \nabla_{\perp}^2 \delta\phi$, where the ion velocity is given as the perturbed $E \times B$ velocity $\delta \mathbf{u} = \hat{\mathbf{b}}_0 \times \nabla \delta\phi$, or by using the quasineutrality condition and the charge conservation law $\nabla \cdot \delta \mathbf{J} = 0$, which follows from the gyrokinetic Ampère equation (4), where the net perturbed current $\delta \mathbf{J} = \delta I_{\parallel} \hat{\mathbf{b}} + \delta \mathbf{J}_{\perp}$ is defined in terms of the parallel current contribution $\delta I_{\parallel} = -\nabla_{\perp}^2 \delta A_{\parallel}$ and the perpendicular current $\delta \mathbf{J}_{\perp}$ expressed as the sum of the perturbed guiding-center current, the perturbed magnetization current, and the polarization current.

Reduced nonlinear fluid models such as the Hasegawa-Mima equation (62) and the Hasegawa-Wakatani equations (63)–(66) have proved successful in the analysis of turbulent magnetized plasmas in the presence of drift-wave (electrostatic) fluctuations and drift-Alfvén fluctuations, respectively. Both sets presented here contain the polarization nonlinearity $\{\delta\phi, \nabla_{\perp}^2 \delta\phi\}_{\mathbf{x}}$ that plays a crucial role in spectral transfer and possibly in self-sustaining turbulence (Scott, 1990). For the zonal (poloidally and toroidally symmetric) part of the potential evolution, this term describes the vorticity transport, which is responsible for zonal flow amplification (Diamond *et al.*, 1993; Diamond, Lebedev, *et al.*, 1994).

These nonlinear reduced fluid models can be obtained from the nonlinear gyrokinetic Vlasov-Maxwell equations based on Eqs. (26), (27), (29), and (31). We consider moments of the gyrokinetic Vlasov equation (Brizard, 1992)

$$\frac{\partial \|\chi\|}{\partial t} + \nabla \cdot \|\dot{\mathbf{X}}\chi\| = \left\| \frac{d\chi}{dt} \right\|, \quad (67)$$

where the gyrocenter phase-space functions $\chi = (1, v_{\parallel}, \mu B, m(v_{\parallel} - U_{\parallel})^2, \dots)$ are used to calculate the gyrofluid moments $\|\chi\| = (N, N U_{\parallel}, P_{\perp}, P_{\parallel}, \dots)$. To lowest order in the FLR expansion and in the absence of compressional magnetic fluctuations ($\delta B_{\parallel} = 0$), the electron gyrofluid moments are identical to the electron fluid moments (i.e., $N_e = n_e$, $U_{\parallel e} = u_{\parallel e}$, ...), and the normalized electron fluid equations are

$$\begin{aligned} \frac{d\delta n_e}{dt} = & -\nabla_{\parallel}\delta u_{\parallel e} + [\hat{\omega}_* - (\hat{\omega}_{\nabla} + \hat{\omega}_{\kappa})]\delta\phi \\ & + (\hat{\omega}_{\nabla}\delta p_{\perp e} + \hat{\omega}_{\kappa}\delta p_{\parallel e}), \end{aligned} \quad (68)$$

$$\frac{\partial\delta A_{\parallel}}{\partial t} = -\nabla_{\parallel}(\delta\phi - \delta p_{\parallel e}) + \hat{\omega}_*^{p(e)}\delta A_{\parallel}, \quad (69)$$

$$\frac{d\delta p_e}{dt} = \hat{\omega}_*^{p(e)}\delta\phi - \nabla_{\parallel}\delta u_{\parallel e}. \quad (70)$$

The compressibility of the perturbed $E \times B$ velocity is represented by the term $(\hat{\omega}_{\nabla} + \hat{\omega}_{\kappa})\delta\phi$ in the perturbed electron continuity equation (68), where the last term represents the divergence of the electron diamagnetic current, and the electron parallel velocity equation (in the absence of electron inertia, $m_e \rightarrow 0$) yields the parallel Ohm's law (68).

The ion gyrofluid equations for the perturbed ion gyrofluid density δN_i and the perturbed gyrofluid pressures $\delta P_{\perp i}$ and $\delta P_{\parallel i}$ can be written in terms of the perturbed ion fluid density δn_i and the perturbed isotropic ion fluid pressure $\delta p_{\perp i} = \delta p_{\parallel i} = \delta p_i$ by making use of the normalized FLR expansions

$$\begin{aligned} \delta N_i = & \delta n_i - \nabla_{\perp}^2 \left(\delta\phi + \frac{1}{2}\delta p_i \right), \\ \delta P_{\perp i} = & \delta p_i - \nabla_{\perp}^2(2\delta\phi + 2\delta p_i - \delta n_i), \\ \delta P_{\parallel i} = & \delta p_i - \nabla_{\perp}^2 \left(\delta\phi + \delta p_i - \frac{1}{2}\delta n_i \right), \end{aligned} \quad (71)$$

where the term $\nabla_{\perp}^2\delta\phi$ denotes the first-order FLR correction associated with polarization effects. Several reduced fluid models have neglected the perturbed ion parallel dynamics (i.e., $\delta u_{\parallel i} \equiv 0$) and the normalized Ampère equation is

$$\nabla_{\perp}^2\delta A_{\parallel} = -\delta J_{\parallel} \equiv \delta u_{\parallel e}. \quad (72)$$

Hence, the ion gyrofluid continuity equation becomes

$$\begin{aligned} \frac{d\delta n_i}{dt} = & [\hat{\omega}_* - (\hat{\omega}_{\nabla} + \hat{\omega}_{\kappa})]\delta\phi - (\hat{\omega}_{\nabla}\delta p_{\perp i} + \hat{\omega}_{\kappa}\delta p_{\parallel i}) \\ & + \left(\frac{d}{dt} + \hat{\omega}_*^{p(i)} \right) \nabla_{\perp}^2\delta\phi - \{\nabla_{\perp}^2\delta\phi, \delta p_{\perp i}\}_{\mathbf{x}}, \end{aligned} \quad (73)$$

while the incompressible isotropic ion pressure equation becomes

$$\frac{d\delta p_i}{dt} = \hat{\omega}_*^{p(i)}\delta\phi. \quad (74)$$

The gyrofluid equations (68)–(74) were presented in greater generality in Brizard (1992), where the perturbed ion parallel dynamics and compressional magnetic fluctuations were retained.

The Hasegawa-Mima equation (62) and the Hasegawa-Wakatani equations (63)–(66) are contained in the gyrofluid equations (68)–(70), (73), and (74). By

combining the quasineutrality condition $\delta n_i = \delta n_e$ with the electron adiabatic response $\delta n_e = \delta\phi$ (in the presence of a neutralizing cold-ion fluid in a uniform magnetic field), the ion gyrofluid continuity equation (73) yields the Hasegawa-Mima equation (62), where $d\delta n_i/dt = d\delta\phi/dt$. To obtain the Hasegawa-Wakatani vorticity equation (63), the electron continuity equation (68) is subtracted from the ion continuity equation (73) and we make use of the quasineutrality condition $\delta n_i = \delta n_e$ and the perturbed Ampère's law, $\nabla_{\perp}^2\delta A_{\parallel} = \delta u_{\parallel e}$. The remaining Hasegawa-Wakatani equations are found in Eqs. (69), (70), and (74).

There are other reduced nonlinear fluid equations, such as the four-field model of Hazeltine *et al.* (1985) and the Yagi-Horton (1994) model that generalize the nonlinear reduced MHD equations by capturing the difference between ion and electron dynamics (i.e., a two-fluid approach), but with less emphasis on the finite ion-gyroradius effects. The hierarchy of various nonlinear governing equations is briefly summarized in a recent review by Diamond *et al.* (2005). Roughly speaking, while fluid approaches (other than Landau-closure-based models for core turbulence applications) are justified by taking collisions into account (Braginskii, 1965), the presence of strong magnetic fields makes fluid descriptions justifiable for the dynamics across the magnetic field. Nonlinear fluid models have been extended to the long-mean-free-path “banana” collisionality regime as well; these models are called neoclassical MHD models (Callen and Shaing, 1985; Connor and Chen, 1985), and have been very useful as a starting point of the studies on neoclassical tearing modes (which are driven by the expansion free energy in the pressure gradient). A useful discussion of the physics and derivation of the MHD equations has been given by Kulsrud (1983). Since the energy-conservation property of a nonlinear reduced fluid model is often a critical component of its usefulness in simulations of turbulent magnetized plasmas, the derivation of energy-conserving reduced fluid models is a topic of active research (Strintzi and Scott, 2004; Brizard, 2005b; Strintzi *et al.*, 2005).

IV. LIE-TRANSFORM PERTURBATION THEORY

After having presented some simple forms of the nonlinear gyrokinetic equations in the previous section, this section focuses on the transformations from particle to gyrocenter phase-space coordinates that allow the dynamical reduction of the original Vlasov-Maxwell equations to generate energy-conserving nonlinear gyrokinetic Vlasov-Maxwell equations.

We begin with a brief introduction to the extended phase-space Lagrangian formulation of charged-particle dynamics in a time-dependent electromagnetic field. The electromagnetic field is represented by the potentials $A^{\mu} = (\phi, \mathbf{A})$, while the eight-dimensional extended phase-space noncanonical coordinates $\mathcal{Z}^a = (x^{\mu}; p^{\mu}) \equiv (ct, \mathbf{x}; w/c, \mathbf{p})$ include the position \mathbf{x} of a charged particle (mass m and charge e), its kinetic momentum \mathbf{p}

$=mv$, and the canonically conjugate energy-time (w,t) coordinates. We use the convenient Minkowski space-time metric $g=\text{diag}(-1,1,1,1)$ whenever we need a concise covariant expression.¹² The use of an eight-dimensional representation of phase space is motivated by that, in the presence of time-dependent electromagnetic fields, the energy of a charged particle is no longer conserved but instead changes according to an additional Hamilton's equation $dw/dt=e\partial\psi/\partial t$, where $\psi=\phi-\mathbf{A}\cdot\mathbf{v}/c=-A^\mu v_\mu/c$ is the effective electromagnetic potential. By introducing the canonical pair (w,t) , where the energy coordinate $w=E$ is equal to the conserved energy in the time-independent case, new extended Hamilton's equations for charged-particle motion in time-dependent electromagnetic fields can be written.

The complete representation of the Hamiltonian dynamics of a charged particle in an electromagnetic field (represented by the four-potentials A^μ) is expressed in terms of a Hamiltonian function H and a Poisson-bracket structure $\{\cdot,\cdot\}$ that satisfies the following properties (valid for arbitrary functions f , g , and h): antisymmetry property

$$\{f,g\} = -\{g,f\}, \quad (75)$$

the Leibnitz rule

$$\{f,(gh)\} = \{f,g\}h + g\{f,h\}, \quad (76)$$

and the Jacobi identity

$$0 = \{f,\{g,h\}\} + \{g,\{h,f\}\} + \{h,\{f,g\}\}. \quad (77)$$

We introduce the general form for the Poisson bracket $\{f,g\}$,

$$\{f,g\} \equiv \frac{\partial f}{\partial \mathcal{Z}^a} J^{ab} \frac{\partial g}{\partial \mathcal{Z}^b}, \quad (78)$$

where J^{ab} denotes the components of the Poisson tensor. The bilinear form of the Poisson bracket (78) automatically satisfies the Leibnitz rule (76), the antisymmetry property (75) requires that the Poisson tensor be anti-symmetric $J^{ba}=-J^{ab}$, and the Jacobi identity (77) requires that

$$0 = J^{a\ell} \partial_\ell J^{bc} + J^{b\ell} \partial_\ell J^{ca} + J^{c\ell} \partial_\ell J^{ab}, \quad (79)$$

where $\partial_\ell \equiv \partial/\partial \mathcal{Z}^\ell$. We note that the extended canonical Poisson tensor

$$\mathbf{J}_{\text{can}} = \left(\begin{array}{cc|cc} \mathbf{0} & \mathbf{0} & g & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & g \\ \hline -g & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -g & \mathbf{0} & \mathbf{0} \end{array} \right)$$

immediately satisfies all three properties.

¹²Summation over repeated indices is, henceforth, implied and latin letters a,b,c,\dots go from 1 to 8 while greek letters μ,ν,\dots go from 0 to 3.

Hamilton's equations are then expressed as $\dot{\mathcal{Z}}^a = \{\mathcal{Z}^a, \mathcal{H}\} = J^{ab} \partial_b \mathcal{H}$ in terms of the extended-phase-space Hamiltonian

$$\mathcal{H}(\mathcal{Z}) = \frac{|\mathbf{p}|^2}{2m} + e\phi - w \equiv H(\mathbf{z},t) - w, \quad (80)$$

where $H(\mathbf{z},t)$ denotes the standard time-dependent Hamiltonian and the physical single-particle motion takes place on the subspace

$$\mathcal{H}(\mathcal{Z}) = H(\mathbf{z},t) - w = 0. \quad (81)$$

Within the canonical formalism, the Poisson-bracket structure is independent of the electromagnetic field and the Hamiltonian depends explicitly on the electromagnetic potentials (ϕ, \mathbf{A}) . Within the noncanonical formalism, however, the Hamiltonian only retains its dependence on the electrostatic potential ϕ and derivatives of the magnetic potential \mathbf{A} appear in the Poisson-bracket structure.

A. Single-particle extended Lagrangian dynamics

The extended-phase-space Lagrangian, or Poincaré-Cartan differential one-form (Arnold, 1998), for a charged particle in eight-dimensional extended phase space is expressed in noncanonical form as

$$\begin{aligned} \hat{\Gamma} &= \left(\frac{e}{c} \mathbf{A} + \mathbf{p} \right) \cdot d\mathbf{x} - w dt - \mathcal{H} d\tau \\ &\equiv \Gamma_a(\mathcal{Z}) d\mathcal{Z}^a - \mathcal{H}(\mathcal{Z}) d\tau, \end{aligned} \quad (82)$$

where Γ_a are the symplectic components of the extended-phase-space Lagrangian $\hat{\Gamma}$ and τ is the Hamiltonian orbit parameter. In Eq. (82), d denotes an *exterior* derivative with the property

$$d^2 f = d(\partial_a f d\mathcal{Z}^a) = \partial_{ab}^2 f d\mathcal{Z}^a \wedge d\mathcal{Z}^b = 0, \quad (83)$$

which holds for any scalar field f , where the wedge product \wedge is antisymmetric (i.e., $df \wedge dg = -dg \wedge df$); we will use the standard-derivative notation d whenever the exterior-derivative properties are not involved (see Appendix A for further details). As a result of property (83), an arbitrary gauge term dS may be added to the extended-phase-space Lagrangian (82) without modifying the Hamiltonian dynamics.

To obtain the extended Hamilton's equations of motion from the phase-space Lagrangian (82), we introduce the single-particle action integral

$$S = \int \hat{\Gamma} = \int_{\tau_1}^{\tau_2} \left(\Gamma_a \frac{d\mathcal{Z}^a}{d\tau} - \mathcal{H} \right) d\tau, \quad (84)$$

where the end points τ_1 and τ_2 are fixed. Hamilton's principle $\delta S = \int \delta \hat{\Gamma} = 0$ for single-particle motion in extended phase space yields

$$\begin{aligned} 0 &= \int \left(\delta Z^a \frac{\partial \Gamma_b}{\partial Z^a} dZ^b + \Gamma_a d\delta Z^a - \delta Z^a \frac{\partial \mathcal{H}}{\partial Z^a} d\tau \right) \\ &= \int \delta Z^a \left[\omega_{ab} dZ^b - \frac{\partial \mathcal{H}}{\partial Z^a} d\tau \right], \end{aligned} \quad (85)$$

where

$$\omega_{ab} \equiv \frac{\partial \Gamma_b}{\partial Z^a} - \frac{\partial \Gamma_a}{\partial Z^b} \quad (86)$$

denotes a component of the 8×8 antisymmetric Lagrange two-form $\omega \equiv d\Gamma$ (Goldstein *et al.*, 2002), and integration by parts of the second term was performed (with the usual assumption of virtual displacements δZ^a vanishing at the end points). Hence, stationarity of the particle action (84) yields the extended phase-space Euler-Lagrange equations

$$\omega_{ab} \frac{dZ^b}{d\tau} = \frac{\partial \mathcal{H}}{\partial Z^a}. \quad (87)$$

For regular (nonsingular) Lagrangian systems, the Lagrange matrix ω is invertible. The components of the inverse of the Lagrange matrix $J \equiv \omega^{-1}$, known as the antisymmetric Poisson matrix, are the fundamental Poisson brackets

$$(\omega^{-1})^{ab} \equiv \{Z^a, Z^b\} = J^{ab}(Z). \quad (88)$$

Using the identity relation

$$J^{ca} \omega_{ab} = \delta_b^c, \quad (89)$$

the Euler-Lagrange equations (87) become the extended Hamilton's equations

$$\frac{dZ^a}{d\tau} = J^{ab} \frac{\partial \mathcal{H}}{\partial Z^b} = \{Z^a, \mathcal{H}\}. \quad (90)$$

Using the identity (89), it can be shown that the Jacobi identity (79) holds if the Lagrange matrix satisfies the identity $d\omega = 0$, or

$$\partial_a \omega_{bc} + \partial_b \omega_{ca} + \partial_c \omega_{ab} = 0, \quad (91)$$

which is automatically satisfied since $\omega \equiv d\Gamma$ is an exact two-form (i.e., $\omega_{ab} = \partial_a \Gamma_b - \partial_b \Gamma_a$). Hence, any Poisson bracket derived through the sequence $\Gamma \rightarrow \omega = d\Gamma \rightarrow J = \omega^{-1}$ automatically satisfies the Jacobi identity (77).

Using the symplectic part of the extended phase-space Lagrangian (82), the Lagrange two-form $\omega \equiv d\Gamma$ is

$$\omega = dp_\mu \wedge dx^\mu + \frac{e}{2c} \epsilon_{ijk} B^k dx^i \wedge dx^j - \frac{e}{c} \frac{\partial A_i}{\partial t} dx^i \wedge dt, \quad (92)$$

from which, using the inverse relation (88), we construct the extended *noncanonical* Poisson bracket

$$\begin{aligned} \{f, g\} &= \frac{\partial f}{\partial x^\mu} \frac{\partial g}{\partial p_\mu} - \frac{\partial f}{\partial p_\mu} \frac{\partial g}{\partial x^\mu} + \frac{e \mathbf{B}}{c} \cdot \frac{\partial f}{\partial \mathbf{p}} \times \frac{\partial g}{\partial \mathbf{p}} \\ &\quad - \frac{e}{c} \frac{\partial \mathbf{A}}{\partial t} \cdot \left(\frac{\partial f}{\partial w} \frac{\partial g}{\partial \mathbf{p}} - \frac{\partial f}{\partial \mathbf{p}} \frac{\partial g}{\partial w} \right). \end{aligned} \quad (93)$$

The Hamiltonian dynamics in extended phase space is expressed as

$$\frac{d\mathbf{x}}{dt} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}} = \mathbf{v},$$

$$\frac{d\mathbf{p}}{dt} = -\nabla \mathcal{H} + \frac{e}{c} \left(\frac{\partial \mathbf{A}}{\partial t} \frac{\partial \mathcal{H}}{\partial w} + \frac{\partial \mathcal{H}}{\partial \mathbf{v}} \times \mathbf{B} \right) = e \left(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B} \right),$$

$$\frac{dw}{dt} = \frac{\partial \mathcal{H}}{\partial t} - \frac{e}{mc} \frac{\partial \mathbf{A}}{\partial t} \cdot \frac{\partial \mathcal{H}}{\partial \mathbf{v}} = e \left(\frac{\partial \phi}{\partial t} - \frac{\mathbf{v}}{c} \cdot \frac{\partial \mathbf{A}}{\partial t} \right),$$

where the Hamilton equation $dt/d\tau = \{t, \mathcal{H}\} = +1$ was used to substitute the orbit parameter τ with time t .

B. Perturbation theory in extended phase space

A variational formulation of single-particle perturbation theory, where the small dimensionless ordering parameter ϵ is used as a measure of the amplitude of the fluctuating electromagnetic fields, can be introduced through the new phase-space Lagrangian one-form (Brizard, 2001)

$$\hat{\Gamma}' \equiv \Gamma_a dZ^a - \mathcal{H} d\tau - \mathcal{S} d\epsilon, \quad (94)$$

where the symplectic components Γ_a and the Hamiltonian \mathcal{H} now depend on the perturbation parameter ϵ (e.g., either ϵ_B or ϵ_δ) and the scalar field \mathcal{S} is the generating function for an infinitesimal canonical transformation that smoothly deforms a particle's extended phase-space orbit from a reference orbit (at $\epsilon=0$) to a perturbed orbit (for $\epsilon \neq 0$). From the phase-space Lagrangian (94), we construct the action path integral $S'_C = \int_C \hat{\Gamma}'$ evaluated along a fixed path C in the (τ, ϵ) parameter space.

The modified principle of least action for perturbed single-particle motion in extended phase space,

$$0 = \int \delta Z^a \left[\omega_{ab} dZ^b - \frac{\partial \mathcal{H}}{\partial Z^a} d\tau - \left(\frac{\partial \mathcal{S}}{\partial Z^a} + \frac{\partial \Gamma_a}{\partial \epsilon} \right) d\epsilon \right],$$

whose derivation is similar to Eq. (85), now yields the extended perturbed Hamilton's equations,

$$\frac{dZ^a}{d\tau} = \{Z^a, \mathcal{H}\}, \quad (95)$$

$$\frac{dZ^a}{d\epsilon} = \{Z^a, \mathcal{S}\} - \frac{\partial \Gamma_b}{\partial \epsilon} \{Z^b, Z^a\}, \quad (96)$$

where Eq. (95) is identical to Eq. (90) except that the extended Hamiltonian \mathcal{H} and symplectic components Γ_a now depend on the perturbation parameter ϵ , while Eq.

(96) determines how particle orbits evolve under the perturbation ϵ flow.

The order of time evolution (τ flow) and perturbation evolution (ϵ flow) is not physically relevant. The commutativity of the two Hamiltonian (τ, ϵ) flows leads to the path independence of the action integral $\int \hat{\Gamma}'$ in the two-dimensional (τ, ϵ) orbit-parameter space. Considering two arbitrary paths C and \bar{C} with identical end points on the (τ, ϵ) parameter space and calculating the action path integrals $S'_C = \int_C \hat{\Gamma}'$ and $S'_{\bar{C}} = \int_{\bar{C}} \hat{\Gamma}'$, the path-independence condition $S'_C = S'_{\bar{C}}$ leads, by applying Stokes's theorem for differential one-forms (Flanders, 1989), to the condition

$$0 = \int_C \hat{\Gamma}' - \int_{\bar{C}} \hat{\Gamma}' \equiv \oint_{\partial D} \hat{\Gamma}' = \int_D d\hat{\Gamma}',$$

where D is the area enclosed by the closed path $\partial D \equiv C - \bar{C}$. Here the two-form $d\hat{\Gamma}'$ on the (τ, ϵ) parameter space is

$$\begin{aligned} d\hat{\Gamma}' &= d\epsilon \wedge d\tau \left[\frac{d\mathcal{Z}^a}{d\epsilon} \omega_{ab} \frac{d\mathcal{Z}^b}{d\tau} - \left(\frac{\partial \mathcal{H}}{\partial \epsilon} + \frac{\partial \mathcal{H}}{\partial \mathcal{Z}^a} \frac{d\mathcal{Z}^a}{d\epsilon} \right) \right. \\ &\quad \left. + \left(\frac{\partial \mathcal{S}}{\partial \mathcal{Z}^a} + \frac{\partial \Gamma_a}{\partial \epsilon} \right) \frac{d\mathcal{Z}^a}{d\tau} \right] \\ &= d\epsilon \wedge d\tau \left(\{\mathcal{S}, \mathcal{H}\} - \frac{\partial \mathcal{H}}{\partial \epsilon} + \frac{\partial \Gamma_a}{\partial \epsilon} \{\mathcal{Z}^a, \mathcal{H}\} \right), \end{aligned}$$

where Eqs. (95) and (96) were used. The condition of path independence requires that $d\hat{\Gamma}' = 0$, yielding the Hamiltonian perturbation equation

$$\frac{d\mathcal{S}}{d\tau} = \{\mathcal{S}, \mathcal{H}\} = \frac{\partial \mathcal{H}}{\partial \epsilon} - \frac{\partial \Gamma_a}{\partial \epsilon} \{\mathcal{Z}^a, \mathcal{H}\}, \quad (97)$$

relating the generating scalar field \mathcal{S} to the perturbation-parameter dependence of the extended Hamiltonian ($\partial_\epsilon \mathcal{H} = \phi_1 + \dots$) and Poisson bracket ($\partial_\epsilon \Gamma_a = \mathbf{A}_1 \cdot \partial \mathbf{x} / \partial \mathcal{Z}^a + \dots$). The perturbed evolution operator $d/d\tau = d_0/d\tau + \dots$ and the generating function $\mathcal{S} = S_1 + \dots$ are expanded in powers of ϵ , with the lower-order operator $d_0/d\tau$ considered to be explicitly integrable. In practice, the first-order term S_1 is solved explicitly as

$$S_1 = \left(\frac{d_0}{d\tau} \right)^{-1} \left[e\phi_1 - e\mathbf{A}_1 \cdot \frac{\mathbf{v}_0}{c} \right], \quad (98)$$

where $\mathbf{v}_0 \equiv d_0 \mathbf{x} / d\tau = \{\mathbf{x}, \mathcal{H}_0\}$ denotes the particle's unperturbed velocity. In order to determine higher-order S_n terms (for $n \geq 2$), however, a more systematic approach, based on applications of the Lie-transform perturbation method, is required.

C. Near-identity phase-space transformations

The Hamiltonian perturbation equation (97) arises naturally within the context of the dynamical reduction of single-particle Hamilton equations (95) through the

decoupling of fast *orbital* time scales from the relevant electromagnetic fluctuation time scales. The most efficient method for deriving reduced Hamilton equations is based on the Hamiltonian (Cary and Kaufman, 1981; Lichtenberg and Lieberman, 1984) or the phase-space Lagrangian (Cary and Littlejohn, 1983) Lie-transform perturbation methods.¹³

The process by which a fast time scale is removed from Hamilton's equations $\dot{\mathcal{Z}}^a = \{\mathcal{Z}^a, \mathcal{H}\}$ involves a near-identity transformation on extended particle phase space (Littlejohn, 1982a),

$$\mathcal{T}_\epsilon: \mathcal{Z} \rightarrow \bar{\mathcal{Z}}(\mathcal{Z}; \epsilon) \equiv \mathcal{T}_\epsilon \mathcal{Z}, \quad (99)$$

with $\bar{\mathcal{Z}}(\mathcal{Z}; 0) = \mathcal{Z}$, and its inverse

$$\mathcal{T}_\epsilon^{-1}: \bar{\mathcal{Z}} \rightarrow \mathcal{Z}(\bar{\mathcal{Z}}; \epsilon) \equiv \mathcal{T}_\epsilon^{-1} \bar{\mathcal{Z}}, \quad (100)$$

with $\mathcal{Z}(\bar{\mathcal{Z}}; 0) = \bar{\mathcal{Z}}$, where $\epsilon \ll 1$ denotes a dimensionless ordering parameter. By adopting the techniques of Lie-transform perturbation theory, these phase-space transformations are expressed in terms of generating vector fields $(\mathbf{G}_1, \mathbf{G}_2, \dots)$ as

$$\mathcal{T}_\epsilon^{\pm 1} \equiv \exp \left(\pm \sum_{n=1}^{\infty} \epsilon^n \mathbf{G}_n \cdot d \right), \quad (101)$$

where the n th-order generating vector field \mathbf{G}_n is chosen to remove the fast time scale at order ϵ^n from the perturbed Hamiltonian dynamics. The near-identity transformations (99) and (100) are explicitly written as

$$\bar{\mathcal{Z}}^a(\mathcal{Z}, \epsilon) = \mathcal{Z}^a + \epsilon G_1^a + \epsilon^2 \left(G_2^a + \frac{1}{2} G_1^b \frac{\partial G_1^a}{\partial \mathcal{Z}^b} \right) + \dots \quad (102)$$

and

$$\mathcal{Z}^a(\bar{\mathcal{Z}}, \epsilon) = \bar{\mathcal{Z}}^a - \epsilon G_1^a - \epsilon^2 \left(G_2^a - \frac{1}{2} G_1^b \frac{\partial G_1^a}{\partial \bar{\mathcal{Z}}^b} \right) + \dots \quad (103)$$

up to second order in the perturbation analysis. The new extended phase-space coordinates include the pair of fast action-angle coordinates $(\bar{J}_g \equiv \bar{\mu}B/\Omega, \bar{\zeta})$ and the reduced phase-space coordinates $\bar{\mathcal{Z}}_R$ such that the magnetic moment $\bar{\mu} = \bar{\mu}_0 + \epsilon \bar{\mu}_1 + \dots$ is an exact invariant of the reduced Hamiltonian dynamics and the Hamiltonian dynamics of the reduced coordinates $\bar{\mathcal{Z}}_R$ is independent of the fast angle $\bar{\zeta}$.

Using the transformation (99), the pull-back operator on scalar fields (Abraham and Marsden, 1978; Littlejohn, 1982a) induced by the near-identity transformation is defined as (99)

¹³Hamiltonian Lie-transform perturbation theory is a special case of phase-space Lagrangian Lie-transform perturbation theory, in which the Poisson-bracket—or symplectic—structure is unperturbed.

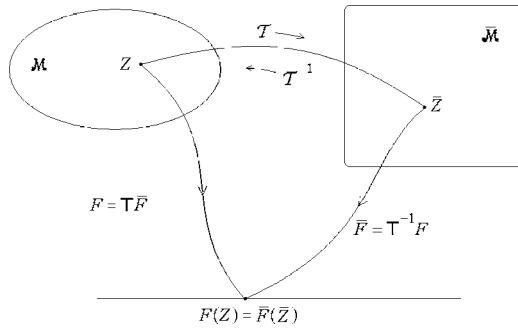


FIG. 5. The phase-space transformation $\bar{Z} = T\mathcal{Z}$ and its inverse $\mathcal{Z} = T^{-1}\bar{Z}$ induce a pull-back operator $\mathcal{F} = T\bar{\mathcal{F}}$ and a push-forward operator $\bar{\mathcal{F}} = T^{-1}\mathcal{F}$.

$$T_\epsilon \cdot \bar{\mathcal{F}} \rightarrow \mathcal{F} \equiv T_\epsilon \bar{\mathcal{F}}, \quad (104)$$

i.e., the pull-back operator T_ϵ transforms a scalar field $\bar{\mathcal{F}}$ on the phase space with coordinates \bar{Z} into a scalar field \mathcal{F} on the phase space with coordinates \mathcal{Z} ,

$$\mathcal{F}(\mathcal{Z}) = T_\epsilon \bar{\mathcal{F}}(\mathcal{Z}) = \bar{\mathcal{F}}(T_\epsilon \mathcal{Z}) = \bar{\mathcal{F}}(\bar{Z}).$$

Using the inverse transformation (100), the push-forward operator is defined as (Littlejohn, 1982a)

$$T_\epsilon^{-1} \cdot \mathcal{F} \rightarrow \bar{\mathcal{F}} \equiv T_\epsilon^{-1} \mathcal{F}, \quad (105)$$

i.e., the push-forward operator T_ϵ^{-1} transforms a scalar field \mathcal{F} on the phase space with coordinates \mathcal{Z} into a scalar field $\bar{\mathcal{F}}$ on the phase space with coordinates \bar{Z} ,

$$\bar{\mathcal{F}}(\bar{Z}) = T_\epsilon^{-1} \mathcal{F}(\bar{Z}) = \mathcal{F}(T_\epsilon^{-1} \bar{Z}) = \mathcal{F}(\mathcal{Z}).$$

The pull-back and push-forward operators are illustrated in Fig. 5.

The pull-back and push-forward operators can be used to transform an arbitrary operator $\mathcal{C}: F(\mathcal{Z}) \rightarrow \mathcal{C}[F](\mathcal{Z})$ acting on the extended Vlasov distribution function F . First, since $\mathcal{C}[\mathcal{F}](\mathcal{Z})$ is a scalar field, it transforms to $T_\epsilon^{-1} \mathcal{C}[\mathcal{F}](\bar{Z})$ with the help of the push-forward operator (105). Next, we replace the extended Vlasov distribution function F with its pull-back representation $F = T_\epsilon \bar{F}$ and define the transformed operator \mathcal{C}_ϵ as

$$\mathcal{C}_\epsilon[\bar{\mathcal{F}}] \equiv T_\epsilon^{-1}(\mathcal{C}[T_\epsilon \bar{F}]). \quad (106)$$

This induced transformation is applied to the Vlasov equation in extended phase space,

$$\frac{d\mathcal{F}}{d\tau} \equiv \{\mathcal{F}, \mathcal{H}\}_{\mathcal{Z}} = 0, \quad (107)$$

where $d/d\tau$ defines the total derivative along a particle orbit in extended phase space and $\{\cdot, \cdot\}_{\mathcal{Z}}$ denotes the extended Poisson bracket on the original extended phase space (with coordinates \mathcal{Z}). Hence, the transformed Vlasov equation is written as

$$0 = \frac{d_\epsilon \bar{\mathcal{F}}}{d\tau} \equiv T_\epsilon^{-1} \left(\frac{d}{d\tau} T_\epsilon \bar{\mathcal{F}} \right) = \{\bar{\mathcal{F}}, \bar{\mathcal{H}}\}_{\bar{Z}}, \quad (108)$$

where the total derivative along the transformed particle orbit $d_\epsilon/d\tau$ is defined in terms of the transformed Poisson bracket $\{\cdot, \cdot\}_{\bar{Z}}$ and the transformed Hamiltonian

$$\bar{\mathcal{H}} \equiv T_\epsilon^{-1} \mathcal{H}. \quad (109)$$

The transformation of the Poisson bracket by Lie-transform methods is performed through the transformation of the extended phase-space Lagrangian, expressed as

$$\bar{\Gamma} = T_\epsilon^{-1} \Gamma + d\mathcal{S}, \quad (110)$$

where \mathcal{S} denotes a (canonical) scalar field used to simplify the transformed phase-space Lagrangian (110), i.e., it has no impact on the new Poisson-bracket structure

$$\bar{\omega} = d\bar{\Gamma} = d(T_\epsilon^{-1} \Gamma) = T_\epsilon^{-1} d\Gamma \equiv T_\epsilon^{-1} \omega, \quad (111)$$

since $d^2\mathcal{S}=0$ (i.e., $\partial_{ab}^2 \mathcal{S} - \partial_{ba}^2 \mathcal{S} = 0$) and T_ϵ^{-1} commutes with d (see Appendix A).

Note that the extended-Hamiltonian transformation (109) may be reexpressed in terms of the regular Hamiltonians H and \bar{H} as

$$\bar{H} = T_\epsilon^{-1} H - \frac{\partial \mathcal{S}}{\partial t}, \quad (112)$$

where \mathcal{S} is the canonical scalar field introduced in Eq. (110); note the similarity with Eq. (25). The new extended phase-space coordinates are chosen so that $d_\epsilon \bar{Z}^a/d\tau = \{\bar{Z}^a, \bar{\mathcal{H}}\}_{\bar{Z}}$ are independent of the fast angle $\bar{\zeta}$ and the adiabatic invariant $\bar{\mu}$ satisfies the exact equation $d_\epsilon \bar{\mu}/d\tau = 0$. The dynamical reduction of single-particle Hamiltonian dynamics has been successfully achieved by phase-space transformation via the construction of a fast invariant $\bar{\mu}$ with its canonically conjugate fast angle $\bar{\zeta}$ becoming an ignorable coordinate.

D. Lie-transform methods

In Lie-transform perturbation theory (Littlejohn, 1982a), the pull-back and push-forward operators (104) and (105) are expressed as Lie transforms,

$$T_\epsilon^{\pm 1} \equiv \exp \left(\pm \sum_{n=1} \epsilon^n \mathcal{L}_n \right), \quad (113)$$

where \mathcal{L}_n denotes the Lie derivative generated by the n th-order vector field G_n (Abraham and Marsden, 1978). A Lie derivative is a special differential operator that preserves the tensorial nature of the object it operates on (see Appendix A for more details). In Eq. (109), for example, the Lie derivative $\mathcal{L}_n \mathcal{H}$ of the scalar field \mathcal{H} is defined as the scalar field

$$\mathcal{L}_n \mathcal{H} \equiv G_n^a \partial_a \mathcal{H}. \quad (114)$$

In Eq. (110) the Lie derivative $\mathcal{L}_n \Gamma$ of a one-form $\Gamma \equiv \Gamma_a d\mathcal{Z}^a$ is defined as the one-form (Abraham and Marsden, 1978)

$$\mathcal{L}_n \Gamma \equiv \mathbf{G}_n \cdot d\Gamma + d(\mathbf{G}_n \cdot \Gamma) = [G_n^a \omega_{ab} + \partial_b(G_n^a \Gamma_a)] d\mathcal{Z}^b, \quad (115)$$

where $\omega_{ab} \equiv \partial_a \Gamma_b - \partial_b \Gamma_a$ are the components of the two-form $\omega \equiv d\Gamma$. At each order ϵ^n , the terms $d(\mathbf{G}_n \cdot \Gamma)$ can be absorbed in the gauge term dS_n in Eq. (110).

1. Transformed extended Poisson-bracket structure

We now write the extended phase-space Lagrangian $\Gamma \equiv \Gamma_0 + \epsilon \Gamma_1$ and the extended Hamiltonian $\mathcal{H} \equiv \mathcal{H}_0 + \epsilon \mathcal{H}_1$ in terms of an unperturbed (zeroth-order) part and a perturbation (first-order) part. The Lie-transform relations associated with Eq. (110) are expressed (up to second order in ϵ) as $\bar{\Gamma}_{0a} \equiv \Gamma_{0a}$ and

$$\bar{\Gamma}_1 = \Gamma_1 - \mathbf{G}_1 \cdot \omega_0 + dS_1, \quad (116)$$

$$\bar{\Gamma}_2 = -\mathbf{G}_2 \cdot \omega_0 - \frac{1}{2}\mathbf{G}_1 \cdot (\omega_1 + \bar{\omega}_1) + dS_2. \quad (117)$$

A general form for the new Poisson bracket $\{ , \}_{\bar{\mathcal{Z}}}$ is obtained by allowing the new phase-space Lagrangian to retain symplectic perturbation terms $\bar{\Gamma} \equiv \bar{\Gamma}_0 + \epsilon \bar{\Gamma}_1$. By choosing a specific form for the perturbed gyrocenter symplectic structure $\bar{\Gamma}_1$, Eqs. (116) and (117) can be solved for the generating vector field $(\mathbf{G}_1, \mathbf{G}_2)$ expressed in terms of the scalar fields (S_1, S_2) . The new phase-space Lagrangian $\bar{\Gamma} \equiv \bar{\Gamma}_R + \bar{J}_g d\bar{\zeta}$, where the reduced phase-space Lagrangian $\bar{\Gamma}_R$ is independent of the fast angle $\bar{\zeta}$ and, by application of the Noether theorem (Cary, 1977), the canonically conjugate action \bar{J}_g is an invariant (i.e., $d\bar{J}_g/dt = \{\bar{J}_g, \bar{\mathcal{H}}\}_{\bar{\mathcal{Z}}} = 0$).

The first-order generating vector field \mathbf{G}_1 needed to obtain the gyrocenter extended phase-space Lagrangian (116) is

$$G_1^a = \{S_1, \mathcal{Z}^a\}_0 + (\Gamma_{1b} - \bar{\Gamma}_{1b}) J_0^{ba}, \quad (118)$$

where $\{ , \}_0$ is the Poisson-bracket structure associated with the unperturbed Poisson matrix J_0^{ab} . The generating vector field (118) is divided into two parts: a canonical part generated by the gauge function S_1 and a symplectic part generated by the difference $\Delta \Gamma_{1b} \equiv \Gamma_{1b} - \bar{\Gamma}_{1b}$ between the old and new phase-space Lagrangian symplectic components.

For the second-order generating vector field \mathbf{G}_2 , the condition $\bar{\Gamma}_2 \equiv 0$ yields the following solution for \mathbf{G}_2 in terms of the scalar field S_2 :

$$G_2^a = \{S_2, \mathcal{Z}^a\}_0 - \frac{1}{2}G_1^b(\omega_{1bc} + \bar{\omega}_{1bc})J_0^{ca}, \quad (119)$$

where ω_{1bc} and $\bar{\omega}_{1bc}$ are components of the first-order perturbed Lagrange matrices. The second-order gener-

ating field (119) is, once again, divided into a canonical part (generated by S_2) and a symplectic part (generated by Γ_{1b} and $\bar{\Gamma}_{1b}$).

The near-identity extended-phase-space transformation (102) is expressed in terms of the asymptotic expansion

$$\bar{\mathcal{Z}}^a = \mathcal{Z}^a + \epsilon(\{S_1, \mathcal{Z}^a\}_0 + \Delta \Gamma_{1b} J_0^{ba}) + O(\epsilon^2), \quad (120)$$

and its explicit expression requires a solution of the scalar fields (S_1, \dots) ; for most practical applications, however, only the first-order function S_1 is needed.

2. Transformed extended Hamiltonian

By substituting the generating vector fields (118) and (119) into the Lie-transform relations associated with Eq. (109),

$$\bar{H}_1 = H_1 - \mathbf{G}_1 \cdot d\mathcal{H}_0, \quad (121)$$

$$\bar{H}_2 = -\mathbf{G}_2 \cdot d\mathcal{H}_0 - \frac{1}{2}\mathbf{G}_1 \cdot d(H_1 + \bar{H}_1), \quad (122)$$

the first-order and second-order terms in the transformed extended Hamiltonian are obtained,

$$\begin{aligned} \bar{H}_1 &= H_1 - (\Gamma_{1a} - \bar{\Gamma}_{1a}) \dot{\mathcal{Z}}_0^a - \{S_1, \mathcal{H}_0\}_0 \\ &\equiv (K_1 + \bar{\Gamma}_{1a} \dot{\mathcal{Z}}_0^a) - \{S_1, \mathcal{H}_0\}_0 \end{aligned} \quad (123)$$

and

$$\begin{aligned} \bar{H}_2 &= -\{S_2, \mathcal{H}_0\}_0 - \frac{1}{2}G_1^a \partial_a (K_1 + \bar{K}_1) \\ &\quad - \frac{1}{2}G_1^b \{(\Gamma_{1b} + \bar{\Gamma}_{1b}), \mathcal{H}_0\}_0 \\ &\quad - \frac{1}{2}(\Gamma_{1b} + \bar{\Gamma}_{1b})(G_1^a \partial_a \dot{\mathcal{Z}}_0^b), \end{aligned} \quad (124)$$

where we have used the Poisson-bracket properties (75)–(77). In Eqs. (123) and (124), $\dot{\mathcal{Z}}_0^a \equiv \{\mathcal{Z}^a, \mathcal{H}_0\}_0$ denotes the zeroth-order Hamilton equations and

$$K_1 \equiv H_1 - \Gamma_{1a} \dot{\mathcal{Z}}_0^a \quad (125)$$

denotes the effective first-order Hamiltonian (and $\bar{K}_1 \equiv \bar{H}_1 - \bar{\Gamma}_{1a} \dot{\mathcal{Z}}_0^a$). The choice of $\bar{\Gamma}_1$, that is relevant only for magnetic perturbations affects both the new Poisson-bracket structure $\{ , \}_{\bar{\mathcal{Z}}}$ and the new Hamiltonian $\bar{\mathcal{H}} \equiv \bar{H} - \bar{w}$.

The two Hamiltonian relations (123) and (124) contain terms on the right side that exhibit both fast and slow time-scale dependence: the slow-time-scale terms are explicitly identified with the new Hamiltonian term \bar{H}_n on the left side, while the fast-time-scale terms are used to define the gauge function S_n . The solution for the new first-order Hamiltonian (123) is expressed in terms of the fast-angle averaging operation $\langle \cdots \rangle$ as

$$\bar{H}_1 \equiv \langle K_1 \rangle + \bar{\Gamma}_{1a} \dot{\mathcal{Z}}_0^a, \quad (126)$$

where the Poisson bracket $\{ , \}$ henceforth denotes the zeroth-order Poisson bracket $\{ , \}_0$ (unless otherwise

noted) and S_1 can be chosen such that $\langle S_1 \rangle = 0$. The first-order gauge function S_1 is determined from the perturbation equation

$$\frac{d_0 S_1}{d\tau} \equiv \{S_1, \mathcal{H}_0\} = \tilde{K}_1 \equiv K_1 - \langle K_1 \rangle, \quad (127)$$

whose solution is $S_1 \equiv (d_0/d\tau)^{-1} \tilde{K}_1$, where $(d_0/d\tau)^{-1}$ denotes integration along an unperturbed extended Hamiltonian orbit. This formal solution is identical to the solution (98) obtained using variational methods. To lowest order in the fast orbital time scale, the unperturbed integration

$$S_1 = (d_0/d\tau)^{-1} \tilde{K}_1 \equiv \Omega^{-1} \int \tilde{K}_1 d\zeta \quad (128)$$

involves an indefinite fast-angle integration, where $\Omega \equiv d_0 \zeta / dt$ denotes the fast-angle frequency; note that the solution (128) for the first-order gauge function S_1 does not depend on the choice of $\bar{\Gamma}_1$.

The solution for the new second-order Hamiltonian (124) yields the fast-angle-averaged expression

$$\begin{aligned} \bar{H}_2 &= -\frac{1}{2} \langle \{S_1, \{S_1, \mathcal{H}_0\}\} \rangle - \langle \Delta \Gamma_{1a} \{\mathcal{Z}^a, K_1\} \rangle \\ &\quad - \frac{1}{2} \langle \Delta \Gamma_{1a} J_0^{ab} \{(\Gamma_{1b} + \bar{\Gamma}_{1b}), \mathcal{H}_0\} \rangle \\ &\quad - \frac{1}{2} \langle \Delta \Gamma_{1a} \{\mathcal{Z}^a, \dot{\mathcal{Z}}_0^b\} (\Gamma_{1b} + \bar{\Gamma}_{1b}) \rangle, \end{aligned} \quad (129)$$

where the first term corresponds to the standard quadratic ponderomotive Hamiltonian (Cary and Kaufman, 1981). The remaining terms (which depend on the symplectic choice $\bar{\Gamma}_1$) will be discussed below. The second-order gauge function S_2 appearing in Eq. (124) is not needed in what follows since the phase-space transformation from guiding-center to gyrocenter coordinates is only needed to first order in ϵ_s .

E. Reduced Vlasov-Maxwell equations

Having derived an expression for the reduced Hamiltonian, $\bar{\mathcal{H}} = (\bar{H}_0 - \bar{w}) + \epsilon \bar{H}_1 + \epsilon^2 \bar{H}_2 + \dots$, where the perturbation terms are given in Eqs. (126) and (129), expressions for the self-consistent reduced Vlasov-Maxwell equations are now derived.

The extended Vlasov equation (107) may be converted into the regular Vlasov equation. In order to satisfy the physical constraint (81), the extended Vlasov distribution is

$$\mathcal{F}(\mathcal{Z}) \equiv c \delta[w - H(\mathbf{z}, t)] f(\mathbf{z}, t), \quad (130)$$

where $f(\mathbf{z}, t)$ denotes the time-dependent Vlasov distribution on regular phase space $\mathbf{z} = (\mathbf{x}, \mathbf{p})$. By integrating the extended Vlasov equation (107) over the energy coordinate w (and using $d\tau = dt$), the regular Vlasov equation is obtained

$$0 = \frac{df}{dt} \equiv \frac{\partial f}{\partial t} + \frac{d\mathbf{z}}{dt} \cdot \frac{\partial f}{\partial \mathbf{z}}. \quad (131)$$

The push-forward transformation of the extended Vlasov distribution (130) yields the reduced extended Vlasov distribution,

$$\bar{\mathcal{F}}(\bar{\mathcal{Z}}) \equiv c \delta[\bar{w} - \bar{H}(\bar{\mathbf{z}}, t)] \bar{f}(\bar{\mathbf{z}}, t), \quad (132)$$

where the reduced extended Hamiltonian $\bar{\mathcal{H}} \equiv \bar{H}(\bar{\mathbf{z}}, t) - \bar{w}$ is defined in Eq. (109). The extended reduced Vlasov equation

$$\frac{d_\epsilon \bar{\mathcal{F}}}{d\tau} \equiv \{\bar{\mathcal{F}}, \bar{\mathcal{H}}\}_\epsilon = 0 \quad (133)$$

can then be converted into the regular reduced Vlasov equation by integrating it over the reduced energy coordinate \bar{w} , yielding the reduced Vlasov equation

$$0 = \frac{d_\epsilon \bar{f}}{dt} \equiv \frac{\partial \bar{f}}{\partial t} + \frac{d_\epsilon \bar{\mathbf{z}}}{dt} \cdot \frac{\partial \bar{f}}{\partial \bar{\mathbf{z}}}, \quad (134)$$

where $\bar{f}(\bar{\mathbf{z}}, t)$ is the time-dependent reduced Vlasov distribution on the new reduced phase space. The pull-back and push-forward operators play a fundamental role in the transformation of the Vlasov equation to the reduced Vlasov equation.

We then investigate how the pull-back and push-forward operators (113) are used in the transformation of Maxwell's equations,

$$\nabla \cdot \mathbf{E} = 4\pi\rho, \quad (135)$$

$$\nabla \times \mathbf{B} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} = \frac{4\pi}{c} \mathbf{J}, \quad (136)$$

where the charge-current densities

$$\begin{pmatrix} \rho \\ \mathbf{J} \end{pmatrix} = \sum e \int d^4 p \mathcal{F} \begin{pmatrix} 1 \\ \mathbf{v} \end{pmatrix} \quad (137)$$

are defined in terms of the extended Vlasov distribution \mathcal{F} (with $d^4 p = c^{-1} dw d^3 p$) and the electric and magnetic fields $\mathbf{E} \equiv -\nabla\phi - c^{-1} \partial \mathbf{A} / \partial t$ and $\mathbf{B} \equiv \nabla \times \mathbf{A}$ satisfy the constraints $\nabla \cdot \mathbf{B} = 0$ and $\nabla \times \mathbf{E} + c^{-1} \partial_t \mathbf{B} = 0$.

The charge-current densities (137) can be expressed in terms of the general expression

$$\begin{aligned} \|v^\mu\|(\mathbf{r}) &\equiv \int d^3 p v^\mu f = \int d^4 p v^\mu \mathcal{F} \\ &= \int d^3 x \int d^4 p v^\mu \delta^3(\mathbf{x} - \mathbf{r}) \mathcal{F}, \end{aligned} \quad (138)$$

where time dependence is omitted for clarity (since time itself is unaffected by the transformations considered here), $v^\mu = (c, \mathbf{v})$, and the delta function $\delta^3(\mathbf{x} - \mathbf{r})$ means that only particles whose positions \mathbf{x} coincide with the field position \mathbf{r} contribute to the moment $\|v^\mu\|(\mathbf{r})$. By applying the extended (time-dependent) phase-space transformation $\mathcal{T}_\epsilon: \mathcal{Z} \rightarrow \bar{\mathcal{Z}}$ on the right side of Eq. (138), we

obtain the push-forward representation for the fluid moments $\|\nu^\mu\|$,

$$\begin{aligned} \|\nu^\mu\|(\mathbf{r}) &= \int d^3\bar{x} \int d^4\bar{p} (\mathbf{T}_\epsilon^{-1}\nu^\mu) \delta^3(\bar{\mathbf{x}} + \rho_\epsilon - \mathbf{r}) \bar{\mathcal{F}} \\ &= \int d^3\bar{p} e^{-\rho_\epsilon \nabla} [(\mathbf{T}_\epsilon^{-1}\nu^\mu) \bar{f}], \end{aligned} \quad (139)$$

where \bar{w} integration was performed, $\mathbf{T}_\epsilon^{-1}\nu^\mu = (c, \mathbf{T}_\epsilon^{-1}\mathbf{v})$ is the push-forward of the particle four-velocity ν^μ , and

$$\rho_\epsilon \equiv \mathbf{T}_\epsilon^{-1}\mathbf{x} - \bar{\mathbf{x}} = -\epsilon G_1^x - \epsilon^2 \left(G_2^x - \frac{1}{2} \mathbf{G}_1 \cdot dG_1^x \right) + \dots \quad (140)$$

is the displacement between the push-forward $\mathbf{T}_\epsilon^{-1}\mathbf{x}$ of the particle position \mathbf{x} and the (new) reduced position $\bar{\mathbf{x}}$ that is defined in terms of the generating vector fields (118) and (119).¹⁴

The push-forward representation for the charge-current densities introduces polarization and magnetization effects into the Maxwell equations that transforms the microscopic Maxwell's equations (135) and (136) into the macroscopic (reduced) equations

$$\nabla \cdot \mathbf{D} = 4\pi\bar{\rho}, \quad (141)$$

$$\nabla \times \mathbf{H} - \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t} = \frac{4\pi}{c} \bar{\mathbf{J}}, \quad (142)$$

where the reduced charge-current densities $(\bar{\rho}, \bar{\mathbf{J}})$ are defined as moments of the reduced Vlasov distribution $\bar{\mathcal{F}}$,

$$\begin{pmatrix} \bar{\rho} \\ \bar{\mathbf{J}} \end{pmatrix} = \sum e \int d^4\bar{p} \bar{\mathcal{F}} \begin{pmatrix} 1 \\ \bar{\mathbf{v}} \end{pmatrix}, \quad (143)$$

and the microscopic electric and magnetic fields \mathbf{E} and \mathbf{B} are replaced by the macroscopic fields (Jackson, 1975)

$$\begin{aligned} \mathbf{D} &= \mathbf{E} + 4\pi\mathbf{P}_\epsilon, \\ \mathbf{H} &= \mathbf{B} - 4\pi\mathbf{M}_\epsilon. \end{aligned} \quad (144)$$

Here \mathbf{P}_ϵ and \mathbf{M}_ϵ denote the polarization and magnetization vectors associated with the dynamical reduction introduced by the phase-space transformation (102). The relation between the particle charge-current densities (ρ, \mathbf{J}) and the reduced charge-current densities $(\bar{\rho}, \bar{\mathbf{J}})$,

$$\rho \equiv \bar{\rho} - \nabla \cdot \mathbf{P}_\epsilon, \quad (145)$$

$$\mathbf{J} \equiv \bar{\mathbf{J}} + \frac{\partial \mathbf{P}_\epsilon}{\partial t} + c \nabla \times \mathbf{M}_\epsilon, \quad (146)$$

defines the polarization density $\rho_{\text{pol}} \equiv -\nabla \cdot \mathbf{P}_\epsilon$, the polarization current $\mathbf{J}_{\text{pol}} \equiv \partial \mathbf{P}_\epsilon / \partial t$, and the magnetization cur-

rent $\mathbf{J}_{\text{mag}} \equiv c \nabla \times \mathbf{M}_\epsilon$. The derivation of the polarization and magnetization vectors \mathbf{P}_ϵ and \mathbf{M}_ϵ is done either directly by the push-forward method (139) or by variational method

$$(\mathbf{D}, \mathbf{H}) \equiv 4\pi \left(\frac{\partial \bar{\mathcal{L}}}{\partial \mathbf{E}}, - \frac{\partial \bar{\mathcal{L}}}{\partial \mathbf{B}} \right), \quad (147)$$

where $\bar{\mathcal{L}}$ is the Lagrangian density for the reduced Vlasov-Maxwell equations. While the direct push-forward method is relatively straightforward to use (see Appendix C for details), the variational method allows a direct derivation of the exact conservation laws (e.g., energy) for the reduced Vlasov-Maxwell equations (see Secs. III and IV.B).

V. NONLINEAR GYROKINETIC VLASOV EQUATION

We now apply the methods of Lie-transform perturbation theory presented in Sec. IV to the dynamical reduction associated with the perturbed dynamics of charged particles (mass m and charge e) moving in a background time-independent magnetic field $\mathbf{B}_0 = \nabla \times \mathbf{A}_0$ in the presence of low-frequency electromagnetic fluctuations represented by the perturbation four-potential $\delta A^\mu = (\delta\phi, \delta\mathbf{A})$, whose amplitude is ordered with a dimensionless parameter $\epsilon_\delta \ll 1$. We focus our attention on deriving the nonlinear gyrocenter Hamiltonian and the associated nonlinear gyrokinetic Vlasov equation and postpone the derivation of the self-consistent gyrokinetic Maxwell equations and the gyrokinetic energy conservation law to Sec. VI.

The eight-dimensional extended phase-space dynamics is expressed in terms of the extended phase-space Lagrangian $\Gamma = \Gamma_0 + \epsilon_\delta \Gamma_1$, where $\Gamma_0 \equiv [(e/c)\mathbf{A}_0 + \mathbf{p}] \cdot d\mathbf{x} - wdt$ and $\Gamma_1 \equiv (e/c)\delta\mathbf{A} \cdot d\mathbf{x}$, and the extended phase-space Hamiltonian $\mathcal{H} = \mathcal{H}_0 + \epsilon_\delta H_1$, where $\mathcal{H}_0 = |\mathbf{p}|^2/2m - w$ and $H_1 = e\delta\phi$. While electrostatic fluctuations perturb the Hamiltonian alone, full electromagnetic fluctuations perturb both the Hamiltonian (H_1) and the symplectic one-form (Γ_1).

The standard gyrokinetic analysis for magnetized plasmas perturbed by low-frequency electromagnetic fluctuations (Brizard, 1989a) proceeds by a sequence of two near-identity phase-space transformations: a time-independent guiding-center phase-space transformation and a time-dependent gyrocenter phase-space transformation. This two-step decoupling procedure removes, first, the fast gyromotion space-time scales associated with the (unperturbed) background magnetic field (first step, guiding-center transformation with ordering parameter ϵ_B) and, second, the fast gyromotion time scale associated with the perturbation electromagnetic fields (second step, gyrocenter transformation with ordering parameters ϵ_δ , ϵ_ω , and ϵ_\perp).

¹⁴We note the similarity between the general form (139) for the push-forward representation of fluid moments and the Frieman-Chen expression (18) for the perturbed plasma density δn .

A. Unperturbed guiding-center Hamiltonian dynamics

The guiding-center phase-space transformation involves an asymptotic expansion, with a small dimensionless parameter $\epsilon_B \equiv \rho_{\text{th}}/L_B \ll 1$ defined as the ratio of the thermal gyroradius ρ_{th} and the background magnetic-field length scale L_B . This transformation is designed to remove the fast gyromotion time scale associated with the time-independent background magnetic field \mathbf{B}_0 associated with an unperturbed magnetized plasma (Littlejohn, 1983). In previous work (Brizard, 1995), this transformation was carried out to second order in ϵ_B with the scalar potential Φ_0 ordered at zeroth order in ϵ_B ; in this section, the equilibrium scalar potential is set equal to zero and issues associated with an inhomogeneous equilibrium electric field are discussed in Appendix D.1.

The results of the guiding-center analysis presented by Littlejohn (1983) are summarized as follows (further details are presented in Appendix B). The guiding-center transformation yields the following guiding-center coordinates $(\mathbf{X}, p_{\parallel}, \mu, \zeta, w, t) \equiv \mathcal{Z}_{\text{gc}}$, where \mathbf{X} is the guiding-center position, p_{\parallel} is the guiding-center kinetic momentum parallel to the unperturbed magnetic field, μ is the guiding-center magnetic moment, ζ is the gyroangle, and (w, t) are the canonically conjugate guiding-center energy-time coordinates (time is unaffected by the transformation while the guiding-center kinetic energy is equal to the particle kinetic energy). The unperturbed guiding-center extended phase-space Lagrangian is

$$\Gamma_{\text{gc}} \equiv \frac{e}{c\epsilon_B} \mathbf{A}_0^* \cdot d\mathbf{X} + \epsilon_B \mu (mc/e) d\zeta - w dt, \quad (148)$$

where $\mathbf{A}_0^* \equiv \mathbf{A}_0 + \epsilon_B (c/e) p_{\parallel} \hat{\mathbf{b}}_0 + \mathcal{O}(\epsilon_B^2)$ is the effective unperturbed vector potential, with $\hat{\mathbf{b}}_0 \equiv \mathbf{B}_0/B_0$ and higher-order correction terms are omitted (see Appendix B for further details); we omit displaying the dimensionless guiding-center parameter ϵ_B for simplicity. The unperturbed extended phase-space guiding-center Hamiltonian is

$$\mathcal{H}_{\text{gc}} = \frac{p_{\parallel}^2}{2m} + \mu B_0 - w \equiv H_{\text{gc}} - w. \quad (149)$$

Last, from the unperturbed guiding-center phase-space Lagrangian (148), the unperturbed guiding-center Poisson bracket $\{ \cdot, \cdot \}_{\text{gc}}$ is obtained, given here in terms of two arbitrary functions \mathcal{F} and \mathcal{G} on extended guiding-center phase space as (Littlejohn, 1983)

$$\begin{aligned} \{\mathcal{F}, \mathcal{G}\}_{\text{gc}} &\equiv \frac{e}{mc} \left(\frac{\partial \mathcal{F}}{\partial \zeta} \frac{\partial \mathcal{G}}{\partial \mu} - \frac{\partial \mathcal{F}}{\partial \mu} \frac{\partial \mathcal{G}}{\partial \zeta} \right) \\ &+ \frac{\mathbf{B}_0^*}{B_{0\parallel}} \cdot \left(\nabla \mathcal{F} \frac{\partial \mathcal{G}}{\partial p_{\parallel}} - \frac{\partial \mathcal{F}}{\partial p_{\parallel}} \nabla \mathcal{G} \right) \\ &- \frac{c \hat{\mathbf{b}}_0}{e B_{0\parallel}} \cdot \nabla \mathcal{F} \times \nabla \mathcal{G} + \left(\frac{\partial \mathcal{F}}{\partial w} \frac{\partial \mathcal{G}}{\partial t} - \frac{\partial \mathcal{F}}{\partial t} \frac{\partial \mathcal{G}}{\partial w} \right), \end{aligned} \quad (150)$$

where $\mathbf{B}_0^* \equiv \nabla \times \mathbf{A}_0^*$ and $B_{0\parallel}^* \equiv \hat{\mathbf{b}}_0 \cdot \mathbf{B}_0^*$ are defined as

$$\mathbf{B}_0^* = \mathbf{B}_0 + (c/e) p_{\parallel} \nabla \times \hat{\mathbf{b}}_0,$$

$$B_{0\parallel}^* = B_0 + (c/e) p_{\parallel} \hat{\mathbf{b}}_0 \cdot \nabla \times \hat{\mathbf{b}}_0. \quad (151)$$

The Jacobian of the guiding-center transformation is $J_{\text{gc}} = m B_{0\parallel}^*$ (i.e., $d^3 x d^3 p = J_{\text{gc}} d^3 X d\mu d\zeta$) and the background magnetic field is assumed to be a time-independent field (e.g., on time scales shorter than collisional time scales) so that the time derivative $\partial \mathbf{A}_0 / \partial t$ is absent from the Poisson bracket (150). The unperturbed guiding-center Hamiltonian dynamics is expressed in terms of the Hamiltonian (149) and the Poisson bracket (150) as $\mathcal{Z}^a = \{\mathcal{Z}^a, \mathcal{H}_{\text{gc}}\}_{\text{gc}}$. The conservation law $\dot{\mu} = 0$ for the guiding-center magnetic moment follows from the fact that the guiding-center Hamiltonian (149) is independent of the fast gyroangle ζ (to arbitrary order in ϵ_B). The first three terms in the guiding-center Poisson bracket (150) are arranged in order of increasing powers of $\epsilon_B \sim e^{-1}$: at order $\epsilon_B^{-1} \sim e$, the first term represents the fast gyromotion; at order $\epsilon_B^0 \sim e^0$, the next term represents the bounce/transit motion parallel to the magnetic-field line; at order $\epsilon_B \sim e^{-1}$, the third term represents the slow drift motion across magnetic-field lines.

B. Perturbed guiding-center Hamiltonian dynamics

We now consider how the guiding-center Hamiltonian system $(\mathcal{H}_{\text{gc}}, \{ \cdot, \cdot \}_{\text{gc}})$ is affected by the introduction of low-frequency electromagnetic field fluctuations $(\delta\phi, \delta\mathbf{A})$ satisfying the low-frequency gyrokinetic orderings (6)–(10). Under the electromagnetic potential perturbations $(\delta\phi, \delta\mathbf{A})$, the guiding-center phase-space Lagrangian (148) and Hamiltonian (149) become

$$\Gamma'_{\text{gc}} \equiv \Gamma_{0\text{gc}} + \epsilon_{\delta} \Gamma_{1\text{gc}},$$

$$\mathcal{H}'_{\text{gc}} \equiv \mathcal{H}_{0\text{gc}} + \epsilon_{\delta} \mathcal{H}_{1\text{gc}}, \quad (152)$$

where the zeroth-order guiding-center phase-space Lagrangian $\Gamma_{0\text{gc}}$ and Hamiltonian $\mathcal{H}_{0\text{gc}}$ are given by Eqs. (148) and (149), respectively. In what follows, although the parameters $(\epsilon_B, \epsilon_{\delta}, \epsilon_w)$ may be of the same order in the conventional nonlinear gyrokinetic ordering (Friedman and Chen, 1982), we keep them independent to emphasize their different physical origins and to retain more flexibility in the perturbative analysis of reduced Hamiltonian dynamics in various situations. An outstanding example, in which this ordering flexibility is necessary, is the case with strong $E \times B$ flow shear as discussed in Appendix D.1.

In Eq. (152), the first-order guiding-center phase-space Lagrangian Γ_{gc1} and Hamiltonian H_{1gc} are

$$\begin{aligned}\Gamma_{1gc} &= \frac{e}{c} \delta\mathbf{A}(\mathbf{X} + \boldsymbol{\rho}, t) \cdot d(\mathbf{X} + \boldsymbol{\rho}) \\ &\equiv \frac{e}{c} \delta\mathbf{A}_{gc}(\mathbf{X}, t; \mu, \zeta) \cdot d(\mathbf{X} + \boldsymbol{\rho})\end{aligned}\quad (153)$$

and

$$H_{1gc} = e\delta\phi(\mathbf{X} + \boldsymbol{\rho}, t) \equiv e\delta\phi_{gc}(\mathbf{X}, t; \mu, \zeta), \quad (154)$$

where $\delta\mathbf{A}_{gc}(\mathbf{X}, t; \mu, \zeta)$ and $\delta\phi_{gc}(\mathbf{X}, t; \mu, \zeta)$ are perturbation potentials evaluated at a particle's position $\mathbf{x} = \mathbf{X} + \boldsymbol{\rho}$ expressed in terms of the guiding-center position \mathbf{X} and the gyroangle-dependent gyroradius vector $\boldsymbol{\rho}(\mu, \zeta)$; to lowest order in ϵ_B , ignoring the spatial dependence of $\boldsymbol{\rho}$ (although these terms can be kept to arbitrary orders in ϵ_B).

Because of the gyroangle dependence in the guiding-center perturbation potentials ($\delta\phi_{gc}, \delta\mathbf{A}_{gc}$), the guiding-center magnetic moment μ is no longer conserved by the perturbed guiding-center equations of motion, i.e., $\dot{\mu} = \mathcal{O}(\epsilon_\delta)$. To remove the gyroangle dependence from the perturbed guiding-center phase-space Lagrangian and Hamiltonian (153) and (154), we proceed with the time-dependent gyrocenter phase-space transformation,

$$\bar{\mathcal{Z}} \equiv (\bar{\mathbf{X}}, p_{\parallel}, \mu, \zeta, w, t) \rightarrow \bar{\mathcal{Z}} \equiv (\bar{\mathbf{X}}, \bar{p}_{\parallel}, \bar{\mu}, \bar{\zeta}, \bar{w}, t),$$

where $\bar{\mathcal{Z}}$ denotes the gyrocenter (gy) extended phase-space coordinates. The nature of the gyrocenter parallel momentum \bar{p}_{\parallel} depends on the choice of representation used for gyrocenter Hamiltonian dynamics (as discussed below) and the time coordinate t is not affected by this transformation.

The results of the nonlinear Hamiltonian gyrocenter perturbation analysis (Brizard, 1989a) are summarized as follows. To first order in the small-amplitude parameter ϵ_δ and zeroth order in the space-time-scale parameters ($\epsilon_\omega, \epsilon_B$), this transformation is represented in terms of generating vector fields ($\mathbf{G}_1, \mathbf{G}_2, \dots$) as

$$\bar{\mathcal{Z}}^a \equiv \mathcal{Z}^a + \epsilon_\delta G_1^a + \dots \quad (155)$$

We wish to construct a new gyrocenter Hamiltonian system in which the new gyrocenter extended phase-space Lagrangian is

$$\begin{aligned}\bar{\Gamma} &= \left[\frac{e}{c} (\mathbf{A}_0 + \epsilon_\delta \delta\mathbf{A}_{gy}) + \bar{p}_{\parallel} \hat{\mathbf{b}}_0 \right] \cdot d\bar{\mathbf{X}} + \frac{mc}{e} \bar{\mu} d\bar{\zeta} - \bar{w} dt \\ &\equiv \bar{\Gamma}_0 + \epsilon_\delta \bar{\Gamma}_1,\end{aligned}\quad (156)$$

where the gyrocenter symplectic-perturbation term $\delta\mathbf{A}_{gy}$ is defined as

$$\delta\mathbf{A}_{gy} \equiv \alpha \langle \delta\mathbf{A}_{\perp gc} \rangle + \beta \langle \delta\mathbf{A}_{\parallel gc} \rangle \hat{\mathbf{b}}_0. \quad (157)$$

The model parameters (α, β) determine the form of the nonlinear gyrocenter model:

Gyrocenter model	α	β	\bar{p}_{\parallel}
Hamiltonian	0	0	canonical
Symplectic	1	1	kinetic
\perp -symplectic	1	0	canonical
\parallel -symplectic	0	1	kinetic

The Hamiltonian gyrocenter model ($\alpha=0=\beta$) and the symplectic gyrocenter model ($\alpha=1=\beta$) were derived by Brizard (1989a), while the parallel-symplectic gyrocenter model ($\beta=1, \alpha=0$) was used by Brizard (1992) to derive the so-called nonlinear electromagnetic gyrofluid equations.

The Jacobian for the transformation from particle to gyrocenter phase space is $\mathcal{J}=m^2 B_{\parallel}^*$, where

$$B_{\parallel}^* \equiv B_{0\parallel}^* + \epsilon_\delta \langle \delta B_{\parallel gc} \rangle, \quad (158)$$

while the general form for the gyrocenter Poisson bracket is

$$\begin{aligned}\{\mathcal{F}, \mathcal{G}\} &= \frac{e}{mc} \left(\frac{\partial \mathcal{F}}{\partial \zeta} \frac{\partial \mathcal{G}}{\partial \mu^*} - \frac{\partial \mathcal{F}}{\partial \mu^*} \frac{\partial \mathcal{G}}{\partial \zeta} \right) \\ &\quad + \frac{\mathbf{B}^*}{B_{\parallel}^*} \cdot \left(\nabla^* \mathcal{F} \frac{\partial \mathcal{G}}{\partial p_{\parallel}} - \frac{\partial \mathcal{F}}{\partial p_{\parallel}} \nabla^* \mathcal{G} \right) \\ &\quad - \frac{c \hat{\mathbf{b}}_0}{e B_{\parallel}^*} \cdot \nabla^* \mathcal{F} \times \nabla^* \mathcal{G} + \left(\frac{\partial \mathcal{F}}{\partial w} \frac{\partial \mathcal{G}}{\partial t} - \frac{\partial \mathcal{F}}{\partial t} \frac{\partial \mathcal{G}}{\partial w} \right),\end{aligned}\quad (159)$$

where $\mathbf{B}^* \equiv \mathbf{B}_0^* + \epsilon_\delta \delta\mathbf{B}_{gy}$, with $\delta\mathbf{B}_{gy} \equiv \nabla \times \delta\mathbf{A}_{gy}$, and

$$\nabla^* \mathcal{F} \equiv \nabla \mathcal{F} - \epsilon_\delta \frac{e}{c} \left(\frac{\partial \delta\mathbf{A}_{gy}}{\partial t} \frac{\partial \mathcal{F}}{\partial w} - \frac{\Omega}{B} \frac{\partial \delta\mathbf{A}_{gy}}{\partial \mu} \frac{\partial \mathcal{F}}{\partial \zeta} \right),$$

$$\frac{\partial \mathcal{F}}{\partial \mu^*} \equiv \frac{\partial \mathcal{F}}{\partial \mu} - \epsilon_\delta^2 \left(\frac{e}{c} \frac{\partial \delta\mathbf{A}_{gy}}{\partial t} \cdot \frac{\partial \delta\mathbf{A}_{gy}}{\partial \mu} \times \hat{\mathbf{b}}_0 \right) \frac{\partial \mathcal{F}}{\partial w}.$$

The guiding-center Poisson bracket (150) is recovered from Eq. (159) with the Hamiltonian gyrocenter model ($\delta\mathbf{A}_{gy}=0$).

The nonlinear gyrocenter Hamilton's equations are

$$\dot{\bar{\mathbf{X}}} = \frac{c \hat{\mathbf{b}}_0}{e B_{\parallel}^*} \times \left(\bar{\nabla} \bar{H} + \epsilon_\delta \frac{e}{c} \frac{\partial \delta\mathbf{A}_{gy}}{\partial t} \right) + \frac{\partial \bar{H}}{\partial \bar{p}_{\parallel}} \frac{\mathbf{B}^*}{B_{\parallel}^*}, \quad (160)$$

$$\dot{\bar{p}}_{\parallel} = - \frac{\mathbf{B}^*}{B_{\parallel}^*} \cdot \left(\bar{\nabla} \bar{H} + \epsilon_\delta \frac{e}{c} \frac{\partial \delta\mathbf{A}_{gy}}{\partial t} \right), \quad (161)$$

where the gyrocenter Hamiltonian $\bar{H} = \bar{H}_0 + \epsilon_\delta \bar{H}_1 + \epsilon_\delta^2 \bar{H}_2$ is derived in the next sections. The gyrocenter Hamilton's equations (160) and (161) satisfy the gyrocenter Liouville theorem

$$0 = \frac{\partial B_{\parallel}^*}{\partial t} + \nabla \cdot (B_{\parallel}^* \dot{\bar{\mathbf{X}}}) + \frac{\partial}{\partial \bar{p}_{\parallel}} (B_{\parallel}^* \dot{\bar{p}}_{\parallel}). \quad (162)$$

The fact that the gyrocenter Hamilton equations (160) and (161) satisfy the Liouville theorem (162) ensures that gyrocenter phase-space volume (as well as other Poincaré invariants) is conserved, a universal property of all Hamiltonian systems.

C. Nonlinear gyrocenter Hamiltonian dynamics

We now briefly review the first-order and second-order perturbation analysis leading to the derivation of the nonlinear gyrocenter Hamiltonian.

We begin with the first-order analysis. From Eq. (118), with $\bar{\Gamma}_1 \equiv (e/c) \delta \mathbf{A}_{\text{gy}} \cdot d\bar{\mathbf{X}}$, the first-order generating vector field for the gyrocenter phase-space transformation is

$$\begin{aligned} G_1^a &= \{S_1, \mathcal{Z}^a\}_0 + \frac{e}{c} \delta \mathbf{A}_{\text{gc}} \cdot \{\mathbf{X} + \boldsymbol{\rho}, \mathcal{Z}^a\}_0 \\ &\quad - \frac{e}{c} \delta \mathbf{A}_{\text{gy}} \cdot \{\mathbf{X}, \mathcal{Z}^a\}_0, \end{aligned} \quad (163)$$

or its components can be explicitly given as

$$\begin{aligned} G_1^x &= -\hat{\mathbf{b}}_0 \frac{\partial S_1}{\partial p_{\parallel}} - \frac{c\hat{\mathbf{b}}_0}{eB_0} \times \nabla S_1 \\ &\quad + (\delta \mathbf{A}_{\text{gc}} - \alpha \langle \delta \mathbf{A}_{\perp \text{gc}} \rangle) \times \frac{\hat{\mathbf{b}}_0}{B_0}, \end{aligned} \quad (164)$$

$$G_1^p = \hat{\mathbf{b}}_0 \cdot \nabla S_1 + \frac{e}{c} (\delta A_{\parallel \text{gc}} - \beta \langle \delta A_{\parallel \text{gc}} \rangle), \quad (165)$$

$$G_1^{\mu} = \frac{e}{mc} \left(\frac{e}{c} \delta \mathbf{A}_{\perp \text{gc}} \cdot \frac{\partial \boldsymbol{\rho}}{\partial \zeta} + \frac{\partial S_1}{\partial \zeta} \right), \quad (166)$$

$$G_1^{\xi} = -\frac{e}{mc} \left(\frac{e}{c} \delta \mathbf{A}_{\perp \text{gc}} \cdot \frac{\partial \boldsymbol{\rho}}{\partial \mu} + \frac{e}{mc} \frac{\partial S_1}{\partial \mu} \right), \quad (167)$$

$$G_1^w = -\frac{\partial S_1}{\partial t}, \quad (168)$$

where effects due to background magnetic-field nonuniformity are omitted. The first-order generating vector field (163) satisfies the identity $\partial_a (B_0 G_1^a) \equiv -\alpha \langle \delta B_{\parallel \text{gc}} \rangle$, showing that the gyrocenter phase-space Jacobian is different from the guiding-center phase-space Jacobian for gyrocenter models with $\alpha=1$ (i.e., the symplectic and \perp -symplectic gyrocenter models). The gyrocenter parallel momentum \bar{p}_{\parallel} is

$$\bar{p}_{\parallel} = p_{\parallel} + \epsilon \frac{e}{c} (\delta A_{\parallel \text{gc}} - \beta \langle \delta A_{\parallel \text{gc}} \rangle) + \dots, \quad (169)$$

showing that the gyrocenter parallel momentum coordinate \bar{p}_{\parallel} is a canonical momentum for gyrocenter models with $\beta=0$ (i.e., the Hamiltonian and \perp -symplectic gyrocenter models).

The first-order gyrocenter Hamiltonian is determined from the first-order Lie-transform equation (123) as

$$\bar{H}_1 \equiv e \delta \psi_{\text{gc}} - \{S_1, \mathcal{H}_0\}_0,$$

where the effective first-order potential is defined as

$$\delta \psi_{\text{gc}} \equiv \delta \phi_{\text{gc}} - \delta \mathbf{A}_{\text{gc}} \cdot \frac{\mathbf{v}}{c} + \beta \frac{v_{\parallel}}{c} \langle \delta A_{\parallel \text{gc}} \rangle. \quad (170)$$

The gyroangle-averaged part of this first-order equation yields

$$\begin{aligned} \bar{H}_1 &\equiv e \langle \delta \psi_{\text{gc}} \rangle = e \left\langle \delta \phi_{\text{gc}} - \frac{\mathbf{v}_{\perp}}{c} \cdot \delta \mathbf{A}_{\perp \text{gc}} \right\rangle \\ &\quad - \frac{ev_{\parallel}}{c} (1 - \beta) \langle \delta A_{\parallel \text{gc}} \rangle, \end{aligned} \quad (171)$$

while the solution for the scalar field S_1 is

$$S_1 = \frac{e}{\Omega_0} \int \delta \tilde{\psi}_{\text{gc}} d\zeta \equiv \frac{e}{\Omega_0} \delta \tilde{\Psi}_{\text{gc}}, \quad (172)$$

where $\delta \tilde{\psi}_{\text{gc}} \equiv \delta \psi_{\text{gc}} - \langle \delta \psi_{\text{gc}} \rangle$ is the gyroangle-dependent part of the first-order effective potential (170).

While the (linear) first-order gyrocenter Hamiltonian (171) is sufficient for linear gyrokinetic theory (i.e., in the absence of polarization and magnetization effects in Maxwell's equations), it must be supplemented by a (nonlinear) second-order gyrocenter Hamiltonian \bar{H}_2 for two important reasons. First, the second-order gyrocenter Hamiltonian \bar{H}_2 is needed in order to obtain polarization and magnetization effects that, within the variational formulation of self-consistent gyrokinetic Vlasov-Maxwell theory presented here, have variational definitions expressed in terms of the partial derivatives $\partial \bar{H}_2 / \partial \mathbf{E}_1$ and $\partial \bar{H}_2 / \partial \mathbf{B}_1$, respectively (see Sec. III). Second, once polarization and magnetization effects are included in the gyrokinetic Maxwell equations, the second-order gyrocenter Hamiltonian \bar{H}_2 must be kept in the gyrokinetic Vlasov Lagrangian density to obtain an exact energy conservation law (derived using the Noether method) for the gyrokinetic Vlasov-Maxwell equations.

The general expression for the second-order gyrocenter Hamiltonian is obtained from Eq. (129), using Eqs. (163), (171), and (172), as

$$\begin{aligned} \bar{H}_2 &= -\frac{e^2}{2\Omega_0} \langle \{\delta \tilde{\Psi}_{\text{gc}}, \delta \tilde{\psi}_{\text{gc}}\}_0 \rangle + \frac{e^2}{2mc^2} \langle |\delta \mathbf{A}_{\text{gc}}|^2 \rangle \\ &\quad - \beta \langle \delta A_{\parallel \text{gc}} \rangle^2 + \alpha \langle \delta \mathbf{A}_{\perp \text{gc}} \rangle \cdot \frac{\hat{\mathbf{b}}_0}{B_0} \times \nabla \bar{H}_1, \end{aligned} \quad (173)$$

where the first term describes low-frequency ponderomotive effects associated with the elimination of the fast gyromotion time scale while the remaining terms involve magnetic perturbations and the choice of gyrocenter-model parameters (α, β) .

D. Nonlinear gyrokinetic Vlasov equation

Once the linear gyrocenter Hamiltonian (171) and the nonlinear gyrocenter Hamiltonian (173) are obtained, it

is simple to derive the corresponding nonlinear gyrokinetic Vlasov equation for the gyrocenter Vlasov distribution \bar{F} ,

$$0 = \frac{\partial \bar{F}}{\partial t} + \{\bar{F}, \bar{H}_{\text{gy}}\}, \quad (174)$$

where, in order to simplify this presentation, (\cdot, \cdot) denotes the guiding-center Poisson bracket (159) in the Hamiltonian gyrocenter model ($\alpha=0=\beta$) and the nonlinear gyrocenter Hamiltonian is $\bar{H}_{\text{gy}}=\bar{H}_{\text{gc}}+e\delta\Psi_{\text{gy}}$. The (unperturbed) guiding-center Hamiltonian is $\bar{H}_{\text{gc}}=\bar{p}_{\parallel}^2/2m+\bar{\mu}B$ and, up to second order in the amplitude parameter ϵ_{δ} , the gyrocenter perturbation potential is

$$\begin{aligned} e\delta\Psi_{\text{gy}} &\equiv \epsilon_{\delta}e\langle\delta\psi_{\text{gc}}\rangle + \frac{\epsilon_{\delta}^2e^2}{2mc^2}\langle|\delta\mathbf{A}_{\text{gc}}|^2\rangle \\ &\quad - \frac{\epsilon_{\delta}^2e^2}{2\Omega}\langle\{\delta\tilde{\Psi}_{\text{gc}}, \delta\tilde{\psi}_{\text{gc}}\}\rangle. \end{aligned} \quad (175)$$

The gyrocenter perturbation potentials presented in Sec. III were all derived in various limits from Eq. (175).

We have obtained a reduced (gyroangle-independent) gyrocenter Hamiltonian description of charged-particle motion in nonuniform magnetized plasmas perturbed by low-frequency electromagnetic fluctuations. At this level, the nonlinear gyrokinetic Vlasov equation can be used to study the evolution of a distribution of test gyrocenters in the presence of low-frequency electromagnetic fluctuations. For a self-consistent treatment that includes an electromagnetic field response to the gyrocenter Hamiltonian dynamics, a set of low-frequency Maxwell equations with charge and current densities expressed in terms of moments of the gyrocenter Vlasov distribution is required.

E. Pull-back representation of the perturbed Vlasov distribution

Before proceeding to the variational derivation of the gyrokinetic Maxwell equations and the exact gyrokinetic energy invariant, the connection between the particle Vlasov distribution f and the gyrocenter Vlasov distribution \bar{F} is investigated.

The perturbed Vlasov distribution is traditionally decomposed into its adiabatic and nonadiabatic components (Antonsen and Lane, 1980; Catto *et al.*, 1981; Brizard, 1994a, 1994b) following an iterative solution of the perturbed guiding-center Vlasov equation. We assume that the magnetic field is uniform and the pull-back transformation from the guiding-center Vlasov distribution F to the particle Vlasov distribution f is

$$f \equiv \mathbf{T}_{\text{gc}}F = e^{-\rho\mathbf{v}}F. \quad (176)$$

The pull-back transformation from the gyrocenter Vlasov distribution \bar{F} to the guiding-center Vlasov distribution F is

$$F = \mathbf{T}_{\text{gy}}\bar{F} = \bar{F} + \epsilon_{\delta}\{S_1, \bar{F}\} + \epsilon_{\delta}\frac{e}{c}\delta\mathbf{A}_{\text{gc}} \cdot \{\mathbf{X} + \boldsymbol{\rho}, \bar{F}\}. \quad (177)$$

No information is lost in transforming the Vlasov equation in particle phase space to the gyrokinetic Vlasov equation in gyrocenter phase space since

$$\frac{df}{dt} = \frac{d}{dt}(\mathbf{T}_{\epsilon}\bar{F}) = \mathbf{T}_{\epsilon}\left[\left(\mathbf{T}_{\epsilon}^{-1}\frac{d}{dt}\mathbf{T}_{\epsilon}\right)\bar{F}\right] = \mathbf{T}_{\epsilon}\left(\frac{d\epsilon\bar{F}}{dt}\right), \quad (178)$$

so that the Vlasov equation $df/dt=0$ is satisfied for the particle Vlasov distribution $f=\mathbf{T}_{\epsilon}\bar{F}=\mathbf{T}_{\text{gc}}(\mathbf{T}_{\text{gy}}\bar{F})$ if the gyrokinetic Vlasov equation $d_{\epsilon}\bar{F}/dt=0$ is satisfied for the gyrocenter Vlasov distribution \bar{F} .

To compare the nonlinear gyrokinetic Vlasov equation directly with the Frieman-Chen gyrokinetic equation (17), the guiding-center Poisson bracket associated with the new guiding-center coordinates $(\mathbf{X}, \mathcal{E}, \mu, \zeta)$ is introduced,

$$\begin{aligned} \{F, G\} &= \Omega\left[\frac{\partial F}{\partial \zeta}\left(\frac{\partial G}{\partial \mathcal{E}} + \frac{1}{B}\frac{\partial G}{\partial \mu}\right) - \left(\frac{\partial F}{\partial \mathcal{E}} + \frac{1}{B}\frac{\partial F}{\partial \mu}\right)\frac{\partial G}{\partial \zeta}\right] \\ &\quad + \mathbf{v}_{\text{gc}} \cdot \left(\nabla F \frac{\partial G}{\partial \mathcal{E}} - \frac{\partial F}{\partial \mathcal{E}} \nabla G\right) - \frac{c\hat{\mathbf{b}}}{eB} \cdot \nabla F \times \nabla G, \end{aligned} \quad (179)$$

where $\mathbf{v}_{\text{gc}}=v_{\parallel}\hat{\mathbf{b}}$ in the absence of magnetic-field nonuniformity and the gyrocenter kinetic energy \mathcal{E} is used instead of the parallel guiding-center velocity v_{\parallel} to simplify the comparison with the Frieman-Chen gyrokinetic formalism. By combining the guiding-center and gyrocenter pull-backs, the pull-back transformation from the gyrocenter Vlasov distribution \bar{F} and the particle Vlasov distribution f is found

$$\begin{aligned} f &= e^{-\rho\mathbf{v}}\left[\bar{F} - e\langle\delta\psi_{\text{gc}}\rangle\left(\frac{\partial F}{\partial \mathcal{E}} + \frac{1}{B}\frac{\partial F}{\partial \mu}\right)\right] + e\delta\phi\frac{\partial\bar{F}}{\partial \mathcal{E}} \\ &\quad + \frac{e}{B}\left(\delta\phi - \frac{v_{\parallel}}{c}\delta A_{\parallel}\right)\frac{\partial\bar{F}}{\partial \mu} + \delta\mathbf{A} \times \frac{\hat{\mathbf{b}}}{B} \cdot \nabla\bar{F}, \end{aligned} \quad (180)$$

where the last three terms represent the adiabatic components of the perturbed particle Vlasov distribution while the first two terms represent the guiding-center pull-back of the gyrocenter Vlasov distribution \bar{F} and the nonadiabatic component of the perturbed particle Vlasov distribution.

By comparing the pull-back decomposition (180) with the Frieman-Chen decomposition (15), a relation between the first-order correction \bar{F}_1 to the gyrocenter distribution $\bar{F}=\bar{F}_0+\epsilon_{\delta}\bar{F}_1$ and the nonadiabatic part \bar{G}_1 is obtained,

$$\bar{F}_1 \equiv \bar{G}_1 + e\langle\delta\psi_{\text{gc}}\rangle\frac{\partial\bar{F}_0}{\partial \mathcal{E}}. \quad (181)$$

Substituting this relation into the nonlinear gyrokinetic Vlasov equation (174),

$$0 = \frac{d_{\text{gy}}\bar{F}}{dt} \equiv \frac{d_{\text{gc}}\bar{F}}{dt} + \epsilon_\delta e\{\bar{F}, \langle \delta\psi_{\text{gc}} \rangle\},$$

with the gyrocenter Hamiltonian truncated at first order

$$\bar{H}_{\text{gy}} = \bar{H}_{\text{gc}} + \epsilon_\delta e\langle \delta\psi_{\text{gc}} \rangle,$$

$$\begin{aligned} 0 &= \frac{d_{\text{gy}}}{dt} \left[\bar{F}_0 + \epsilon_\delta \left(\bar{G}_1 + e\langle \delta\psi_{\text{gc}} \rangle \frac{\partial \bar{F}_0}{\partial \mathcal{E}} \right) \right] \\ &= \epsilon_\delta \{(\bar{F}_0 + \epsilon_\delta \bar{G}_1), e\langle \delta\psi_{\text{gc}} \rangle\} \\ &\quad + \epsilon_\delta \left(\frac{d_{\text{gc}}\bar{G}_1}{dt} + e \frac{d_{\text{gc}}\langle \delta\psi_{\text{gc}} \rangle}{dt} \frac{\partial \bar{F}_0}{\partial \mathcal{E}} \right). \end{aligned} \quad (182)$$

Using the guiding-center Poisson bracket (179), we find

$$\begin{aligned} \{\bar{F}_0, e\langle \delta\psi_{\text{gc}} \rangle\} &= \frac{c\hat{\mathbf{b}}}{B} \times \nabla\langle \delta\psi_{\text{gc}} \rangle \cdot \nabla\bar{F}_0 \\ &\quad - (e\mathbf{v}_{\text{gc}} \cdot \nabla\langle \delta\psi_{\text{gc}} \rangle) \frac{\partial \bar{F}_0}{\partial \mathcal{E}} \end{aligned}$$

and the nonlinear gyrokinetic Vlasov equation (182) becomes the Frieman-Chen nonlinear gyrokinetic Vlasov,

$$\begin{aligned} \frac{d_{\text{gc}}\bar{G}_1}{dt} &= - \left(e \frac{\partial\langle \delta\psi_{\text{gc}} \rangle}{\partial t} \frac{\partial \bar{F}_0}{\partial \mathcal{E}} + \frac{c\hat{\mathbf{b}}}{B} \times \nabla\langle \delta\psi_{\text{gc}} \rangle \cdot \nabla\bar{F}_0 \right) \\ &\quad - \frac{c\hat{\mathbf{b}}}{B} \times \nabla\langle \delta\psi_{\text{gc}} \rangle \cdot \nabla\bar{G}_1, \end{aligned} \quad (183)$$

where higher-order terms (e.g., $\epsilon_B\epsilon_\delta^2$) were omitted. The Frieman-Chen nonlinear gyrokinetic Vlasov equation is contained in the nonlinear gyrokinetic Vlasov equation (174). While the adiabatic and nonadiabatic decompositions have at times appeared mysterious, they naturally appear in the context of the action of the pull-back operators used in the derivation of the nonlinear gyrokinetic Vlasov equation. The physical interpretation of the pull-back operator is that it performs a partial solution of the Vlasov equation associated with fast-time gyromotion dynamics. The truncation of the gyrocenter Hamiltonian at first order in ϵ_δ implies that the gyrokinetic energy conservation law will break down unless the gyrokinetic polarization and magnetization terms are also dropped in the gyrokinetic Maxwell equations (as discussed in Sec. III and further discussed in the next section).

VI. GYROKINETIC VARIATIONAL FORMULATION

After having derived various expressions for the nonlinear gyrocenter Hamiltonian and its associated nonlinear gyrokinetic Vlasov equation, we now derive self-consistent expressions for the gyrokinetic Maxwell's equations, in which gyrocenter polarization and magnetization effects appear. Once a set of self-consistent nonlinear gyrokinetic Vlasov-Maxwell equations is derived, the exact energy conservation law these nonlinear gyrokinetic equations satisfy will be derived. These two tasks

are simultaneously performed in this section using a variational formulation for the nonlinear gyrokinetic Vlasov-Maxwell equations (Brizard, 2000a, 200b).

A. Nonlinear gyrokinetic Vlasov-Maxwell equations

The nonlinear self-consistent gyrokinetic Vlasov-Maxwell equations (using the Hamiltonian gyrocenter model) are derived from a reduced variational principle that will also be used to derive an exact energy conservation law for the gyrokinetic Vlasov-Maxwell equations. The reduced action functional for the low-frequency gyrokinetic Vlasov-Maxwell equations (Brizard, 2000b; Sugama, 2000) is

$$\begin{aligned} \mathcal{A}_{\text{gy}} &= - \int d^8\mathcal{Z} \mathcal{F}(\mathcal{Z}) \mathcal{H}(\mathcal{Z}; A_{1\mu}, F_{1\mu\nu}) \\ &\quad + \int \frac{d^4x}{8\pi} (|\nabla\Phi|^2 - |\mathbf{B}|^2), \end{aligned} \quad (184)$$

where \mathcal{H} is the nonlinear gyrocenter Hamiltonian (175), and we use the notation

$$\Phi = \epsilon\phi_1 \quad \text{and} \quad \mathbf{B} = \mathbf{B}_0 + \epsilon\nabla \times \mathbf{A}_1.$$

We omit the overbar to denote gyrocenter coordinates and functions on extended gyrocenter phase space (and set $\epsilon \equiv \epsilon_\delta$) and summation over species is implied whenever appropriate. The absence of the inductive part $-c^{-1}\partial_t\mathbf{A}_1$ of the perturbed electric field \mathbf{E}_1 in the Maxwell part of the reduced action functional (184) means that the displacement current $\partial_t\mathbf{E}_1$ will be absent from Ampère's equation. This is consistent with the low-frequency approximation ($\epsilon_\omega \ll 1$) used in the nonlinear gyrokinetic ordering (6) and (7).

The variational principle $\delta\mathcal{A}_{\text{gy}} = \int \delta\mathcal{L}_{\text{gy}} d^4x = 0$ for the nonlinear low-frequency gyrokinetic Vlasov-Maxwell equations is based on Eulerian variations for $\mathcal{F}(\mathcal{Z})$ and (ϕ_1, \mathbf{A}_1) . Variation of \mathcal{A}_{gy} with respect to $\delta\mathcal{F}$ and $\delta A_1^\mu(\mathbf{x}) = (\delta\phi_1, \delta\mathbf{A}_1)$ yields

$$\begin{aligned} \delta\mathcal{A}_{\text{gy}} &= \int \frac{d^4x}{4\pi} (\epsilon \nabla \delta\phi_1 \cdot \nabla\Phi - \epsilon \nabla \times \delta\mathbf{A}_1 \cdot \mathbf{B}) \\ &\quad - \int d^8\mathcal{Z} \left[\delta\mathcal{F}(\mathcal{Z}) \mathcal{H} + \mathcal{F}(\mathcal{Z}) \right. \\ &\quad \left. \times \int d^3x \left(\delta A_{1\mu}(\mathbf{x}) \frac{\delta H}{\delta A_{1\mu}(\mathbf{x})} \right) \right]. \end{aligned} \quad (185)$$

The Eulerian variation $\delta\mathcal{F}$ is constrained to be of the form

$$\delta\mathcal{F} \equiv \{\mathcal{S}, \mathcal{F}\}, \quad (186)$$

where $\{\cdot, \cdot\}$ is the extended guiding-center Poisson bracket (150). The functional derivatives $\delta H / \delta A_{1\mu}(\mathbf{x})$ in Eq. (185) are evaluated using the gyrocenter Hamiltonian (175) (to second order in ϵ) as

$$\frac{\delta H}{\delta A_{1\mu}(\mathbf{x})} \equiv -\epsilon e \left\langle \mathsf{T}_{\text{gy}}^{-1} \left(\frac{v^\mu}{c} \delta_{\text{gc}}^3 \right) \right\rangle, \quad (187)$$

where $\delta_{\text{gc}}^3 \equiv \delta^3(\mathbf{x} - \mathbf{X} - \boldsymbol{\rho})$, $\mathsf{T}_{\text{gy}}^{-1}$ is the gyrocenter push-forward operator, and we have used the identity

$$A_{1\mu}(\mathbf{X} + \boldsymbol{\rho}) = \int d^3x \delta^3(\mathbf{x} - \mathbf{X} - \boldsymbol{\rho}) A_{1\mu}(\mathbf{x})$$

so that

$$\frac{\delta A_{1\mu}^{\text{gc}}}{\delta A_{1\nu}(\mathbf{x})} = \delta^{\mu\nu} \delta_{\text{gc}}^3.$$

After rearranging terms and integrating by parts, the variation (185) becomes

$$\begin{aligned} \delta \mathcal{A}_{\text{gy}} &= \int d^4x (\partial \cdot \mathcal{J}_{\text{gy}}) - \int d^8Z \mathcal{S}\{\mathcal{F}, \mathcal{H}\} \\ &\quad - \int d^4x \left\{ \epsilon \delta \phi_1 \left[\frac{\nabla^2 \Phi}{4\pi} + e \int d^6Z F \langle \mathsf{T}_{\text{gy}}^{-1} \delta_{\text{gc}}^3 \rangle \right] \right. \\ &\quad \left. + \epsilon \delta \mathbf{A}_1 \cdot \left[\frac{\nabla \times \mathbf{B}}{4\pi} - e \int d^6Z F \left\langle \mathsf{T}_{\text{gy}}^{-1} \left(\frac{\mathbf{v}}{c} \delta_{\text{gc}}^3 \right) \right\rangle \right] \right\}, \end{aligned} \quad (188)$$

where the first term on the right side of Eq. (188) involves the exact space-time divergence

$$\begin{aligned} \partial \cdot \mathcal{J}_{\text{gy}} &\equiv \frac{\partial}{\partial x^\mu} \left(\int d^4p \mathcal{S} \mathcal{F} \dot{X}^\mu \right) \\ &\quad + \nabla \cdot \left(\epsilon \frac{\delta \phi_1}{4\pi} \nabla \Phi - \epsilon \frac{\delta \mathbf{A}_1}{4\pi} \times \mathbf{B} \right), \end{aligned} \quad (189)$$

with $\dot{X}^\mu \equiv \{X^\mu, \mathcal{H}\}$ denoting the lowest-order gyrocenter four-velocity. Since Eq. (189) is an exact space-time divergence, it does not contribute to the reduced variational principle $\delta \mathcal{A}_{\text{gy}} \equiv 0$.

By requiring that the gyrokinetic action functional \mathcal{A}_{gy} be stationary with respect to an arbitrary variation \mathcal{S} (that vanishes on the integration boundaries), the nonlinear gyrokinetic Vlasov equation is

$$0 = \{\mathcal{F}, \mathcal{H}\}, \quad (190)$$

which, when integrated over the energy coordinate w , yields the standard nonlinear gyrokinetic Vlasov equation

$$\frac{\partial F}{\partial t} + \left(\frac{\mathbf{B}^*}{B_\parallel^*} \frac{\partial H}{\partial p_\parallel} + \frac{c\hat{\mathbf{b}}}{eB_\parallel^*} \times \nabla H \right) \cdot \nabla F - \left(\frac{\mathbf{B}^*}{B_\parallel^*} \cdot \nabla H \right) \frac{\partial F}{\partial p_\parallel} = 0. \quad (191)$$

Stationarity of the gyrokinetic action functional with respect to arbitrary variations δA_1^μ yields the gyrokinetic Maxwell equations: the gyrokinetic Poisson equation

$$\begin{aligned} \nabla^2 \Phi(\mathbf{x}) &= -4\pi e \int d^6Z F \langle \mathsf{T}_{\text{gy}}^{-1} \delta_{\text{gc}}^3 \rangle \\ &\equiv -4\pi e \int d^3p \langle e^{-\boldsymbol{\rho} \cdot \nabla} (\mathsf{T}_{\text{gy}} F) \rangle \end{aligned} \quad (192)$$

and the gyrokinetic Ampère equation

$$\begin{aligned} \nabla \times \mathbf{B}(\mathbf{x}) &= \frac{4\pi e}{c} \int d^6Z F(Z) \langle \mathsf{T}_{\text{gy}}^{-1} (\mathbf{v} \delta_{\text{gc}}^3) \rangle \\ &\equiv \frac{4\pi e}{c} \int d^3p \langle e^{-\boldsymbol{\rho} \cdot \nabla} (\mathbf{v} \mathsf{T}_{\text{gy}} F) \rangle, \end{aligned} \quad (193)$$

that are valid for all gyrocenter models discussed in Sec. V.B. The gyrocenter pull-back of the gyrocenter Vlasov distribution $\mathsf{T}_{\text{gy}} F$ is

$$\begin{aligned} \mathsf{T}_{\text{gy}} F &= F + \epsilon_\delta \{S_1, F\} + \epsilon_\delta \frac{e}{c} \delta \mathbf{A}_{\text{gc}} \cdot \{\mathbf{X} + \boldsymbol{\rho}, F\} \\ &\quad - \epsilon_\delta \left(\alpha \langle \delta \mathbf{A}_{\perp \text{gc}} \rangle \cdot \frac{\hat{\mathbf{b}}_0}{B_0} \times \nabla F + \frac{e}{c} \beta \langle \delta A_{\parallel \text{gc}} \rangle \frac{\partial F}{\partial p_\parallel} \right). \end{aligned}$$

The nonlinear gyrokinetic equations (191)–(193), with the gyrocenter Hamiltonian (175), are the self-consistent nonlinear gyrokinetic Vlasov-Maxwell equations in general magnetic-field geometry (Brizard, 1989a).

B. Gyrokinetic energy conservation law

We now apply the Noether method on the gyrokinetic action functional (184) to derive an exact gyrokinetic energy conservation law. By substituting Eqs. (190), (192), and (193) into Eq. (188), the variational equation $\delta \mathcal{A}_{\text{gy}} \equiv \int \delta \mathcal{L}_{\text{gy}} d^4x$ yields the Noether equation

$$\delta \mathcal{L}_{\text{gy}} \equiv \partial \cdot \mathcal{J}_{\text{gy}}. \quad (194)$$

Using the Noether method, the variations $(\mathcal{S}, \delta A_1^\mu, \delta \mathcal{L}_{\text{gy}})$ are expressed in terms of generators for infinitesimal translations in space or time.

Following a translation in time $t \rightarrow t + \delta t$, the variations \mathcal{S} , $\delta \phi_1$, $\delta \mathbf{A}_1$, and $\delta \mathcal{L}_{\text{gy}}$ become, respectively,

$$\mathcal{S} = -w \delta t,$$

$$\delta \phi_1 = -\delta t \partial_t \phi_1,$$

$$\delta \mathbf{A}_1 = -\delta t \partial_t \mathbf{A}_1 \equiv c \delta t (\mathbf{E}_1 + \nabla \phi_1),$$

$$\delta \mathcal{L}_{\text{gy}} = -\delta t \partial_t \mathcal{L}_{\text{gy}}. \quad (195)$$

In the last expression in Eq. (195), the gyrokinetic Vlasov-Maxwell Lagrangian density is $\mathcal{L}_{\text{gy}} = (|\nabla \Phi|^2 - |\mathbf{B}|^2)/8\pi$ after the physical constraint $\mathcal{H}=0$ is imposed in the space-time integrand of the reduced action functional (184). By combining Eq. (195) with Eqs. (189) and (194), we obtain, after rearranging and canceling some

terms (Brizard, 2000b), the local gyrokinetic energy conservation law,¹⁵

$$\frac{\partial \mathcal{E}_{\text{gy}}}{\partial t} + \nabla \cdot \mathbf{S}_{\text{gy}} = 0, \quad (196)$$

where the gyrokinetic energy density is

$$\mathcal{E}_{\text{gy}} = \int d^3 p F(H - e\langle \mathbf{T}_{\text{gy}}^{-1} \Phi_{\text{gc}} \rangle) + \frac{1}{8\pi} (|\nabla \Phi|^2 + |\mathbf{B}|^2) \quad (197)$$

while the gyrokinetic energy density flux is

$$\begin{aligned} \mathbf{S}_{\text{gy}} = & \int d^3 p F(H \dot{\mathbf{X}} - e\langle \mathbf{T}_{\text{gy}}^{-1} \mathbf{v} \Phi_{\text{gc}} \rangle) \\ & + \frac{\epsilon}{4\pi} \left(c \mathbf{E}_1 \times \mathbf{B} - \Phi \nabla \frac{\partial \phi_1}{\partial t} \right). \end{aligned} \quad (198)$$

We also obtain an expression for the global gyrokinetic energy conservation law $dE/dt=0$, where the global gyrokinetic energy is

$$E = \int \frac{d^3 x}{8\pi} (|\nabla \Phi|^2 + |\mathbf{B}|^2) + \int d^6 Z F(H - e\langle \mathbf{T}_{\text{gy}}^{-1} \Phi_{\text{gc}} \rangle). \quad (199)$$

The existence of an exact energy conservation law for nonlinear gyrokinetic equations provides a stringent test on simulations based on nonlinear electrostatic (Dubin *et al.*, 1983; Hahm, 1988) and electromagnetic (Hahm *et al.*, 1988; Brizard, 1989a) gyrokinetic equations.

VII. SUMMARY

The foundations of modern nonlinear gyrokinetic theory are based on three pillars: (i) a gyrokinetic Vlasov equation written in terms of a gyrocenter Hamiltonian with quadratic low-frequency ponderomotive-like terms, (ii) a set of gyrokinetic Maxwell equations written in terms of the gyrokinetic Vlasov distribution that contain low-frequency polarization and magnetization terms derived from the quadratic nonlinearities in the gyrocenter Hamiltonian, and (iii) an exact energy conservation law for the gyrokinetic Vlasov-Maxwell equations that includes all relevant linear and nonlinear coupling terms.

These three pillars were emphasized in Sec. III, where simplified forms of the nonlinear gyrokinetic equations were presented for the cases of electrostatic fluctuations, shear-Alfvénic fluctuations, and compressional magnetic fluctuations. In the full electromagnetic case, the gyrocenter polarization and magnetization vectors were defined in terms of derivatives of the effective gyrocenter

perturbation potential with respect to the perturbed electric and magnetic fields, respectively. Section III also showed the connection between the nonlinear gyrokinetic description of turbulent magnetized plasmas and several nonlinear reduced fluid models such as the Hasegawa-Mima equation for electrostatic drift-wave turbulence.

Through the use of Lie-transform perturbation methods on extended particle phase space, the derivation of a set of nonlinear low-frequency gyrokinetic Vlasov-Maxwell equations describing the reduced Hamiltonian description of gyrocenter dynamics in a time-independent background magnetic field perturbed by low-frequency electromagnetic fluctuations was shown. A self-consistent treatment is obtained through a low-frequency gyrokinetic variational principle and an exact gyrokinetic energy conservation law is obtained by applying the Noether method. Physical motivations for nonlinear gyrokinetic equations and various applications in theory and simulations thereof were discussed.

ACKNOWLEDGMENTS

We acknowledge Liu Chen and W. W. Lee for their pioneering work in introducing the concepts of nonlinear gyrokinetics to the plasma physics community and to thank them for their continuing support of our research. Next, we thank P. H. Diamond for crucial advice on applications to various nonlinear problems and W. M. Tang for his leadership in advancing the kinetic approach and large-scale gyrokinetic computation in the plasma physics community. We acknowledge J. A. Krommes for emphasizing the powerfulness of the Hamiltonian approach and Lie-transform perturbation methods in the rigorous derivation of nonlinear gyrokinetic equations, as well as A. N. Kaufman and R. G. Littlejohn for their inspiring groundbreaking applications of mathematical methods in plasma physics. We have also benefited from collaborations with Z. Lin in the course of research related to the subject of this review. We are grateful for fruitful discussions with K. H. Burrell, J. R. Cary, A.A. Chan, C. S. Chang, Y. Chen, C. Z. Cheng, B. I. Cohen, J. W. Connor, A. M. Dimits, D. H. E. Dubin, R. J. Fonck, X. Garbet, Ph. Ghendrih, G. W. Hammett, F. L. Hinton, W. Horton, Y. Idomura, K. Itoh, S. I. Itoh, Y. Kishimoto, R. M. Kulsrud, J. N. Leboeuf, N. Mattor, E. Mazzucato, H. Naitou, Y. Nishimura, C. Oberman, M. Ottaviani, W. Park, S. E. Parker, F. W. Perkins, H. Qin, G. Rewoldt, Y. Sarazin, B. D. Scott, H. Sugama, R. D. Sydora, L. Villard, W. Wang, T. H. Watanabe, M. Yagi, and F. Zonca. We also acknowledge S. Zweber for travel support. This work was supported by the U.S. Department of Energy Contracts No. DE-AC03-76SFOO098 (A.J.B.) and No. DE-AC02-CHO-3073 (T.S.H.), and partly by the National Science Foundation under Grant No. DMS-0317339 (A.J.B.) and the U.S. Department of Energy SciDAC Center for Gyrokinetic Particle Simulation of Turbulent Transport in Burning Plasmas (T.S.H.).

¹⁵Note that the gyrokinetic energy conservation law (196) is invariant under the transformation $\mathcal{E}_{\text{gy}} \rightarrow \mathcal{E}_{\text{gy}} + \nabla \cdot \mathbf{C}$ and $\mathbf{S}_{\text{gy}} \rightarrow \mathbf{S}_{\text{gy}} - \partial \mathbf{C} / \partial t + \nabla \times \mathbf{D}$, where \mathbf{C} and \mathbf{D} are arbitrary vector fields. The gyrokinetic energy density (197) is written in a form that explicitly yields $E \equiv \int d^3 x \mathcal{E}_{\text{gy}}$ in Eq. (199).

APPENDIX A: MATHEMATICAL PRIMER

This appendix presents a brief summary of the differential geometric foundations of Lie-transform perturbation methods used in deriving the nonlinear gyrokinetic equations presented in Sec. V.

Differential k -forms (Flanders, 1989)

$$\omega_k = \frac{1}{k!} \omega_{i_1 i_2 \cdots i_k} dz^{i_1} \wedge dz^{i_2} \wedge \cdots \wedge dz^{i_k}$$

are fundamental objects in the differential geometry of n -dimensional space (with coordinates \mathbf{z}), where the components $\omega_{i_1 i_2 \cdots i_k}$ are antisymmetric with respect to interchange of two adjacent indices since the wedge product \wedge is skew symmetric (i.e., $dz^a \wedge dz^b = -dz^b \wedge dz^a$) with respect to the exterior derivative d (which has properties similar to the standard derivative d).

Note that the exterior derivative $d\omega_k$ of a differential k -form (or k -form for short) ω_k is a $(k+1)$ -form. For example, the exterior derivative of a zero-form f is defined as

$$df \equiv \partial_a f dz^a, \quad (\text{A1})$$

and df is a differential one-form; note that its components are the components of ∇f . The exterior derivative of a one-form Γ is a two-form,

$$d\Gamma \equiv d\Gamma_b \wedge dz^b = \partial_a \Gamma_b dz^a \wedge dz^b,$$

which, as a result of the skew symmetry of the wedge product \wedge , may be expressed as

$$d\Gamma = \frac{1}{2} (\partial_a \Gamma_b - \partial_b \Gamma_a) dz^a \wedge dz^b \equiv \frac{1}{2} \omega_{ab} dz^a \wedge dz^b, \quad (\text{A2})$$

where $\omega_{ab} = -\omega_{ba}$ is the antisymmetric components of the two-form $\omega \equiv d\Gamma$.

An important difference between the exterior derivative d and the standard derivative d comes from the property that $d^2\omega_k = d(d\omega_k) = 0$ for any k -form ω_k . Indeed, for a zero-form we find

$$d^2f = \partial_{ab}^2 f dz^a \wedge dz^b = 0,$$

since $\partial_{ab}^2 f$ is symmetric with respect to interchange $a \leftrightarrow b$ while \wedge is antisymmetric. For a one-form we find

$$d^2\Gamma = \frac{1}{3!} (\partial_a \omega_{bc} + \partial_b \omega_{ca} + \partial_c \omega_{ab}) dz^a \wedge dz^b \wedge dz^c = 0.$$

A k -form ω_k is said to be *closed* if its exterior derivative is $d\omega_k = 0$, while a k -form ω_k is said to be *exact* if it can be written in terms of a $(k-1)$ -form Γ_{k-1} as $\omega_k = d\Gamma_{k-1}$. Poincaré's lemma states that all closed k -forms are exact (as can easily be verified), while its converse states that all exact k -forms are closed. For example, the infinitesimal volume element in three-dimensional space with curvilinear coordinates $\mathbf{u} = (u^1, u^2, u^3)$ and Jacobian \mathcal{J} ,

$$\Omega \equiv \mathcal{J}(\mathbf{u}) du^1 \wedge du^2 \wedge du^3,$$

is a closed three-form since $d\Omega = 0$. Hence, according to the converse of Poincaré's lemma, there exists a two-form σ such that $\Omega = d\sigma$, where

$$\sigma \equiv \frac{1}{2} \epsilon_{ijk} \sigma^k(\mathbf{u}) du^i \wedge du^j$$

is the infinitesimal area two-form, with the Jacobian defined as $\mathcal{J} = \partial_a(\mathbf{u}) / \partial u^i$.

We now introduce the inner-product operation involving a vector field \mathbf{v} and a k -form ω_k , denoted here as $\mathbf{v} \cdot \omega_k$, which produces a $(k-1)$ -form. For example, for a one-form it is defined as $\mathbf{v} \cdot \Gamma = v^a \Gamma_a$, while for a two-form, it is defined as

$$\mathbf{v} \cdot \omega \equiv \frac{1}{2} (v^a \omega_{ab} dz^b - \omega_{ab} v^b dz^a) = v^a \omega_{ab} dz^b.$$

Note that $d(\mathbf{v} \cdot \Omega) = \mathcal{J}^{-1} \partial_a(\mathcal{J} v^a) \Omega \equiv (\nabla \cdot \mathbf{v}) \Omega$, which can be used to derive the divergence of any vector field expressed in arbitrary curvilinear coordinates.

The Lie derivative \mathcal{L}_v along the vector field \mathbf{v} of a k -form ω_k is defined in terms of the homotopy formula (Abraham and Marsden, 1978),

$$\mathcal{L}_v \omega_k = \mathbf{v} \cdot d\omega_k + d(\mathbf{v} \cdot \omega_k). \quad (\text{A3})$$

The Lie derivative of a k -form is itself a k -form. For example, the Lie derivative of a zero-form f along the vector field \mathbf{v} is $\mathcal{L}_v f \equiv \mathbf{v} \cdot df = v^a \partial_a f$ (i.e., the directional derivative $\mathbf{v} \cdot \nabla f$), while the Lie derivative of a one-form $\Gamma = \Gamma_a dz^a$ is

$$\mathcal{L}_v \Gamma = [v^a \omega_{ab} + \partial_b(\mathbf{v} \cdot \Gamma)] dz^b.$$

Note that the Lie derivative satisfies the Leibnitz property $\mathcal{L}_v(fg) = (\mathcal{L}_v f)g + f(\mathcal{L}_v g)$ and \mathcal{L}_v commutes with d . The Lie transform generated by the vector field \mathbf{v} as $T = \exp \mathcal{L}_v$ such that T is distributive $T(fg) = (Tf)(Tg)$ and T commutes with d . For example, the pull-back operator $T_\epsilon = \exp(\epsilon \mathcal{L}_\xi)$ associated with the nonuniform space transformation

$$\mathbf{X} = \mathbf{x} + \epsilon \xi + \frac{\epsilon^2}{2} \xi \cdot \nabla \xi + \cdots,$$

which is generated by the vector field $\epsilon \xi$, yields the scalar-invariance identity

$$\begin{aligned} f(\mathbf{x}) &= F(\mathbf{X}) = F\left(\mathbf{x} + \epsilon \xi + \frac{\epsilon^2}{2} \xi \cdot \nabla \xi + \cdots\right) \\ &= F(\mathbf{x}) + \epsilon \xi \cdot \nabla F(\mathbf{x}) + \frac{\epsilon^2}{2} \xi \cdot \nabla(\xi \cdot \nabla F) + \cdots \\ &= \exp(\epsilon \xi \cdot \nabla) F(\mathbf{x}) \equiv \exp(\epsilon \mathcal{L}_\xi) F(\mathbf{x}). \end{aligned} \quad (\text{A4})$$

The pull-back operator associated with a near-identity transformation is expressed as a Lie-transform operation along the vector fields that generate the transformation.

A near-identity phase-space transformation $\mathbf{z} \rightarrow \bar{\mathbf{z}} = T_\epsilon \mathbf{z}$ carried out by Lie-transform methods (i.e., generated by the vector fields $\mathbf{G}_1, \mathbf{G}_2, \dots$) induces a transformation on Jacobians,

$$\mathcal{J} \rightarrow \bar{\mathcal{J}} \equiv \mathcal{J} - \epsilon \partial_a(\mathcal{J} G_1^a) + \cdots.$$

The phase-space transformation is said to be canonical if $\partial_a(\mathcal{J} G_1^a) + \cdots = 0$.

APPENDIX B: UNPERTURBED GUIDING-CENTER HAMILTONIAN DYNAMICS

In this appendix, derivations of guiding-center and bounce-averaged guiding-center (bounce-center) Hamiltonian dynamics for the case of a strong (static) magnetic field with weak spatial inhomogeneities in the absence of a background electric field are summarized.

1. Guiding-center phase-space transformation

Under the time-independent guiding-center transformation $(\mathbf{x}, \mathbf{p}) \rightarrow (\mathbf{X}, p_{\parallel}, \mu, \zeta)$, the particle phase-space Lagrangian, $\Gamma = (\mathbf{p} + e\mathbf{A}/c) \cdot d\mathbf{x} - (|\mathbf{p}|^2/2m)dt$ is transformed into the guiding-center phase-space Lagrangian,

$$\begin{aligned} \Gamma_{gc} = & \left[\epsilon^{-1} \frac{e}{c} \mathbf{A} + p_{\parallel} \hat{\mathbf{b}} - \epsilon \left(\frac{mc}{e} \right) \mu \mathbf{R}^* \right] \cdot d\mathbf{X} + \epsilon \left(\frac{mc}{e} \right) \mu d\zeta \\ & - H_{gc} dt, \end{aligned} \quad (B1)$$

where $\epsilon \equiv \epsilon_B$ is the ratio of the characteristic gyroradius to the magnetic-field gradient length scale, the vector \mathbf{R}^* is defined below, and the guiding-center Hamiltonian is

$$H_{gc} = \frac{p_{\parallel}^2}{2m} + \mu B. \quad (B2)$$

The guiding-center phase-space Lagrangian (B1) and guiding-center Hamiltonian (B2) were originally derived by Littlejohn (1983) using Lie-transform methods in the form of asymptotic expansions $Z_{gc}^{\alpha} = Z_0^{\alpha} + \epsilon G_1^{\alpha} + \dots$, where $Z_0^{\alpha} \equiv (\mathbf{x}, p_{\parallel 0}, \mu_0, \zeta_0)$ are local particle phase-space coordinates ($p_{\parallel 0} = \mathbf{p} \cdot \hat{\mathbf{b}}$ is the kinetic momentum parallel to the magnetic field, $\mu_0 = |\mathbf{p}_{\perp}|^2/2mB$ is the lowest-order magnetic moment, and $\zeta_0 = \arctan [(-\mathbf{p} \cdot \hat{1})/(-\mathbf{p} \cdot \hat{2})]$ is the gyration angle) and the components of the first-order generating vector field are

$$G_1^{\mathbf{x}} = -\boldsymbol{\rho}_0 = -\frac{mc}{e} \sqrt{\frac{2\mu}{mB}} \hat{\rho}, \quad (B3)$$

$$G_1^p_{\parallel} = (mc/e)\mu(\mathbf{a}_1 \cdot \nabla \hat{\mathbf{b}} + \hat{\mathbf{b}} \cdot \nabla \times \hat{\mathbf{b}}) - p_{\parallel} \boldsymbol{\rho}_0 \cdot (\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}}), \quad (B4)$$

$$\begin{aligned} G_1^{\mu} = & \boldsymbol{\rho}_0 \cdot \left(\mu \nabla \ln B + \frac{mv_{\parallel}^2}{B} \hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} \right) \\ & - \mu \frac{v_{\parallel}}{\Omega} (\mathbf{a}_1 \cdot \nabla \hat{\mathbf{b}} + \hat{\mathbf{b}} \cdot \nabla \times \hat{\mathbf{b}}), \end{aligned} \quad (B5)$$

$$\begin{aligned} G_1^{\zeta} = & -\boldsymbol{\rho}_0 \cdot \mathbf{R} + \frac{\partial \boldsymbol{\rho}_0}{\partial \zeta} \cdot \nabla \ln B + \frac{v_{\parallel}}{\Omega} \mathbf{a}_2 \cdot \nabla \hat{\mathbf{b}} \\ & + \frac{mv_{\parallel}^2}{2\mu B} \left(\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}} \cdot \frac{\partial \boldsymbol{\rho}_0}{\partial \zeta} \right). \end{aligned} \quad (B6)$$

Here we use the rotating (right-handed) unit vectors $(\hat{\mathbf{b}}, \hat{\perp}, \hat{\rho})$,

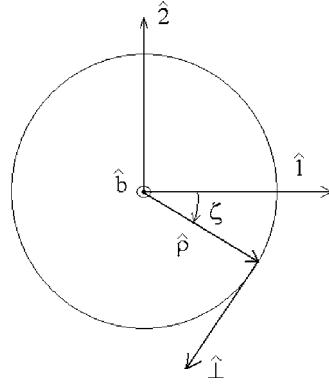


FIG. 6. Fixed unit vectors $(\hat{1}, \hat{2}, \hat{b})$ and rotating unit vectors $(\hat{\perp}, \hat{\rho}, \hat{b})$. Gyrogauge invariance involves an arbitrary rotation of the perpendicular unit vectors $\hat{1}$ and $\hat{2}$ about the parallel unit vector \hat{b} .

$$\hat{\perp} = -\hat{1} \sin \zeta - \hat{2} \cos \zeta = \frac{\partial \hat{\rho}}{\partial \zeta},$$

$$\hat{\rho} = \hat{1} \cos \zeta - \hat{2} \sin \zeta = -\frac{\partial \hat{\perp}}{\partial \zeta},$$

which are defined in terms of the fixed (local) unit vectors $\hat{1} \times \hat{2} = \hat{b}$ (see Fig. 6), the vector field $\mathbf{R} = \nabla \hat{\perp} \cdot \hat{\rho} = \nabla \hat{1} \cdot \hat{2}$ denotes Littlejohn's gyrogauge vector field (Littlejohn, 1983, 1988), which is used to define the gradient operator

$$\nabla^* \equiv \nabla + \mathbf{R}^* \frac{\partial}{\partial \zeta}, \quad (B7)$$

where $\mathbf{R}^* \equiv \mathbf{R} + (\hat{b} \cdot \nabla \times \hat{b}) \hat{b}/2$, and the gyroangle-dependent dyadic matrices $(\mathbf{a}_1, \mathbf{a}_2)$ are defined as

$$\begin{aligned} \mathbf{a}_1 = & -\frac{1}{2} (\hat{\rho} \hat{\perp} + \hat{\perp} \hat{\rho}) = \frac{\partial \mathbf{a}_2}{\partial \zeta}, \\ \mathbf{a}_2 = & \frac{1}{4} (\hat{\perp} \hat{\perp} - \hat{\rho} \hat{\rho}) = -\frac{1}{4} \frac{\partial \mathbf{a}_1}{\partial \zeta}. \end{aligned}$$

Gyrogauge invariance is defined in terms of the requirement that the guiding-center Hamiltonian dynamics be not only independent of the gyroangle ζ but also how it is measured. By introducing the gyrogauge transformation $\zeta' = \zeta + \chi(\mathbf{X})$, the perpendicular unit vectors $(\hat{1}, \hat{2})$ are transformed as $\hat{1}' = \hat{1} \cos \chi + \hat{2} \sin \chi$ and $\hat{2}' = -\hat{1} \sin \chi + \hat{2} \cos \chi$ so that the vector \mathbf{R} transforms as $\mathbf{R}' = \mathbf{R} + \nabla \chi$. For the guiding-center Hamiltonian dynamics to be gyrogauge invariant, the guiding-center phase-space Lagrangian (B1) must contain the gyrogauge-invariant term $d\zeta - \mathbf{R} \cdot d\mathbf{X} \equiv d\zeta' - \mathbf{R}' \cdot d\mathbf{X}$.

The guiding-center kinetic energy $\mathcal{E} = p_{\parallel}^2/2m + \mu B$ is identical to the particle kinetic energy (to first order) since the energy component of the first-order generating vector field,

$$G_1^{\mathcal{E}} = (p_{\parallel}/m) G_1^{p_{\parallel}} + B G_1^{\mu} + G_1^{\mathbf{x}} \cdot \boldsymbol{\mu} \nabla B \equiv 0,$$

is identically zero when the components (B3)–(B5) of the first-order generating vector field are used.

2. Guiding-center Hamiltonian dynamics

The Jacobian $\mathcal{J}=mB_{\parallel}^{*}$ for the guiding-center transformation $(\mathbf{x}, \mathbf{p}) \rightarrow (\mathbf{X}, p_{\parallel}, \mu, \zeta)$ is defined in terms of the guiding-center phase-space function $B_{\parallel}^{*}=\hat{\mathbf{b}} \cdot \mathbf{B}^{*}$ derived from the generalized magnetic field,

$$\mathbf{B}^{*} \equiv \mathbf{B} + \epsilon \left(\frac{cp_{\parallel}}{e} \right) \nabla \times \hat{\mathbf{b}},$$

where the second-order gyrogauge-invariant term $\epsilon^2(mc^2/e^2)\mu \nabla \times \mathbf{R}^{*}$ is omitted (note that while the vector \mathbf{R} is gyrogauge-dependent, its curl $\nabla \times \mathbf{R}$ is not). The components (B3)–(B6) of the first-order guiding-center generating vector field G_1^{α} satisfy the identity $B_{\parallel}^{*} \equiv B - \partial_{\alpha}(BG_1^{\alpha})$, which shows how the Jacobian mB for the local particle phase-space coordinates is transformed into the Jacobian mB_{\parallel}^{*} for the guiding-center phase-space coordinates.

The guiding-center Poisson bracket is constructed from the guiding-center phase-space Lagrangian (B1) and is expressed in terms of two arbitrary functions F and G of $(\mathbf{X}, p_{\parallel}, \mu, \zeta)$ as

$$\begin{aligned} \{F, G\}_{\text{gc}} &= \epsilon^{-1} \frac{e}{mc} \left(\frac{\partial F}{\partial \zeta} \frac{\partial G}{\partial \mu} - \frac{\partial F}{\partial \mu} \frac{\partial G}{\partial \zeta} \right) + \frac{\mathbf{B}^{*}}{B_{\parallel}^{*}} \cdot \left(\nabla^{*} F \frac{\partial G}{\partial p_{\parallel}} \right. \\ &\quad \left. - \frac{\partial F}{\partial p_{\parallel}} \nabla^{*} G \right) - \epsilon \frac{c \hat{\mathbf{b}}}{e B_{\parallel}^{*}} \cdot \nabla^{*} F \times \nabla^{*} G, \end{aligned} \quad (\text{B8})$$

where ∇^{*} is the gradient operator (B7) and the ϵ -ordering clearly separates the fast gyromotion time scale (ϵ^{-1}), the intermediate parallel time scale (ϵ^0), and the slow drift-motion time scale (ϵ).

The equations of guiding-center motion $\dot{Z}^{\alpha} = \{Z^{\alpha}, H_{\text{gc}}\}_{\text{gc}}$ are expressed in terms of the guiding-center Poisson bracket (B8) and the guiding-center Hamiltonian (B2) as

$$\begin{aligned} \dot{\mathbf{X}} &= v_{\parallel} \frac{\mathbf{B}^{*}}{B_{\parallel}^{*}} + \epsilon \frac{c \hat{\mathbf{b}}}{e B_{\parallel}^{*}} \times \mu \nabla B \\ &\equiv \mathbf{v}_{\text{gc}} = v_{\parallel} \hat{\mathbf{b}} + \epsilon \frac{c \hat{\mathbf{b}}}{e B_{\parallel}^{*}} \times (\mu \nabla B + m v_{\parallel}^2 \hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}}), \end{aligned} \quad (\text{B9})$$

$$\dot{p}_{\parallel} = - \frac{\mathbf{B}^{*}}{B_{\parallel}^{*}} \cdot \mu \nabla B, \quad (\text{B10})$$

while $\dot{\mu} \equiv -(\Omega/B) \partial H_{\text{gc}} / \partial \zeta \equiv 0$ and

$$\dot{\zeta} = \epsilon^{-1} \Omega + v_{\parallel} \hat{\mathbf{b}} \cdot \mathbf{R}^{*} + \mathcal{O}(\epsilon). \quad (\text{B11})$$

Note that the guiding-center Poisson bracket (B8) satisfies the Liouville identities

$$\epsilon \nabla \times \left(\frac{c \hat{\mathbf{b}}}{e} \right) - \frac{\partial \mathbf{B}^{*}}{\partial p_{\parallel}} = \mathbf{0} \quad \text{and} \quad \nabla \cdot \mathbf{B}^{*} = 0,$$

from which the guiding-center Liouville theorem is derived

$$\nabla \cdot \left(B_{\parallel}^{*} \frac{d\mathbf{X}}{dt} \right) + \frac{\partial}{\partial p_{\parallel}} \left(B_{\parallel}^{*} \frac{dp_{\parallel}}{dt} \right) = 0. \quad (\text{B12})$$

Moreover, by substituting $B_{\parallel}^{*} \rightarrow B$ in Eqs. (B9) and (B10), the guiding-center equations of Northrop (1963) are recovered but we lose the guiding-center Liouville property (B12).

3. Guiding-center pull-back transformation

The guiding-center pull-back transformation T_{gc} relates the guiding-center Vlasov distribution F to the particle Vlasov distribution $f = T_{\text{gc}} F$, expanded to first order in gradient length scale as

$$f = F - \boldsymbol{\rho}_0 \cdot \nabla F + \epsilon G_1^{p_{\parallel}} \frac{\partial F}{\partial p_{\parallel}} + \epsilon G_1^{\mu} \frac{\partial F}{\partial \mu}, \quad (\text{B13})$$

where the second term on the right side is also ordered at ϵ . The Vlasov equation in (local) particle phase space $(\mathbf{x}, p_{0\parallel}, \mu_0, \zeta_0)$ is expressed as

$$0 = \frac{df}{dt} \equiv \frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f + \dot{z}_0^i \frac{\partial f}{\partial z_0^i}, \quad (\text{B14})$$

where the velocity-space equations of motion $\dot{z}_0^i = (\dot{p}_{0\parallel}, \dot{\mu}_0, \dot{\zeta}_0)$ are expressed in terms of the first-order generating-field components (B4)–(B6) as

$$\dot{p}_{0\parallel} = - \epsilon \left(\mu_0 \hat{\mathbf{b}} \cdot \nabla B + \Omega \frac{\partial G_1^{p_{\parallel}}}{\partial \zeta} \right), \quad (\text{B15})$$

$$\dot{\mu}_0 = - \epsilon \Omega \frac{\partial G_1^{\mu}}{\partial \zeta}, \quad (\text{B16})$$

$$\dot{\zeta}_0 = (\Omega - \epsilon \boldsymbol{\rho}_0 \cdot \nabla \Omega) + \epsilon \left(v_{\parallel} \hat{\mathbf{b}} \cdot \mathbf{R}^{*} - \Omega \frac{\partial G_1^{\zeta}}{\partial \zeta} \right). \quad (\text{B17})$$

We now show that Eq. (B13) is a solution of the Vlasov equation (B14) provided the guiding-center Vlasov distribution F satisfies the guiding-center Vlasov equation,

$$0 = \frac{d_{\text{gc}} F}{dt} \equiv \frac{\partial F}{\partial t} + \mathbf{v}_{\text{gc}} \cdot \nabla F + \dot{p}_{\parallel} \frac{\partial F}{\partial p_{\parallel}}, \quad (\text{B18})$$

where $\partial F / \partial \zeta \equiv 0$ and $\dot{\mu} \equiv 0$; here we use $\mathbf{v}_{\text{gc}} = v_{\parallel} \hat{\mathbf{b}}$ and $\dot{p}_{\parallel} = -\epsilon \mu \hat{\mathbf{b}} \cdot \nabla B$. We write

$$\begin{aligned} \frac{df}{dt} &= \frac{\partial F}{\partial t} + \epsilon \left(\mathbf{v} - \Omega \frac{\partial \boldsymbol{\rho}_0}{\partial \zeta} \right) \cdot \nabla F + \left(\dot{p}_{0\parallel} + \epsilon \Omega \frac{\partial G_1^{p_{\parallel}}}{\partial \zeta} \right) \frac{\partial F}{\partial p_{\parallel}} \\ &\quad + \left(\dot{\mu}_0 + \epsilon \Omega \frac{\partial G_1^{\mu}}{\partial \zeta} \right) \frac{\partial F}{\partial \mu}, \end{aligned} \quad (\text{B19})$$

where we have used the fact that F is independent of the

gyroangle ζ . By inserting definitions (B15) and (B16) for $(\dot{p}_{0||}, \dot{\mu}_0)$, we find $d\dot{f}/dt = d_{gc}F/dt$, so that the particle Vlasov equation (B14) is satisfied if the guiding-center Vlasov equation (B18) is satisfied.

We conclude that the pull-back operator T_{gc} provides a partial solution of the particle Vlasov equation by integrating the fast-time-scale particle dynamics. Note that the guiding-center pull-back transformation (B13) is normally derived directly from the iterative solution of the particle Vlasov equation.

4. Bounce-center Hamiltonian dynamics

When the characteristic time scale τ is much longer than the bounce period (i.e., when the guiding center has executed many bounce cycles during time τ), the fast bounce angle can be asymptotically removed from the guiding center's orbital dynamics, and a corresponding adiabatic invariant (the longitudinal or bounce action $J \equiv J_b$) can be constructed. The resulting *bounce-averaged guiding-center* dynamics takes place in a reduced two-dimensional phase space with spatial (magnetic) coordinates (y^1, y^2) , where each coordinate y^a (with $a=1$ or 2) satisfies the condition $\mathbf{B} \cdot \nabla y^a = 0$. Using the notation $\bar{\mathbf{x}} \equiv (\mathbf{y}, s)$, with $\mathbf{y} \equiv (\alpha, \beta)$ used for coordinates in the space of field-line labels (i.e., each magnetic field line is represented as a point in \mathbf{y} space) and s is the parallel spatial coordinate (where $\hat{\mathbf{b}} \equiv \partial \mathbf{x} / \partial s$), the magnetic vector potential is

$$\mathbf{A} \equiv \frac{1}{2} \sum_{a,b} \eta_{ab} y^a \nabla y^b, \quad (B20)$$

where η_{ab} is antisymmetric in its indices (with $\eta_{12} = +1 = -\eta_{21}$). Using the orthogonality relations

$$\nabla \bar{x}^i \cdot \frac{\partial \mathbf{x}}{\partial \bar{x}^j} = \delta_j^i \quad (B21)$$

between the contravariant and covariant basis vectors, we obtain the following expression for ∇s :

$$\nabla s \equiv \hat{\mathbf{b}} - \sum_a \mathcal{R}_a \nabla y^a, \quad (B22)$$

where

$$\mathcal{R}_a \equiv \hat{\mathbf{b}} \cdot \frac{\partial \mathbf{x}}{\partial y^a}, \quad (B23)$$

while for $\partial \mathbf{x} / \partial y^a$ we find

$$\frac{\partial \mathbf{x}}{\partial y^a} \equiv \mathcal{R}_a \hat{\mathbf{b}} + \sum_b \eta_{ab} \nabla y^b \times \frac{\mathbf{B}}{B^2}. \quad (B24)$$

It is now quite simple to check that the sets $(\nabla y^a, \nabla s)$ and $(\partial \mathbf{x} / \partial y^a, \partial \mathbf{x} / \partial s)$ satisfy the relations (B21). Bounce-averaged guiding-center dynamics in static magnetic fields has been shown to possess a canonical Hamiltonian structure (Littlejohn, 1982b).

We begin with the unperturbed guiding-center phase-space Lagrangian (B1) written in magnetic coordinates $\bar{\mathbf{x}}$ as

$$\begin{aligned} \Gamma_0 &= \left(\frac{e}{2c\epsilon_d} \eta_{ab} y^a + p_{||} \mathcal{R}_b \right) dy^b + p_{||} ds - \left(\mu B + \frac{p_{||}^2}{2m} \right) dt \\ &\equiv \mathcal{F}_b dy^b + p_{||} ds - H_0 dt, \end{aligned} \quad (B25)$$

where the gyromotion dynamics ($\propto \mu d\zeta$) has been removed and the dimensionless parameter $\epsilon_d \ll 1$ is introduced as an ordering parameter representing the ratio of the fast bounce time scale to the slow drift time scale.

To lowest order in the drift ϵ_d ordering, the fast guidingcenter motion is described by the quasiperiodic bounce motion,

$$\dot{s} = v_{||} \quad \text{and} \quad \dot{v}_{||} = -(\mu/m) \partial_B B, \quad (B26)$$

i.e., the motion is taking place along a magnetic-field line (labeled by \mathbf{y}) and drift motion is absent (to lowest order). Following a standard procedure in classical mechanics (Goldstein *et al.*, 2002), one constructs action-angle canonical variables associated with this periodic motion. The action-angle coordinates (J, Θ) associated with periodic bounce motion have the following lowest-order expressions: for the bounce action $J \equiv J_b = J_0 + \dots$, we find (Northrop, 1963; Littlejohn, 1982b)

$$\begin{aligned} J_0(\mathcal{E}, \mu; \mathbf{y}) &\equiv \frac{1}{2\pi} \oint p_{||}(s, \mathcal{E}, \mu; \mathbf{y}) ds \\ &= \frac{1}{\pi} \int_{s_0}^{s_1} \sqrt{2m[\mathcal{E} - \mu B(s; \mathbf{y})]} ds, \end{aligned} \quad (B27)$$

where (s_0, s_1) are the turning points where $v_{||}$ vanishes, while for the bounce angle $\Theta = \Theta_0 + \dots$ we find (Littlejohn, 1982b)

$$\Theta_0(s, \mathcal{E}, \mu; \mathbf{y}) \equiv \pi \pm \omega_b \int_{s_0}^s \frac{ds'}{\sqrt{\frac{2}{m}[\mathcal{E} - \mu B(s'; \mathbf{y})]}}, \quad (B28)$$

where \pm denotes the sign of $v_{||}$ and $\Theta(s=s_0) \equiv \pi$ for both branches. The bounce frequency ω_b is defined from Eq. (B27) as

$$\omega_b(\mathbf{y}; \mathcal{E}, \mu) \equiv \left(\frac{\partial J}{\partial \mathcal{E}} \right)^{-1} = 2\pi \left(\oint \frac{ds}{v_{||}} \right)^{-1}. \quad (B29)$$

We now proceed with the substitution $(s, p_{||}) \rightarrow (J, \Theta)$ in the guiding-center phase-space Lagrangian (B25). The transformation $(s, p_{||}) \rightarrow (J, \Theta) \equiv \mathbf{u}$ is canonical since $dp_{||} \wedge ds \equiv dJ \wedge d\Theta$. In the guiding-center phase-space Lagrangian (B25), the differential ds becomes $ds = \partial_\alpha s du^\alpha$ and we have

$$\begin{aligned} \Gamma_0 &\equiv \left(\frac{q}{2c\epsilon_d} \eta_{ab} y^a + p_{||} \mathcal{R}_b \right) dy^b + \left(p_{||} \frac{\partial s}{\partial u^\alpha} \right) du^\alpha \\ &\quad - H_0(\mathbf{y}, \mathbf{u}) dt, \end{aligned} \quad (B30)$$

where $H_0(\mathbf{y}, \mathbf{u}) \equiv \mu B(\mathbf{y}; s(\mathbf{u})) + [p_{||}(\mathbf{y}, \mathbf{u})]^2/2m$ is the lowest-order unperturbed guiding-center Hamiltonian

and explicit bounce-angle dependence now appears in the guiding-center phase-space Lagrangian (B30). Because of its dependence on the field-line labels \mathbf{y} , the bounce action (B27) is not conserved at order ϵ_d [i.e., $dJ/dt = \mathcal{O}(\epsilon_d)$]. To remove the bounce-angle dependence in the guiding-center phase-space Lagrangian (B30) and construct an asymptotic expansion for the bounce-action adiabatic invariant, we proceed by performing an infinitesimal transformation $(\mathbf{y}, \mathbf{u}) \rightarrow (\bar{\mathbf{y}}, \bar{\mathbf{u}})$, where the relation between the guiding-center coordinates (\mathbf{y}, \mathbf{u}) and the bounce-guiding-center coordinates $(\bar{\mathbf{y}}, \bar{\mathbf{u}})$ is given in terms of the asymptotic expansions

$$\begin{aligned}\bar{y}^a &\equiv y^a + \epsilon_d G_1^a + \dots, \\ \bar{u}^\alpha &\equiv u^\alpha + \epsilon_d G_1^\alpha + \dots,\end{aligned}\quad (B31)$$

where the components G_n^a and G_n^α of the n th-order generating vector field are constructed so that the bounce action $\bar{J} = J + \sum_{k=1}^n \epsilon_d^k G_k^J$ is conserved at the n th order, i.e., $d\bar{J}/dt = \mathcal{O}(\epsilon_d^{n+1})$. The \mathbf{y} components of the first-order generating vector are (Littlejohn, 1982)

$$G_1^a = -\eta^{ab} \frac{c}{e} \left(\frac{\partial S_1}{\partial \bar{y}^b} + p_\parallel \mathcal{R}_b \right), \quad (B32)$$

where $\eta^{ab} = -\eta_{ab}$ and the gauge function $S_1(\bar{\mathbf{y}}, \bar{\mathbf{u}})$ is defined from

$$\frac{\partial S_1}{\partial \bar{u}^\beta} = -\frac{\eta_{\alpha\beta}}{2} \bar{u}^\alpha - p_\parallel(\bar{\mathbf{y}}, \bar{\mathbf{u}}) \frac{\partial S(\bar{\mathbf{u}})}{\partial \bar{u}^\beta}, \quad (B33)$$

with $\eta_{\alpha\beta}$ antisymmetric in its indices (with $\eta_{12} = +1$).

The purpose of the transformation (B31) is to remove the bounce-angle dependence at all orders in ϵ_d . The unperturbed bounce-averaged guiding-center (or bounce-center) phase-space Lagrangian becomes

$$\bar{\Gamma}_0 \equiv \epsilon_d^{-1} \frac{e}{2c} \eta_{ab} \bar{y}^a d\bar{y}^b + \bar{J} d\bar{\Theta} - \bar{H}_0(\bar{\mathbf{y}}, \bar{J}; \epsilon_d) dt, \quad (B34)$$

and the unperturbed bounce-center Hamiltonian is (Littlejohn, 1982b)

$$\bar{H}_0 \equiv H_0 - \frac{\epsilon_d}{2} \left(\omega_b \eta_{ab} \left\langle G_1^a \frac{\partial G_1^b}{\partial \bar{\Theta}} \right\rangle_b \right), \quad (B35)$$

where $\langle \rangle_b$ denotes averaging with respect to $\bar{\Theta}$. The unperturbed bounce guiding-center Poisson bracket is defined in terms of two arbitrary functions \mathcal{F} and \mathcal{G} on bounce guiding-center phase space $(\bar{\mathbf{y}}, \bar{\mathbf{u}})$ as

$$\{\mathcal{F}, \mathcal{G}\} = \frac{\partial \mathcal{F}}{\partial \bar{u}^\alpha} \eta^{\alpha\beta} \frac{\partial \mathcal{G}}{\partial \bar{u}^\beta} + \epsilon_d \frac{c}{e} \frac{\partial \mathcal{F}}{\partial \bar{y}^a} \eta^{ab} \frac{\partial \mathcal{G}}{\partial \bar{y}^b}, \quad (B36)$$

where $\eta^{\alpha\beta} \equiv \eta_{\alpha\beta}^{-1} = -\eta_{\alpha\beta}$ and the first term on the right represents the bounce motion while the second term represents the bounce-averaged drift motion.

The bounce-guiding-center position \bar{y}^a is the (bounce-motion) time-averaged position of the guiding-center position y^a , i.e., $\bar{y}^a \equiv \langle y^a \rangle_b$, and thus

$$\Lambda_b^a \equiv y^a - \langle y^a \rangle_b = -\epsilon_d G_1^a(\bar{\mathbf{y}}, \bar{\mathbf{u}}) \quad (B37)$$

represents the bounce-angle-dependent bounce radius. In tokamak geometry (Boozer, 2004), we find that $\Lambda_b^\psi \neq 0$ implies that trapped-particle orbits have finite banana widths, while in an axisymmetric magnetic-dipole field $\mathbf{B} = \nabla\psi \times \nabla\varphi$, where the azimuthal angle φ is an ignorable angle, we find $\Lambda_b^\psi \equiv 0$ implies that trapped-particle orbits remain on the same magnetic surface ψ .

APPENDIX C: PUSH-FORWARD REPRESENTATION OF FLUID MOMENTS

1. Push-forward representation of fluid moments

Applications of Lie-transform methods in plasma physics include the transformation of an arbitrary fluid moment on particle phase space into a fluid moment on the transformed phase space. With the help of the push-forward representation of arbitrary fluid moments, we uncover several polarization and magnetization effects in Maxwell's equations that are related to the phase-space transformation itself.

We start with the push-forward representation (139) for the moment $\|v^\mu\|$, where $v^\mu = (c, \mathbf{v})$, and expand it to first order in the displacement $\boldsymbol{\rho}_\epsilon$, defined in terms of the generating vector fields (G_1, G_2, \dots) by Eq. (140), so that we obtain

$$\|v^\mu\| = \int d^4\bar{p} (\mathbf{T}_\epsilon^{-1} v^\mu) \bar{\mathcal{F}} - \nabla \cdot \left[\int d^4\bar{p} \boldsymbol{\rho}_\epsilon (\mathbf{T}_\epsilon^{-1} v^\mu) \bar{\mathcal{F}} + \dots \right], \quad (C1)$$

where integration by parts was performed to obtain the second term, and terms omitted inside the divergence include higher-order multipole moments (e.g., electric and magnetic quadrupole moments).

We now derive the push-forward representation for the four-current $J^\mu = (c\rho, \mathbf{J}) \equiv \Sigma e \|v^\mu\|$. First, we derive the push-forward expression for the charge density (145), $\rho = \bar{\rho} - \nabla \cdot \mathbf{P}_\epsilon$, where $\bar{\rho} \equiv \Sigma e \int d^4\bar{p} \bar{\mathcal{F}}$ denotes the reduced charge density and the polarization vector is defined as

$$\mathbf{P}_\epsilon \equiv \sum e \int d^4\bar{p} \left[\boldsymbol{\rho}_\epsilon \bar{\mathcal{F}} - \nabla \cdot \left(\frac{\boldsymbol{\rho}_\epsilon \boldsymbol{\rho}_\epsilon}{2} \bar{\mathcal{F}} + \dots \right) \right], \quad (C2)$$

where $\mathbf{T}_\epsilon^{-1} v^0 \equiv \mathbf{T}_\epsilon^{-1} c = c$ was used in Eq. (C1) and $\pi_\epsilon \equiv e \boldsymbol{\rho}_\epsilon$ is the lowest-order electric-dipole moment associated with the charge separation induced by the phase-space transformation. Second, we derive the push-forward expression for the current density (146), where the push-forward of the particle velocity $\mathbf{v} = d\mathbf{x}/dt$ (using its Lagrangian representation),

$$\mathbf{T}_\epsilon^{-1} \mathbf{v} = \mathbf{T}_\epsilon^{-1} \frac{d\mathbf{x}}{dt} = \left[\mathbf{T}_\epsilon^{-1} \frac{d}{dt} \mathbf{T}_\epsilon \right] (\mathbf{T}_\epsilon^{-1} \mathbf{x}) \equiv \frac{d_\epsilon \bar{\mathbf{x}}}{dt} + \frac{d_\epsilon \boldsymbol{\rho}_\epsilon}{dt}, \quad (C3)$$

is expressed in terms of the reduced total time derivative d_ϵ/dt . Here $d_\epsilon \bar{\mathbf{x}}/dt$ denotes the reduced (e.g., guiding-center) velocity and $d_\epsilon \boldsymbol{\rho}_\epsilon/dt$ denotes the particle polarization velocity, which includes the perpendicular par-

ticle velocity and the standard polarization drift velocity (Sosenko *et al.*, 2001). We replace the term $d_{\epsilon}\boldsymbol{\rho}_{\epsilon}/dt$ in Eq. (C3), which contains the polarization-drift velocity, by using the following identity based on the expression (C2) for the reduced polarization vector:

$$\begin{aligned} \frac{\partial \mathbf{P}_{\epsilon}}{\partial t} &= \sum e \int d^4 \bar{p} \left[\left(\frac{\partial \boldsymbol{\rho}_{\epsilon}}{\partial t} \bar{\mathcal{F}} + \boldsymbol{\rho}_{\epsilon} \frac{\partial \bar{\mathcal{F}}}{\partial t} \right) \right. \\ &\quad \left. - \nabla \cdot \left(\frac{1}{2} \frac{\partial (\boldsymbol{\rho}_{\epsilon} \boldsymbol{\rho}_{\epsilon})}{\partial t} \bar{\mathcal{F}} + \frac{\boldsymbol{\rho}_{\epsilon} \boldsymbol{\rho}_{\epsilon}}{2} \frac{\partial \bar{\mathcal{F}}}{\partial t} \right) + \dots \right] \\ &= \sum e \int d^4 \bar{p} \left\{ \left(\frac{d_{\epsilon} \boldsymbol{\rho}_{\epsilon}}{dt} \right) \bar{\mathcal{F}} - \nabla \cdot \left[\bar{\mathcal{F}} \left(\frac{d_{\epsilon} \bar{\mathbf{x}}}{dt} \boldsymbol{\rho}_{\epsilon} \right. \right. \right. \\ &\quad \left. \left. \left. + \frac{1}{2} \frac{d_{\epsilon} (\boldsymbol{\rho}_{\epsilon} \boldsymbol{\rho}_{\epsilon})}{dt} \right) + \dots \right] \right\}, \end{aligned} \quad (\text{C4})$$

where the reduced Vlasov equation (108) was used and integration by parts was performed. The push-forward representation of the current density is

$$\begin{aligned} \mathbf{J} &= \bar{\mathbf{J}} + \frac{\partial \mathbf{P}_{\epsilon}}{\partial t} + \nabla \times \left[\sum e \int d^4 \bar{p} \boldsymbol{\rho}_{\epsilon} \right. \\ &\quad \left. \times \left(\frac{d_{\epsilon} \bar{\mathbf{x}}}{dt} + \frac{1}{2} \frac{d_{\epsilon} \boldsymbol{\rho}_{\epsilon}}{dt} \right) \bar{\mathcal{F}} \right] \\ &\equiv \bar{\mathbf{J}} + \mathbf{J}_{\text{pol}} + \mathbf{J}_{\text{mag}}, \end{aligned} \quad (\text{C5})$$

where $\bar{\mathbf{J}} \equiv \sum e \int d^4 \bar{p} (d_{\epsilon} \bar{\mathbf{x}}/dt) \bar{\mathcal{F}}$ is the reduced current density, $\mathbf{J}_{\text{pol}} \equiv \partial \mathbf{P}_{\epsilon} / \partial t$ is the polarization current, and $\mathbf{J}_{\text{mag}} \equiv c \nabla \times \mathbf{M}_{\epsilon}$ is the divergenceless magnetization current, with magnetization vector

$$\begin{aligned} \mathbf{M}_{\epsilon} &= \sum \frac{e}{c} \int d^4 \bar{p} \boldsymbol{\rho}_{\epsilon} \times \left(\frac{d_{\epsilon} \bar{\mathbf{x}}}{dt} + \frac{1}{2} \frac{d_{\epsilon} \boldsymbol{\rho}_{\epsilon}}{dt} \right) \bar{\mathcal{F}} \\ &\equiv \sum \int d^4 \bar{p} \bar{\mathcal{F}} \left(\boldsymbol{\mu}_{\epsilon} + \boldsymbol{\pi}_{\epsilon} \times \frac{1}{c} \frac{d_{\epsilon} \bar{\mathbf{x}}}{dt} \right) \end{aligned} \quad (\text{C6})$$

expressed in terms of the intrinsic magnetic-dipole contribution

$$\boldsymbol{\mu}_{\epsilon} \equiv \frac{e}{2c} \boldsymbol{\rho}_{\epsilon} \times \frac{d_{\epsilon} \boldsymbol{\rho}_{\epsilon}}{dt} \quad (\text{C7})$$

and the moving electric-dipole contribution ($\boldsymbol{\pi}_{\epsilon} \times d_{\epsilon} \bar{\mathbf{x}}/dt$) (Jackson, 1975).

We have shown that Lie-transform methods offer a powerful approach for introducing reduced polarization and magnetization effects into the reduced Maxwell equations, which are associated with a near-identity phase-space transformation designed to eliminate fast degrees of freedom from the Vlasov-Maxwell equations.

2. Push-forward representation of gyrocenter fluid moments

The push-forward representation of fluid moments (C1) is now used to derive gyrocenter polarization and magnetization effects, which can then be compared with the variational expressions presented in Sec. III.

Based on Lie-transform perturbation analysis presented in Sec. V, the gyrocenter displacement (gyroradius) vector is defined as

$$\boldsymbol{\rho}_{\text{gy}} \equiv \boldsymbol{\rho}_{\text{gc}} - \epsilon_{\delta} \mathbf{G}_1^* + \dots, \quad (\text{C8})$$

where $\boldsymbol{\rho}_{\text{gc}} = \boldsymbol{\rho}_0 + \dots$ denotes the gyroangle-dependent gyroradius and the effective first-order gyrocenter vector field \mathbf{G}_1^* is [see Eqs. (164), (166), and (167)]

$$\begin{aligned} \mathbf{G}_1^* &= G_1^x + G_1^{\mu} \frac{\partial \boldsymbol{\rho}_0}{\partial \mu} + G_1^{\zeta} \frac{\partial \boldsymbol{\rho}_0}{\partial \zeta} \\ &= \{S_1, \mathbf{X} + \boldsymbol{\rho}_0\}_0 + \alpha \frac{\hat{\mathbf{b}}_0}{B_0} \times \langle \delta \mathbf{A}_{\perp \text{gc}} \rangle, \end{aligned} \quad (\text{C9})$$

where S_1 is defined in Eq. (172) and other definitions are given in Sec. V.

We begin with the gyrocenter electric-dipole moment $\boldsymbol{\pi}_{\text{gy}} \equiv e \langle \boldsymbol{\rho}_{\text{gy}} \rangle$, where

$$\langle \boldsymbol{\rho}_{\text{gy}} \rangle = - \frac{e}{B_0} \frac{\partial}{\partial \mu} \langle \delta \tilde{\psi}_{\text{gc}} \boldsymbol{\rho}_0 \rangle - \alpha \frac{\hat{\mathbf{b}}_0}{B_0} \times \langle \delta \mathbf{A}_{\perp \text{gc}} \rangle. \quad (\text{C10})$$

In the zero-Larmor-radius limit, we find

$$\boldsymbol{\pi}_{\text{gy}} = - \frac{mc^2}{B_0^2} \left(\nabla_{\perp} \delta \phi - \frac{p_{\parallel}}{mc} \nabla_{\perp} \delta A_{\parallel} \right) + (1 - \alpha) \frac{e \hat{\mathbf{b}}_0}{B_0} \times \delta \mathbf{A}_{\perp}, \quad (\text{C11})$$

which, in the gyrocenter Hamiltonian model ($\alpha=0$), is identical to Eq. (59) derived by the variational method. Extension of the gyrokinetic polarization density to gyrobounce-kinetics leads to the neoclassical polarization density (Fong and Hahm, 1999; Brizard, 2000c) as discussed in Appendix D.2.

The intrinsic gyrocenter magnetic-dipole moment is

$$\boldsymbol{\mu}_{\text{gy}} \equiv \frac{e}{2c} \left\langle \boldsymbol{\rho}_{\text{gy}} \times \frac{d_{\text{gy}} \boldsymbol{\rho}_{\text{gy}}}{dt} \right\rangle = - \mu \hat{\mathbf{b}}_0 + \dots, \quad (\text{C12})$$

which is identical to Eq. (60) derived by the variational method; by taking into account the background magnetic-field nonuniformity, a moving-electric-dipole contribution is added to the guiding-center magnetization current (Kaufman, 1986).

The push-forward and variational methods yield identical expressions for the polarization and magnetization effects appearing in the reduced Maxwell's equations.

APPENDIX D: EXTENSIONS OF NONLINEAR GYROKINETIC EQUATIONS

In this appendix, two extensions of nonlinear gyrokinetic theory are presented. First, extension of the nonlinear gyrokinetic equations presented in the text by introducing the effects of an inhomogeneous equilibrium electric field are discussed. Two new ordering parameters must first be introduced: the dimensionless parameter ϵ_E represents the strength of the equilibrium $E \times B$ velocity (e.g., compared to the ion thermal velocity),

while the dimensionless parameter ϵ_S represents the gradient-length scale of the $E \times B$ shear flow (e.g., compared to the ion thermal gyroradius).

The second extension of the nonlinear gyrokinetic equations presented in this appendix involves the derivation of nonlinear bounce-kinetic equations, in which the fast bounce-motion time scale of trapped guiding centers is asymptotically removed by Lie-transform perturbation methods.

1. Strong $E \times B$ flow shear

In Sec. II the various expansion parameters appearing in nonlinear gyrokinetic theory originate from different physical reasons, and the standard nonlinear gyrokinetic ordering is not a unique ordering. In this appendix, an example in which a further ordering consideration is necessary is presented. This example not only demonstrates the flexibility and the power of the modern Lie-transform perturbation approach, but also addresses highly relevant forefront research issues in magnetically confined plasmas. While the nonlinear gyrokinetic theory based on the standard ordering captures most of the essential physics associated with tokamak core turbulence, significant experimental progress in reducing turbulence and transport in the past decade has demonstrated that a new parameter regime characterized by a strong shear in the $E \times B$ flow, a steep pressure gradient, and a low fluctuation level can be reproduced routinely. This motivates further improvement of the standard nonlinear gyrokinetic ordering.

The analytic nonlinear theories of the $E \times B$ shear decorrelation of turbulence (Biglari *et al.*, 1990) and of transition dynamics (Carreras *et al.*, 1994; Diamond, Liang, *et al.*, 1994) in cylindrical geometry have demonstrated a possible important role of the $E \times B$ shear in the (low-confinement) L -mode to (high-confinement) H -mode transition (Wagner *et al.*, 1982; Burrell, 1997). Consequent generalization of the $E \times B$ shearing rate to toroidal geometry (Hahm, 1994; Hahm and Burrell, 1995) with a proper dependence on the poloidal magnetic field B_θ has made this hypothesis applicable to core transport barriers in reversed-shear plasmas (Mazzucato *et al.*, 1996; Burrell, 1997; Synakowski *et al.*, 1997) and has been utilized in analytical threshold calculations for transport bifurcation (Diamond *et al.*, 1997; Lebedev and Diamond, 1997).

While there has been significant progress in both shear-flow physics [see, for example, Diamond and Kim (1991); Terry (2000)] and transport-barrier physics [see, for example, theory reviews by Connor and Wilson (2000) and by Hahm (2002), and experimental reviews by Burrell (1997) and by Synakowski *et al.* (1997)], nonlinear gyrokinetic simulations are desirable for more quantitative comparisons to experimental data and extrapolation to future machines. The existing nonlinear gyrokinetic formalism in the absence of the equilibrium radial electric field ($E_r=0$) needs to be improved further for an accurate description of plasma turbulence in a

core transport barrier region with significant E_r shear. We note that many previous works, which contain the modification of the gyrokinetic Vlasov equation due to plasma flow (Bernstein and Catto, 1985; Hahm, 1992; Artun and Tang, 1994; Brizard, 1995), consider a situation in which the toroidal flow of ions is the dominant contributor to the radial electric field (Hinton *et al.*, 1994). Therefore, those equations cannot be applied to some core transport barriers where either the poloidal or diamagnetic flow plays a dominant role (Bell *et al.*, 1998; Cromb   *et al.*, 2005). Furthermore, since the individual guiding-center motion is determined by the electromagnetic field rather than by the equilibrium mass-flow velocity, it is natural to develop a gyrokinetic theory in terms of E_r (Hahm, 1996) in the laboratory frame. This approach is also conceptually simpler than a formulation in terms of the relative velocity in the frame moving with the mass flow (Hahm, 1992; Artun and Tang, 1994; Brizard, 1995) because one can formally treat the guiding-center motion part separately from the equilibrium mass-flow issue, which is related to the determination of the ion distribution function from neoclassical theory.

A general formulation can be pursued with $u_E/v_{th} \sim 1$, in addition to the standard gyrokinetic ordering $\omega/\Omega \sim e\delta\phi/T_i \sim \rho_i k_\parallel \sim \epsilon$ and $k_\perp \rho_i \sim 1$. Here u_E is the equilibrium $E \times B$ velocity.¹⁶ Only electrostatic fluctuations are considered in this appendix.

We begin with the unperturbed guiding-center phase-space Lagrangian,

$$\Gamma_0 \equiv \left(\frac{e}{c} \mathbf{A} + m \mathbf{u}_E + mv_\parallel \hat{\mathbf{b}} \right) \cdot d\mathbf{X} + \frac{\mu B}{\Omega} d\zeta - H_0 dt, \quad (D1)$$

where the equilibrium $E \times B$ velocity $\mathbf{u}_E \equiv (c\hat{\mathbf{b}}/B) \times \nabla\Phi$ (Littlejohn, 1981) is associated with the equilibrium potential Φ , and the guiding-center Hamiltonian is

$$H_0 = e\Phi + \mu(B + B_E) + \frac{m}{2}(v_\parallel^2 + |\mathbf{u}_E|^2), \quad (D2)$$

with

$$B_E \equiv \frac{\mathbf{B}}{2\Omega} \cdot \nabla \times \mathbf{u}_E = \frac{c}{2} \left[\nabla \cdot \left(\frac{\nabla\Phi}{\Omega} \right) + \frac{\nabla\Phi}{\Omega} \cdot (\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}}) \right] \quad (D3)$$

the finite-Larmor-orbit-average reduction of the equilibrium potential (Brizard, 1995). While the term μB_E in the guiding-center Hamiltonian (D2) might be smaller than $m|\mathbf{u}_E|^2/2$, we choose to keep it because of its clear physical meaning.

Introducing the electrostatic perturbation $\delta\phi(\mathbf{x}, t)$, the Lie-transform perturbation analysis can be carried out as described in Sec. V and further details can be found

¹⁶A tokamak-specific ordering, $B_\theta/B = r/qR \ll 1$, with further subsidiary orderings, simplifies the formulation for applications. This exemplifies the nonuniqueness of the standard nonlinear gyrokinetic ordering; details can be found in Hahm (1996) and Hahm, Lin, *et al.* (2004).

in Hahm (1996). Perturbation analysis up to the second order is required for energy conservation up to $O(\epsilon_\phi^2)$ in the formulation in terms of the total distribution function (Dubin *et al.* 1983; Hahm, 1988; Brizard, 1989a). The total phase-space Lagrangian is given up to the second order by

$$\bar{\Gamma} = \left(\frac{e}{c} \mathbf{A} + m \mathbf{u}_E + m \bar{v}_\parallel \hat{\mathbf{b}} \right) \cdot d\bar{\mathbf{X}} + \frac{\bar{\mu} B}{\Omega} d\bar{\zeta} - (H_0 + e \delta\Psi_{gy}) dt, \quad (\text{D4})$$

where the effective gyrocenter perturbation potential is

$$\delta\Psi_{gy} \equiv \langle \delta\phi_{gc} \rangle - \frac{e}{2B} \frac{\partial}{\partial \bar{\mu}} \langle \delta\tilde{\phi}_{gc}^2 \rangle.$$

The corresponding Euler-Lagrange equation is

$$-\frac{e \mathbf{B}^*}{c} \times \frac{d\bar{\mathbf{X}}}{dt} - m \hat{\mathbf{b}} \frac{d\bar{v}_\parallel}{dt} = \bar{\nabla}(H_0 + e \delta\Psi_{gy}), \quad (\text{D5})$$

which can be decomposed into the following gyrocenter equations of motion:

$$\begin{aligned} \frac{d\bar{\mathbf{X}}}{dt} &= \bar{v}_\parallel \frac{\mathbf{B}^*}{B_\parallel^*} + \frac{c \hat{\mathbf{b}}}{e B_\parallel^*} \times \left[e \bar{\nabla}(\Phi + \delta\Psi_{gy}) + \bar{\mu} \bar{\nabla}(B + B_E) \right. \\ &\quad \left. + \frac{m}{2} \bar{\nabla} |\mathbf{u}_E|^2 \right]. \end{aligned} \quad (\text{D6})$$

and

$$\begin{aligned} \frac{d\bar{v}_\parallel}{dt} &= -\frac{\mathbf{B}^*}{m B_\parallel^*} \cdot \left[e \bar{\nabla}(\Phi + \delta\Psi_{gy}) + \bar{\mu} \bar{\nabla}(B + B_E) \right. \\ &\quad \left. + \frac{m}{2} \bar{\nabla} |\mathbf{u}_E|^2 \right]. \end{aligned} \quad (\text{D7})$$

Although Eqs. (D6) and (D7) are mathematically concise, they can be written in the following form, which is closer to the results of previous work in terms of the mass flow (Artun and Tang, 1994; Brizard, 1995):

$$\begin{aligned} \frac{d\bar{\mathbf{X}}}{dt} &= \mathbf{u}_E + \bar{v}_\parallel \hat{\mathbf{b}} + \frac{c \hat{\mathbf{b}}}{e B_\parallel^*} \times [e \bar{\nabla} \delta\Psi_{gy} + \bar{\mu} \bar{\nabla}(B + B_E) \\ &\quad + m(\mathbf{u}_E + \bar{v}_\parallel \hat{\mathbf{b}}) \cdot \bar{\nabla}(\mathbf{u}_E + \bar{v}_\parallel \hat{\mathbf{b}})] \end{aligned} \quad (\text{D8})$$

and

$$\begin{aligned} \frac{d\bar{v}_\parallel}{dt} &= -\frac{\mathbf{B}^{*(0)}}{m B_\parallel^{*(0)}} \cdot [e \bar{\nabla}(\Phi_1 + \delta\Psi_{gy}) + \bar{\mu} \bar{\nabla}(B + B_E^{(0)}) \\ &\quad + m(\mathbf{u}_E^{(0)} + \bar{v}_\parallel \hat{\mathbf{b}}) \cdot \bar{\nabla}(\mathbf{u}_E^{(0)} + \bar{v}_\parallel \hat{\mathbf{b}})]. \end{aligned} \quad (\text{D9})$$

Here $\mathbf{u}_E^{(0)} \equiv \hat{\mathbf{b}} \times \nabla \Phi^{(0)} / B$, $\mathbf{B}^{*(0)} \equiv \mathbf{B} + (m/e) \nabla \times (\mathbf{u}_E^{(0)} + v_\parallel \hat{\mathbf{b}})$, and $B_\parallel^{*(0)} \equiv \hat{\mathbf{b}} \cdot \mathbf{B}^{*(0)}$. Although Eq. (D8) is valid for an arbitrary Φ , Eq. (D9) can be only obtained from Eq. (D7) via a perturbative analysis (Brizard, 1995). The equilibrium electrostatic potential, in general, consists of two parts $\Phi \equiv \Phi_0 + \Phi_1$. In most cases, Φ can be approximated by a flux function $\Phi_0(\psi)$ satisfying $\hat{\mathbf{b}} \cdot \nabla \Phi_0 = 0$. The poloidal-angle-dependent $\Phi_1(\psi, \theta)$ can be produced, for

instance, by the centrifugal-force-driven charge separation in strongly rotating plasmas (Hinton and Wong, 1985; Connor *et al.*, 1987). According to the ordering in this section, $\Phi_0 = O(\epsilon_E^{-1})$ and $\Phi_1 = O(1)$. The theory of $E \times B$ flow-shear suppression of turbulence has also been extended to include the poloidal-angle-dependent potential $\Phi_1(\psi, \theta)$ (Hahm and Burrell, 1996), exhibiting the tensor nature of the shearing process, which can also occur when small eddies (e.g., from ETG turbulence) are sheared by large convective cells (e.g., from ITG turbulence) (Holland and Diamond, 2004).

With Eqs. (D8) and (D9), one can write explicitly the gyrokinetic Vlasov equation for the gyrocenter distribution function $\bar{F}(\bar{\mathbf{X}}, \bar{\mu}, \bar{v}_\parallel, t)$,

$$\frac{\partial \bar{F}}{\partial t} + \frac{d\bar{\mathbf{X}}}{dt} \cdot \bar{\nabla} \bar{F} + \frac{d\bar{v}_\parallel}{dt} \frac{\partial \bar{F}}{\partial \bar{v}_\parallel} = 0. \quad (\text{D10})$$

Note that $d\bar{\mu}/dt \equiv 0$ and $\partial F/\partial \bar{\zeta} \equiv 0$ have been used. The accompanying gyrokinetic Poisson equation expressed in terms of the gyrocenter distribution function $\bar{F}(\bar{\mathbf{X}}, \bar{\mu}, \bar{v}_\parallel, t)$ is (Hahm, 1996)

$$\nabla^2(\Phi + \delta\phi) = 4\pi e(n_e - \bar{N}_i), \quad (\text{D11})$$

where the ion gyrofluid density

$$\bar{N}_i \equiv \int d^3 \bar{p} \left\langle e^{-\rho \cdot \nabla} \left(F + \frac{e \delta\tilde{\phi}_{gc}}{B} \frac{\partial F}{\partial \bar{\mu}} \right) \right\rangle$$

includes the ion polarization density, and the electron density n_e can be obtained from the drift-kinetic equation ($\rho_e \cdot \nabla \rightarrow 0$). The invariant energy for Eqs. (D10) and (D11) is obtained by transforming the energy constant of the original Vlasov-Poisson system,

$$\begin{aligned} E &= \int d^6 \bar{Z} \bar{F}_i \left[\bar{\mu}(B + B_E) + \frac{m}{2} (|\mathbf{u}_E|^2 + \bar{v}_\parallel^2) \right] \\ &\quad + \int d^6 z f_e \left(\frac{m_e}{2} v^2 \right) + \int \frac{d^3 x}{8\pi} |\mathbf{E}|^2 \\ &\quad + \frac{e^2}{2B} \int d^6 \bar{Z} \bar{F}_i \left(\frac{\partial}{\partial \bar{\mu}} \langle \delta\tilde{\phi}_{gc}^2 \rangle \right), \end{aligned} \quad (\text{D12})$$

where $\mathbf{E} \equiv -\nabla(\Phi + \delta\phi)$ is the total electric field. In this total- F formulation, the second-order nonlinear correction to the effective potential should be kept alongside the sloshing energy in order to ensure energy conservation.

In a related subject, namely, extension of nonlinear gyrokinetic formulations to edge turbulence, a different ordering is desirable due to the high relative fluctuation amplitude in low-confinement L -mode plasmas and the strong $E \times B$ flow shear in high-confinement (H -mode) plasmas (Hahm, Lin, *et al.*, 2004). Gyrokinetic simulations of edge turbulence began to appear recently (Scott, 2006). With a rigorous derivation additional terms (other than the radially dependent Doppler-shift-like term) appear in the gyrocenter equations of motion. Some of these terms are kept in comprehensive gyrokinetic sta-

bility analysis addressing the $E \times B$ shear effects (Re-woldt *et al.*, 1998; Peeters and Strintzi, 2004).

For nonlinear gyrokinetic simulations of turbulence much of the emphasis in the past decade has been concentrated on the study of zonal flows that are spontaneously generated by turbulence (Diamond *et al.*, 2005). The self-generated zonal flows are radially localized ($k_r L_F \gg 1$), axisymmetric ($k_\varphi = 0$), and mainly poloidal $E \times B$ flows.

There have been early indications from a fluid simulation (Hasegawa and Wakatani, 1987) that self-generated zonal flows can be important in drift-wave turbulence. In the 1990s, the importance of Reynolds's stress in zonal-flow generation was recognized (Diamond and Kim, 1991; Diamond *et al.*, 1993; Diamond, Liang, *et al.*, 1994) while nonlinear gyrofluid simulations (Dorland, 1993; Hammett *et al.*, 1993; Waltz *et al.*, 1994; Beer, 1995) have shown that zonal flows can regulate the ion-temperature-gradient turbulence and its associated transport. Based on nonlinear gyrokinetic simulations (Lin *et al.*, 1998; Dimits *et al.*, 2000) with a proper treatment of undamped zonal flows in collisionless toroidal geometry (Rosenbluth and Hinton, 1998), it is now widely recognized that understanding zonal-flow dynamics in regulating turbulence is essential in predicting transport in magnetically confined plasmas quantitatively. The important role of zonal flows has been recognized in nearly all cases and regimes of plasma turbulence so that the plasma microturbulence problem can be referred to as the “drift wave–zonal flow problem,” thereby emphasizing the two-component nature of the self-regulating system. Both nonlinear gyrokinetic simulations and theories have made essential contributions to this paradigm shift as recently reviewed (Diamond *et al.*, 2005; Itoh *et al.*, 2006), and have influenced experiments. For instance, characterization of the experimentally testable features of zonal-flow properties from nonlinear gyrokinetic simulations (Hahm *et al.*, 2000) have motivated some experimental measurements (see, for example, McKee *et al.*, 2003; Conway *et al.*, 2005).

One important effect of zonal flows on drift wave turbulence is the shearing of turbulent eddies. While the shearing due to mean $E \times B$ flow is well understood and pedagogical explanations are available, the complex spatiotemporal behavior of zonal flows introduces two important modifications. The first one is the time variation of zonal flows. High k_r components of zonal flows can vary on the eddy turnover time scale (Beer, 1995), unlike externally driven macroscopic $E \times B$ flows, which vary on a much slower time scale. It has been shown that fast time-varying components of zonal flows are less effective in shearing turbulence eddies (Hahm *et al.*, 1999).

The fundamental reason for this is that the zonal-flow shear pattern changes before the eddies can be completely torn apart. The turbulent eddies can then recover some of their original shape, and the shearing effect is reduced. This effect was first characterized via the “effective shearing rate” (Hamh *et al.*, 1999). Later, this trend was confirmed in particular turbulence models (Kim *et al.*, 2004). This is also the reason why the geo-

desic acoustic mode (GAM) (Winsor *et al.*, 1968), with $\omega_{\text{GAM}} \sim v_{\text{th}}/R$, does not reduce the ambient turbulence significantly for typical core parameters (Miyato *et al.*, 2004; Angelino, 2006). At the edge, sharp pressure gradients make the diamagnetic drift frequency at the relevant long wavelengths closer to the GAM frequency, i.e., $\omega_* / \omega_{\text{GAM}} \sim (k_\theta R) \rho_i / L_p \sim 1$. Therefore, the GAM can possibly affect the edge ambient turbulence (Hillatschek and Biskamp, 2001; Scott, 2003). The second reason is the chaotic pattern of the zonal flows. As a result, the shearing due to zonal flows is better characterized by a random diffraction derived from statistical approaches (Diamond *et al.*, 1998, 2001; Diamond, Lebedev, *et al.*, 1994) rather than the coherent stretching, which is applicable to the shearing due to the mean $E \times B$ shear. It has also been shown that the evolution of the mean $E \times B$ flow and that of the zonal flow can be quite different during transport-barrier formation (Kim and Diamond, 2002).

2. Bounce-center-kinetic Vlasov equation

As a second example of the extension of the nonlinear gyrokinetic formalism presented in the text, we derive the bounce-center-kinetic Vlasov equation. A conventional derivation with an emphasis on application to trapped-particle-driven turbulence can be found in Gang and Diamond (1990). We construct nonlinear Hamilton equations for charged particles in the presence of low-frequency electromagnetic fluctuations with characteristic mode frequency ω such that

$$\omega_d, \omega \ll \omega_b \ll \Omega, \quad (\text{D13})$$

where ω_b and ω_d denote the bounce and drift frequencies of a trapped guiding-center particle. This new time-scale ordering allows the removal of the fast gyration and bounce angles, i.e., the new reduced dynamics preserves the invariance of the magnetic moment μ and the bounce action $J=J_b$. While this represents a typical situation in magnetically confined plasmas, one can consider a more general case without a specific ordering between ω_b and Ω ; this interesting example has been studied previously by Dubin and Krommes (1982).

In deriving these reduced equations, we ignore finite-Larmor-radius effects associated with the electromagnetic field perturbations (i.e., we consider the long-wavelength limit $k_\perp^2 \rho_i^2 \ll 1$), and refrain from ordering the perpendicular and parallel wave numbers since k_\parallel/k_\perp may not be very small for some macroscopic instabilities.

In the presence of electromagnetic field fluctuations, the background magnetic field becomes perturbed. Depending on the characteristic time scales of the fluctuating fields, this situation typically may lead to the destruction of the guiding-center adiabatic invariants μ and/or \bar{J} . The electromagnetic field fluctuations are represented by the perturbed scalar potential $\delta\phi$, the parallel component of the perturbed vector potential $\delta\mathbf{A}_\parallel$ ($\equiv \hat{\mathbf{b}}_0 \cdot \delta\mathbf{A}$), and the parallel component of the perturbed magnetic

field δB_{\parallel} ($\equiv \hat{\mathbf{b}}_0 \cdot \nabla \times \delta \mathbf{A}$). We assume that the characteristic mode frequency ω is much smaller than the bounce frequency ω_b , i.e.,

$$\frac{\omega}{\omega_b} \sim \epsilon_{\omega}, \quad (\text{D14})$$

where ϵ_{ω} is a small ordering parameter; we henceforth set ϵ_d equal to 1 for clarity.

The perturbed guiding-center phase-space Lagrangian can be written as

$$\bar{\Gamma} = \bar{\Gamma}_0 + \epsilon_{\delta} \bar{\Gamma}_1, \quad (\text{D15})$$

where the first-order guiding-center phase-space Lagrangian is (Brizard, 1989a)

$$\bar{\Gamma}_1 \equiv \frac{e}{c} \left(\delta A_{\parallel bc} \frac{\partial s}{\partial \bar{u}^{\alpha}} \right) d\bar{u}^{\alpha}, \quad (\text{D16})$$

and the first-order guiding-center Hamiltonian is

$$\bar{H}_1 \equiv e \delta \phi_{gc} + \bar{\mu} \delta B_{\parallel bc}. \quad (\text{D17})$$

Dependence on the fast bounce-angle $\bar{\Theta}$ is reintroduced in $\bar{\Gamma}_n$ ($n \geq 1$) because the perturbation fields $(\delta \phi_{gc}, \delta A_{\parallel bc}, \delta B_{\parallel bc})$ depend on $\bar{\Theta}$ through $s(\bar{\mathbf{u}}) \equiv s(\mathbf{u})$ (to lowest order in ϵ_d) and $\mathbf{y} \equiv \bar{\mathbf{y}} + \Lambda_b$. For example, the perturbed scalar potential $\delta \phi_{gc}(\bar{\mathbf{y}}, \bar{\mathbf{u}})$ is defined as

$$\delta \phi_{gc}(\bar{\mathbf{y}}, \bar{\mathbf{u}}; t) \equiv \delta \phi(\bar{\mathbf{y}} + \Lambda_b, s(\bar{\mathbf{u}}); t). \quad (\text{D18})$$

In what follows, no assumptions about the orderings of the parallel and perpendicular wave numbers are made. In Eq. (D17), the gyrocenter magnetic moment

$$\bar{\mu} \equiv \mu + \epsilon_{\delta} \left[\frac{e \boldsymbol{\rho}}{B_0} \cdot \bar{\nabla} \left(\delta \phi_{gc} - \frac{v_{\parallel}}{c} \delta A_{\parallel bc} \right) - \mu \frac{\delta B_{\parallel bc}}{B_0} \right] + \dots \quad (\text{D19})$$

is an adiabatic invariant for the low-frequency nonlinear gyrocenter Hamiltonian dynamics (Brizard, 1989a), while μ is the (unperturbed) guiding-center magnetic moment and $\boldsymbol{\rho}$ is the gyroradius. The second-order gyrocenter Hamiltonian (in the limit $\rho^2 k_{\perp}^2 \ll 1$) is

$$\begin{aligned} \bar{H}_2 \equiv & -\frac{mc^2}{2B_0^2} \left| \bar{\nabla}_{\perp} \left(\delta \phi_{gc} - \frac{v_{\parallel}}{c} \delta A_{\parallel bc} \right) \right|^2 \\ & - e \delta \mathbf{A}_{\perp bc} \cdot \frac{\hat{\mathbf{b}}_0}{B_0} \times \bar{\nabla}_{\perp} \left(\delta \phi_{gc} - \frac{v_{\parallel}}{c} \delta A_{\parallel bc} \right), \end{aligned} \quad (\text{D20})$$

The new bounce-gyrocenter phase-space Lagrangian is chosen to be

$$\hat{\Gamma} \equiv \frac{e}{2c} \eta_{ab} \hat{y}^a d\hat{y}^b + \hat{J} d\hat{\Theta} - \hat{w} dt, \quad (\text{D21})$$

i.e., all electromagnetic perturbation effects have been transferred to the bounce-gyrocenter Hamiltonian

$$\hat{H}(\hat{\mathbf{y}}, t; \hat{J}) \equiv \hat{H}_0 + \epsilon_{\delta} \hat{H}_1 + \epsilon_{\delta}^2 \hat{H}_2, \quad (\text{D22})$$

where the second-order bounce-center Hamiltonian contains low-frequency ponderomotive terms associated with the asymptotic decoupling of the bounce-motion time scale. The first-order bounce-center Hamiltonian is

$$\hat{H}_1 \equiv e \langle \delta \psi_{bc} \rangle_b = \left\langle e \left(\delta \phi_{gc} - \frac{v_{\parallel}}{c} \delta A_{\parallel bc} \right) \right\rangle_b + \bar{\mu} \langle \delta B_{\parallel bc} \rangle_b, \quad (\text{D23})$$

where the bounce-angle averaging with respect to $\bar{\Theta}$ is denoted $\langle \cdot \rangle_b$. The second-order bounce-center Hamiltonian is

$$\hat{H}_2 \equiv \langle \bar{H}_2 \rangle_b + \frac{e^2}{2mc^2} \langle (\delta A_{\parallel bc})^2 \rangle_b - \frac{e^2}{2\omega_b} \langle \{ \delta \tilde{\Psi}_{bc}, \delta \tilde{\psi}_{bc} \}_{bc} \rangle_b, \quad (\text{D24})$$

where $\delta \tilde{\Psi}_{bc} \equiv \int \delta \tilde{\psi}_{bc} d\bar{\Theta}$ and $\{ \cdot, \cdot \}_{bc}$ denotes the unperturbed bounce-center Poisson bracket.

The nonlinear bounce-gyrocenter Hamiltonian is

$$\begin{aligned} \hat{H} \equiv & \hat{H}_0 + \epsilon_{\delta} \left\langle e \left(\delta \phi_{gc} - \frac{v_{\parallel}}{c} \delta A_{\parallel bc} \right) + \bar{\mu} \delta B_{\parallel bc} \right\rangle_b \\ & + \epsilon_{\delta}^2 \left[\langle \bar{H}_2 \rangle_b + \frac{e^2}{2mc^2} \langle (\delta A_{\parallel bc})^2 \rangle_b \right. \\ & \left. - \frac{e^2}{2\omega_b} \langle \{ \delta \tilde{\Psi}_{bc}, \delta \tilde{\psi}_{bc} \}_{bc} \rangle_b \right]. \end{aligned} \quad (\text{D25})$$

This expression generalizes the previous works of Gang and Diamond (1990) and Fong and Hahm (1999), who considered electrostatic perturbations only. The nonlinear bounce-gyrocenter Hamilton equations presented here contain terms associated with full electromagnetic perturbations and include classical ($\langle \bar{H}_2 \rangle_b$) and neoclassical ($\langle \{ \delta \tilde{\Psi}_{bc}, \delta \tilde{\psi}_{bc} \}_{bc} \rangle_b$) terms. The bounce-center-kinetic Vlasov equation for the distribution \hat{F} of bounce centers is expressed as in terms of the nonlinear bounce-gyrocenter Hamiltonian (D25) and the bounce-center Poisson bracket (B36) as

$$\frac{\partial \hat{F}}{\partial t} + \frac{c}{e} \frac{\partial \hat{F}}{\partial \hat{y}^a} \eta^{ab} \frac{\partial \hat{H}}{\partial \hat{y}^a} = 0, \quad (\text{D26})$$

where the trapped-particle $E \times B$ nonlinearity $[(d\delta\hat{y}^a/dt)\partial\delta\hat{F}/\partial\hat{y}^a]$ plays a crucial role in the nonlinear saturation of the trapped-electron mode (Gang et al., 1991; Hahm and Tang, 1991) and the trapped-ion mode.

The bounce-center phase-space transformation $(\bar{\mathbf{y}}, \bar{\mathbf{u}}) \rightarrow (\hat{\mathbf{y}}, \hat{\mathbf{u}})$ is defined up to first order in ϵ_{δ} as

$$\hat{y}^a = \bar{y}^a + \epsilon_{\delta} \{ \bar{S}_1, \bar{y}^a \}_{bc},$$

$$\hat{u}^{\alpha} = \bar{u}^{\alpha} + \epsilon_{\delta} \{ \bar{S}_1, \bar{u}^{\alpha} \}_{bc} + \epsilon_{\delta} (e/c) \delta A_{\parallel bc} \{ \bar{s}, \bar{u}^{\alpha} \}_{bc}, \quad (\text{D27})$$

where $\bar{S}_1 \equiv e \delta \tilde{\Psi}_{bc} / \omega_b$. The neoclassical polarization density can be defined in terms of the push-forward expres-

sion $\rho_{\text{neopol}} \equiv -\Sigma e \nabla \cdot \|\boldsymbol{\rho}_e\|$, where $\|\cdot\|$ denotes a momentum integration over the bounce-center distribution function and

$$\rho_e^a \equiv -\epsilon_\delta \{\bar{S}_1, \bar{y}^a + \Lambda_b^a\}_{bc}.$$

The neoclassical polarization density ρ_{neopol} accounts for the difference between the bounce-center density and the gyrocenter density (Fong and Hahm, 1999). This is related, via a continuity equation, to the neoclassical polarization current (Hinton and Robertson, 1984), which plays a crucial role in zonal-flow evolution (Rosenbluth and Hinton, 1998). By definition, the bounce-center moment $\|\rho_e^a\|$ involves a bounce-angle average and to lowest order in the bounce-kinetic ordering we find

$$\langle \rho_e^a \rangle_b = \epsilon_\delta \frac{e}{\omega_b} \frac{\partial}{\partial \hat{J}} \langle \delta \tilde{\psi}_{bc} \Lambda_b^a \rangle_b,$$

where Λ_b^a denotes the bounce radius defined in Eq. (B37). We see that each asymptotic decoupling of a fast time scale introduces a corresponding ponderomotive-like nonlinear term in the reduced Hamiltonian. These ponderomotive-like terms, in turn, are used to introduce polarization and magnetization effects into the reduced Maxwell's equations.

REFERENCES

- Abraham, R., and J. E. Marsden, 1978, *Foundations of Mechanics*, 2nd ed. (Benjamin/Cummings, Reading, MA).
- Angelino, P., A. Bottino, R. Hatzky, S. Jolliet, O. Sauter, T. M. Tran, and L. Villard, 2006, *Plasma Phys. Controlled Fusion* **48**, 557.
- Antonsen, T. M., and B. Lane, 1980, *Phys. Fluids* **23**, 1205.
- Arnold, V. I., 1989, *Mathematical Methods of Classical Mechanics*, 2nd ed. (Springer, New York).
- Artun, M., and W. M. Tang, 1994, *Phys. Plasmas* **1**, 2682.
- Beer, M. A., 1995, Ph.D. thesis (Princeton University).
- Bell, R. E., F. M. Levinton, S. H. Batha, E. J. Synakowski, and M. C. Zarnstorff, 1998, *Phys. Rev. Lett.* **81**, 1429.
- Bernstein, I. B., and P. J. Catto, 1985, *Phys. Fluids* **28**, 1342.
- Biglari, H., P. H. Diamond, and P. W. Terry, 1990, *Phys. Fluids B* **2**, 1.
- Boozer, A. H., 2004, *Rev. Mod. Phys.* **76**, 1071.
- Braginskii, S. I., 1965, in *Reviews of Plasma Physics*, edited by M. A. Leontovich (Consultants Bureau, New York), Vol. 1, pp. 205–311.
- Briguglio, S., L. Chen, J. Q. Dong, G. Fogaccia, R. A. Santoro, G. Vlad, and F. Zonca, 2000, *Nucl. Fusion* **40**, 701.
- Briguglio, S., F. Zonca, and G. Vlad, 1998, *Phys. Plasmas* **5**, 3287.
- Brizard, A. J., 1989a, *J. Plasma Phys.* **41**, 541.
- Brizard, A. J., 1989b, *Phys. Fluids B* **1**, 1381.
- Brizard, A. J., 1990, Ph.D. thesis (Princeton University).
- Brizard, A. J., 1992, *Phys. Fluids B* **4**, 1213.
- Brizard, A. J., 1994a, *Phys. Plasmas* **1**, 2460.
- Brizard, A. J., 1994b, *Phys. Plasmas* **1**, 2473.
- Brizard, A. J., 1995, *Phys. Plasmas* **2**, 459.
- Brizard, A. J., 2000a, *Phys. Rev. Lett.* **84**, 5768.
- Brizard, A. J., 2000b, *Phys. Plasmas* **7**, 4816.
- Brizard, A. J., 2000c, *Phys. Plasmas* **7**, 3238.
- Brizard, A. J., 2001, *Phys. Lett. A* **291**, 146.
- Brizard, A. J., 2004, *Phys. Plasmas* **11**, 4429.
- Brizard, A. J., 2005a, *J. Plasma Phys.* **71**, 225.
- Brizard, A. J., 2005b, *Phys. Plasmas* **12**, 092302.
- Brizard, A. J., and A. A. Chan, 1999, *Phys. Plasmas* **6**, 4548.
- Burrell, K. H., 1997, *Phys. Plasmas* **4**, 1499.
- Callen, J. D., and K. C. Shaing, 1985, *Phys. Fluids* **28**, 1845.
- Candy, J., and R. E. Waltz, 2006, *21st IAEA Fusion Energy Conference* (International Atomic Energy Agency, Vienna).
- Carreras, B. A., D. E. Newman, P. H. Diamond, and Y. M. Liang, 1994, *Phys. Plasmas* **1**, 4014.
- Cary, J. R., 1977, *J. Math. Phys.* **18**, 2432.
- Cary, J. R., and A. N. Kaufman, 1981, *Phys. Fluids* **24**, 1238.
- Cary, J. R., and R. G. Littlejohn, 1983, *Ann. Phys. (N.Y.)* **151**, 1.
- Catto, P. J., W. M. Tang, and D. E. Baldwin, 1981, *Phys. Fluids* **23**, 639.
- Chang, C. S., and F. L. Hinton, 1982, *Phys. Fluids* **25**, 1493.
- Chen, L., 1994, *Phys. Plasmas* **1**, 1519.
- Chen, L., 1999, *J. Geophys. Res.* **104**, 2421.
- Chen, L., and A. Hasegawa, 1994, *J. Geophys. Res.* **99**, 179.
- Chen, L., Z. Lin, and R. B. White, 2000, *Phys. Plasmas* **7**, 3129.
- Chen, L., Z. Lin, R. B. White, and F. Zonca, 2001, *Nucl. Fusion* **41**, 747.
- Chen, L., R. B. White, and F. Zonca, 2004, *Phys. Rev. Lett.* **92**, 075004.
- Chen, L., F. Zonca, and Z. Lin, 2005, *Plasma Phys. Controlled Fusion* **47**, B71.
- Chen, Y., and S. E. Parker, 2001, *Phys. Plasmas* **8**, 2095.
- Chen, Y., S. E. Parker, B. I. Cohen, A. M. Dimits, W. M. Nevins, D. Shumaker, V. K. Decyk, and J. N. Leboeuf, 2003, *Nucl. Fusion* **43**, 1121.
- Connor, J. W., and L. Chen, 1985, *Phys. Fluids* **28**, 2201.
- Connor, J. W., S. C. Cowley, R. J. Hastie, and L. R. Pan, 1987, *Plasma Phys. Controlled Fusion* **29**, 919.
- Connor, J. W., and H. R. Wilson, 1994, *Plasma Phys. Controlled Fusion* **36**, 719.
- Connor, J. W., and H. R. Wilson, 2000, *Plasma Phys. Controlled Fusion* **42**, R1.
- Conway, G. D., B. D. Scott, J. Schirmer, M. Reich, A. Kendl, and the ASDEX Upgrade Team, 2005, *Plasma Phys. Controlled Fusion* **47**, 1165.
- Cromb  , K., Y. Andrew, M. Brix, C. Giroud, S. Hacquin, N. C. Hawkes, A. Murari, M. F. F. Nave, J. Ongena, V. Parail, G. Van Oost, I. Voitsekhovitch, and K. D. Zastrow, 2005, *Phys. Rev. Lett.* **95**, 155003.
- Cummings, J. C., 1995, Ph.D. thesis (Princeton University).
- Dannert, T., and F. Jenko, 2005, *Phys. Plasmas* **12**, 072309.
- Depret, G., X. Garbet, P. Bertrand, and A. Ghizzo, 2000, *Plasma Phys. Controlled Fusion* **42**, 949.
- Diamond, P. H., S. Champeaux, M. Malkov, A. Das, I. Gruzinov, M. N. Rosenbluth, C. Holland, B. Wecht, A. I. Smolyakov, F. L. Hinton, Z. Lin, and T. S. Hahm, 2001, *Nucl. Fusion* **41**, 1067.
- Diamond, P. H., and T. S. Hahm, 1995, *Phys. Plasmas* **2**, 3640.
- Diamond, P. H., S. I. Itoh, K. Itoh, and T. S. Hahm, 2005, *Plasma Phys. Controlled Fusion* **47**, R35.
- Diamond, P. H., and Y.-B. Kim, 1991, *Phys. Fluids B* **3**, 1626.
- Diamond, P. H., V. B. Lebedev, Y. M. Liang, A. V. Gruzinov, I. Gruzinova, M. Medvedev, B. A. Carreras, D. E. Newman, L. Charlton, K. L. Sidikman, D. B. Batchelor, E. F. Jaeger, C. Y. Wang, G. G. Craddock, N. Mattor, T. S. Hahm, M. Ono, B. Leblanc, H. Biglari, F. Y. Gang, and D. J. Sigmar, 1994, *15th*

- IAEA Fusion Energy Conference* (International Atomic Energy Agency, Vienna), Vol. 3, pp. 323–353.
- Diamond, P. H., V. B. Lebedev, D. E. Newman, B. A. Carreras, T. S. Hahm, W. M. Tang, G. Rewoldt, and K. Avinash, 1997, Phys. Rev. Lett. **78**, 1472.
- Diamond, P. H., Y. M. Liang, B. A. Carreras, and P. W. Terry, 1994, Phys. Rev. Lett. **72**, 2565.
- Diamond, P. H., M. N. Rosenbluth, F. L. Hinton, M. Malkov, J. Fleischer, and A. Smolyakov, 1998, *17th IAEA Fusion Energy Conference* (International Atomic Energy Agency, Vienna), Vol. 4, pp. 1421–1428.
- Diamond, P. H., V. Shapiro, V. Shevchenko, Y. B. Kim, M. N. Rosenbluth, B. A. Carreras, K. L. Sidikman, V. E. Lynch, L. Garcia, P. W. Terry, and R. Z. Sagdeev, 1992, *14th IAEA Fusion Energy Conference* (International Atomic Energy Agency, Vienna), Vol. 2, pp. 97–113.
- 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.
- Dimits, A. M., L. L. Lodestro, and D. H. E. Dubin, 1992, Phys. Fluids B **4**, 274.
- Dimits, A. M., T. J. Williams, J. A. Byers, and B. I. Cohen, 1996, Phys. Rev. Lett. **77**, 71.
- Dimotakis, P. E., 2000, J. Fluid Mech. **409**, 69.
- Dorland, W., 1993, Ph.D. thesis (Princeton University).
- Dorland, W., and G. W. Hammett, 1993, Phys. Fluids B **5**, 812.
- Dorland, W., F. Jenko, M. Kotschenreuther, and B. N. Rogers, 2000, Phys. Rev. Lett. **85**, 5579.
- Dubin, D. H. E., and J. A. Krommes, 1982, *Long Time Prediction in Dynamics* (Wiley, New York), pp. 251–280.
- Dubin, D. H. E., J. A. Krommes, C. Oberman, and W. W. Lee, 1983, Phys. Fluids **26**, 3524.
- Dupree, T. H., 1972, Phys. Fluids **15**, 334.
- Dupree, T. H., 1982, Phys. Fluids **25**, 277.
- Endler, M., H. Niedermeyer, L. Giannone, E. Kolzhauer, A. Rudyj, G. Theimer, and N. Tsois, 1995, Nucl. Fusion **35**, 1307.
- Ernst, D. R., P. T. Bonoli, P. J. Catto, W. Dorland, C. L. Fiore, R. S. Granetz, M. Greenwald, A. E. Hubbard, M. Porkolab, M. H. Redi, J. E. Rice, K. Zhurovich, and the Alcator C-Mod Group, 2004, Phys. Plasmas **11**, 2637.
- Flanders, H., 1989, *Differential Forms with Applications to the Physical Sciences* (Dover, New York).
- Fonck, R. J., G. Cosby, R. D. Durst, S. F. Paul, N. Bretz, S. Scott, E. Synakowski, and G. Taylor, 1993, Phys. Rev. Lett. **70**, 3736.
- Fong, B. H., and T. S. Hahm, 1999, Phys. Plasmas **6**, 188.
- Frieman, E. A., and L. Chen, 1982, Phys. Fluids **25**, 502.
- Frisch, U., 1995, *Turbulence* (Cambridge University Press, Cambridge).
- Fu, G. Y., and W. Park, 1995, Phys. Rev. Lett. **74**, 1594.
- Gang, F. Y., and P. H. Diamond, 1990, Phys. Fluids B **2**, 2976.
- Gang, F. Y., P. H. Diamond, and M. N. Rosenbluth, 1991, Phys. Fluids B **3**, 68.
- Garbet, X., L. Laurent, A. Samain, and J. Chinardet, 1994, Nucl. Fusion **34**, 963.
- Garbet, X., and R. E. Waltz, 1998, Phys. Plasmas **5**, 2836.
- Goldstein, H., C. Poole, and J. Safko, 2002, *Classical Mechanics*, 3rd ed. (Addison-Wesley, San Francisco).
- Gurcan, O. D., P. H. Diamond, and T. S. Hahm, 2006a, Phys. Plasmas **13**, 052306.
- Gurcan, O. D., P. H. Diamond, and T. S. Hahm, 2006b, Phys. Rev. Lett. **97**, 024502.
- Gurcan, O. D., P. H. Diamond, T. S. Hahm, and Z. Lin, 2005, Phys. Plasmas **12**, 032303.
- Hagan, W. K., and E. A. Frieman, 1985, Phys. Fluids **28**, 2641.
- Hahm, T. S., 1988, Phys. Fluids **31**, 2670.
- Hahm, T. S., 1992, Phys. Fluids B **4**, 2801.
- Hahm, T. S., 1994, Phys. Plasmas **1**, 2940.
- Hahm, T. S., 1996, Phys. Plasmas **3**, 4658.
- Hahm, T. S., 2002, Plasma Phys. Controlled Fusion **44**, A87.
- Hahm, T. S., M. A. Beer, Z. Lin, G. W. Hammett, W. W. Lee, and W. M. Tang, 1999, Phys. Plasmas **6**, 922.
- Hahm, T. S., and K. H. Burrell, 1995, Phys. Plasmas **2**, 1648.
- Hahm, T. S., and K. H. Burrell, 1996, Plasma Phys. Controlled Fusion **38**, 1427.
- Hahm, T. S., K. H. Burrell, Z. Lin, R. Nazikian, and E. J. Synakowski, 2000, Plasma Phys. Controlled Fusion **42**, A205.
- Hahm, T. S., P. H. Diamond, Z. Lin, K. Itoh, and S. I. Itoh, 2004, Plasma Phys. Controlled Fusion **46**, A323.
- Hahm, T. S., P. H. Diamond, Z. Lin, G. Rewoldt, O. D. Gurcan, and S. Ethier, 2005, Phys. Plasmas **12**, 090903.
- Hahm, T. S., W. W. Lee, and A. J. Brizard, 1988, Phys. Fluids **31**, 1940.
- Hahm, T. S., Z. Lin, P. H. Diamond, G. Rewoldt, W. X. Wang, S. Ethier, O. Gurcan, W. W. Lee, and W. M. Tang, 2004, *20th IAEA Fusion Energy Conference* (International Atomic Energy Agency, Vienna), IAEA-CN-116/TH/1-4.
- Hahm, T. S., and W. M. Tang, 1990, Phys. Fluids B **2**, 1815.
- Hahm, T. S., and W. M. Tang, 1991, Phys. Fluids B **3**, 989.
- Hahm, T. S., and W. M. Tang, 1996, Phys. Plasmas **3**, 242.
- Hallatschek, K., and D. Biskamp, 2001, Phys. Rev. Lett. **86**, 1223.
- Hammett, G. W., 2001, private communication.
- Hammett, G. W., M. A. Beer, W. Dorland, S. C. Cowley, and S. A. Smith, 1993, Plasma Phys. Controlled Fusion **35**, 973.
- Hasegawa, A., and K. Mima, 1978, Phys. Fluids **21**, 87.
- Hasegawa, A., and M. Wakatani, 1983, Phys. Fluids **26**, 2770.
- Hasegawa, A., and M. Wakatani, 1987, Phys. Rev. Lett. **59**, 1581.
- Hastie, R. J., J. B. Taylor, and F. A. Haas, 1967, Ann. Phys. (N.Y.) **41**, 302.
- Hatzky, R., T. M. Tran, A. Koenies, R. Kleiber, and S. J. Allfrey, 2002, Phys. Plasmas **9**, 898.
- Hazeltine, R. D., and F. L. Hinton, 2005, Phys. Plasmas **12**, 102506.
- Hazeltine, R. D., M. Kotschenreuther, and P. J. Morrison, 1985, Phys. Fluids **28**, 2466.
- Hinton, F. L., and R. D. Hazeltine, 1976, Rev. Mod. Phys. **48**, 239.
- Hinton, F. L., J. Kim, Y. B. Kim, A. Brizard, and K. H. Burrell, 1994, Phys. Rev. Lett. **72**, 1216.
- Hinton, F. L., and J. A. Robertson, 1984, Phys. Fluids **27**, 1243.
- Hinton, F. L., and S. K. Wong, 1985, Phys. Fluids **28**, 3082.
- Holland, C., and P. H. Diamond, 2004, Phys. Plasmas **11**, 1043.
- Horton, W., 1999, Rev. Mod. Phys. **71**, 735.
- Horton, W., P. Zhu, G. T. Hoang, T. Aniel, M. Ottaviani, and X. Garbet, 2000, Phys. Plasmas **7**, 1494.
- Howes, G. G., S. C. Cowley, W. Dorland, G. W. Hammett, E. Quataert, and A. A. Schekochihin, 2006, Astrophys. J. **651**, 590.
- Idomura, Y., 2006, Phys. Plasmas **13**, 080701.
- Idomura, Y., M. Wakatani, and S. Tokuda, 2000, Phys. Plasmas **7**, 3551.

- Itoh, S. I., and K. Itoh, 2001, *Plasma Phys. Controlled Fusion* **43**, 1055.
- Itoh, K., S. I. Itoh, P. H. Diamond, T. S. Hahm, A. Fujisawa, G. R. Tynan, M. Yagi, and Y. Nagashima, 2006, *Phys. Plasmas* **13**, 055502.
- Itoh, K., S. I. Itoh, T. S. Hahm, and P. H. Diamond, 2005, *J. Phys. Soc. Jpn.* **74**, 2001.
- Jackson, J. D., 1975, *Classical Electrodynamics*, 2nd ed. (Wiley, New York).
- Jenko, F., W. Dorland, M. Kotschenreuther, and B. N. Rogers, 2000, *Phys. Plasmas* **7**, 1904.
- Jenko, F., and B. D. Scott, 1999, *Phys. Plasmas* **6**, 2705.
- Kaufman, A. N., 1986, *Phys. Fluids* **29**, 1736.
- Kim, E. J., and P. H. Diamond, 2002, *Phys. Plasmas* **9**, 71.
- Kim, E. J., P. H. Diamond, and T. S. Hahm, 2004, *Phys. Plasmas* **11**, 4554.
- Kim, E. J., P. H. Diamond, M. Malkov, T. S. Hahm, K. Itoh, S. I. Itoh, S. Champeaux, I. Gruzinov, O. Gurcan, C. Holland, M. N. Rosenbluth, and A. Smolyakov, 2003, *Nucl. Fusion* **43**, 961.
- Kim, E. J., C. Holland, and P. H. Diamond, 2003, *Phys. Rev. Lett.* **91**, 075003.
- Kniep, J. C., J. N. Leboeuf, and V. K. Decyk, 2003, *Comput. Phys. Commun.* **164**, 98.
- Krommes, J. A., 1993a, *Phys. Rev. Lett.* **70**, 3067.
- Krommes, J. A., 1993b, *Phys. Fluids B* **5**, 2405.
- Krommes, J. A., 2002, *Phys. Rep.* **360**, 1.
- Krommes, J. A., 2006, in *Turbulence and Coherent Structures in Fluids, Plasmas and Nonlinear Media*, edited by M. Shats and H. Punzmann (World Scientific, Singapore).
- Krommes, J. A., and C. B. Kim, 1988, *Phys. Fluids* **31**, 869.
- Krommes, J. A., W. W. Lee, and C. Oberman, 1986, *Phys. Fluids* **29**, 2421.
- Krommes, J. A., C. Oberman, and R. G. Kleva, 1983, *J. Plasma Phys.* **30**, 11.
- Kruskal, M., 1962, *J. Math. Phys.* **3**, 806.
- Kulsrud, R. M., 1983, in *Handbook of Plasma Physics*, edited by M. N. Rosenbluth and R. Z. Sagdeev (North-Holland, New York), Vol. I, pp. 115–145.
- Lashmore-Davies, C. N., and R. O. Dendy, 1989, *Phys. Fluids B* **1**, 1565.
- Lebedev, V. B., and P. H. Diamond, 1997, *Phys. Plasmas* **4**, 1087.
- Lee, W. W., 1983, *Phys. Fluids* **26**, 556.
- Lee, W. W., 1987, *J. Comput. Phys.* **72**, 243.
- Lee, W. W., J. A. Krommes, C. R. Oberman, and R. A. Smith, 1984, *Phys. Fluids* **27**, 2652.
- Lee, W. W., J. L. V. Lewandowski, T. S. Hahm, and Z. Lin, 2001, *Phys. Plasmas* **8**, 4435.
- Lee, W. W., and W. M. Tang, 1988, *Phys. Fluids* **31**, 612.
- Lee, X. S., J. R. Myra, and P. J. Catto, 1983, *Phys. Fluids* **26**, 223.
- Lichtenberg, A. J., and M. A. Lieberman, 1984, *Regular and Stochastic Motion* (Springer, New York).
- Liewer, P. C., 1985, *Nucl. Fusion* **25**, 543.
- Lin, Y., X. Wang, Z. Lin, and L. Chen, 2005, *Plasma Phys. Controlled Fusion* **47**, 657.
- Lin, Z., L. Chen, and F. Zonca, 2005, *Phys. Plasmas* **12**, 056125.
- Lin, Z., S. Ethier, T. S. Hahm, and W. M. Tang, 2002, *Phys. Rev. Lett.* **88**, 195004.
- Lin, Z., and T. S. Hahm, 2004, *Phys. Plasmas* **11**, 1099.
- Lin, Z., T. S. Hahm, W. W. Lee, W. M. Tang, and P. H. Diamond, 1999, *Phys. Rev. Lett.* **83**, 3645.
- Lin, Z., T. S. Hahm, W. W. Lee, W. M. Tang, and R. B. White, 1998, *Science* **281**, 1835.
- 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* **2/I**, 119.
- Littlejohn, R. G., 1983, *J. Plasma Phys.* **29**, 111.
- Littlejohn, R. G., 1988, *Phys. Rev. A* **38**, 6034.
- Matsumoto, T., H. Naitou, S. Tokuda, and Y. Kishimoto, 2005, *Nucl. Fusion* **45**, 1264.
- Matsumoto, T., S. Tokuda, Y. Kishimoto, and H. Naitou, 2003, *Phys. Plasmas* **10**, 195.
- Mattor, N., 1992, *Phys. Fluids B* **4**, 3952.
- Mattor, N., and P. H. Diamond, 1989, *Phys. Fluids B* **1**, 1980.
- Mazzucato, E., 1982, *Phys. Rev. Lett.* **48**, 1828.
- Mazzucato, E., S. H. Batha, M. Beer, M. Bell, R. E. Bell, R. V. Budny, C. Bush, T. S. Hahm, G. W. Hammett, F. M. Levinton, R. Nazikian, H. Park, G. Rewoldt, G. L. Schmidt, E. J. Synkowski, W. M. Tang, G. Taylor, and M. C. Zarnstorff, 1996, *Phys. Rev. Lett.* **77**, 3145.
- McKee, G. R., R. J. Fonck, M. Jakubowski, K. H. Burrell, K. Hallatschek, R. A. Moyer, D. L. Rudakov, W. Nevins, G. D. Porter, P. Schoch, and X. Xu, 2003, *Phys. Plasmas* **10**, 1712.
- Mishchenko, A., R. Hatzky, and A. Könies, 2004, *Phys. Plasmas* **11**, 5480.
- Miyato, N., Y. Kishimoto, and J. Li, 2004, *Phys. Plasmas* **11**, 5557.
- Naitou, H., K. Tsuda, W. W. Lee, and R. D. Sydora, 1995, *Phys. Plasmas* **2**, 4257.
- Naulin, V., A. H. Neilsen, and J. J. Rasmussen, 2005, *Phys. Plasmas* **12**, 122306.
- Naulin, V., J. Nycander, and J. J. Rasmussen, 1998, *Phys. Rev. Lett.* **81**, 4148.
- Nazikian, R., K. Shinohara, G. J. Kramer, E. Valeo, K. Hill, T. S. Hahm, G. Rewoldt, S. Ide, Y. Koide, Y. Oyama, H. Shirai, and W. Tang, 2005, *Phys. Rev. Lett.* **94**, 135002.
- Newman, D. E., B. A. Carreras, P. H. Diamond, and T. S. Hahm, 1996, *Phys. Plasmas* **3**, 1858.
- Northrop, T. G., 1963, *Adiabatic Motion of Charged Particles* (Wiley, New York).
- Northrop, T. G., and E. Teller, 1960, *Phys. Rev.* **117**, 215.
- Ono, M., et al., 2003, *Plasma Phys. Controlled Fusion* **45**, A335.
- Otsuka, F., and T. Hada, 2003, *Space Sci. Rev.* **107**, 499.
- Park, W., S. E. Parker, H. Biglari, M. Chance, L. Chen, C. Z. Cheng, T. S. Hahm, W. W. Lee, R. Kulsrud, D. Monticello, L. Sugiyama, and R. White, 1992, *Phys. Fluids A* **4**, 2033.
- Parker, S. E., Y. Chen, W. Wan, B. I. Cohen, and W. M. Nevins, 2004, *Phys. Plasmas* **11**, 2594.
- Parker, S. E., and W. W. Lee, 1993, *Phys. Fluids B* **5**, 77.
- Parker, S. E., W. W. Lee, and R. A. Santoro, 1993, *Phys. Rev. Lett.* **71**, 2042.
- Peeters, A. G., and D. Strintzi, 2004, *Phys. Plasmas* **11**, 3748.
- Politzer, P. A., M. E. Austin, M. Gilmore, G. R. McKee, T. L. Rhodes, C. X. Yu, E. J. Doyle, T. E. Evans, and R. A. Moyer, 2002, *Phys. Plasmas* **9**, 1962.
- Qin, H., and W. M. Tang, 2004, *Phys. Plasmas* **11**, 1052.
- Qin, H., W. M. Tang, and W. W. Lee, 2000, *Phys. Plasmas* **7**, 4433.
- Qin, H., W. M. Tang, W. W. Lee, and G. Rewoldt, 1999, *Phys. Plasmas* **6**, 1575.
- Rechester, A. B., and M. N. Rosenbluth, 1978, *Phys. Rev. Lett.* **40**, 38.
- Rewoldt, G., M. A. Beer, M. S. Chance, T. S. Hahm, Z. Lin,

- and W. M. Tang, 1998, Phys. Plasmas **5**, 1815.
- Roach, C. M., D. J. Applegate, J. W. Connor, S. C. Cowley, W. D. Dorland, R. J. Hastie, N. Joiner, S. Saarelma, A. A. Schekochihin, R. J. Akers, C. Brickley, A. R. Field, M. Valovic, and the MAST Team, 2005, Plasma Phys. Controlled Fusion **47**, B323.
- Rosenbluth, M. N., R. D. Hazeltine, and F. L. Hinton, 1972, Phys. Fluids **15**, 116.
- Rosenbluth, M. N., and F. L. Hinton, 1998, Phys. Rev. Lett. **80**, 724.
- Rutherford, P. H., and E. A. Frieman, 1968, Phys. Fluids **11**, 569.
- Sagdeev, R. Z., and A. A. Galeev, 1969, *Nonlinear Plasma Theory* (Benjamin, New York).
- Santoro, R. A., and L. Chen, 1996, Phys. Plasmas **6**, 2349.
- Sarazin, Y., and Ph. Ghendrih, 1998, Phys. Plasmas **5**, 4214.
- Sarazin, Y., V. Grandgirard, E. Fleurence, X. Garbet, P. Ghendrih, P. Bertrand, and G. Depret, 2005, Plasma Phys. Controlled Fusion **47**, 1817.
- Schulz, M., and L. J. Lanzerotti, 1974, *Particle Diffusion in the Radiation Belts* (Springer, New York).
- Scott, B. D., 1990, Phys. Rev. Lett. **65**, 3289.
- Scott, B. D., 1997, Plasma Phys. Controlled Fusion **39**, 1635.
- Scott, B. D., 2003, Phys. Lett. A **320**, 53.
- Scott, B. D., 2005, New J. Phys. **7**, 1.
- Scott, B. D., 2006, Plasma Phys. Controlled Fusion **48**, A387.
- Similon, P. L., and P. H. Diamond, 1984, Phys. Fluids **27**, 916.
- Smith, R. A., J. A. Krommes, and W. W. Lee, 1985, Phys. Fluids **28**, 1069.
- Snyder, P. B., and G. W. Hammett, 2001, Phys. Plasmas **8**, 744.
- Sosenko, P. P., P. Bertrand, and V. K. Decyk, 2001, Phys. Scr. **64**, 264.
- Stern, D. P., 1970, Am. J. Phys. **38**, 494.
- Strauss, H. R., 1976, Phys. Fluids **19**, 134.
- Strauss, H. R., 1977, Phys. Fluids **20**, 1354.
- Strintzi, D., and B. D. Scott, 2004, Phys. Plasmas **11**, 5452.
- Strintzi, D., B. D. Scott, and A. J. Brizard, 2005, Phys. Plasmas **12**, 052517.
- Sugama, H., 2000, Phys. Plasmas **7**, 466.
- Sugama, H., M. Okamoto, W. Horton, and M. Wakatani, 1996, Phys. Plasmas **3**, 2379.
- Sugama, H., T. H. Watanabe, and W. Horton, 2001, Phys. Plasmas **8**, 2617.
- Sydora, R. D., 1990, Phys. Fluids B **2**, 1455.
- Sydora, R. D., 2001, Phys. Plasmas **8**, 1929.
- Sydora, R. D., V. K. Decyk, and J. M. Dawson, 1996, Plasma Phys. Controlled Fusion **12**, A281.
- Sydora, R. D., T. S. Hahm, W. W. Lee, and J. M. Dawson, 1990, Phys. Rev. Lett. **64**, 2015.
- Sykes, A., J. W. Ahn, R. J. Akers, E. Arends, P. G. Carolan, G. F. Counsell, S. J. Fielding, M. Gryaznevich, R. Martin, M. Price, C. Roach, V. Shevchenko, M. Tournianski, M. Valovic, M. J. Walsh, and H. R. Wilson, 2001, Phys. Plasmas **8**, 2101.
- Synakowski, E. J., S. H. Batha, M. Beer, M. G. Bell, R. E. Bell, R. V. Budny, C. E. Bush, P. C. Efthimion, T. S. Hahm, G. W. Hammett, B. Leblanc, F. Levinton, E. Mazzucato, H. Park, A. T. Ramsey, G. Schmidt, G. Rewoldt, S. D. Scott, G. Taylor, and M. C. Zarnstorff, 1997, Phys. Plasmas **4**, 1736.
- Tang, W. M., 1978, Nucl. Fusion **18**, 1089.
- Tang, W. M., 2002, Phys. Plasmas **9**, 1856.
- Tang, W. M., and V. S. Chan, 2005, Plasma Phys. Controlled Fusion **47**, R1.
- Tang, W. M., J. W. Connor, and R. J. Hastie, 1980, Nucl. Fusion **20**, 1439.
- Taylor, J. B., 1967, Phys. Fluids **10**, 1357.
- Terry, P. W., 2000, Rev. Mod. Phys. **72**, 109.
- Terry, P. W., P. H. Diamond, and T. S. Hahm, 1990, Phys. Fluids B **2**, 2048.
- Todo, Y., H. L. Berk, and B. N. Breizman, 2003, Phys. Plasmas **10**, 2888.
- Tsai, S. T., J. W. Van Dam, and L. Chen, 1984, Plasma Phys. Controlled Fusion **26**, 907.
- Villard, L., S. J. Allfrey, A. Bottino, M. Brunetti, G. L. Falchetto, P. Angelino, V. Grandgirard, R. Hatzky, J. Nuhrenberg, A. G. Peeters, O. Sauter, S. Sorge, and J. Vaclavik, 2004, Nucl. Fusion **44**, 172.
- Villard, L., P. Angelino, A. Bottino, S. J. Allfrey, R. Hatzky, Y. Idomura, O. Sauter, M. Brunetti, and T. M. Tran, 2004, Plasma Phys. Controlled Fusion **46**, B51.
- Vlad, G., F. Zonca, and S. Briguglio, 1999, Riv. Nuovo Cimento **22**, 1.
- Wagner, F., et al., 1982, Phys. Rev. Lett. **49**, 1408.
- Waltz, R. E., and J. M. Candy, 2005, Phys. Plasmas **12**, 072303.
- Waltz, R. E., J. M. Candy, and M. N. Rosenbluth, 2002, Phys. Plasmas **9**, 1938.
- Waltz, R. E., G. D. Kerbel, and J. Milovich, 1994, Phys. Plasmas **1**, 2229.
- Wang, W. X., T. S. Hahm, G. Rewoldt, J. Manickam, and W. M. Tang, 2006, *21st IAEA Fusion Energy Conference* (International Atomic Energy Agency, Vienna).
- Watanabe, T. H., and H. Sugama, 2006, Nucl. Fusion **46**, 24.
- Winsor, N., J. L. Johnson, and J. M. Dawson, 1968, Phys. Fluids **11**, 2448.
- Wong, K. L., N. Bretz, T. S. Hahm, and E. J. Synakowski, 1997, Phys. Lett. A **236**, 339.
- Wootton, A. J., B. A. Carreras, H. Matsumoto, K. McGuire, W. A. Peebles, Ch. P. Ritz, P. W. Terry, and S. J. Zweben, 1990, Phys. Fluids B **2**, 2879.
- Yagi, M., and W. Horton, 1994, Phys. Plasmas **1**, 2135.
- Yagi, M., S. I. Itoh, A. Fukuyama, and M. Azumi, 1995, Phys. Plasmas **2**, 4140.
- Yagi, M., T. Ueda, S. I. Itoh, M. A. Azumi, K. Itoh, P. H. Diamond, and T. S. Hahm, 2006, Plasma Phys. Controlled Fusion **48**, A409.
- Yang, S. C., and D. Choi, 1985, Phys. Lett. **108A**, 25.
- Zonca, F., S. Briguglio, L. Chen, S. Detrick, G. Fogaccia, D. Testa, and G. Vlad, 2002, Phys. Plasmas **9**, 4939.
- Zonca, F., S. Briguglio, L. Chen, G. Fogaccia, and G. Vlad, 2005, Nucl. Fusion **45**, 477.
- Zonca, F., R. B. White, and L. Chen, 2004, Phys. Plasmas **11**, 2488.
- Zweben, S. J., and S. S. Medley, 1989, Phys. Fluids B **1**, 2058.