Princeton Plasma Physics Laboratory

PPPL-4153 PPPL-4153

Foundations of Nonlinear Gyrokinetic Theory

A.J. Brizard and T.S. Hahm

February 2006

PPPL
PRINCETON PLASMA
PHYSICS LABORATORY

Prepared for the U.S. Department of Energy under Contract DE-AC02-76CH03073.

February 21, 2006

Foundations of Nonlinear Gyrokinetic Theory

A. J. Brizard*

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

T. S. Hahm

Princeton University, Plasma Physics Laboratory, Princeton, NJ 08543, USA

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 important pillars: (1) a gyrokinetic Vlasov equation written in terms of a gyrocenter Hamiltonian with quadratic low-frequency ponderomotive-like terms; (2) a set of gyrokinetic Maxwell equations written in terms of the gyrocenter Vlasov distribution that contain low-frequency polarization and magnetization terms (derived from the quadratic nonlinearities in the Hamiltonian); and (3) 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 the rigorous applications of Lagrangian and Hamiltonian methods used in the variational derivation of nonlinear gyrokinetic Vlasov-Maxwell equations. The physical motivations and applications of the nonlinear gyrokinetic equations, which describe the turbulent evolution of low-frequency electromagnetic fluctuations in a nonuniform magnetized plasmas with arbitrary magnetic geometry, are also discussed.

Contents

*Electronic address: abrizard@smcvt.edu

I. INTRODUCTION

Magnetically-confined plasmas, found either in fusion-research devices or nature, exhibit a wide range of spatial and temporal scales. Consider, for example, the behavior of single charged particles in a plasma confined by a strong magnetic field (Northrop, 1963), represented in divergenceless form as

$$\mathbf{B} \;=\; \nabla\alpha \times \nabla\beta \;\equiv\; B(\alpha,\beta,s)\,\widehat{\mathsf{b}}(\alpha,\beta,s)$$

where the Euler potentials (α,β) are time-dependent magnetic-line labels, s denotes the parallel coordinate along a field line, and $\widehat{\mathsf{b}} \equiv \partial\mathbf{x}/\partial s$ denotes the unit vector along the magnetic-field line. Here, the primary terms associated with magnetic-field inhomogeneity are represented by the parallel gradient $\widehat{\mathsf{b}}\cdot\nabla\ln B$ and the perpendicular gradient $\widehat{\mathsf{b}}\times\nabla\ln B$ as well as the magnetic curvature $\widehat{\mathsf{b}}\cdot\nabla\widehat{\mathsf{b}}$.[1] Charged particles confined in a strong magnetic field with weak inhomogeneity execute three types of quasi-periodic motion (Kruskal, 1962): (1) a rapid gyro-motion about a single magnetic field line, with a short gyroperiod $\tau_{\rm g}(\mathbf{x})$ that depends on the particle's spatial position $\mathbf{x} = (\alpha,\beta,s)$; (2) an intermediate bounce (or transit) motion along a magnetic field line (driven by the parallel gradient), with an intermediate characteristic time scale $\tau_{\rm b}(\mathbf{y};\mathcal{E},J_{\rm g})$ that depends on the particle's energy $\mathcal{E}$ and its gyro-action $J_{\rm g} \equiv (mc/e)\,\mu$ (where μ denotes the magnetic moment) as well as the field-line labels $\mathbf{y} \equiv (\alpha,\beta)$; and (3) 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 $\tau_{\rm d}(\mathcal{E},J_{\rm g},J_{\rm b})$ that depends on the bounce-action $J_{\rm b}$ as well as the energy $\mathcal{E}$ and the gyro-action $J_{\rm g}$. In general, the orbital time scales are well separated for charged particles magnetically-confined in a strong magnetic field with weak inhomogeneity: $\tau_{\rm g} \ll \tau_{\rm b} \ll \tau_{\rm d}$. For example, the orbital time scales of a 10-keV proton equatorially mirroring at geosynchronous orbit (Schulz and Lanzerotti, 1974) are $\tau_{\rm g} \sim 0.33\,{\rm sec} \ll \tau_{\rm b} \sim 33\,{\rm sec} \ll \tau_{\rm d} \sim 10^5\,{\rm sec}$. It has, thus, 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 $(J_{\rm g},J_{\rm b},J_{\rm d})$, where the drift-action $J_{\rm d} \equiv (e/c)\,\Phi_{\rm B}$ is defined in terms of the magnetic flux $\Phi_{\rm B}$ enclosed by the bounce-averaged precession motion of the magnetically-trapped charged particles.

The typical energy-confinement time $\tau_{\rm E}$ in high-temperature magnetized plasmas (which is of great interest in the development of fusion energy) generally satisfies the condition $\tau_{\rm E} \gg \tau_{\rm b} \gg \tau_{\rm g}$ and, thus, the time scales associated with charge particle's gyro-motion and bounce/transit motion are much larger than the transport time scale of interest. Moreover, the observed anomalous transport associated with present magnetic confinement devices is thought to be intimately related to the fluctuation-induced transport processes due to saturated (finite-amplitude) plasma turbulence, whose characteristic time scale is much longer than the gyroperiod.

Experimental observations of magnetically-confined plasmas indicate that such magnetized plasmas represent strongly turbulent systems; see, for example, Liewer (1985) or Wootton (1990). The observed turbulence in high-temperature magnetized plasmas is characterized by fluctuation spectra exhibiting the following features: (1) a broad frequency spectrum ($\Delta\omega \sim \omega$) at fixed wavevector $\mathbf{k}$; (2) the characteristic (mean) frequency ($\omega \sim \omega_*$) and perpendicular wavelength ($\lambda_\perp \sim \rho_{\rm s}$) of the fluctuation spectrum are typical of drift-wave turbulence theories (Horton, 1999); (3) fluctuations in density, temperature, electrostatic potential, and magnetic field, with each fluctuating quantity having its own spatial profile across the plasma discharge; and (4) a fixed-frequency fluctuation spectrum that is highly anisotropic in wavevector ($k_\| \ll k_\perp$). Lastly, plasma turbulence is believed to originate from collective instabilities driven by the expansion free energy associated with radial gradients in temperature or density (Horton, 1999; Tang, 1978).

Understanding the nonlinear dynamics of magnetically-confined plasmas is a formidable task for the following reasons. First, there exists a wide variety of instabilities in inhomogeneous magnetically-confined plasmas whose nonlinear behavior is, in general, different from the corresponding linear behavior. Secondly, there is no clear separation in plasma turbulence between the "inertial" range and the "dissipation" range, in contrast to the inertial range in fluid turbulence (Frisch, 1995), which exists over several decades in wave-vector $\mathbf{k}$ and for which the Reynolds number R (a dimensionless number characterizing the ratio between nonlinear coupling and classical dissipation) satisfies the condition $R \gg 1$. Furthermore, plasma turbulence involves a plethora of additional dimensionless parameters associated with the orbital dynamics of magnetically-confined charged particles not present in fluid turbulence. Thirdly, many aspects of the nonlinear dynamics involved in the evolution toward such a saturated state, which often also exhibits self-organized large-scale motion, are not yet well understood. Lastly, it is important to note that many plasmas of interest in magnetic fusion and in astrophysics are "collisionless" in particle dynamics and turbulence time scales, since typical charged particles can execute many gyro-motions and bounce/transit motions and collective waves can oscillate many times before particles suffer a 90^o Coulomb collision. Therefore, a collisionless kinetic description is desirable for such plasmas.

---

[1] Additional terms associated with magnetic-field inhomogeneity involve $\widehat{\mathsf{b}}\cdot\nabla\times\widehat{\mathsf{b}}$, which is related to the plasma current flowing along magnetic-field lines, and $\mathbf{R} \equiv \nabla\widehat{1}\cdot\widehat{2}$, where $\widehat{1} \equiv \nabla\alpha/|\nabla\alpha|$ and $\widehat{2} \equiv \widehat{\mathsf{b}}\times\widehat{1}$; these secondary terms appear in the Hamiltonian guiding-center theory (Littlejohn, 1983).

The development of gyrokinetic theory was initially motivated by the need to describe complex plasma dynamics over time scales that are long compared to the short gyro-motion time scale. Thus, gyrokinetic theory was built upon a generalization of guiding-center theory; see, for example, Northrop (1963) and 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 $\overline{\mu}$) can be constructed as an asymptotic expansion in powers of the amplitude of the perturbation field. This early result indicated that gyrokinetic theory was to be built upon an additional transformation beyond the guiding-center phase-space coordinates, thereby constructing new *gyrocenter* phase-space coordinates, which describe gyroangle-averaged perturbed guiding-center dynamics. This additional step, however, was not considered as 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.

Nonlinear gyrokinetic theory focuses its attention on the low-frequency electromagnetic fluctuations that are observed in inhomogeneous magnetized plasmas; see, for example, Frieman and Chen (1982), Dubin *et al.* (1983), Hahm *et al.* (1988), Hahm (1988), Brizard (1989a), and Hahm (1996). Microturbulence and its associated anomalous transport are subject of active research and wide interest over many years. The following review papers have addressed this subject with different emphasis: 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 are based on 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 equations 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 self-consistent gyrokinetic Vlasov-Maxwell equations that includes all the relevant linear and nonlinear coupling terms.

I. Nonlinear Gyrokinetic Vlasov Equation

Particle Hamiltonian Dynamics
$\Downarrow$
Guiding-center Hamiltonian Dynamics
$\Downarrow$
Gyrocenter Hamiltonian Dynamics
$\Downarrow$
Gyrokinetic Vlasov Equation

Our derivation of the nonlinear gyrokinetic Vlasov equation proceeds in two steps; each step involves the asymptotic decoupling of the fast gyro-motion time scale from a set of Hamilton equations by Lie-transform methods (I). The first step is concerned with the derivation of the guiding-center 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 $\mu = \mu_0 + \cdots$ (where $\mu_0 \equiv m|\mathbf{v}_\perp|^2/2B$ denotes the lowest-order term) is treated as a dynamical invariant within 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 into the perturbed guiding-center Hamiltonian system. In the second step, we derive a new set of gyrocenter Hamiltonian equations through the elimination of the gyroangle from the perturbed guiding-center equations. As a result of the gyrocenter transformation, the new gyrocenter magnetic moment $\overline{\mu} = \mu + \cdots$ is constructed as the new adiabatic invariant and the gyrocenter gyroangle $\overline{\zeta}$ is an ignorable coordinate. Within the gyrocenter Hamiltonian formalism, the gyrokinetic Vlasov equation, thus, expresses the fact that the gyrocenter Vlasov distribution $\overline{F}(\overline{\mathbf{X}}, \overline{v}_\parallel, t; \overline{\mu})$ is constant along a gyrocenter orbit in gyrocenter phase space $(\overline{\mathbf{X}}, \overline{v}_\parallel; \overline{\mu}, \overline{\zeta})$:

$$\frac{\partial \overline{F}}{\partial t} \;+\; \frac{d\overline{\mathbf{X}}}{dt}\cdot\overline{\nabla F} \;+\; \frac{d\overline{v}_\parallel}{dt}\,\frac{\partial \overline{F}}{\partial \overline{v}_\parallel} \;=\; 0, \qquad (1)$$

where $d\overline{\mu}/dt \equiv 0$ and $\partial\overline{F}/\partial\overline{\zeta} \equiv 0$. Here, $\overline{\mathbf{X}}$ denotes the gyrocenter position, $\overline{v}_\parallel \equiv \widehat{\mathsf{b}}\cdot d\overline{\mathbf{X}}/dt$ denotes the gyrocenter parallel velocity, and the equations of motion $(d\overline{\mathbf{X}}/dt, d\overline{v}_\parallel/dt)$ are independent of the gyrocenter gyroangle $\overline{\zeta}$ (explicit expressions are given below).

II. Gyrokinetic Maxwell Equations

Maxwell Equations
$\Downarrow$
Gyrokinetic Maxwell Equations

The self-consistent description of the low-frequency electromagnetic fluctuations, which are 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 gyrocenter phase-space transformations) and each step introduces a polarization density and a magnetization current in the gyrokinetic Maxwell equations:

$$\nabla\cdot(\mathbf{E}\,+\,\delta\mathbf{E}) \;=\; 4\pi\;(\overline{\rho}\,+\,\rho_{\rm pol}), \tag{2}$$

$$\nabla\times(\mathbf{B}\,+\,\delta\mathbf{B}) \;=\; \frac{4\pi}{c}\;\left(\overline{\mathbf{J}}\,+\,\mathbf{J}_{\rm mag}\right), \tag{3}$$

where $\overline{\rho}$ and $\overline{\mathbf{J}}$ denote charge and current densities evaluated as moments of the gyrocenter Vlasov distribution $\overline{F}$, while the polarization density $\rho_{\rm pol} \equiv -\nabla\cdot\mathbf{P}_{\rm gy}$ and the magnetization current $\mathbf{J}_{\rm mag} \equiv c\,\nabla\times\mathbf{M}_{\rm gy}$ are defined in terms the gyrocenter polarization vector $\mathbf{P}_{\rm gy}$ and the gyrocenter magnetization vector $\mathbf{M}_{\rm gy}$.[2] For example, the guiding-center magnetization current $\mathbf{M}_{\rm gc} \equiv -\,\|\mu\|_{\rm gc}\,\widehat{\mathsf{b}}$ (where $\|\cdots\|_{\rm gc}$ denotes a moment with respect to the guiding-center Vlasov distribution) explains the difference between the particle current and the guiding-center current. Gyrocenter polarization and magnetization effects, on the other hand, involve expressions for $\rho_{\rm pol}$ and $\mathbf{J}_{\rm mag}$ 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).

III. Gyrokinetic Energy Conservation Law

Gyrokinetic Variational Formulation $\Downarrow$ Gyrokinetic Vlasov-Maxwell Equations $\Downarrow$ Gyrokinetic Energy Conservation Law (Noether)

The polarization and magnetization effects appearing in the gyrokinetic Maxwell equations (2)-(3) can be computed either directly by *push-forward* method, which involves transforming particle moments into guiding-center moments and then into gyrocenter moments, or by variational method from a nonlinear low-frequency gyrokinetic action functional. While the direct approach has the advantage of being the simplest derivation method to use (Brizard, 1989a,b, 1990; Dubin *et al.*, 1983; Hahm *et al.*, 1988), the variational approach (Brizard, 2000a,b) 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).

The purpose of the present paper is to review the modern foundations of nonlinear gyrokinetic theory by presenting the Lagrangian and Hamiltonian methods used in the derivation of 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.

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

II. BASIC PROPERTIES OF NONLINEAR GYROKINETIC EQUATIONS

A. Physical Motivations and Nonlinear Gyrokinetic Orderings

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

By decoupling the information with the nearly-circular gyro-motion (Kruskal, 1962), one can derive the gyrokinetic equation (1) which describes the spatio-temporal evolution of the gyrocenter distribution function defined over a *reduced* (4+1)-dimensional gyrocenter phase space $(\overline{\mathbf{X}}, \overline{v}_\parallel; \overline{\mu})$, a key feature of the modern nonlinear gyrokinetic approach. In simulating strongly magnetized plasmas, one can, thus, save enormous amount of computing time by having a time step greater than the gyroperiod, and by reducing the number of dynamical variables by one.

[2] The polarization current $\mathbf{J}_{\rm pol} \equiv \partial\mathbf{P}_{\rm gy}/\partial t$ is not shown in Eq. (3) because its gyrokinetic ordering is typically higher than the other terms.

An excellent example where the nonlinear gyrokinetic formulations have applied well, and have made a deep and long-lasting impact, is the theoretical study of microturbulence in tokamak devices. Experimental measurements over the last three decades have indicated that, in the absence of macroscopic magnetohydrodynamic (MHD) instabilities,[3] tokamak microturbulence is believed to be responsible for the 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 (Chang and Hinton, 1982; Connor *et al.*, 1987; Hinton and Hazeltine, 1976; Hinton and Wong, 1985; Rosenbluth *et al.*, 1972).

From experimental observations (see references listed in Wootton (1990)), the typical fluctuation frequency spectrum is found to be broadband ($\Delta\omega \sim \omega_{\rm k}$) 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 $\omega_* \equiv \mathbf{k} \cdot \mathbf{v}_{\rm D} \ll \Omega$, where the diamagnetic velocity $\mathbf{v}_{\rm D}$ [$\equiv (cT/eB)\,\widehat{\mathsf{b}} \times \nabla \ln P$] is caused by a perpendicular gradient in plasma pressure P and $\Omega \equiv eB/m_{\rm i}c$ denotes the ion gyrofrequency. Here, 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 $\rho_{\rm i} \sim 0.2\,$cm and the frequency ratio $\omega_*/\Omega \equiv (k_\theta \rho_{\rm i})\,\rho_{\rm i}/L \sim 10^{-3}$, where $k_\theta \sim 1\,{\rm cm}^{-1}$ denotes the poloidal component of the wave vector $\mathbf{k}$ (see Figure 1). Its correlation lengths in both radial and poloidal directions, on the other hand, are on the order of several ion gyroradii, which are much shorter than the macroscopic gradient-scale length L. Its wavelength (or correlation length) along the equilibrium magnetic field is rarely measured, in particular inside the last closed magnetic surface. But some measurements at the scrape-off layer indicate that it is much less than the connection length (Endler *et al.*, 1995; Zweben and Medley, 1989). Lastly, the relative density fluctuation level $\delta n/n_0$ ranges typically from well under 1 % at the core (near the magnetic axis) to the order of 10 % at the edge (see Figure 2). Fluctuations in electric field and magnetic field in the interior of tokamaks are also rarely measured, but estimates indicate that $e\,\delta\phi/T_{\rm e} \sim \delta n/n_0$, and $|\delta\mathbf{B}|/B_0 \sim 10^{-4}$.

From these spatio-temporal scales of tokamak microturbulence, one can make a very rough estimate of transport coefficient $D_{\rm turb}$ using a dimensional analysis based on a random-walk argument:

$$D_{\rm turb} \;\sim\; \frac{(\Delta r)^2}{\Delta t} \;\sim\; \frac{\Delta\omega}{k_r^2} \;\propto\; \frac{\omega_*}{k_r^2} \;\sim\; \left(\frac{k_\theta}{k_r^2 \rho_{\rm i}}\right) \frac{\rho_{\rm i}}{L} \;\cdot\; \frac{cT_{\rm i}}{eB}.$$

If we further assume that $k_\theta \sim k_r \propto \rho_{\rm i}^{-1}$ as observed

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

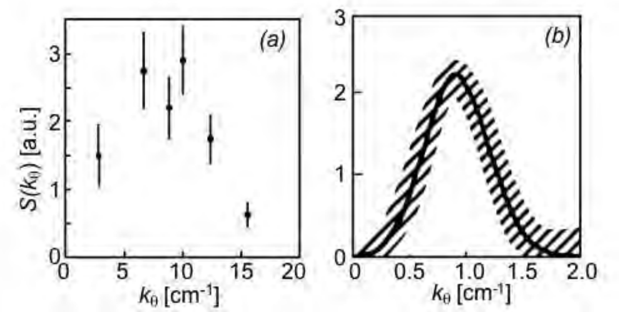

FIG. 1 Spatial wave-number spectra obtained from spatial correlation coefficients in the poloidal direction for (a) the Adiabatic Toroidal Compression tokamak (Mazzucato, 1982) and (b) the Tokamak Fusion Test Reactor (Fonck *et al.*, 1993).

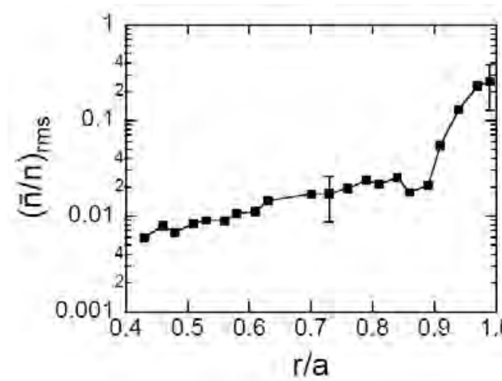

FIG. 2 Spatial profile of the total rms density-fluctuation amplitude obtained by Beam Emission Spectroscopy on the Tokamak Fusion Test Reactor (Fonck *et al.*, 1993).

in gyrokinetic simulations with self-generated zonal flows (see Table I) and in some experiments (Fonck *et al.*, 1993; McKee *et al.*, 2003), we obtain

$$D_{\rm turb} \;\sim\; \frac{\rho_{\rm i}}{L} \;\cdot\; \frac{cT_{\rm i}}{eB}.$$

This scaling is called gyroBohm because the Bohm scaling ($\sim\; cT_{\rm i}/eB$) is reduced by a factor $\rho_{\rm i}/L$ involving the ratio of gyroradius to a macroscopic length scale. This gyroBohm transport scaling is expected when local physics dominates. Sometimes, it can be modified due to a variety of mesoscale phenomena (Itoh and Itoh, 2001) such as turbulence spreading (Chen *et al.*, 2004; Garbet *et al.*, 1994; Gurcan *et al.*, 2005, 2006; Hahm *et al.*, 2004a, 2005; Itoh *et al.*, 2005; Kim *et al.*, 2003; Lin and Hahm, 2004; Naulin *et al.*, 2005; Villard *et al.*, 2004b; Waltz and Candy, 2005; Zonca *et al.*, 2004) and avalanches (Diamond and Hahm, 1995; Garbet and Waltz, 1998; Naulin *et al.*, 1998; Newman *et al.*, 1996; Politzer *et al.*, 2002; Sarazin and Ghendrih, 1998).

On the other hand, some early simulations without self-generated zonal flows (see Table I for examples) have reported $\Delta r \propto \sqrt{L\,\rho_{\rm i}}$, with $k_\theta \propto \rho_{\rm i}^{-1}$, from which we obtain the Bohm scaling $D_{\rm turb} \sim cT_{\rm i}/eB$. While the distinction between Bohm scaling and gyroBohm scaling may seem quite simple and obvious, it is complicated by many subtle issues (Lin *et al.*, 2002; Waltz *et al.*, 2002).

As stated above, the nonlinear gyrokinetic formalism pursues a *dynamical reduction* of the original Vlasov-Maxwell equations for both computational and analytic feasibility while keeping the general description of the relevant physical phenomena of interest intact. In this Section, we describe the standard nonlinear gyrokinetoc ordering (Frieman and Chen, 1982) with an emphasis on physical motivations. We note that our understanding of microturbulence based on experimental observations was incomplete when the earlier versions of nonlinear gyrokinetic equations were being developed. Therefore, the original motivation of the ordering might have been somewhat different from our interpretation here. The adiabatic invariance of the new magnetic moment $\overline{\mu}$ is, however, established on the fundamental fluctuation-based space-time orderings $\omega \ll \Omega$ and $|\mathbf{k}_\perp| \gg L^{-1}$, which have broad experimental basis for the most important 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, in turn, 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 the fluctuation fields (represented by the perturbed Vlasov distribution δf and the perturbed electric and magnetic fields $\delta\mathbf{E}$ and $\delta\mathbf{B}$):

$$
\begin{aligned}
\omega &= \text{characteristic fluctuation frequency} \\
k_\| &= \text{characteristic fluctuation parallel wavenumber} \\
|\mathbf{k}_\perp| &= \text{characteristic fluctuation perpendicular} \\
&\quad\ \text{wavenumber} \\
\Omega &= \text{ion cyclotron frequency} \\
v_{\rm th} &= \sqrt{T_{\rm i}/m_{\rm i}} \;=\; \text{ion thermal speed} \\
\rho_{\rm i} &= v_{\rm th}/\Omega \;=\; \text{ion thermal gyroradius} \\
L_{\rm B} &= \text{characteristic background magnetic-field} \\
&\quad\ \text{nonuniformity length scale} \\
L_{\rm F} &= \text{characteristic background plasma density} \\
&\quad\ \text{and temperature nonuniformity length scale}
\end{aligned}
$$

First, the background plasma is described in terms of the (guiding-center) small parameter $\epsilon_{\rm B} \equiv \rho_{\rm i}/L_{\rm B}$ as

$$
|\rho_{\rm i}\,\nabla\ln B| \;\sim\; \epsilon_{\rm B} \quad\text{and}\quad \left|\frac{1}{\Omega}\,\frac{\partial\ln B}{\partial t}\right| \;\sim\; \epsilon_{\rm B}^3, \tag{4}
$$

where the background time-scale ordering ($\epsilon_{\rm B}^3$) is consistent with the transport time-scale ordering (Hinton and Hazeltine, 1976); note that the background Vlasov distribution F satisfies a similar space-time ordering[4] with $\epsilon_{\rm F} \equiv \rho_{\rm i}/L_{\rm F}$. The background magnetized plasma is, therefore, 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.

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

$$
|\mathbf{k}_\perp|\,\rho_{\rm i} \;\equiv\; \epsilon_\perp \;\sim\; 1 \;\;\text{and}\;\; \frac{\omega}{\Omega} \;\sim\; \epsilon_\omega \;\ll\; 1. \tag{5}
$$

Note that, while microturbulence spectra of present-day high-temperature plasmas typically peak around $\epsilon_\perp \sim 0.1 - 0.2 < 1$ at nonlinear saturation, the linear growth rates are typically highest at $\epsilon_\perp \sim 1$. Hence, 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 ($\epsilon_\perp \sim 1$) as well. Second, 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 $\omega \sim |k_\||v_{\rm th}$ and, thus, we also have the ordering

$$
\frac{|k_\||}{|\mathbf{k}_\perp|} \;\sim\; \frac{\epsilon_\omega}{\epsilon_\perp}. \tag{6}
$$

Note that the most dangerous plasma instabilities in a strong magnetic field tend to satisfy this parallel ordering. Third, the relative fluctuation levels are described in terms of the amplitude ordering parameter ϵ_δ:

$$
\left|\frac{\delta f}{F}\right| \;\sim\; \frac{c|\delta\mathbf{E}_\perp|}{B\,v_{\rm th}} \;\sim\; \frac{|\delta\mathbf{B}|}{B} \;\sim\; \epsilon_\delta \;\ll\; 1. \tag{7}
$$

The electric fluctuation ordering

$$
\epsilon_\delta \;\sim\; \frac{c|\delta\mathbf{E}_\perp|}{B\,v_{\rm th}} \;\sim\; \epsilon_\perp\,\frac{e\,\delta\phi}{T_{\rm i}} \tag{8}
$$

implies that, for $\epsilon_\perp \sim 1$ and $T_{\rm e} \sim T_{\rm i}$, we have $e\,\delta\phi/T_{\rm e} \sim \epsilon_\delta$. In addition, a covariant description of electromagnetic fluctuations requires that the parallel component $\delta A_\| \equiv \widehat{\mathsf{b}}\cdot\delta\mathbf{A}$ of the perturbed vector potential satifies the amplitude ordering $(v_\|/c)\,\delta A_\| \sim \delta\phi$, where the parallel particle velocity is $v_\| \sim v_{\rm th}$, so that

$$
\epsilon_\delta \;\sim\; \frac{v_\|}{c}\,\frac{e\,\delta A_\|}{T_{\rm i}} \;\sim\; \frac{|\delta\mathbf{B}_\perp|}{\epsilon_\perp\,B}, \tag{9}
$$

[4] Note that the ratio $\epsilon_{\rm B}/\epsilon_{\rm F} < 1$ 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, we assume that $\epsilon_{\rm F} \sim \epsilon_{\rm B}$.

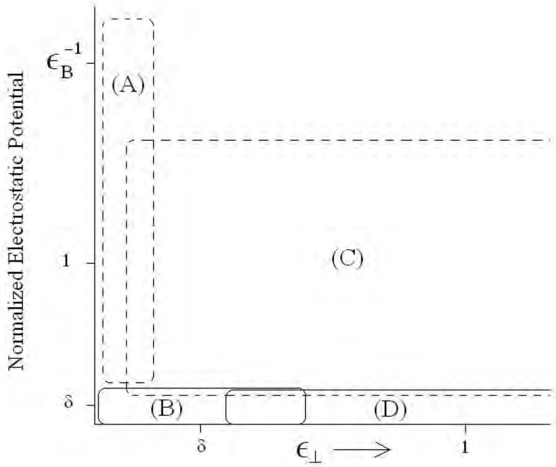

FIG. 3 Regimes of validity of nonlinear (A) and linear (B) drift-kinetic and nonlinear (C) and linear (D) gyrokinetic theories displayed on a plot of normalized electrostatic potential $(L/\rho_{\rm i})\, e\,\delta\phi/T_{\rm e} \sim \epsilon_\delta/(\epsilon_\perp\epsilon_{\rm B})$ versus $\epsilon_\perp$, where the ordering parameters $(\epsilon_\delta, \epsilon_\perp, \epsilon_{\rm B})$ 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.

which implies that, for $\epsilon_\perp \sim 1$, we have $|\delta\mathbf{B}_\perp|/B \sim \epsilon_\delta$. Hence, the orderings (8) and (9) imply that the terms $|k_\parallel|\delta\phi$ and $(\omega/c)\,\delta A_\parallel$ in the parallel perturbed electric field $\delta E_\parallel$ 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 $|\delta E_\parallel|/|\delta\mathbf{E}_\perp| \sim |k_\parallel|/|\mathbf{k}_\perp| \sim \epsilon_\omega/\epsilon_\perp \ll 1$. Lastly, for a fully electromagnetic gyrokinetic ordering (and high-β plasmas, where $\beta \equiv 8\pi P/B^2$), we also require that $|\delta B_\parallel|/B \sim \beta\,\epsilon_\delta$ (Brizard, 1992)[5]; note, here, that we use the gyrokinetic gauge condition $\nabla_\perp\cdot\delta\mathbf{A}_\perp = 0$, which represents the low-frequency gyrokinetic limit of the Lorentz 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.

Lastly, the regimes of validity of various drift-kinetic and gyrokinetic theories are summarized in Fig. 3 in terms of the normalized electrostatic potential $(L/\rho_{\rm i})\, e\,\delta\phi/T_{\rm e} \sim \epsilon_\delta/(\epsilon_\perp\epsilon_{\rm B})$ (Dimits *et al.*, 1992). 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 $\epsilon_\delta \sim \epsilon_{\rm B}$. The latter ordering implies that the linear drive term ($\sim \nabla\delta\phi\times\widehat{\mathsf{b}}\cdot\nabla F$) is of the same order as the nonlinear $E\times B$ coupling term ($\sim \nabla\delta\phi\times\widehat{\mathsf{b}}\cdot\nabla\delta f$); this is a generic situation for strong turbulence, which yields a nonlinear saturation roughly at a *mixing length* level $\delta n \sim (\rho_{\rm i}/L)\,n$ (nonetheless, with a subsidiary ordering, the nonlinear gyrokinetic equations can describe weak turbulence as well).

[5] While $e\,\delta\phi/T_{\rm e} \gg |\delta\mathbf{B}_\perp|/B \gg |\delta B_\parallel|/B$ for typical low-to-modest β tokamak plasmas, the fluctuation ordering (7) is retained for its generality (which, thus, makes it applicable to high-β devices such as spherical tori).

B. Frieman-Chen Nonlinear Gyrokinetic Equation

The first significant work on nonlinear gyrokinetic equations in general magnetic geometry was presented by Frieman and Chen (1982), who used a conventional approach based on a maximal multi-scale-ordering expansion involving a single ordering parameter ϵ ($\epsilon \sim \epsilon_{\rm B} \sim \epsilon_\omega \sim \epsilon_\delta$). Here, the linear-physics-drive terms are ordered at $\epsilon_\omega\,\epsilon_\delta$ and $\epsilon_{\rm B}\,\epsilon_\delta$ (which recognizes the crucial role played by the background magnetic-field nonuniformity), 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 indeed 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 minimum at least in the context of nonlinear gyrokinetics (we elaborate 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

$$\frac{df}{dt} \equiv \frac{\partial f}{\partial t} + \frac{d\mathbf{z}}{dt}\cdot\frac{\partial f}{\partial\mathbf{z}} = 0, \tag{10}$$

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 (10) 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 an 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 $d\mathbf{z}/dt = d\mathbf{Z}/dt + \epsilon_\delta\, d\delta\mathbf{z}/dt$, 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 presence 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 $\overline{f} \equiv F$ and $\overline{(d\mathbf{z}/dt)} \equiv d\mathbf{Z}/dt$. Hence, the short-space-scale average of the Vlasov equation (10) yields

$$\frac{\partial F}{\partial t} + \frac{d\mathbf{Z}}{dt}\cdot\frac{\partial F}{\partial\mathbf{z}} \equiv \frac{DF}{Dt} = -\,\epsilon_{\delta}^{2}\,\overline{\left(\frac{d\delta\mathbf{z}}{dt}\cdot\frac{\partial\delta f}{\partial\mathbf{z}}\right)}, \quad (11)$$

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 $F = F_0 + \epsilon_{\rm B}\,F_1 + \cdots$ for the background Vlasov equation (11) up to first order in $\epsilon_{\rm B}$ and without the nonlinear fluctuation-driven term ($\epsilon_{\delta} = 0$), where the gyroangle-dependent part of the first-order correction

$$\epsilon_{\rm B}\,\widetilde{F}_1 = -\left(\int\frac{d\zeta}{\Omega}\,\dot{\mu}_0\right)\frac{\partial F_0}{\partial\mu} \quad (12)$$

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., $\widehat{\mathsf{b}}\cdot\nabla F_0 = 0$), the (lowest-order) guiding-center kinetic energy $\mathcal{E}$ (with $\dot{\mathcal{E}}_0 \equiv 0$), and the (lowest-order) guiding-center magnetic moment μ. In Eq. (12), $\dot{\mu}_0$ denotes the time derivative of the lowest-order magnetic moment, which is ordered at $\epsilon_{\rm B}$ for a time-independent magnetic field and is explicitly gyroangle-dependent.[6]

By subtracting the averaged Vlasov equation (11) from the Vlasov equation (10), we obtain the *fluctuating* Vlasov equation

$$\frac{D\delta f}{Dt} = -\,\frac{d\delta\mathbf{z}}{dt}\cdot\frac{\partial F}{\partial\mathbf{z}} - \frac{d\delta\mathbf{z}}{dt}\cdot\frac{\partial\delta f}{\partial\mathbf{z}} + \overline{\left(\frac{d\delta\mathbf{z}}{dt}\cdot\frac{\partial\delta f}{\partial\mathbf{z}}\right)}, \quad (13)$$

where the left side contains terms of order $\epsilon_\omega\epsilon_\delta$ and $\epsilon_{\rm B}\epsilon_\delta$, while the first term on the right side provides the linear drive for δf (at order $\epsilon_{\rm B}\epsilon_\delta$) 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. (13) 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 (11).

Next, Frieman and Chen (1982) adopt a standard iterative procedure (Hastie *et al.*, 1967) designed to solve the fluctuating Vlasov equation (13) by introducing a decomposition of the perturbed Vlasov distribution δf in terms of its *adiabatic* and *nonadiabatic* components (Antonsen and Lane, 1980; Catto *et al.*, 1981):

$$\delta f \equiv \left[\,e\,\delta\phi\,\frac{\partial}{\partial\mathcal{E}} + \frac{e}{B}\left(\delta\phi - \frac{v_\parallel}{c}\,\delta A_\parallel\right)\frac{\partial}{\partial\mu}\,\right]F_0 + e^{-\boldsymbol{\rho}\cdot\nabla}\left(\delta g - \frac{e\langle\delta\psi_{\rm gc}\rangle}{B}\,\frac{\partial F_0}{\partial\mu}\right). \quad (14)$$

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

$$\langle\delta\psi_{\rm gc}\rangle \equiv \left\langle e^{\boldsymbol{\rho}\cdot\nabla}\left(\delta\phi - \frac{\mathbf{v}}{c}\cdot\delta\mathbf{A}\right)\right\rangle = \left\langle\delta\phi_{\rm gc} - \frac{\mathbf{v}}{c}\cdot\delta\mathbf{A}_{\rm gc}\right\rangle. \quad (15)$$

Note that all terms on the right side of Eq. (14) are ordered at ϵ_δ and an additional adiabatic term $\nabla F_0\cdot\delta\mathbf{A}\times\widehat{\mathsf{b}}/B$ (of order $\epsilon_{\rm B}\epsilon_\delta$) 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 (15), where perturbed potentials are evaluated at the guiding-center position, as the *guiding-center* adiabatic term.

By substituting the nonadiabatic decomposition (14) into the fluctuating Vlasov equation (13), Frieman and Chen (1982) obtain (after a tremendous amount of tedious algebra) the nonlinear gyrokinetic equation for nonadiabatic part δg of the perturbed Vlasov distribution

$$\frac{d_{\rm gc}\delta g}{dt} = -\left(e\,\frac{\partial\langle\delta\psi_{\rm gc}\rangle}{\partial t}\,\frac{\partial}{\partial\mathcal{E}} + \frac{c\widehat{\mathsf{b}}}{B}\times\nabla\langle\delta\psi_{\rm gc}\rangle\cdot\nabla\right)F_0 - \frac{c\widehat{\mathsf{b}}}{B}\times\nabla\langle\delta\psi_{\rm gc}\rangle\cdot\nabla\delta g \quad (16)$$

where $D/Dt \equiv d_{\rm gc}/dt$ denotes the unperturbed averaged Vlasov operator expressed in guiding-center coordinates $(\mathbf{X},\mathcal{E},\mu)$. Here, we note that the time evolution of the nonadiabatic part δg depends on the effective first-order Hamiltonian (15). The terms appearing on the left side of Eq. (16), as well as the F_0-terms on the right side, are ordered at $\epsilon_\omega\epsilon_\delta$ and $\epsilon_{\rm B}\epsilon_\delta$, 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 $(c\widehat{\mathsf{b}}/B)\times\nabla\langle\delta\phi_{\rm gc}\rangle$, the magnetic flutter velocity $(v_\parallel/B)\,\langle\delta\mathbf{B}_{\perp\rm gc}\rangle$, and the perturbed grad-B drift velocity $(-\,\widehat{\mathsf{b}}/B)\times\nabla\langle\mathbf{v}_\perp\cdot\delta\mathbf{A}_{\perp\rm gc}\rangle$. The nonlinear gyrokinetic Vlasov equation (48) derived by Frieman and

[6] Using $(\mathcal{E},\mu)$ rather than $(v_\parallel,\mu)$ 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 $(v_\parallel,\mu)$-formulation. However, for more complex realistic nonlinear applications, we find the $(v_\parallel,\mu)$-formulation more straightforward in describing physics.

Chen (1982) contains additional terms, defined in their equation (45), that are subsequently omitted in their final equation (50).

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 (14). For example, using the nonadiabatic decomposition (14), the perturbed particle fluid density $\delta n \equiv \int d^3p\,\delta f$ is expressed as

$$\delta n \;=\; \int d^3P\,\left\langle e^{-\boldsymbol{\rho}\cdot\nabla}\left(\delta g \;-\; \frac{e\langle\delta\psi_{\rm gc}\rangle}{B}\,\frac{\partial F_0}{\partial\mu}\right)\right\rangle, \quad (17)$$

where d^3P denotes the momentum-space integral in guiding-center coordinates (which involves a gyroangle integration that is explicitly represented, here, by the gyroangle average $\langle\;\;\rangle$). We immediately note that the particle adiabatic terms in Eq. (14) have cancelled out of Eq. (17) and only the guiding-center adiabatic and nonadiabatic terms contribute to δn. We shall show later on that this guiding-center adiabatic contribution leads to the so-called polarization density (see Sec. C.2); a similar treatment for the perturbed particle moment $\int d^3p\,\mathbf{v}\,\delta f$ 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.

We note that the Frieman-Chen nonlinear gyrokinetic Vlasov (16) is contained in modern versions of the nonlinear gyrokinetic Vlasov equation (Brizard, 1989a). The Frieman-Chen formulation, for example, 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's equation (Lee, 1983) has provided a crucial computational advantage in nonlinear gyrokinetic simulations. It is also a key quantity in relating the nonlinear gyrokinetic approach to reduced magnetohydrodynamics (Hahm *et al.*, 1988).

C. Modern Nonlinear Gyrokinetic Equations

The major difficulties encountered in the conventional Frieman-Chen derivation of Eq. (16) involve (a) inserting the solution (12) for the first-order correction F_1 to the background Vlasov distribution F_0 into the first term on the right of Eq. (13) and (b) constructing a new magnetic moment $\overline{\mu}$ that is invariant at first order in $\epsilon_{\rm B}$ and ϵ_δ. While the Frieman-Chen equations are valid up to order ϵ^2 and should be fine for immediate practical purposes,[7] 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). 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, ignoring the $\mathcal{O}(\epsilon^3)$ nonlinear wave-particle interactions due to parallel-velocity-space nonlinearity, for example, can artificially limit the energy exchange between particles and waves.

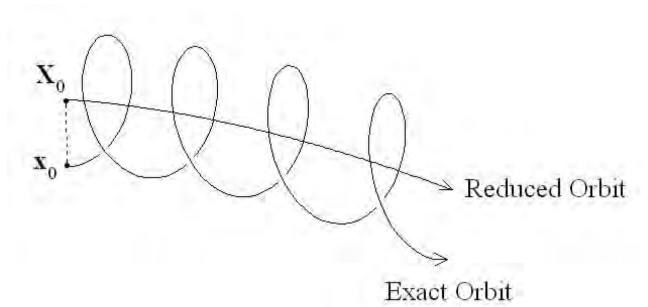

FIG. 4 Exact and reduced single-particle orbits in a magnetic field.

In contrast to conventional methods used for deriving nonlinear gyrokinetic equations, which consist of 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, thus, 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 $\overline{\mathbf{Z}}$ (to be defined later) such that the new *gyrocenter* equations of motion $d\overline{\mathbf{Z}}/dt$ are independent of the fast gyro-motion time scale at arbitrary orders in $\epsilon_{\rm B}$ and ϵ_δ. The purpose of this transformation is to have the fast gyro-motion time scale effectively decoupled from the slow *reduced* time scales. In the course of this derivation, the important underlying symmetry and conservation laws of the original system are kept intact. Moreover, in contrast to the conventional derivation (where different small parameters are lumped together via a particular ordering), various expansion parameters

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

appear at different stages of the modern derivation. This feature makes the modifications of ordering for specific applications more transparent, such as nonlinear gyrokinetic equations with strong $E\times B$ shear flows as described in Appendix E.1.

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

$$\overline{\mathbf{Z}} \;\equiv\; \sum_{n=0} \epsilon_\delta^n\, \overline{\mathbf{Z}}_n(\mathbf{z}), \tag{18}$$

where the lowest-order term $\overline{\mathbf{Z}}_0(\mathbf{z})$ is expressed in terms of an asymptotic expansion in powers of the background-plasma ordering parameter $\epsilon_{\rm B}$ associated with the guiding-center transformation (see Figure 4). Note that the original particle dynamics $d\mathbf{z}/dt$ in Eq. (10) can be represented as a Hamiltonian system $d\mathbf{z}/dt \equiv \{\mathbf{z},\, H\}_{\rm z}$, where $H(\mathbf{z},t)$ denotes the particle Hamiltonian and $\{\;,\;\}_{\rm z}$ denotes the Poisson bracket on particle phase space with coordinates $\mathbf{z}$ (which are, generically, non-canonical). Since Hamiltonian systems have important conservation properties, e.g., the Liouville theorem associated with the invariance of the phase-space volume under Hamiltonian evolution (Goldstein *et al.*, 2002), we, thus, require that the new equations of motion $d\overline{\mathbf{Z}}/dt$ be also expressed as a Hamiltonian system in terms of a new Hamiltonian $\overline{H}(\overline{\mathbf{Z}},t)$ and a new Poisson bracket $\{\;,\;\}_{\overline{\mathbf{Z}}}$ such that $d\overline{\mathbf{Z}}/dt \equiv \{\overline{\mathbf{Z}},\, \overline{H}\}_{\overline{\mathbf{Z}}}$.

We now turn our attention to the impact of the phase-space transformation (18) on the Vlasov equation (10) itself. The phase-space transformation $\mathbf{z} \rightarrow \overline{\mathbf{Z}}$ *induces* a transformation from the (old) particle Vlasov distribution f to a (new) reduced Vlasov distribution $\overline{F}$, subject to the scalar-invariance property $\overline{F}(\overline{\mathbf{Z}}) = f(\mathbf{z})$, such that the new Vlasov distribution $\overline{F}$ is constant along a reduced orbit $\overline{\mathbf{Z}}(t)$. From the scalar-invariance property, the induced transformation $f \rightarrow \overline{F}$ is, therefore, defined as

$$f(\mathbf{z}) \;\equiv\; \overline{F}(\overline{\mathbf{Z}}) \;=\; \overline{F}\left(\sum_{n=0} \epsilon_\delta^n\, \overline{\mathbf{Z}}_n(\mathbf{z})\right), \tag{19}$$

which generates an asymptotic expansion in powers of ϵ_δ:

$$f \;\equiv\; \sum_{n=0} \epsilon_\delta^n\, f_n(\overline{F}), \tag{20}$$

where each term $f_n(\overline{F})$ is expressed in terms of derivatives of the new Vlasov distribution $\overline{F}$. For example, we consider the infinitesimal constant translation $x \rightarrow X = x + \epsilon$ and the induced transformation $f \rightarrow F$:

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

where $f_n(F) \equiv (1/n!)\, d^nF/dx^n$, which clearly shows that the two functions f and F are formally different functions. Hence, we recognize that the nonadiabatic decomposition (14) is, in fact, a similar asymptotic expansion $f = f_0 + \epsilon_\delta\, \delta f + ...$, where f_0 and δf are expressed in terms a *reduced* Vlasov distribution $\overline{F}$.

Lastly, in order for our new Vlasov kinetic theory to be dissipation-free, we require that the phase-space transformation (18) be invertible (i.e., entropy-conserving) so that no information about the fast-time-scale particle dynamics is lost. Hence, we must also define the following inverse relations

$$\mathbf{z} \;\equiv\; \sum_{n=0} \epsilon_\delta^n\, \mathbf{z}_n(\overline{\mathbf{Z}}), \tag{21}$$

$$\overline{F}(\overline{\mathbf{Z}}) \;\equiv\; f(\mathbf{z}) \;=\; f\left(\sum_{n=0} \epsilon_\delta^n\, \mathbf{z}_n(\overline{\mathbf{Z}})\right), \tag{22}$$

$$\overline{F} \;\equiv\; \sum_{n=0} \epsilon_\delta^n\, \overline{F}_n(f), \tag{23}$$

which are justified by the smallness of the ordering parameter $\epsilon_\delta \ll 1$. We note that, within canonical Hamiltonian perturbation theory (Goldstein *et al.*, 2002), for example, the relation between the (old) particle Hamiltonian H and the (new) reduced Hamiltonian $\overline{H}$ is expressed as

$$\overline{H}(\overline{\mathbf{Z}},t) \;\equiv\; H(\mathbf{z},t) \;-\; \frac{\partial S}{\partial t}(\mathbf{z},t), \tag{24}$$

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

The purpose of the modern formulation of the nonlinear gyrokinetic Vlasov theory is to provide powerful algorithms necessary to construct the phase-space transformations (18) and (21) and the induced transformations (20) and (23). These algorithms are based on applications of differential geometric methods associated with Lie-transforms (see Appendix A for a primer on these mathematical methods).

III. SIMPLE FORMS OF NONLINEAR GYROKINETIC EQUATIONS

We now present the simplified forms of the nonlinear gyrokinetic equations that are recommended for simulations and analytic applications. Therefore, this Section will provide 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 systematically derived later in Sections V and VI.

In this review, we do not attempt to cover exhaustively the recent remarkable progress in nonlinear gyrokinetic simulations (Tang, 2002; Tang and Chan, 2005). Instead, we discuss relevant physics issues that arise when nonlinear gyrokinetic equations are simplified and applied to specific collective waves and instabilities in plasmas. 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. It should be noted that some of the early simulations were performed as the modern nonlinear gyrokinetic formulation were being developed. Hence, not all of the gyrokinetic theoretical knowledge discussed in this review article was available then.

A. General Gyrokinetic Vlasov-Maxwell Equations

For nonlinear simulations, the nonlinear gyrokinetic Vlasov equation (1) for the gyrocenter distribution F is written in terms of the Hamiltonian gyrocenter equations of motion

$$\frac{d\mathbf{X}}{dt} \;=\; v_{\|}\,\frac{\mathbf{B}^{*}}{B_{\|}^{*}} \;+\; \frac{c\widehat{\mathsf{b}}}{eB_{\|}^{*}}\times\left(\mu\,\nabla B \;+\; e\,\nabla\delta\Psi_{\rm gy}\right), \quad (25)$$

and

$$\frac{dp_{\|}}{dt} \;=\; -\;\frac{\mathbf{B}^{*}}{B_{\|}^{*}}\boldsymbol{\cdot}\left(\mu\,\nabla B \;+\; e\,\nabla\delta\Psi_{\rm gy}\right) \;-\; \frac{e}{c}\,\frac{\partial\delta A_{\|\rm gy}}{\partial t}, \quad (26)$$

where the overbar notation used to identify the gyrocenter coordinates is omitted for the remainder of this Section. Here, $H_{\rm gy} = p_{\|}^{2}/2m + \mu B + e\,\delta\Psi_{\rm gy}$ denotes the gyrocenter Hamiltonian, where the effective gyrocenter perturbation potential $\delta\Psi_{\rm gy}$ contains terms that are linear in the perturbed electromagnetic potentials $(\delta\phi, \delta\mathbf{A})$ [e.g., the effective linear potential $\langle\delta\psi_{\rm gc}\rangle$ defined in Eq. (15)] and terms that are nonlinear (quadratic) in $(\delta\phi, \delta\mathbf{A})$. Next, the gyrocenter Poisson-bracket (symplectic) structure is represented by the modified magnetic field $\mathbf{B}^{*}$ (with $B_{\|}^{*} \equiv \widehat{\mathsf{b}}\boldsymbol{\cdot}\mathbf{B}^{*}$)

$$\mathbf{B}^{*} \;\equiv\; \mathbf{B} \;+\; (c/e)\,p_{\|}\;\nabla\times\widehat{\mathsf{b}} \;+\; \delta\mathbf{B}_{\rm gy}, \quad (27)$$

where the first term denotes the background magnetic field $\mathbf{B} \equiv B\,\widehat{\mathsf{b}}$, the second term is associated with the guiding-center curvature drift, and the third term represents the symplectic magnetic perturbation $\delta\mathbf{B}_{\rm gy} \equiv \nabla\times\delta\mathbf{A}_{\rm gy}$, which may or may not be present depending on the choice of gyrocenter model adopted (see below). The perturbed linear gyrocenter dynamics contained in Eq. (25) includes the linear perturbed $E\times B$ velocity $\delta\mathbf{u}_{\rm E} = (c\widehat{\mathsf{b}}/B)\times\nabla\delta\phi$, the perturbed magnetic-flutter velocity $v_{\|}\,\delta\mathbf{B}_{\perp}/B$, and the perturbed grad-B velocity $(c\widehat{\mathsf{b}}/eB)\times\mu\,\nabla\delta B_{\|}$. Note that, when magnetic perturbations are present, the gyrocenter parallel momentum $p_{\|}$ appearing in Eq. (26) is either a canonical momentum if the symplectic magnetic perturbation $\delta\mathbf{A}_{\rm gy}$ is chosen so that $\delta A_{\|\rm gy} \equiv \widehat{\mathsf{b}}\boldsymbol{\cdot}\delta\mathbf{A}_{\rm gy} = 0$ or a kinetic momentum (i.e., $p_{\|} = mv_{\|}$) if $\delta A_{\|\rm gy} \neq 0$. Moreover, the magnetic-flutter velocity $v_{\|}\,\delta\mathbf{B}_{\perp}/B$ is either included in $v_{\|}\,\delta\mathbf{B}_{\rm gy}/B$ (in the Symplectic gyrocenter model) or in $(c\widehat{\mathsf{b}}/B)\times\nabla\delta\Psi_{\rm gy}$ (in the Hamiltonian gyrocenter model) (Brizard, 1989a; Hahm *et al.*, 1988), while the inductive part $(\partial_t\delta A_{\|\rm gy})$ of the perturbed electric field appears on the right side of the gyrocenter parallel force equation (26) only in the Symplectic gyrocenter model. Lastly, we note that, for a specific application of nonlinear gyrokinetics, not all terms in Eqs. (25)-(26) are used simultaneously; we have written all the terms, here, for easy reference within this Section.

Next, a closed self-consistent description of the interactions involving the perturbed electromagnetic field and a Vlasov distribution of gyrocenters implies that the gyrokinetic Maxwell's equations should be written with charge-current densities expressed in terms of the gyrocenter distribution function. Hence, the gyrokinetic Poisson equation (2) is written as

$$\nabla^{2}\delta\phi \;=\; -\,4\pi\sum e\int d^{3}p\;\left\langle\, e^{-\boldsymbol{\rho}\boldsymbol{\cdot}\nabla}\left(\mathsf{T}_{\rm gy}F\right)\right\rangle \;\equiv\; -\,4\pi e\,(N_{\rm i} \;-\; n_{\rm e}) \;+\; 4\pi\,\nabla\boldsymbol{\cdot}\mathbf{P}_{\rm gy}, \quad (28)$$

where N_i denotes the ion gyrofluid density, $n_{\rm e}$ denotes the electron drift-fluid density $(k_{\perp}\rho_{\rm e} \to 0)$, and the gyrocenter polarization density is defined as

$$\rho_{\rm pol} \;\equiv\; -\,\nabla\boldsymbol{\cdot}\mathbf{P}_{\rm gy} \;=\; \sum e\int d^{3}p\;\left\langle e^{-\boldsymbol{\rho}\boldsymbol{\cdot}\nabla}\left(\mathsf{T}_{\rm gy}F - F\right)\right\rangle. \quad (29)$$

The gyrokinetic Ampère equation (3), on the other hand, is written as

$$\nabla\times(\mathbf{B}+\delta\mathbf{B}) \;=\; \frac{4\pi}{c}\sum e\int d^{3}p\;\left\langle \mathbf{v}\;\; e^{-\boldsymbol{\rho}\boldsymbol{\cdot}\nabla}\left(\mathsf{T}_{\rm gy}F\right)\right\rangle \;\equiv\; \frac{4\pi}{c}\,(\mathbf{J}_{\rm i} \;-\; \mathbf{j}_{\rm e}) \;+\; 4\pi\,\nabla\times\mathbf{M}_{\rm gy}, \quad (30)$$

where $\mathbf{J}_{\rm i}$ and $\mathbf{j}_{\rm e}$ denote the ion gyrofluid and electron drift-fluid current densities, and the gyrocenter magnetization current is defined as

$$\mathbf{J}_{\rm mag} \;\equiv\; c\,\nabla\times\mathbf{M}_{\rm gy} \;=\; \sum e\int d^{3}p\;\left\langle e^{-\boldsymbol{\rho}\boldsymbol{\cdot}\nabla}\left[\mathbf{v}\;\left(\mathsf{T}_{\rm gy}F - F\right)\right]\right\rangle. \quad (31)$$

Here, the gyrocenter polarization vector $\mathbf{P}_{\rm gy}$ and the gyrocenter magnetization vector $\mathbf{M}_{\rm gy}$ are expressed in terms of the difference $(\mathsf{T}_{\rm gy}F - F)$ between the guiding-center distribution $\mathsf{T}_{\rm gy}F$ (expressed as the *pull-back* of the gyrocenter distribution) and the gyrocenter distribution F, which explicitly involves the perturbed electromagnetic potentials $(\delta\phi, \delta\mathbf{A})$. 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 $e\,\delta\Psi_{\rm gy}$ with respect to the perturbed electric and magnetic fields $\delta\mathbf{E}$ and $\delta\mathbf{B}$, respectively:

$$\mathbf{P}_{\rm gy} \;\equiv\; -\sum e\int d^3p\;F\left(\frac{\partial\delta\Psi_{\rm gy}}{\partial\delta\mathbf{E}}\right) \;=\; \sum\int d^3p\;F\;\boldsymbol{\pi}_{\rm gy}, \tag{32}$$

$$\mathbf{M}_{\rm gy} \;\equiv\; -\sum e\int d^3p\;F\left(\frac{\partial\delta\Psi_{\rm gy}}{\partial\delta\mathbf{B}}\right) \;=\; \sum\int d^3p\;F\;\left(\boldsymbol{\mu}_{\rm gy} \;+\; \boldsymbol{\pi}_{\rm gy}\times\frac{p_\parallel}{mc}\,\widehat{\mathsf{b}}\right), \tag{33}$$

where $\boldsymbol{\pi}_{\rm gy}$ denotes the gyrocenter electric-dipole moment, $\boldsymbol{\mu}_{\rm gy}$ denotes the gyrocenter magnetic-dipole moment, and the gyrocenter magnetization vector (33) includes a moving-electric-dipole contribution $\boldsymbol{\pi}_{\rm gy}\times(p_\parallel/mc)\widehat{\mathsf{b}}$ (Jackson, 1975).

Lastly, the nonlinear gyrokinetic Vlasov-Maxwell equations possess an exact energy conservation law $dE/dt \equiv 0$, where the global gyrokinetic energy integral is

$$E \;=\; \int d^6Z\,F\left(\frac{p_\parallel^2}{2m} + \mu\,B + e\,\delta\Psi_{\rm gy} \;-\; e\left\langle\mathsf{T}_{\rm gy}^{-1}\delta\phi_{\rm gc}\right\rangle\right) \;+\; \int\frac{d^3x}{8\pi}\left(|\nabla\delta\phi|^2 \;+\; |\mathbf{B}+\delta\mathbf{B}|^2\right), \tag{34}$$

where $\mathsf{T}_{\rm gy}^{-1}\delta\phi_{\rm gc}$ denotes the gyrocenter *push-forward* of the perturbed scalar potential. Note that this exact conservation law is either 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, which represent the fundamental tools used in the modern derivation of the nonlinear gyrokinetic equations (1)-(34), will be defined in Section IV. For the remainder of this Section, we discuss various limiting cases of the nonlinear gyrokinetic equations, which are presented in general magnetic geometry (while some applications which we mention, were made in simple geometry).

B. Electrostatic Fluctuations

First, we start the case when only electrostatic fluctuations are present (i.e., $\delta\mathbf{A}\equiv 0$). The electrostatic nonlinear gyrokinetic equations in general geometry (Hahm, 1988) can be used for studies of most drift-wave-type fluctuations driven by the expansion free energy associated with the gradients in density and temperature. Note that sound-wave dynamics as well as linear and nonlinear Landau damping (Sagdeev and Galeev, 1969) are all contained in the nonlinear gyrokinetic formulations. The electrostatic nonlinear gyrokinetic equations can be used for ion dynamics associated with ion-temperature-gradient (ITG) instability, electron drift waves including trapped-electron-mode (TEM), collisionless trapped-ion-modes (TIM), universal and dissipative drift instabilities. These gyrokinetic equations can also be used for electron dynamics of electron-temperature-gradient (ETG) instability. While an unmagnetized "adiabatic" ion response is commonly used for ETG studies, more accurate treatment of ion dynamics associated with ETG instability can be made possible with a gyrokinetic formulation with a proper high-$\mathbf{k}$ behavior. Finally, the nonlinear gyrokinetic formulations can also be used to study zonal flows (Diamond *et al.*, 2005) and Geodesic Acoustic Modes (Winsor *et al.*, 1968), which 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 (27) has the guiding-center form (with $\delta\mathbf{B}_{\rm gy}\equiv 0$), and the effective gyrocenter perturbation potential $\delta\Psi_{\rm gy}$ in the gyrocenter equations of motion (25)-(26) is expressed in simplified form as

$$\delta\Psi_{\rm gy} \;=\; \langle\delta\phi_{\rm gc}\rangle \;-\; \frac{e}{2B}\,\frac{\partial}{\partial\mu}\left\langle\delta\widetilde{\phi}_{\rm gc}^2\right\rangle, \tag{35}$$

which retains full FLR effects in both the linear term and the nonlinear term,[8] where $\delta\widetilde{\phi}_{\rm gc}\equiv\delta\phi_{\rm gc}-\langle\delta\phi_{\rm gc}\rangle$ denotes the gyroangle-dependent part of $\delta\phi_{\rm gc}\equiv\exp(\boldsymbol{\rho}\cdot\nabla)\,\delta\phi$. From gyrokinetic Maxwell's equations, only the gyrokinetic Poisson equation (28) is relevant in the electrostatic limit, where the gyrocenter pull-back $\mathsf{T}_{\rm gy}F$ consistent with the simplified effective gyrocenter perturbation potential (35) is

$$\mathsf{T}_{\rm gy}F \;=\; F \;+\; \frac{e\,\delta\widetilde{\phi}_{\rm gc}}{B}\,\frac{\partial F}{\partial\mu}.$$

Thus, the integrand on the right side of the gyrokinetic Poisson equation (28) includes the polarization term

$$e^{-\boldsymbol{\rho}\cdot\nabla}\left(\mathsf{T}_{\rm gy}F - F\right) = -\,\frac{e\,f_0}{T_\perp}\left(\delta\phi - e^{-\boldsymbol{\rho}\cdot\nabla}\langle\delta\phi_{\rm gc}\rangle\right), \tag{36}$$

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 $T_\perp$). Lastly, the gyrokinetic energy invariant (34) includes the perturbation term

$$e\,\delta\Psi_{\rm gy} \;-\; e\left\langle\mathsf{T}_{\rm gy}^{-1}\delta\phi_{\rm gc}\right\rangle \;=\; \frac{e}{2B}\,\frac{\partial}{\partial\mu}\left\langle\delta\widetilde{\phi}_{\rm gc}^2\right\rangle, \tag{37}$$

which is consistent with the effective gyrocenter perturbation potential (35) and the gyrokinetic Poisson equation (28), with the polarization density (36).

[8] An additional nonlinear term involving the multi-dimensional expression $\widehat{\mathsf{b}}\cdot\langle\nabla\delta\widetilde{\Phi}_{\rm gc}\times\nabla\delta\widetilde{\phi}_{\rm gc}\rangle$, where $\delta\widetilde{\Phi}_{\rm gc}=\int\delta\widetilde{\phi}_{\rm gc}\,d\zeta$, is omitted here for clarity.

The energy-conserving gyrokinetic Vlasov-Poisson equations constructed with the effective gyrocenter perturbation potential (35) and the gyrokinetic polarization density (36) can be written in a more familiar form (Dubin *et al.*, 1983) by Fourier transforming the gyrokinetic Poisson equation (28) into **k**-space:

$$n_0\,\left(|\mathbf{k}|^2\lambda_{\rm Di}^2\right)\,\frac{e\delta\phi_{\mathbf{k}}}{T_{\rm i\perp}} \;=\; \delta n_{\rm i\mathbf{k}} \;-\; \delta n_{\rm e\mathbf{k}}, \tag{38}$$

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

$$\begin{aligned}\delta n_{\rm i\mathbf{k}} \;=\; & \delta N_{\rm i\mathbf{k}} \;-\; n_0\,(1-\Gamma_0)\,\frac{e\delta\phi_{\mathbf{k}}}{T_{\rm i\perp}} \\ & +\; n_0\left[\,\rho_{\rm i}^2\,(i\mathbf{k}_\perp\cdot\nabla\ln n_0)\,(\Gamma_1-\Gamma_0)\,\right]\frac{e\delta\phi_{\mathbf{k}}}{T_{\rm i\perp}}\end{aligned} \tag{39}$$

is expressed in terms of the perturbed ion gyrofluid density $\delta N_{\rm i\mathbf{k}} \equiv \int d^3p\,\left\langle e^{-i\boldsymbol{\rho}\cdot\mathbf{k}_\perp}\right\rangle\delta F_{\rm i\mathbf{k}}$ and $\Gamma_n(b) \equiv I_n(b)e^{-b}$ is expressed in terms of modified Bessel functions I_n (of order n), with $b \equiv |\mathbf{k}_\perp|^2\rho_{\rm i}^2$. Note that, while the last term in Eq. (39) is smaller than the leading term and is neglected by most authors, it is crucial in preserving the correct form of the polarization density (Dubin *et al.*, 1983; Hahm *et al.*, 1988). The invariant energy for these electrostatic gyrokinetic equations is

$$\begin{aligned}E \;=\; & \int d^6Z\,\delta F_{\rm i}\left(\mu B + \frac{m_{\rm i}}{2}\,v_\parallel^2\right) \\ & +\int d^6z\,\delta f_{\rm e}\,\left(\frac{m_{\rm e}}{2}\,v^2\right) \;+\; \int\frac{d^3x}{8\pi}\,|\delta\mathbf{E}|^2 \\ & +\frac{n_0e^2}{2T_{\rm i}}\int\frac{d^3\mathbf{k}}{(2\pi)^3}\,(1-\Gamma_0)|\delta\phi_{\mathbf{k}}|^2,\end{aligned} \tag{40}$$

which provides an accurate linear response for arbitrary $k_\perp\rho_{\rm i}$ and dominant $E\times B$ nonlinearity needed for most applications.

For both simulation and analytic applications, often the distribution function $F = F_0 + \delta F$ is split into the equilibrium part F_0 and the perturbed part δF, with $\delta F/F_0 \sim \epsilon_\delta$. One can also write the equilibrium part and the perturbed part of Eqs. (25)-(26) separately. Then, Eqs. (1)-(26) become

$$\frac{\partial\delta F}{\partial t} + \frac{d\mathbf{Z}}{dt}\cdot\frac{\partial\delta F}{\partial\mathbf{Z}} \;=\; -\,\frac{d\delta\mathbf{Z}}{dt}\cdot\frac{\partial F_0}{\partial\mathbf{Z}}, \tag{41}$$

where the perturbed equations of motion are

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

and the full equations of motion are

$$\frac{d\mathbf{X}}{dt} \;=\; v_\parallel\,\frac{\mathbf{B}^*}{B_\parallel^*} + \frac{c\widehat{\mathsf{b}}}{eB_\parallel^*}\times\Big(\,e\,\nabla\langle\delta\phi_{\rm gc}\rangle \;+\; \mu\,\nabla B\,\Big), \tag{42}$$

and

$$\frac{dv_\parallel}{dt} \;=\; -\frac{\mathbf{B}^*}{mB_\parallel^*}\cdot\Big(\,e\,\nabla\langle\delta\phi_{\rm gc}\rangle \;+\; \mu\,\nabla B\,\Big). \tag{43}$$

Note that the second term on the left hand side of Eq. (41) contains the dominant $E\times B$ nonlinearity and the subdominant velocity-space parallel nonlinearity $[(d\delta v_\parallel/dt)\,\partial\delta F/\partial v_\parallel]$. If we ignore this parallel velocity-space nonlinearity in the last term of Eq. (43)), the physics contained in Eqs. (41)-(43) is essentially the same as the electrostatic limit of the Frieman-Chen gyrokinetic equation (16). One consequence of omitting this term is that now the energy invariant and Eq. (38) are *not* conserved to the same order. While most turbulence simulations have not kept this small subdominant term for simplicity, some simulations (Hatzky *et al.*, 2002; Sydora *et al.*, 1996; Villard *et al.*, 2004b) have kept it. In principle, simulations with this term should have better energy-conservation property and, therefore, less time-accumulated error. This term can, thus, turn out to be crucial in the long-time simulations; this topic is one of active areas of current research.

The nonlinear gyrokinetic Vlasov-Poisson equations (35)-(37) can be further simplified by taking the long-wavelength (drift-kinetic) limit ($k_\perp\rho_{\rm i} \ll 1$) of the nonlinear correction in the effective gyrocenter perturbation potential (35):

$$e\delta\Psi_{\rm gy} \;=\; e\langle\delta\phi_{\rm gc}\rangle \;-\; \frac{m}{2}\,|\delta\mathbf{u}_{\rm E}|^2, \tag{44}$$

where the second term has a definite physical meaning of normalized kinetic energy associated with the perturbed $E\times B$ drift (Scott, 2005). There is a one-to-one correspondence with this term, and the polarization density term in the gyrokinetic Poisson equation, and the sloshing energy term in the energy invariant. Indeed, in the same drift-kinetic limit (Dubin *et al.*, 1983), the linear gyrocenter polarization vector in the gyrokinetic Poisson equation (28) is expressed in terms of the gyrocenter electric-dipole moment

$$\begin{aligned}\boldsymbol{\pi}_{\rm gy} \;\equiv\; & -\,e\,\frac{\partial\delta\Psi_{\rm gy}}{\partial\delta\mathbf{E}_\perp} \\ =\; & -\,\frac{mc^2}{B^2}\,\nabla_\perp\delta\phi \;\equiv\; \frac{c\widehat{\mathsf{b}}}{B}\times\Big(\,m\,\delta\mathbf{u}_{\rm E}\Big),\end{aligned} \tag{45}$$

which is directly related to the nonlinear terms in the effective gyrocenter perturbation potential (44); note that the linear term $\langle\delta\phi_{\rm gc}\rangle$ contains the guiding-center polarization vector, which is automatically included in the definition of the ion gyrofluid density. Note that because the gyrocenter electric-dipole moment (45) is proportional to the particle's mass, the dominant contribution to the polarization density comes from the ion species. Moreover, we note that representing the polarization drift as a shielding term in the gyrokinetic Poisson equation provided one of the principal computational advantages of the gyrokinetic approach.[9] Lastly, the energy invariant

[9] There exists a simple relation between the polarization current $\mathbf{J}_{\rm pol}$ and the polarization density $\rho_{\rm pol}$ (Fong and Hahm, 1999; Krommes, 2002) via continuity equation: $\partial_t\rho_{\rm pol} + \nabla\cdot\mathbf{J}_{\rm pol} = 0$.

consistent with the simplified effective gyrocenter perturbation potential (44) and the gyrokinetic Poisson equation (28), with gyrocenter polarization vector (45), includes the nonlinear term

$$e\delta\Psi_{\rm gy} \;-\; e\left\langle \mathsf{T}_{\rm gy}^{-1}\delta\phi_{\rm gc}\right\rangle \;=\; \frac{m}{2}\,|\delta\mathbf{u}_{\rm E}|^{2}. \qquad (46)$$

The simplified nonlinear gyrokinetic Vlasov-Poisson equations based on Eqs. (44)-(46), thus, highlight the three pillars of nonlinear gyrokinetic theory: a gyrocenter Hamiltonian (44) that contains nonlinear (quadratic) terms, a gyrokinetic Poisson equation that contains a polarization density derived from the nonlinear gyrocenter Hamiltonian (45), and an energy invariant that includes all the relevant nonlinear coupling terms (46). An energy-conserving set of nonlinear electrostatic gyrofluid equations, with full FLR effects retained in the linear terms and the nonlinear terms expressed in the drift-kinetic limit, was derived by Strintzi *et al.* (2005) by variational methods.

While the nonlinear electrostatic gyrokinetic equations 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 percent, and the nonlinear corrections to the effective potential are indeed small. However, these nonlinear corrections may play important roles in edge turbulence where the relative fluctuation amplitude is high, typically greater than 10 percent (see Fig.2).

C. Shear-Alfvenic Magnetic Fluctuations

It has been shown by Hahm *et al.* (1988) that the reduced magnetohydrodynamic (MHD) equations (whose derivation makes use of the ratio $k_\|/k_\perp \ll 1$ as an expansion parameter) can be recovered from the electromagnetic nonlinear gyrokinetic equations. Note, here, that for finite-β plasmas (with $m_{\rm e}/m_{\rm i} < \beta \ll 1$), perpendicular magnetic fluctuations $\delta\mathbf{B}_\perp \equiv \nabla_\perp\delta A_\| \times \widehat{\mathsf{b}}$ become important as the magnetic-flutter $v_\|\,\delta\mathbf{B}_\perp/B$ becomes comparable to the perturbed (linear) $E\times B$ velocity $(c\widehat{\mathsf{b}}/B)\times\nabla_\perp\delta\phi$ (i.e., $\delta A_\|\,v_\|/c \sim \delta\phi$). Therefore, physics associated with shear-Alfven waves and instabilities (which include a wide variety of MHD instabilities) can be studied using the gyrokinetic approach. Early applications to MHD modes consisted of the various hybrid approaches with nonlinear gyrokinetic description of energetic particle dynamics and MHD description of bulk plasmas. The nonlinear gyrokinetic approach has also been applied to the tearing and kink instabilities for which the free energy comes from the radial gradient of 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. It should also be noted that the electromagnetic modifications of drift wave turbulence, which is often referred to as "drift-Alfvén" turbulence, is an outstanding topic in magnetic confinement physics. There have been nonlinear simulations based on nonlinear gyrokinetic formulations (some examples of applications are listed in Table I).

Since the early days of derivation of modern nonlinear gyrokinetic formulation (Hahm *et al.*, 1988), it became apparent that there can be at least two different versions of electromagnetic nonlinear gyrokinetic equations. One version is the Hamiltonian formulation, which uses the parallel canonical momentum $p_\|$ as an independent variable, the other is the Symplectic formulation, where the parallel velocity $v_\|$ is used as an independent variable. Each approach has its own advantages and drawbacks. We confine ourselves only to the case where the nonlinear modifications of perturbed potential are expressed in the drift-kinetic limit, which may turn out to be important in the nonlinear gyrokinetic simulation of edge turbulence as stated before.

1. Hamiltonian ($p_\|$) formulation

In the Hamiltonian formulation, the magnetic perturbation $\delta A_\| \equiv \widehat{\mathsf{b}}\cdot\delta\mathbf{A}$ is treated as part of the gyrocenter Hamiltonian, with all linear and nonlinear $\delta A_\|$-terms entirely included in the effective gyrocenter potential $\delta\Psi_{\rm gy}$ in Eqs. (25)-(26) (i.e., the linear perturbation potential $\delta\psi_{\rm gc} = \delta\phi_{\rm gc} - (p_\|/mc)\,\delta A_{\|{\rm gc}}$ is manifestly covariant) while the symplectic magnetic perturbation is $\delta\mathbf{A}_{\rm gy} \equiv 0$. It is worth noting that the gyrocenter parallel velocity $v_\| \equiv \widehat{\mathsf{b}}\cdot d\overline{\mathbf{X}}/dt$ is expressed in terms of the gyrocenter parallel *canonical* momentum $p_\|$ and the perturbed parallel vector potential $e\,\delta A_\|/c$:

$$v_\| \;\equiv\; \frac{1}{m}\left(p_\| \;-\; \epsilon_\delta\,\frac{e}{c}\,\langle\delta A_{\|{\rm gc}}\rangle + \cdots\right).$$

From this reason, this Hamiltonian formulation is sometimes referred to as the "canonical-momentum" formulation or more casually the P_z-formulation following the terminology from an early work (Hahm *et al.*, 1988) in a staight 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. (26). This feature is computationally desirable (Hahm *et al.*, 1988) and is one of the motivations for the canonical-momentum formulation alongside the explicit manifestation of covariance (Krommes and Kim, 1988). Second, in the perpendicular gyrocenter velocity (25), the second term $\widehat{\mathsf{b}}\times\nabla\langle\delta\psi_{\rm gc}\rangle$ 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 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). While the test-particle transport in the stochastic magnetic fields has been thoroughly studied (Krommes *et al.*, 1983; Rechester and Rosenbluth, 1978), a fully self-consistent calculation of the net transport due to this mechanism in collisionless plasmas is extremely difficult (Krommes and Kim, 1988; Terry *et al.*, 1986, 1988; Thoul *et al.*, 1987, 1988).

For the Hamiltonian gyrocenter formulation of shear-Alfvenic fluctuations, the effective gyrocenter perturbation potential is given in simplified form as (Hahm *et al.*, 1988)

$$e\delta\Psi_{\rm gy} \;=\; e\langle\delta\psi_{\rm gc}\rangle \;+\; \frac{e^2\delta A_\|^2}{2\,mc^2} \;-\; \frac{m}{2}\left|\delta\mathbf{u}_{\rm E} + \frac{p_\|}{m}\,\frac{\delta\mathbf{B}_\perp}{B}\right|^2, \tag{47}$$

where the linear term retains full FLR effects while the nonlinear terms are given in the drift-kinetic limit, with $\delta\mathbf{B}_\perp \equiv \nabla\delta A_\| \times \widehat{\mathsf{b}}$ denoting the perturbed magnetic field. The gyrokinetic Maxwell's equations now consist of the gyrokinetic Poisson equation (28), with the linear gyrocenter polarization vector expressed in terms of the gyrocenter electric-dipole moment

$$\begin{aligned}\boldsymbol{\pi}_{\rm gy} \;&\equiv\; -\,e\,\frac{\partial\delta\Psi_{\rm gy}}{\partial\delta\mathbf{E}_\perp} \qquad (48)\\ &=\; -\,\frac{mc^2}{B^2}\,\left(\nabla_\perp\delta\phi \;-\; \frac{p_\|}{mc}\,\nabla_\perp\delta A_\|\right),\\ &=\; \frac{c\widehat{\mathsf{b}}}{B}\times\left(m\,\delta\mathbf{u}_{\rm E} \;+\; p_\|\,\frac{\delta\mathbf{B}_\perp}{B}\right),\end{aligned}$$

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

$$-\,\nabla_\perp^2\delta A_\| \;=\; -\,\left(\frac{\omega_{\rm p}^2}{c^2}\right)\delta A_\| \;+\; \frac{4\pi}{c}\,(J_{{\rm i}\|} \;-\; j_{{\rm e}\|}) \;+\; 4\pi\,\nabla\cdot\left(\mathbf{M}_{\rm gy}\times\widehat{\mathsf{b}}\right), \tag{49}$$

where the linear gyrocenter magnetization vector

$$\begin{aligned}\mathbf{M}_{\rm gy} \;&\equiv\; -\sum e\int d^3p\,f_0\,\left(\frac{\partial\delta\Psi_{\rm gy}}{\partial\delta\mathbf{B}_\perp}\right) \qquad (50)\\ &=\; \sum\int d^3p\,f_0\,\left(\boldsymbol{\pi}_{\rm gy}\times\frac{p_\|}{mc}\,\widehat{\mathsf{b}}\right)\end{aligned}$$

only displays the moving-electric-dipole contribution. The explicit appearance of the collisionless skin depth $(\omega_{\rm p}/c)$ on the right side of Eq. (49), whose dominant contribution comes from the electron species, suggests a possibility that it can be a characteristic correlation length for electromagnetic turbulence in magnetized plasmas. While this fact alone is not sufficient theoretical evidence, turbulence at the scale of collisonless skin depth has been simulated (Horton *et al.*, 2000; Yagi *et al.*, 1995) and measured from experiments (Wong *et al.*, 1997). This "canonical-momentum" formulation has been widely used for kink mode, tearing mode, and drift-tearing mode nonlinear gyrokinetic simulations as listed in Table I. Lastly, the gyrokinetic energy invariant consistent with the effective gyrocenter perturbation potential (47), the gyrokinetic polarization density (48), and the gyrokinetic parallel magnetization current (50) includes the terms

$$e\delta\Psi_{\rm gy} \;-\; e\langle\mathsf{T}_{\rm gy}^{-1}\delta\phi_{\rm gc}\rangle \;=\; -\,\frac{ep_\|}{mc}\,\langle\delta A_{\|{\rm gc}}\rangle \;+\; \frac{e^2\delta A_\|^2}{2\,mc^2} \;+\; \frac{1}{2}\left(m\,|\delta\mathbf{u}_{\rm E}|^2 \;-\; \frac{p_\|^2}{mB^2}\,|\delta\mathbf{B}_\perp|^2\right). \tag{51}$$

We conclude that, in the Hamiltonian gyrocenter model of shear-Alfvenic fluctuations, the magnetic-flutter perturbed motion changes the gyrocenter polarization density (48), while the perturbed $E\times B$ motion contributes to the gyrocenter magnetization current (50). Note that, if one decides to drop some 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 ($v_\|$) formulation

While the Hamiltonian formulation has some computational advantages (Hahm *et al.*, 1988) and is in a mathematically-concise form readily 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. It is also true that, in the Hamiltonian formulation, it is often inefficient to calculate the relatively large terms such as $p_\|$ and $(e/c)\,\delta A_\|$ which appear explicitly, up to an accuracy which is sufficient for calculation of their difference $mv_\|$ (Chen *et al.*, 2003; Lin and Wang *et al.*, 2005; Mishchenko *et al.*, 2004). In this sense alone, it is more efficient computationally and is easier to understand physics if a smaller term "$v_\|$" is used as an independent variable. In the symplectic formulation, the perturbed parallel vector potential $\delta A_\|$ appears explicitly in the gyrocenter Poisson bracket (where $\delta\mathbf{A}_{\rm gy} = \langle\delta A_{\|{\rm gc}}\rangle\,\widehat{\mathsf{b}}$ and $\delta\mathbf{B}_{\rm gy} = \nabla\times\delta\mathbf{A}_{\rm gy}$), not in the gyrocenter Hamiltonian part. As a consequence, the resulting Euler-Lagrange equations contain the induction part of the electric field with $\partial\langle\delta A_{\|{\rm gc}}\rangle/\partial t$. This is not a computationally attractive feature as stated before.

This symplectic version of the electromagnetic nonlinear gyrokinetic equation is more suitable in showing its relation to various reduced fluid equations by taking moments (Brizard, 1992). For instance, one of the key points in understanding the shear-Alfvén physics in the context of the electromagnetic nonlinear gyrokinetic formulation is to recognize that the "vorticity evolution" in the reduced MHD equation is equivalent to the evolution of the ion polarization density, which is the difference between the ion gyrofluid density and the electron density (Brizard, 1992; Hahm *et al.*, 1988). 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 gyro-center drift due to magnetic field inhomogeneity (for the driving term for ballooning and interchange instability) and the variation of the equilibrium current along the perturbed magnetic field (for the driving term for kink, tearing, and peeling instabilties).

For the Symplectic formulation of shear-Alfvenic fluctuations, the effective gyrocenter perturbation potential is given in simplified form as (Hahm *et al.*, 1988)

$$\begin{aligned} e\delta\Psi_{\rm gy} \;=\; & e\langle\delta\phi_{\rm gc}\rangle \;+\; \frac{\mu}{2B}\,|\delta\mathbf{B}_\perp|^2 \\ & -\;\frac{m}{2}\left|\delta\mathbf{u}_{\rm E} + \frac{p_\|}{m}\,\frac{\delta\mathbf{B}_\perp}{B}\right|^2, \end{aligned} \tag{52}$$

where the linear term, which retains full FLR effects, includes only the perturbed electrostatic potential (the symplectic perturbed magnetic potential $\delta\mathbf{A}_{\rm gy} \equiv \langle\delta A_{\|{\rm gc}}\rangle\,\widehat{\mathsf{b}}$ appears in the modified magnetic field $\mathbf{B}^*$ and the inductive part $\partial_t\langle\delta A_{\|{\rm gc}}\rangle$ of the parallel perturbed electric field) while the nonlinear terms are given in the drift-kinetic limit. The gyrokinetic Maxwell's equations now consist of the gyrokinetic Poisson equation (28), where the gyrocenter electric-dipole moment $\boldsymbol{\pi}_{\rm gy}$ is given by Eq. (48), and the gyrokinetic parallel Ampère equation

$$-\,\nabla_\perp^2\delta A_\| \;=\; \frac{4\pi}{c}\,(J_{{\rm i}\|} \,-\, j_{{\rm e}\|}) \;+\; 4\pi\,\nabla\cdot\left(\mathbf{M}_{\rm gy}\times\widehat{\mathsf{b}}\right), \tag{53}$$

where the linear gyrocenter magnetization vector has the moving-electric-dipole contribution shown in Eq. (50) and an intrinsic gyrocenter magnetic-dipole moment contribution

$$\boldsymbol{\mu}_{\rm gy} \;\equiv\; -\,\mu\,\frac{\delta\mathbf{B}_\perp}{B}. \tag{54}$$

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

$$\begin{aligned} e\delta\Psi_{\rm gy} \;-\; e\langle\mathsf{T}_{\rm gy}^{-1}\delta\phi_{\rm gc}\rangle \;=\; & \frac{\mu}{2B}\,|\delta\mathbf{B}_\perp|^2 \\ +\;\frac{m}{2}\left(|\delta\mathbf{u}_{\rm E}|^2 \;-\; \frac{p_\|^2}{m^2B^2}\,|\delta\mathbf{B}_\perp|^2\right). \end{aligned} \tag{55}$$

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) by 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 $\omega \sim k_\perp v_{\rm A}$, then $\omega/\Omega \sim k_\perp v_{\rm A}/\Omega \sim k_\perp\rho_{\rm i}/\beta^{1/2}$. Therefore, with the full FLR description in gyrokinetics ($k_\perp\rho_{\rm i} \sim 1$), 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. Hence, to describe the compressional Alfvén wave, it is necessary to use a drift-kinetic description ($k_\perp\rho_{\rm i} \ll 1$). It has been shown that it is possible to decouple the gyro-motion from dynamics associated with high-frequency waves with $\omega/\Omega > 1$ and develop a high-frequency linear gyrokinetic equation (Lashmore-Davies and Dendy, 1989; Tsai *et al.*, 1984). It has also been shown that the phase-space Lagrangian and Lie-transform perturbation method can be very 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). However, a satisfactory nonlinear high-frequency gyrokinetic formulation has not been derived to date and we do not discuss the progress in the linear high-frequency gyrokinetic formulation in this review.

Although the compressional Alfvén wave does not exist within the nonlinear (low-frequency) gyrokinetic formulation, the compressional component ($\delta B_\| \equiv \widehat{\mathsf{b}}\cdot\delta\mathbf{B}$) of the perturbed magnetic field $\delta\mathbf{B}$ becomes gradually important as the plasma β value is increased. So one must keep $\delta B_\|$ 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 Section V. Here, we use the Hamiltonian gyrocenter model (with $\delta\mathbf{A}_{\rm gy} \equiv 0$) and express the effective gyrocenter perturbation potential in the drift-kinetic limit as

$$\begin{aligned} e\delta\Psi_{\rm gy} \;=\; & e\left(\delta\phi \;-\; \frac{p_\|}{mc}\,\delta A_\|\right) \;+\; \mu\,\delta B_\| \;-\; \frac{e}{c}\,\delta\mathbf{A}_\perp\cdot\mathbf{v}_{\rm gc} \\ & +\;\frac{e^2\delta A_\|^2}{2\,mc^2} \;-\; \frac{m}{2}\left|\delta\mathbf{u}_{\rm E} \;+\; \frac{p_\|}{m}\,\frac{\delta\mathbf{B}_\perp}{B}\right|^2 \\ & -\;\frac{e}{c}\,\delta\mathbf{A}_\perp\cdot\left(\delta\mathbf{u}_{\rm E} \;+\; \frac{p_\|}{m}\,\frac{\delta\mathbf{B}_\perp}{B}\right), \end{aligned} \tag{56}$$

where we have added the linear term $e\,\delta\mathbf{A}_\perp\cdot\mathbf{v}_{\rm gc}/c$, which includes the perpendicular guiding-center velocity $\mathbf{v}_{\rm gc}$ (involving the grad-B and curvature drifts). The linear perpendicular gyrocenter dynamics is represented by the linear perturbed $E\times B$ velocity ($\widehat{\mathsf{b}}\times\nabla\delta\phi$), the linear magnetic-flutter velocity ($v_\|\,\delta\mathbf{B}_\perp$), and the linear perturbed grad-B drift ($\mu\,\widehat{\mathsf{b}}\times\nabla\delta B_\|$). While in most fusion plasmas (with $\beta < 1$), the radial transport due to this last term is subdominant to the 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 being extended to high-β astrophysical plasmas (Howes *et al.*, 2006). Although the nonlinear terms in the gyrocenter potential (56) are small compared to the linear terms, they nonetheless play an important role in contributing to the gyrocenter polarization and magnetization vectors

$$\boldsymbol{\pi}_{\rm gy} \;=\; \frac{c\widehat{\mathsf{b}}}{B}\times\left(\frac{e}{c}\,\delta\mathbf{A}_\perp \;+\; m\,\delta\mathbf{u}_{\rm E} \;+\; p_\parallel\,\frac{\delta\mathbf{B}_\perp}{B}\right), \tag{57}$$

$$\boldsymbol{\mu}_{\rm gy} \;=\; -\,\mu\,\widehat{\mathsf{b}}. \tag{58}$$

Note, here, that the $\delta\mathbf{A}_\perp$-contribution to the gyrocenter polarization density vanishes if we assume quasineutrality. Moreover, the lowest-order contribution of the intrinsic magnetization current, derived from the gyrocenter magnetic-dipole moment (58), to the perpendicular gyrokinetic Ampère equation yields the perpendicular pressure balance condition $\delta B_\parallel + 4\pi\,\delta P_\perp/B = 0$, where $\delta P_\perp$ denotes the perturbed perpendicular (total) pressure (Brizard, 1992; Tang *et al.*, 1980); a straightforward demonstration of this condition can be found in Roach *et al.* (2005). Lastly, the corresponding energy invariant includes the terms

$$\begin{aligned} e\delta\Psi_{\rm gy} \;-\; e\,\langle\mathsf{T}_{\rm gy}^{-1}\delta\phi_{\rm gc}\rangle \;=\; &-\;\frac{ep_\parallel}{mc}\,\delta A_\parallel \;+\; \mu\,\delta B_\parallel \\ &-\;\frac{e}{c}\,\delta\mathbf{A}_\perp\boldsymbol{\cdot}\left(\mathbf{v}_{\rm gc} \;+\; \frac{p_\parallel}{m}\,\frac{\delta\mathbf{B}_\perp}{B}\right) \\ &+\;\frac{1}{2}\left(m\,|\delta\mathbf{u}_{\rm E}|^2 \;-\; \frac{p_\parallel^2}{mB^2}\,|\delta\mathbf{B}_\perp|^2\right). \end{aligned} \tag{59}$$

IV. LIE-TRANSFORM PERTURBATION THEORY

After having presented simple forms of the nonlinear gyrokinetic equations in the previous Section, we now focus our attention on the transformation from particle phase-space coordinates 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 this Section with a brief introduction to the extended phase-space Lagrangian formulation of charged-particle dynamics in a time-dependent electromagnetic field. Here, 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 $\mathbf{p} = m\mathbf{v}$, and the canonically-conjugate energy-time (w, t) coordinates; in this work, we use the convenient Minkowski space-time metric $\mathsf{g} = \mathrm{diag}(-1, 1, 1, 1)$ whenever we need a concise covariant expression. The use of an eight-dimensional representation of phase space is motivated by the fact 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 $dw/dt \equiv e\,\partial\psi/\partial t$, where $\psi \equiv \phi - \mathbf{A}\boldsymbol{\cdot}\mathbf{v}/c = -\,A^\mu v_\mu/c$ denotes the effective electromagnetic potential. Hence, 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.

First, 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 $\{\ ,\ \}$, which satisfies the following properties (valid for arbitrary functions f, g, and h): the antisymmetry property

$$\{f,\,g\} \;=\; -\,\{g,\,f\}, \tag{60}$$

the Leibnitz rule

$$\{f,\,(g\,h)\} \;=\; \{f,\,g\}\,h \;+\; g\,\{f,\,h\}, \tag{61}$$

and the Jacobi identity

$$0 \;=\; \{f,\,\{g,\,h\}\} \;+\; \{g,\,\{h,\,f\}\} \;+\; \{h,\,\{f,\,g\}\}. \tag{62}$$

To be more specific, we introduce the general form for the Poisson bracket $\{f,\,g\}$:

$$\{f,\,g\} \;\equiv\; \frac{\partial f}{\partial\mathcal{Z}^a}\,J^{ab}\,\frac{\partial g}{\partial\mathcal{Z}^b}, \tag{63}$$

where J^{ab} denotes the components of the Poisson tensor and summation over repeated indices is, henceforth, implied (here, latin letters $a, b, c, ...$ go from 1 to 8 while greek letters $\mu, \nu, ...$ go from 0 to 3). The bilinear form of the Poisson bracket (63) automatically satisfies the Leibnitz rule (61), the antisymmetry property (60) requires that the Poisson tensor be antisymmetric, $J^{ba} = -\,J^{ab}$, and the Jacobi identity (62) requires that

$$0 \;=\; J^{a\ell}\,\partial_\ell J^{bc} \;+\; J^{b\ell}\,\partial_\ell J^{ca} \;+\; J^{c\ell}\,\partial_\ell J^{ab}, \tag{64}$$

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

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

immediately satisfies all three properties.

Next, Hamilton's equations are 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

$$\mathcal{H}(\mathcal{Z}) \;=\; \frac{|\mathbf{p}|^2}{2m} \;+\; e\,\phi \;-\; w \;\equiv\; H(\mathbf{z}, t) \;-\; w, \tag{65}$$

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

$$\mathcal{H}(\mathcal{Z}) \;=\; H(\mathbf{z},t) \;-\; w \;=\; 0. \tag{66}$$

Note that, within the canonical formalism, the Poisson-bracket structure is independent of the electromagnetic field and the Hamiltonian depends explicitly on the electromagnetic potentials $(\phi, \mathbf{A})$. Within the noncanonical formalism, however, the Hamiltonian retains only 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, 1989), for a charged particle in eight-dimensional extended phase space is expressed in noncanonical form as

$$\begin{aligned}\widehat{\Gamma} \;&=\; \left(\frac{e}{c}\,\mathbf{A} \;+\; \mathbf{p}\right)\boldsymbol{\cdot}\mathsf{d}\mathbf{x} \;-\; w\,\mathsf{d}t \;-\; \mathcal{H}\,\mathsf{d}\tau \\ &\equiv\; \Gamma_a(\mathcal{Z})\,\mathsf{d}\mathcal{Z}^a \;-\; \mathcal{H}(\mathcal{Z})\,\mathsf{d}\tau,\end{aligned} \tag{67}$$

where Γ_a are known as the *symplectic* components of the extended-phase-space Lagrangian $\widehat{\Gamma}$, and τ denotes the Hamiltonian orbit parameter. In Eq. (67), the symbol d denotes an *exterior* derivative with the property

$$\mathsf{d}^2 f \;=\; \mathsf{d}\left(\partial_a f\,\mathsf{d}\mathcal{Z}^a\right) \;=\; \partial^2_{ab} f\,\mathsf{d}\mathcal{Z}^a\wedge\mathsf{d}\mathcal{Z}^b \;=\; 0, \tag{68}$$

which holds for any scalar field f, where the *wedge* product $\wedge$ is antisymmetric (i.e., $\mathsf{d}f\wedge\mathsf{d}g = -\,\mathsf{d}g\wedge\mathsf{d}f$); we will use the standard-derivative notation d whenever the exterior-derivative properties are not involved (see Appendix A for further details). Note that, as a result of property (68), we may add an arbitrary gauge term $\mathsf{d}\mathcal{S}$ to the extended-phase-space Lagrangian (67) without modifying the Hamiltonian dynamics.

Next, to obtain the extended Hamilton's equations of motion from the phase-space Lagrangian (67), we introduce the single-particle action integral

$$S \;=\; \int\widehat{\Gamma} \;=\; \int_{\tau_1}^{\tau_2}\left(\Gamma_a\,\frac{d\mathcal{Z}^a}{d\tau} \;-\; \mathcal{H}\right)d\tau, \tag{69}$$

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

$$\begin{aligned}0 \;&=\; \int\left(\delta\mathcal{Z}^a\,\frac{\partial\Gamma_b}{\partial\mathcal{Z}^a}\,d\mathcal{Z}^b \;+\; \Gamma_a\,d\delta\mathcal{Z}^a \;-\; \delta\mathcal{Z}^a\,\frac{\partial\mathcal{H}}{\partial\mathcal{Z}^a}\,d\tau\right) \\ &=\; \int\delta\mathcal{Z}^a\left[\,\omega_{ab}\,d\mathcal{Z}^b \;-\; \frac{\partial\mathcal{H}}{\partial\mathcal{Z}^a}\,d\tau\,\right]\end{aligned} \tag{70}$$

where

$$\omega_{ab} \;\equiv\; \frac{\partial\Gamma_b}{\partial\mathcal{Z}^a} \;-\; \frac{\partial\Gamma_a}{\partial\mathcal{Z}^b} \tag{71}$$

denotes a component of the 8×8 antisymmetric Lagrange two-form $\boldsymbol{\omega}\equiv\mathsf{d}\Gamma$ (Goldstein *et al.*, 2002), and integration by parts of the second term was performed (with the usual assumption of virtual displacements $\delta\mathcal{Z}^a$ vanishing at the end points). Hence, stationarity of the particle action (69) yields the extended phase-space Euler-Lagrange equations

$$\omega_{ab}\,\frac{d\mathcal{Z}^b}{d\tau} \;=\; \frac{\partial\mathcal{H}}{\partial\mathcal{Z}^a}. \tag{72}$$

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

$$(\boldsymbol{\omega}^{-1})^{ab} \;\equiv\; \{\mathcal{Z}^a,\;\mathcal{Z}^b\} \;=\; J^{ab}(\mathcal{Z}). \tag{73}$$

Using the identity relation

$$J^{ca}\,\omega_{ab} \;=\; \delta^c{}_b, \tag{74}$$

the Euler-Lagrange equations (72) become the extended Hamilton's equations

$$\frac{d\mathcal{Z}^a}{d\tau} \;=\; J^{ab}\,\frac{\partial\mathcal{H}}{\partial\mathcal{Z}^b} \;=\; \{\mathcal{Z}^a,\;\mathcal{H}\}. \tag{75}$$

By using the identity (74), it can be shown that the Jacobi identity (64) holds if the Lagrange matrix satisfies the identity $\mathsf{d}\boldsymbol{\omega}=0$, or

$$\partial_a\,\omega_{bc} \;+\; \partial_b\,\omega_{ca} \;+\; \partial_c\,\omega_{ab} \;=\; 0, \tag{76}$$

which is automatically satisfied since $\omega\equiv\mathsf{d}\Gamma$ 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 $\Gamma\rightarrow\omega = \mathsf{d}\Gamma\rightarrow J = \omega^{-1}$ automatically satisfies the Jacobi identity (62).

Using the symplectic part of the extended phase-space Lagrangian (67), the Lagrange two-form $\boldsymbol{\omega}\equiv\mathsf{d}\Gamma$ is

$$\begin{aligned}\boldsymbol{\omega} \;&=\; \mathsf{d}p_\mu\,\wedge\,\mathsf{d}x^\mu \;+\; \frac{e}{2c}\,\epsilon_{ijk}\,B^k\,\mathsf{d}x^i\,\wedge\,\mathsf{d}x^j \\ &\quad-\;\frac{e}{c}\,\frac{\partial A_i}{\partial t}\,\mathsf{d}x^i\,\wedge\,\mathsf{d}t,\end{aligned} \tag{77}$$

from which, using the inverse relation (73), we construct the extended *noncanonical* Poisson bracket

$$\begin{aligned}\{f,\;g\} \;&=\; \frac{\partial f}{\partial x^\mu}\,\frac{\partial g}{\partial p_\mu} \;-\; \frac{\partial f}{\partial p_\mu}\,\frac{\partial g}{\partial x^\mu} \;+\; \frac{e\mathbf{B}}{c}\boldsymbol{\cdot}\frac{\partial f}{\partial\mathbf{p}}\times\frac{\partial g}{\partial\mathbf{p}} \\ &\quad-\;\frac{e}{c}\,\frac{\partial\mathbf{A}}{\partial t}\boldsymbol{\cdot}\left(\frac{\partial f}{\partial w}\,\frac{\partial g}{\partial\mathbf{p}} \;-\; \frac{\partial f}{\partial\mathbf{p}}\,\frac{\partial g}{\partial w}\right).\end{aligned} \tag{78}$$

The Hamiltonian dynamics in extended phase space is, thus, expressed as

$$\begin{aligned}\frac{d\mathbf{x}}{dt} \;&=\; \frac{\partial\mathcal{H}}{\partial\mathbf{p}} \;=\; \mathbf{v}, \\ \frac{d\mathbf{p}}{dt} \;&=\; -\,\nabla\mathcal{H} \;+\; \frac{e}{c}\left(\frac{\partial\mathbf{A}}{\partial t}\,\frac{\partial\mathcal{H}}{\partial w} \;+\; \frac{\partial\mathcal{H}}{\partial\mathbf{v}}\times\mathbf{B}\right) \\ &=\; e\left(\mathbf{E} \;+\; \frac{\mathbf{v}}{c}\times\mathbf{B}\right), \\ \frac{dw}{dt} \;&=\; \frac{\partial\mathcal{H}}{\partial t} \;-\; \frac{e}{mc}\,\frac{\partial\mathbf{A}}{\partial t}\boldsymbol{\cdot}\frac{\partial\mathcal{H}}{\partial\mathbf{v}} \;=\; e\left(\frac{\partial\phi}{\partial t} \;-\; \frac{\mathbf{v}}{c}\boldsymbol{\cdot}\frac{\partial\mathbf{A}}{\partial t}\right),\end{aligned}$$

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)

$$\widehat{\Gamma}' \;\equiv\; \Gamma_a\,\mathsf{d}\mathcal{Z}^a \;-\; \mathcal{H}\,\mathsf{d}\tau \;-\; \mathcal{S}\,\mathsf{d}\epsilon, \tag{79}$$

where the symplectic components Γ_a and the Hamiltonian $\mathcal{H}$ now depend on the perturbation parameter ϵ (e.g., either $\epsilon_{\rm B}$ 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 (79), we construct the action *path*-integral $S'_C = \int_C \widehat{\Gamma}'$ 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

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

whose derivation is similar to Eq. (70), now yields the extended perturbed Hamilton's equations

$$\frac{d\mathcal{Z}^a}{d\tau} \;=\; \{\mathcal{Z}^a,\, \mathcal{H}\}, \tag{80}$$

$$\frac{d\mathcal{Z}^a}{d\epsilon} \;=\; \{\mathcal{Z}^a,\, \mathcal{S}\} \;-\; \frac{\partial\Gamma_b}{\partial\epsilon}\,\left\{\mathcal{Z}^b,\, \mathcal{Z}^a\right\}, \tag{81}$$

where Eq. (80) is identical to Eq. (75) except that the extended Hamiltonian $\mathcal{H}$ and symplectic components Γ_a now depend on the perturbation parameter ϵ, while Eq. (81) determines how particle orbits evolve under the perturbation ϵ-flow.

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

$$0 \;=\; \int_C \widehat{\Gamma}' \;-\; \int_{\overline{C}} \widehat{\Gamma}' \;\equiv\; \oint_{\partial D} \widehat{\Gamma}' \;=\; \int_D \mathsf{d}\widehat{\Gamma}',$$

where D is the area enclosed by the closed path $\partial D \equiv C - \overline{C}$. Here, the two-form $\mathsf{d}\widehat{\Gamma}'$ on the $(\tau,\, \epsilon)$-parameter space is

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

where Eqs. (80)-(81) were used. The condition of path independence requires that $\mathsf{d}\widehat{\Gamma}' = 0$, which yields the Hamiltonian perturbation equation

$$\frac{d\mathcal{S}}{d\tau} \;\equiv\; \{\mathcal{S},\, \mathcal{H}\} \;=\; \frac{\partial\mathcal{H}}{\partial\epsilon} \;-\; \frac{\partial\Gamma_a}{\partial\epsilon}\,\{\mathcal{Z}^a,\, \mathcal{H}\}, \tag{82}$$

relating the generating scalar field $\mathcal{S}$ to the perturbation-parameter dependence of the extended Hamiltonian ($\partial_\epsilon\mathcal{H} = \phi_1 + \cdots$) and Poisson bracket ($\partial_\epsilon\Gamma_a = \mathbf{A}_1\cdot\partial\mathbf{x}/\partial\mathcal{Z}^a + \cdots$). Here, the perturbed evolution operator $d/d\tau = d_0/d\tau + \cdots$ and the generating function $\mathcal{S} = S_1 + \cdots$ 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

$$S_1 \;\equiv\; \left(\frac{d_0}{d\tau}\right)^{-1}\left[\,e\,\phi_1 \;-\; e\,\mathbf{A}_1\cdot\frac{\mathbf{v}_0}{c}\,\right], \tag{83}$$

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 the higher-order terms S_n (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 (82) arises naturally within the context of the dynamical reduction of single-particle Hamilton's equations (80) through the decoupling of fast *orbital* time scales from the relevant electromagnetic fluctutation time scales. The most efficient method for deriving reduced Hamilton's 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.[10]

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):

$$\mathcal{T}_\epsilon:\; \mathcal{Z} \;\rightarrow\; \overline{\mathcal{Z}}(\mathcal{Z};\epsilon) \;\equiv\; \mathcal{T}_\epsilon\mathcal{Z}, \tag{84}$$

[10] Hamiltonian 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.

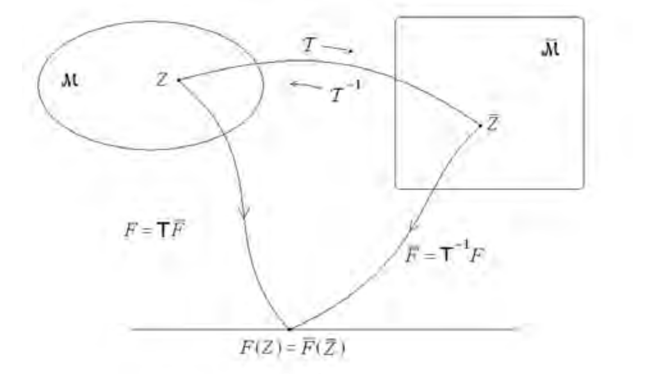

FIG. 5 The phase-space transformation $\overline{\mathcal{Z}} = \mathcal{T}\mathcal{Z}$ and its inverse $\mathcal{Z} = \mathcal{T}^{-1}\overline{\mathcal{Z}}$ induce a pull-back operator $\mathcal{F} = \mathsf{T}\overline{\mathcal{F}}$ and a push-forward operator $\overline{\mathcal{F}} = \mathsf{T}^{-1}\mathcal{F}$.

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

$$\mathcal{T}_\epsilon^{-1} : \overline{\mathcal{Z}} \rightarrow \mathcal{Z}(\overline{\mathcal{Z}};\epsilon) \equiv \mathcal{T}_\epsilon^{-1}\overline{\mathcal{Z}}, \tag{85}$$

with $\mathcal{Z}(\overline{\mathcal{Z}};0) = \overline{\mathcal{Z}}$, 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 $(\mathsf{G}_1, \mathsf{G}_2, ...)$ as

$$\mathcal{T}_\epsilon^{\pm 1} \equiv \exp\left(\pm \sum_{n=1}^{\infty} \epsilon^n \, \mathsf{G}_n \cdot \mathsf{d}\right), \tag{86}$$

where the nth-order generating vector field G_n is chosen to remove the fast time scale at order ϵ^n from the perturbed Hamiltonian dynamics. Here, we explicitly write the near-identity transformations (84) and (85) as

$$\overline{\mathcal{Z}}^a(\mathcal{Z},\epsilon) = \mathcal{Z}^a + \epsilon\, G_1^a + \epsilon^2\left(G_2^a + \frac{1}{2}\, G_1^b \, \frac{\partial G_1^a}{\partial \mathcal{Z}^b}\right) + \cdots, \tag{87}$$

and

$$\mathcal{Z}^a(\overline{\mathcal{Z}},\epsilon) = \overline{\mathcal{Z}}^a - \epsilon\, G_1^a - \epsilon^2\left(G_2^a - \frac{1}{2}\, G_1^b \, \frac{\partial G_1^a}{\partial \overline{\mathcal{Z}}^b}\right) + \cdots, \tag{88}$$

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

Next, using the transformation (84), we define the *pull-back* operator on scalar fields (Abraham and Marsden, 1978; Littlejohn, 1982a) *induced* by the near-identity transformation (84):

$$\mathsf{T}_\epsilon : \overline{\mathcal{F}} \rightarrow \mathcal{F} \equiv \mathsf{T}_\epsilon\overline{\mathcal{F}}, \tag{89}$$

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

$$\mathcal{F}(\mathcal{Z}) = \mathsf{T}_\epsilon\overline{\mathcal{F}}(\mathcal{Z}) = \overline{\mathcal{F}}(\mathcal{T}_\epsilon\mathcal{Z}) = \overline{\mathcal{F}}(\overline{\mathcal{Z}}).$$

Using the inverse transformation (85), we also define the *push-forward* operator (Littlejohn, 1982a):

$$\mathsf{T}_\epsilon^{-1} : \mathcal{F} \rightarrow \overline{\mathcal{F}} \equiv \mathsf{T}_\epsilon^{-1}\mathcal{F}, \tag{90}$$

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 $\overline{\mathcal{F}}$ on the phase space with coordinates $\overline{\mathcal{Z}}$:

$$\overline{\mathcal{F}}(\overline{\mathcal{Z}}) = \mathsf{T}_\epsilon^{-1}\mathcal{F}(\overline{\mathcal{Z}}) = \mathcal{F}(\mathcal{T}_\epsilon^{-1}\overline{\mathcal{Z}}) = \mathcal{F}(\mathcal{Z}).$$

The pull-back and push-forward operators can now 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 $\mathcal{F}$. First, since $\mathcal{C}[\mathcal{F}](\mathcal{Z})$ is a scalar field, it transforms to $\mathsf{T}_\epsilon^{-1}\mathcal{C}[\mathcal{F}](\overline{\mathcal{Z}})$ with the help of the push-forward operator (90). Next, we replace the extended Vlasov distribution function $\mathcal{F}$ with its pull-back representation $\mathcal{F} = \mathsf{T}_\epsilon\overline{\mathcal{F}}$ and define the transformed operator $\mathcal{C}_\epsilon$ as

$$\mathcal{C}_\epsilon[\overline{\mathcal{F}}] \equiv \mathsf{T}_\epsilon^{-1}\left(\mathcal{C}[\mathsf{T}_\epsilon\overline{\mathcal{F}}]\right). \tag{91}$$

We now apply this induced transformation on the Vlasov equation in extended phase space

$$\frac{d\mathcal{F}}{d\tau} \equiv \{\mathcal{F},\, \mathcal{H}\}_{\mathcal{Z}} = 0, \tag{92}$$

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

$$0 = \frac{d_\epsilon\overline{\mathcal{F}}}{d\tau} \equiv \mathsf{T}_\epsilon^{-1}\left(\frac{d}{d\tau}\,\mathsf{T}_\epsilon\overline{\mathcal{F}}\right) = \{\overline{\mathcal{F}},\, \overline{\mathcal{H}}\}_{\overline{\mathcal{Z}}}, \tag{93}$$

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

$$\overline{\mathcal{H}} \equiv \mathsf{T}_\epsilon^{-1}\mathcal{H}. \tag{94}$$

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

$$\overline{\Gamma} = \mathsf{T}_\epsilon^{-1}\Gamma + \mathsf{d}\mathcal{S}, \tag{95}$$

where $\mathcal{S}$ denotes a (canonical) scalar field used to simplify the transformed phase-space Lagrangian (95), i.e., it has no impact on the new Poisson-bracket structure

$$\overline{\omega} = \mathsf{d}\overline{\Gamma} = \mathsf{d}\left(\mathsf{T}_\epsilon^{-1}\Gamma\right) = \mathsf{T}_\epsilon^{-1}\mathsf{d}\Gamma \equiv \mathsf{T}_\epsilon^{-1}\omega, \tag{96}$$

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

Note that the extended-Hamiltonian transformation (94) may be re-expressed in terms of the regular Hamiltonians H and $\overline{H}$ as

$$\overline{H} \;=\; \mathsf{T}_\epsilon^{-1} H \;-\; \frac{\partial \mathcal{S}}{\partial t}, \tag{97}$$

where $\mathcal{S}$ is the canonical scalar field introduced in Eq. (95); note the similarity with Eq. (24). The new extended phase-space coordinates are chosen so that $d_\epsilon\overline{\mathcal{Z}}^a/d\tau = \{\overline{\mathcal{Z}}^a,\ \overline{\mathcal{H}}\}_{\overline{\mathcal{Z}}}$ are independent of the fast angle $\overline{\zeta}$ and the adiabatic invariant $\overline{\mu}$ satisfies the exact equation $d_\epsilon\overline{\mu}/d\tau \equiv 0$. The dynamical reduction of single-particle Hamiltonian dynamics has, thus, been successfully achieved by phase-space transformation via the construction of a fast invariant $\overline{\mu}$ with its canonically-conjugate fast-angle $\overline{\zeta}$ becoming an ignorable coordinate.

D. Lie-transform Methods

In Lie-transform perturbation theory (Littlejohn, 1982a), the pull-back and push-forward operators (89) and (90) are expressed as Lie *transforms*:

$$\mathsf{T}_\epsilon^{\pm 1} \;\equiv\; \exp\left(\pm \sum_{n=1} \epsilon^n\, \pounds_n \right) \tag{98}$$

where $\pounds_n$ denotes the Lie *derivative* generated by the nth-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. (94), for example, the Lie derivative $\pounds_n\mathcal{H}$ of the scalar field $\mathcal{H}$ is defined as the scalar field

$$\pounds_n\mathcal{H} \;\equiv\; G_n^a\,\partial_a\mathcal{H}. \tag{99}$$

In Eq. (95), on the other hand, the Lie derivative $\pounds_n\Gamma$ of a one-form $\Gamma \equiv \Gamma_a\,\mathsf{d}\mathcal{Z}^a$ is defined as the one-form (Abraham and Marsden, 1978)

$$\begin{aligned} \pounds_n\Gamma \;&\equiv\; \mathsf{G}_n\cdot\mathsf{d}\Gamma \;+\; \mathsf{d}\,(\mathsf{G}_n\cdot\Gamma) \\ &=\; [\,G_n^a\,\omega_{ab} \;+\; \partial_b\,(G_n^a\,\Gamma_a)\,]\,\mathsf{d}\mathcal{Z}^b, \end{aligned} \tag{100}$$

where $\omega_{ab} \equiv \partial_a\Gamma_b - \partial_b\Gamma_b$ are the components of the two-form $\boldsymbol{\omega} \equiv \mathsf{d}\Gamma$. Note that, at each order ϵ^n, the terms $\mathsf{d}(\mathsf{G}_n\cdot\Gamma)$ can be absorbed in the gauge term $\mathsf{d}S_n$ in Eq. (95).

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 $\mathcal{H} \equiv \mathcal{H}_0 + \epsilon\,H_1$ in terms of an unperturbed (zeroth-order) part and a perturbation (first-order) part. The Lie-transform relations associated with Eq. (95) are expressed (up to second order in ϵ) as $\overline{\Gamma}_{0a} \equiv \Gamma_{0a}$ and

$$\overline{\Gamma}_1 \;=\; \Gamma_1 \;-\; \mathsf{G}_1\cdot\boldsymbol{\omega}_0 \;+\; \mathsf{d}S_1, \tag{101}$$

$$\overline{\Gamma}_2 \;=\; -\,\mathsf{G}_2\cdot\boldsymbol{\omega}_0 \;-\; \frac{1}{2}\,\mathsf{G}_1\cdot(\boldsymbol{\omega}_1 + \overline{\boldsymbol{\omega}}_1) \;+\; \mathsf{d}S_2. \tag{102}$$

A general form for the new Poisson bracket $\{\ ,\ \}_{\overline{\mathcal{Z}}}$ is obtained by allowing the new phase-space Lagrangian to retain symplectic perturbation terms $\overline{\Gamma} \equiv \overline{\Gamma}_0 + \epsilon\,\overline{\Gamma}_1$. By chossing a specific form for the perturbed gyrocenter symplectic structure $\overline{\Gamma}_1$, we can, thus, solve Eqs. (101) and (102) for the generating vector field $(\mathsf{G}_1, \mathsf{G}_2)$ expressed in terms of the scalar fields (S_1, S_2). Here, we note that the new phase-space Lagrangian $\overline{\Gamma} \equiv \overline{\Gamma}_{\rm R} + \overline{J}_{\rm g}\,d\overline{\zeta}$, where the reduced phase-space Lagrangian $\overline{\Gamma}_{\rm R}$ is independent of the fast angle $\overline{\zeta}$ and, by application of the Noether theorem (Cary, 1977), the canonically-conjugate action $\overline{J}_{\rm g}$ is an invariant (i.e., $d\overline{J}_{\rm g}/dt = \{\overline{J}_{\rm g},\ \overline{H}\}_{\overline{\mathcal{Z}}} \equiv 0$).

The first-order generating vector field G_1 needed to obtain the gyrocenter extended phase-space Lagrangian (101) is

$$G_1^a \;=\; \{S_1,\ \mathcal{Z}^a\}_0 \;+\; \big(\Gamma_{1b} \;-\; \overline{\Gamma}_{1b}\big)\,J_0^{ba} \tag{103}$$

where $\{\ ,\ \}_0$ is the Poisson-bracket structure associated with the unperturbed Poisson matrix J_0^{ab}. Here, we note that the generating vector field (103) is divided into two parts: a canonical part generated by the gauge function S_1 and a symplectic part generated by the difference $\Delta\Gamma_{1b} \equiv \Gamma_{1b} - \overline{\Gamma}_{1b}$ between the old and new phase-space Lagrangian symplectic components.

Next, for the second-order generating vector field G_2, the condition $\overline{\Gamma}_2 \equiv 0$ yields the following solution for G_2 in terms of the scalar field S_2:

$$G_2^a \;=\; \{S_2,\ \mathcal{Z}^a\}_0 \;-\; \frac{1}{2}\,G_1^b\;(\omega_{1bc} \;+\; \overline{\omega}_{1bc})\,J_0^{ca}, \tag{104}$$

where ω_{1bc} and $\overline{\omega}_{1bc}$ denote the components of the first-order perturbed Lagrange matrices. We note that the second-order generating field (104) is, once again, divided into a canonical part (generated by S_2) and a symplectic part (generated by Γ_{1b} and $\overline{\Gamma}_{1b}$).

The near-identity extended-phase-space transformation (87) is, thus, expressed in terms of the asymptotic expansion

$$\overline{\mathcal{Z}}^a \;=\; \mathcal{Z}^a \;+\; \epsilon\,\big(\{S_1,\ \mathcal{Z}^a\}_0 \;+\; \Delta\Gamma_{1b}\,J_0^{ba}\big) \;+\; \mathcal{O}(\epsilon^2), \tag{105}$$

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

2. Transformed extended Hamiltonian

By substituting the generating vector fields (103) and (104) into the Lie-transform relations associated with

Eq. (94):

$$\overline{H}_1 \;=\; H_1 \;-\; \mathsf{G}_1\cdot \mathsf{d}\mathcal{H}_0, \tag{106}$$

$$\overline{H}_2 \;=\; -\,\mathsf{G}_2\cdot \mathcal{H}_0 \;-\; \frac{1}{2}\,\mathsf{G}_1\cdot \mathsf{d}\left(H_1 \;+\; \overline{H}_1\right), \tag{107}$$

we obtain the first-order and second-order terms in the transformed extended Hamiltonian:

$$\begin{aligned}\overline{H}_1 \;=&\; H_1 \;-\; (\Gamma_{1a} \;-\; \overline{\Gamma}_{1a})\,\dot{\mathcal{Z}}_0^a \;-\; \{S_1,\;\mathcal{H}_0\}_0 \\ \equiv&\; \left(K_1 \;+\; \overline{\Gamma}_{1a}\,\dot{\mathcal{Z}}_0^a\right) \;-\; \{S_1,\;\mathcal{H}_0\}_0\,,\end{aligned} \tag{108}$$

and

$$\begin{aligned}\overline{H}_2 \;=&\; -\{S_2,\;\mathcal{H}_0\}_0 \;-\; \frac{1}{2}\;G_1^a\,\partial_a\,(\,K_1 + \overline{K}_1) \\ &-\;\frac{1}{2}\;G_1^b\{\,(\Gamma_{1b} + \overline{\Gamma}_{1b}),\;\mathcal{H}_0\}_0 \\ &-\;\frac{1}{2}\,(\Gamma_{1b} + \overline{\Gamma}_{1b})\;\left(G_1^a\,\partial_a\dot{\mathcal{Z}}_0^b\right),\end{aligned} \tag{109}$$

where we used the Poisson-bracket properties (60)-(62). In Eqs. (108)-(109), $\dot{\mathcal{Z}}_0^a \equiv \{\mathcal{Z}^a,\;\mathcal{H}_0\}_0$ denotes the zeroth-order Hamilton's equations and

$$K_1 \;\equiv\; H_1 \;-\; \Gamma_{1\,a}\,\dot{\mathcal{Z}}_0^a, \tag{110}$$

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

The two Hamiltonian relations (108)-(109) 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 $\overline{H}_n$ 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 (108) is, thus, expressed in terms of the fast-angle averaging operation $\langle\cdots\rangle$ as

$$\overline{H}_1 \;\equiv\; \langle K_1\rangle \;+\; \overline{\Gamma}_{1a}\,\dot{\mathcal{Z}}_0^a, \tag{111}$$

where the Poisson bracket $\{\;,\;\}$, henceforth, denotes the zeroth-order Poisson bracket $\{\;,\;\}_0$ (unless otherwise noted) and S_1 can be chosen such that $\langle S_1\rangle \equiv 0$. The first-order gauge function S_1 is determined from the perturbation equation

$$\frac{d_0 S_1}{d\tau} \;\equiv\; \{S_1,\;\mathcal{H}_0\} \;=\; \widetilde{K}_1 \;\equiv\; K_1 \;-\; \langle K_1\rangle, \tag{112}$$

whose solution is $S_1 \equiv (d_0/d\tau)^{-1}\widetilde{K}_1$, where $(d_0/d\tau)^{-1}$ denotes an integration along an unperturbed extended Hamiltonian orbit; note that this formal solution is identical to the solution (83) obtained by variational methods. To lowest order in the fast orbital time scale, the unperturbed integration

$$S_1 \;=\; (d_0/d\tau)^{-1}\widetilde{K}_1 \;\equiv\; \Omega^{-1}\;\int \widetilde{K}_1 d\overline{\zeta} \tag{113}$$

involves an indefinite fast-angle integration, where $\Omega \equiv d_0\overline{\zeta}/dt$ denotes the fast-angle frequency; note also that the solution (113) for the first-order gauge function S_1 does not depend on the choice of $\overline{\Gamma}_1$.

The solution for the new second-order Hamiltonian (109) yields the fast-angle-averaged expression

$$\begin{aligned}\overline{H}_2 \;=&\; -\,\frac{1}{2}\,\langle\{S_1,\;\{S_1,\;\mathcal{H}_0\}\}\rangle \;-\; \langle\Delta\Gamma_{1a}\;\{\mathcal{Z}^a,\;K_1\}\rangle \\ &-\;\frac{1}{2}\;\left\langle\Delta\Gamma_{1a}\quad J_0^{ab}\{\,(\Gamma_{1b} + \overline{\Gamma}_{1b}),\;\mathcal{H}_0\}\right\rangle \\ &-\;\frac{1}{2}\;\left\langle\Delta\Gamma_{1a}\;\left\{\mathcal{Z}^a,\;\dot{\mathcal{Z}}_0^b\right\}(\Gamma_{1b} + \overline{\Gamma}_{1b})\right\rangle,\end{aligned} \tag{114}$$

where the first term corresponds to the standard quadratic ponderomotive Hamiltonian (Cary and Kaufman, 1981) while the remaining terms (which depend on the symplectic choice $\overline{\Gamma}_1$) will be discussed below. The second-order gauge function S_2 appearing in Eq. (109) is not needed in what follows since the phase-space transformation from guiding-center coordinates to gyrocenter coordinates is only needed to first order in ϵ_δ.

E. Reduced Vlasov-Maxwell Equations

The extended Vlasov equation (92) may be converted into the regular Vlasov equation as follows. First, in order to satisfy the physical constraint (66), the extended Vlasov distribution is expressed as

$$\mathcal{F}(\mathcal{Z}) \;\equiv\; c\,\delta[w - H(\mathbf{z},t)]\;f(\mathbf{z},t), \tag{115}$$

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 (92) over the energy coordinate w (and using $d\tau = dt$), we, thus, obtain the regular Vlasov equation

$$0 \;=\; \frac{df}{dt} \;\equiv\; \frac{\partial f}{\partial t} \;+\; \frac{d\mathbf{z}}{dt}\boldsymbol{\cdot}\frac{\partial f}{\partial \mathbf{z}}. \tag{116}$$

Next, the push-forward transformation of the extended Vlasov distribution (115) yields the reduced extended Vlasov distribution

$$\overline{\mathcal{F}}(\overline{\mathcal{Z}}) \;\equiv\; c\,\delta[\overline{w} - \overline{H}(\overline{\mathbf{z}},t)]\;\overline{f}(\overline{\mathbf{z}},t), \tag{117}$$

where the reduced extended Hamiltonian $\overline{\mathcal{H}} \equiv \overline{H}(\overline{\mathbf{z}},t) - \overline{w}$ is defined in Eq. (94). Lastly, the extended reduced Vlasov equation

$$\frac{d_\epsilon\overline{\mathcal{F}}}{d\tau} \;\equiv\; \{\overline{\mathcal{F}},\;\overline{\mathcal{H}}\}_\epsilon \;=\; 0 \tag{118}$$

can be converted into the regular reduced Vlasov equation by integrating it over the reduced energy coordinate $\overline{w}$, which yields the reduced Vlasov equation

$$0 \;=\; \frac{d_\epsilon\overline{f}}{dt} \;\equiv\; \frac{\partial\overline{f}}{\partial t} \;+\; \frac{d_\epsilon\overline{\mathbf{z}}}{dt}\boldsymbol{\cdot}\frac{\partial\overline{f}}{\partial\overline{\mathbf{z}}}, \tag{119}$$

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

We now investigate how the pull-back and push-forward operators are used in the transformation of Maxwell's equations

$$\nabla\cdot\mathbf{E} \;=\; 4\pi\,\rho, \tag{120}$$

$$\nabla\times\mathbf{B} - \frac{1}{c}\frac{\partial\mathbf{E}}{\partial t} \;=\; \frac{4\pi}{c}\,\mathbf{J}, \tag{121}$$

where the charge-current densities

$$\begin{pmatrix}\rho\\ \mathbf{J}\end{pmatrix} \;=\; \sum e\int d^4p\;\mathcal{F}\begin{pmatrix}1\\ \mathbf{v}\end{pmatrix} \tag{122}$$

are defined in terms of the extended Vlasov distribution $\mathcal{F}$ (with $d^4p = c^{-1}dw\,d^3p$) and $\mathbf{E}\equiv -\nabla\phi - c^{-1}\partial\mathbf{A}/\partial t$ 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 (122) can be expressed in terms of the general expression (where time dependence is omitted for clarity)

$$\begin{aligned}\|v^\mu\|(\mathbf{r}) \;&\equiv\; \int d^3p\,v^\mu\,f \;=\; \int d^4p\,v^\mu\,\mathcal{F}\\ &=\; \int d^3x\int d^4p\;v^\mu\;\delta^3(\mathbf{x}-\mathbf{r})\;\mathcal{F},\end{aligned} \tag{123}$$

where $v^\mu = (c,\mathbf{v})$ 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 $\mathcal{T}_\epsilon:\mathcal{Z}\rightarrow\overline{\mathcal{Z}}$ (where time t itself is unaffected) on the right side of Eq. (123), we obtain the push-forward representation for the fluid moments $\|v^\mu\|$:

$$\begin{aligned}\|v^\mu\|(\mathbf{r}) \;&=\; \int d^3\overline{x}\int d^4\overline{p}\;\left(\mathsf{T}_\epsilon^{-1}v^\mu\right)\;\delta^3(\overline{\mathbf{x}}+\boldsymbol{\rho}_\epsilon-\mathbf{r})\;\overline{\mathcal{F}}\\ &=\; \int d^3\overline{p}\;\;e^{-\boldsymbol{\rho}_\epsilon\cdot\nabla}\left[\,\left(\mathsf{T}_\epsilon^{-1}v^\mu\right)\;\overline{f}\,\right],\end{aligned} \tag{124}$$

where $\mathsf{T}_\epsilon^{-1}v^\mu = (c,\mathsf{T}_\epsilon^{-1}\mathbf{v})$ denotes the push-forward of the four-velocity v^μ and

$$\begin{aligned}\boldsymbol{\rho}_\epsilon \;&\equiv\; \mathsf{T}_\epsilon^{-1}\mathbf{x}-\overline{\mathbf{x}}\\ &=\; -\,\epsilon\,G_1^{\mathbf{x}} \;-\; \epsilon^2\left(G_2^{\mathbf{x}}-\frac{1}{2}\,\mathsf{G}_1\cdot\mathsf{d}G_1^{\mathbf{x}}\right)+\cdots\end{aligned} \tag{125}$$

denotes the displacement between the push-forward $\mathsf{T}_\epsilon^{-1}\mathbf{x}$ of the particle position $\mathbf{x}$ and the (new) reduced position $\overline{\mathbf{x}}$, which is defined in terms of the generating vector fields (103) and (104).[11]

[11] We immediately note the similarity between the general form (124) for the push-forward representation of fluid moments and the Frieman-Chen expression (17) for the perturbed plasma density δn.

The push-forward representation for the charge-current densities, therefore, introduces polarization and magnetization effects into the Maxwell equations, which transforms the microscopic Maxwell's equations (120)-(121) into the macroscopic (reduced) equations

$$\nabla\cdot\mathbf{D} \;=\; 4\pi\,\overline{\rho}, \tag{126}$$

$$\nabla\times\mathbf{H} \;-\; \frac{1}{c}\,\frac{\partial\mathbf{D}}{\partial t} \;=\; \frac{4\pi}{c}\,\overline{\mathbf{J}}, \tag{127}$$

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

$$\begin{pmatrix}\overline{\rho}\\ \overline{\mathbf{J}}\end{pmatrix} \;=\; \sum e\int d^4\overline{p}\;\overline{\mathcal{F}}\begin{pmatrix}1\\ \overline{\mathbf{v}}\end{pmatrix}, \tag{128}$$

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

$$\left.\begin{aligned}\mathbf{D} \;&=\; \mathbf{E} \;+\; 4\pi\,\mathbf{P}_\epsilon\\ \mathbf{H} \;&=\; \mathbf{B} \;-\; 4\pi\,\mathbf{M}_\epsilon\end{aligned}\right\}. \tag{129}$$

Here, $\mathbf{P}_\epsilon$ and $\mathbf{M}_\epsilon$ denote the polarization and magnetization vectors associated with the dynamical reduction introduced by the phase-space transformation (87). Lastly, the relation between the particle charge-current densities $(\rho,\mathbf{J})$ and the reduced charge-current densities $(\overline{\rho},\overline{\mathbf{J}})$:

$$\rho \;\equiv\; \overline{\rho} \;-\; \nabla\cdot\mathbf{P}_\epsilon, \tag{130}$$

$$\mathbf{J} \;\equiv\; \overline{\mathbf{J}} \;+\; \frac{\partial\mathbf{P}_\epsilon}{\partial t} \;+\; c\,\nabla\times\mathbf{M}_\epsilon, \tag{131}$$

defines the polarization density $\rho_{\rm pol}\equiv -\nabla\cdot\mathbf{P}_\epsilon$, the polarization current $\mathbf{J}_{\rm pol}\equiv\partial\mathbf{P}_\epsilon/\partial t$, and the magnetization current $\mathbf{J}_{\rm mag}\equiv c\,\nabla\times\mathbf{M}_\epsilon$. The derivation of the polarization and magnetization vectors $\mathbf{P}_\epsilon$ and $\mathbf{M}_\epsilon$ is done either directly by the push-forward method (124) or by variational method

$$(\mathbf{D},\;\mathbf{H}) \;\equiv\; 4\pi\left(\frac{\partial\overline{\mathcal{L}}}{\partial\mathbf{E}},\;-\,\frac{\partial\overline{\mathcal{L}}}{\partial\mathbf{B}}\right), \tag{132}$$

where $\overline{\mathcal{L}}$ denotes 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 Section VI.C).

F. Example: Oscillation-center Hamiltonian Dynamics

Before proceeding with the derivation of the nonlinear gyrokinetic equations, we consider the Lie-transform derivation of the oscillation-center (ponderomotive) Hamiltonian dynamics for charged particles moving in high-frequency, short-wavelength electromagnetic (wave) fields. The derivation of the ponderomotive

Hamiltonian is the paradigm for the application of Lie-transform methods in plasma physics (Cary and Kaufman, 1981), and it introduces important concepts that translate well to the low-frequency limit appropriate for gyrokinetics. Another example of a successful application of Lie-transform perturbation theory involves the perturbation analysis of a charged-particle beam orbit in a wiggler magnetic field (Kishimoto *et al.*, 1995). For the sake of simplicity of presentation, however, we consider a weakly-inhomogeneous unmagnetized plasma background in the presence of a weakly-inhomogeneous electric potential Φ_0, so that the wave amplitudes are also weakly-varying in space and time, and we ignore the self-consistent back-reaction Vlasov-Maxwell response, since we focus our attention on deriving the reduced Hamiltonian dynamics.

In order to separate the space-time scales appearing in this problem, we introduce the *eikonal* representation for the wave fields:

$$\begin{pmatrix} \phi_1 \\ \mathbf{A}_1 \end{pmatrix} \equiv \begin{pmatrix} \widetilde{\phi}_1 \\ \widetilde{\mathbf{A}}_1 \end{pmatrix} e^{i\Theta/\epsilon_0} + \begin{pmatrix} \widetilde{\phi}_1^* \\ \widetilde{\mathbf{A}}_1^* \end{pmatrix} e^{-i\Theta/\epsilon_0}, \quad (133)$$

where the wave amplitudes (denoted, here, by a tilde) are weakly-varying space-time functions and derivatives of the eikonal phase $\Theta(\epsilon_0\mathbf{r}, \epsilon_0 t)$:

$$\left. \begin{array}{rcl} \epsilon_0^{-1}\nabla\Theta(\epsilon_0\mathbf{r}, \epsilon_0 t) & = & \mathbf{k}(\epsilon_0\mathbf{r}, \epsilon_0 t) \\ \epsilon_0^{-1}\partial_t\Theta(\epsilon_0\mathbf{r}, \epsilon_0 t) & = & -\,\omega(\epsilon_0\mathbf{r}, \epsilon_0 t) \end{array} \right\} \quad (134)$$

define the weakly-varying wavevector $\mathbf{k}$ and wave frequency ω. Here, the eikonal parameter ϵ_0 is defined as $\epsilon_0 \sim k^{-1}|\nabla\ln\Phi_0| \sim \omega^{-1}|\partial_t\ln\Phi_0|$ in terms of the background electric potential Φ_0.

The extended-phase-space Hamiltonian dynamics of charged particles in such electromagnetic-wave fields is expressed in terms of the extended phase-space Lagrangian

$$\begin{aligned} \Gamma &= \left(\mathbf{p} + \epsilon_\delta \frac{e}{c}\mathbf{A}_1\right)\cdot d\mathbf{x} - w\,dt \\ &\equiv \Gamma_0 + \epsilon_\delta\,\Gamma_1, \end{aligned} \quad (135)$$

where $\mathbf{p} = m\,\mathbf{v}$ denotes the kinetic momentum of the charged particle, and the extended Hamiltonian

$$\begin{aligned} \mathcal{H} &= \frac{1}{2m}|\mathbf{p}|^2 + e\,(\Phi_0 + \epsilon_\delta\,\phi_1) - w \\ &\equiv \mathcal{H}_0 + \epsilon_\delta\,H_1. \end{aligned} \quad (136)$$

Here, we see that both the Hamiltonian and the symplectic (Poisson-bracket) structure exhibit explicit dependence on the eikonal phase Θ and, thus, a particle orbit exhibits both fast-wave space-time scales and slow-background space-time scales.

By definition, the oscillation-center Hamiltonian dynamics must be expressed in terms of an extended-phase-space Lagrangian $\overline{\Gamma}$ and Hamiltonian $\overline{\mathcal{H}}$ that are explicitly independent of the eikonal phase Θ. Hence, we choose the Hamiltonian representation $\overline{\Gamma} \equiv \overline{\Gamma}_0$ (i.e., $\overline{\Gamma}_n \equiv 0$ for $n \geq 1$ and all magnetic perturbations are transfered to the oscillation-center Hamiltonian), so that the first-order generating vector field is obtained from Eq. (103) as

$$G_1^a = \{S_1,\, \mathcal{Z}^a\}_0 + \frac{e}{c}\mathbf{A}_1\cdot\{\mathbf{x},\, \mathcal{Z}^a\}_0. \quad (137)$$

Here, we note that the symplectic part (second term) of the generating vector field (137) was used to remove the wave-field perturbation on the Poisson-bracket structure (135) and we, henceforth, use the notation $\{\ ,\ \} \equiv \{\ ,\ \}_0$.

Next, the second-order generating vector field is obtained from Eq. (104) as

$$G_2^a = \{S_2,\, \mathcal{Z}^a\} - \frac{1}{2}G_1^b\,\omega_{1bc}\,J_0^{ca}, \quad (138)$$

where (assuming that $G_n^t \equiv 0$ so that particle and oscillation-center times are identical)

$$\begin{aligned} G_1^b\,\omega_{1bc}\,J_0^{ca} &= -\frac{e}{c}G_1^{\mathbf{x}}\cdot\mathbf{B}_1\times\{\mathbf{x},\, \mathcal{Z}^a\} \\ &\quad - \frac{e}{c}G_1^{\mathbf{x}}\cdot\frac{\partial\mathbf{A}_1}{\partial t}\,\{t,\, \mathcal{Z}^a\}. \end{aligned}$$

The near-identity extended-phase-space transformation from particle coordinates $\mathcal{Z}^a = (\mathbf{x}, \mathbf{p}; w, t)$ to oscillation-center coordinates $\overline{\mathcal{Z}}^a = (\overline{\mathbf{x}}, \overline{\mathbf{p}}; \overline{w}, t)$ is, thus, expressed (up to first order in ϵ_δ) as

$$\overline{\mathcal{Z}}^a = \mathcal{Z}^a + \epsilon_\delta\left(\{S_1,\, \mathcal{Z}^a\} + \frac{e}{c}\mathbf{A}_1\cdot\{\mathbf{x},\, \mathcal{Z}^a\}\right), \quad (139)$$

and its explicit expression requires a solution for the scalar fields $(S_1, ...)$.

By substituting the generating vector fields (137) and (138) into Eqs. (108)-(109), we obtain the following Hamiltonian relations

$$\begin{aligned} \overline{H}_1 &= e\left(\phi_1 - \frac{\mathbf{v}}{c}\cdot\mathbf{A}_1\right) - \{S_1,\, \mathcal{H}_0\} \\ &\equiv e\,\psi_1 - \{S_1,\, \mathcal{H}_0\}, \end{aligned} \quad (140)$$

and

$$\begin{aligned} \overline{H}_2 &= -\left\{\left(S_2 - \frac{e}{2c}\mathbf{A}_1\cdot\{\mathbf{x},\, S_1\}\right),\, \mathcal{H}_0\right\} \\ &\quad + \frac{e^2\,|\mathbf{A}_1|^2}{2\,mc^2} - \frac{e}{2}\,\{S_1,\, \psi_1\}. \end{aligned} \quad (141)$$

First, we note that, since the wave fields (133) are explicitly eikonal-dependent, the gauge function S_1 may be chosen to be explicitly Θ-dependent:

$$S_1 = \widetilde{S}_1\,e^{i\Theta/\epsilon_0} + \widetilde{S}_1^*\,e^{-i\Theta/\epsilon_0},$$

and, thus, we may define the first-order oscillation-center Hamiltonian to be identically zero: $\overline{H}_1 \equiv 0$. By substituting the eikonal representation into the equation $\{S_1,\ \mathcal{H}_0\} = e\,\psi_1$, we readily find $\widetilde{S}_1 = i\,e\widetilde{\psi}_1/\omega'$, where

$\omega' \equiv \omega - \mathbf{k}\cdot\mathbf{v}$ denotes the Doppler-shifted wave frequency. Hence, the extended-phase-space transformation $(\mathbf{x},\mathbf{p},w,t) \rightarrow (\overline{\mathbf{x}},\overline{\mathbf{p}},\overline{w},t)$ is defined (to first order in ϵ_δ) as

$$\begin{aligned}\overline{\mathbf{x}} &= \mathbf{x} - \epsilon_\delta\,\frac{\partial S_1}{\partial\mathbf{p}} + \cdots,\\ \overline{\mathbf{p}} &= \mathbf{p} + \epsilon_\delta\left(\nabla S_1 + \frac{e}{c}\,\mathbf{A}_1\right) + \cdots,\\ \overline{w} &= w - \epsilon_\delta\,\frac{\partial S_1}{\partial t} + \cdots,\end{aligned}$$

where the presence of the perturbed vector potential $\mathbf{A}_1$ in the definition of the oscillation-center momentum $\overline{\mathbf{p}}$ implies that it is a *canonical* momentum (in contrast to the kinetic momentum $\mathbf{p} = m\mathbf{v}$). In what follows, it is useful to define the eikonal-dependent displacement $\boldsymbol{\xi} \equiv \partial S_1/\partial\mathbf{p}$ between the particle position $\mathbf{x}$ and the oscillation-center position $\overline{\mathbf{x}}$ (Hatori and Washimi, 1981). Here, the eikonal amplitude $\widetilde{\boldsymbol{\xi}}$ is expressed as

$$\widetilde{\boldsymbol{\xi}} \equiv -\,\frac{e}{m\omega'^2}\left(\widetilde{\mathbf{E}}_1 + \frac{\overline{\mathbf{v}}}{c}\times\widetilde{\mathbf{B}}_1\right) \tag{142}$$

for the first-order oscillation-center phase-space *displacement*. By substituting the eikonal solution $\widetilde{S}_1$ into the second-order eikonal-averaged Hamiltonian (141), we obtain the second-order oscillation-center Hamiltonian

$$\begin{aligned}\overline{H}_2 &= \frac{e^2\,|\widetilde{\mathbf{A}}_1|^2}{mc^2} - e\left\{\epsilon_0^{-1}\Theta,\ \mathrm{Im}\left(\widetilde{S}_1^*\,\widetilde{\psi}_1\right)\right\}\\ &\equiv m\,|\omega'\,\widetilde{\boldsymbol{\xi}}|^2 = -\,e\,\widetilde{\boldsymbol{\xi}}^*\boldsymbol{\cdot}\left(\widetilde{\mathbf{E}}_1 + \frac{\overline{\mathbf{v}}}{c}\times\widetilde{\mathbf{B}}_1\right),\end{aligned} \tag{143}$$

from which we obtain the oscillation-center velocity

$$\overline{\mathbf{v}} \equiv \{\overline{\mathbf{x}},\ \overline{\mathcal{H}}\} = \mathbf{v} + \epsilon_\delta^2\,\frac{\partial\overline{H}_2}{\partial\overline{\mathbf{p}}}, \tag{144}$$

where the ponderomotive Hamiltonian $\overline{H}_2$ introduces nonlinear effects into the oscillation-center Hamiltonian dynamics.

Lastly, the oscillation-center transformation (139) introduces polarization and magnetization effects into Maxwell's equations. The oscillation-center (ponderomotive) polarization and magnetization vectors are derived in Appendix C by the direct push-forward method as

$$\mathbf{P}_{\rm osc} = \epsilon_\delta^2\sum\int d^3\overline{p}\,\overline{f}\,\overline{\boldsymbol{\pi}}_2, \tag{145}$$

$$\mathbf{M}_{\rm osc} = \epsilon_\delta^2\sum\int d^3\overline{p}\,\overline{f}\,\left(\overline{\boldsymbol{\mu}}_2 + \overline{\boldsymbol{\pi}}_2\times\frac{\overline{\mathbf{v}}}{c}\right), \tag{146}$$

where $\overline{f}$ denotes the oscillation-center Vlasov distribution, $\overline{\boldsymbol{\pi}}_2 = e\,\mathbf{k}\times(i\,\widetilde{\boldsymbol{\xi}}\times\widetilde{\boldsymbol{\xi}}^*)$ denotes the second-order ponderomitive electric-dipole moment, $\overline{\boldsymbol{\mu}}_2 = (e/c)\,\omega'\ (i\,\widetilde{\boldsymbol{\xi}}\times\widetilde{\boldsymbol{\xi}}^*)$ denotes the second-order ponderomotive magnetic-dipole moment, and a moving-electric-dipole contribution (Jackson, 1975) also appears in Eq. (146).

These expressions show that nonlinear ponderomotive terms in the oscillation-center Hamiltonian can not only modify the oscillation-center Hamiltonian dynamics but also introduce back-reaction effects into Maxwell's equations that generate second-order eikonal-independent electromagnetic fields.

V. NONLINEAR GYROKINETIC VLASOV EQUATION

We are now ready to apply the methods of Lie-transform perturbation theory presented in Section 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 small parameter $\epsilon_\delta \ll 1$. Here, 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 Section 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 $\Gamma_0 \equiv [(e/c)\,\mathbf{A}_0 + \mathbf{p}]\cdot d\mathbf{x} - w\,dt$ and $\Gamma_1 \equiv (e/c)\,\delta\mathbf{A}\boldsymbol{\cdot}d\mathbf{x}$, and the extended phase-space Hamiltonian $\mathcal{H} = \mathcal{H}_0 + \epsilon_\delta\,H_1$, where $\mathcal{H}_0 \equiv |\mathbf{p}|^2/2m - w$ and $H_1 \equiv e\,\delta\phi$. Note that, 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 gyro-motion space-time scales associated with the (unperturbed) background magnetic field (first step = guiding-center transformation with ordering parameter $\epsilon_{\rm B}$) and, second, the fast gyro-motion time scale associated with the perturbation electromagnetic fields (second step = gyrocenter transformation with ordering parameters ϵ_δ, ϵ_ω, and $\epsilon_\perp$).

A. Unperturbed Guiding-center Hamiltonian Dynamics

The *guiding-center* phase-space transformation involves an asymptotic expansion, with a small dimensionless parameter $\epsilon_{\rm B} \equiv \rho_{\rm th}/L_{\rm B} \ll 1$ defined as the ratio of the thermal gyroradius $\rho_{\rm th}$ and the background magnetic-field length scale $L_{\rm B}$. This transformation is designed to remove the fast gyro-motion 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 $\epsilon_{\rm B}$ with the scalar potential Φ_0 ordered at zeroth order in $\epsilon_{\rm B}$; in this Section, we set the equilibrium scalar potential equal to zero and discuss issues associated with an inhomogeneous equilibrium electric field in Appendix E.1.

The results of the guiding-center analysis presented by Littlejohn (1983) are summarized as follows (further details are presented in Appendix B). First, the guiding-center transformation yields the following guiding-center coordinates $(\mathbf{X}, p_\|, \mu, \zeta, w, t) \equiv \mathcal{Z}_{\rm gc}$, where $\mathbf{X}$ is the guiding-center position, $p_\|$ 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 (here, time is unaffected by the transformation while the guiding-center energy is chosen to be equal to the particle energy). Next, the unperturbed guiding-center extended phase-space Lagrangian is

$$\Gamma_{\rm gc} \;\equiv\; \frac{e}{c}\,\mathbf{A}_0^*\cdot d\mathbf{X} \;+\; \mu\,(mc/e)\,d\zeta \;-\; w\;dt, \qquad (147)$$

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

$$\mathcal{H}_{\rm gc} \;=\; \frac{p_\|^2}{2m} \;+\; \mu\,B_0 \;-\; w \;\equiv\; H_{\rm gc} - w. \qquad (148)$$

Lastly, from the unperturbed guiding-center phase-space Lagrangian (147), we obtain the unperturbed guiding-center Poisson bracket $\{\;,\;\}_{\rm gc}$, given here in terms of two arbitrary functions $\mathcal{F}$ and $\mathcal{G}$ on extended guiding-center phase space as (Littlejohn, 1983)

$$\begin{aligned}\{\mathcal{F},\,\mathcal{G}\}_{\rm gc} \;\equiv\;& \frac{e}{mc}\left(\frac{\partial\mathcal{F}}{\partial\zeta}\,\frac{\partial\mathcal{G}}{\partial\mu} - \frac{\partial\mathcal{F}}{\partial\mu}\,\frac{\partial\mathcal{G}}{\partial\zeta}\right)\\ &+\; \frac{\mathbf{B}_0^*}{B_{0\|}^*}\cdot\left(\nabla\mathcal{F}\,\frac{\partial\mathcal{G}}{\partial p_\|} - \frac{\partial\mathcal{F}}{\partial p_\|}\,\nabla\mathcal{G}\right)\\ &-\; \frac{c\widehat{\mathbf{b}}_0}{eB_{0\|}^*}\cdot\nabla\mathcal{F}\times\nabla\mathcal{G}\\ &+\; \left(\frac{\partial\mathcal{F}}{\partial w}\,\frac{\partial\mathcal{G}}{\partial t} - \frac{\partial\mathcal{F}}{\partial t}\,\frac{\partial\mathcal{G}}{\partial w}\right), \end{aligned}\qquad (149)$$

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

$$\left.\begin{aligned}\mathbf{B}_0^* \;&=\; \mathbf{B}_0 \;+\; (c/e)\,p_\|\;\nabla\times\widehat{\mathsf{b}}_0\\ B_{0\|}^* \;&=\; B_0 \;+\; (c/e)p_\|\;\widehat{\mathsf{b}}_0\cdot\nabla\times\widehat{\mathsf{b}}_0\end{aligned}\right\}. \qquad (150)$$

Note that the Jacobian of the guiding-center transformation is $\mathcal{J}_{\rm gc} = m\,B_{0\|}^*$ (i.e., $d^3x\,d^3p = \mathcal{J}_{\rm gc}\,d^3X\,dp_\|\,d\mu\,d\zeta$) 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 (149). The unperturbed guiding-center Hamiltonian dynamics is, thus, expressed in terms of the Hamiltonian (148) and the Poisson bracket (149) as $\dot{\mathcal{Z}}^a \equiv \{\mathcal{Z}^a,\;\mathcal{H}_{\rm gc}\}_{\rm gc}$. In particular, the conservation law $\dot{\mu} \equiv 0$ for the guiding-center magnetic moment follows from the fact that the guiding-center Hamiltonian (148) is independent of the fast gyroangle ζ (to arbitrary order in $\epsilon_{\rm B}$).

B. Perturbed Guiding-center Hamiltonian Dynamics

We now consider how the guiding-center Hamiltonian system $(\mathcal{H}_{\rm gc};\;\{\;,\;\}_{\rm gc})$ is affected by the introduction of low-frequency electromagnetic field fluctuations $(\delta\phi, \delta\mathbf{A})$ satisfying the low-frequency gyrokinetic orderings (5)-(9). Under the electromagnetic perturbations $(\delta\phi, \delta\mathbf{A})$, the guiding-center phase-space Lagrangian (147) and Hamiltonian (148) become

$$\left.\begin{aligned}\Gamma_{\rm gc}' \;&\equiv\; \Gamma_{\rm 0gc} \;+\; \epsilon_\delta\,\Gamma_{\rm 1gc}\\ \mathcal{H}_{\rm gc}' \;&\equiv\; \mathcal{H}_{\rm 0gc} \;+\; \epsilon_\delta\,H_{\rm 1gc}\end{aligned}\right\}, \qquad (151)$$

where the zeroth-order guiding-center phase-space Lagrangian $\Gamma_{\rm 0gc}$ and Hamiltonian $\mathcal{H}_{gc0}$ are given by (147) and (148), respectively. In what follows, although the three small parameters $(\epsilon_{\rm B}, \epsilon_\delta, \epsilon_\omega)$ may be of the same order in the conventional nonlinear gyrokinetic ordering (Frieman and Chen, 1982), we keep them independent in order 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 E.1.

In Eq. (151), the first-order guiding-center phase-space Lagrangian Γ_{gc1} and Hamiltonian $H_{\rm 1gc}$ are

$$\begin{aligned}\Gamma_{\rm 1gc} \;&=\; \frac{e}{c}\,\delta\mathbf{A}(\mathbf{X}+\boldsymbol{\rho}, t)\cdot d(\mathbf{X}+\boldsymbol{\rho})\\ &\equiv\; \frac{e}{c}\,\delta\mathbf{A}_{\rm gc}(\mathbf{X}, t; \mu, \zeta)\cdot d(\mathbf{X}+\boldsymbol{\rho}),\end{aligned}\qquad (152)$$

and

$$H_{\rm 1gc} \;=\; e\delta\phi(\mathbf{X}+\boldsymbol{\rho}, t) \;\equiv\; e\delta\phi_{\rm gc}(\mathbf{X}, t; \mu, \zeta), \qquad (153)$$

where $\delta\mathbf{A}_{\rm gc}(\mathbf{X}, t; \mu, \zeta)$ and $\delta\phi_{\rm gc}(\mathbf{X}, t; \mu, \zeta)$ denote perturbation potentials evaluated at a particle's position $\mathbf{x} \equiv \mathbf{X} + \boldsymbol{\rho}$ expressed in terms of the guiding-center position $\mathbf{X}$ and the gyroangle-dependent gyroradius vector $\boldsymbol{\rho}(\mu, \zeta)$; here, to lowest order in $\epsilon_{\rm B}$, we ignore the spatial dependence of $\boldsymbol{\rho}$.

Because of the gyroangle-dependence in the guiding-center perturbation potentials $(\delta\phi_{\rm gc}, \delta\mathbf{A}_{\rm gc})$, 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 (152)-(153), we proceed with the time-dependent *gyrocenter* phase-space transformation

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

where $\overline{\mathcal{Z}}$ denote the *gyrocenter* (gy) extended phase-space coordinates; we note that the nature of the gyrocenter parallel momentum $\overline{p}_\|$ depends on the choice of representation used for gyrocenter Hamiltonian dynamics (as will be 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 $(\epsilon_\omega, \epsilon_{\rm B})$, this transformation is represented in terms of generating vector fields $(\mathsf{G}_1, \mathsf{G}_2, ...)$ as

$$\overline{\mathcal{Z}}^a \;\equiv\; \mathcal{Z}^a \;+\; \epsilon_\delta\, G_1^a \;+\; \cdots \tag{154}$$

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

$$\begin{aligned} \overline{\Gamma} \;&=\; \left[\, \frac{e}{c}\,(\mathbf{A}_0 \;+\; \epsilon_\delta\, \delta\mathbf{A}_{\rm gy}) \;+\; \overline{p}_\|\, \widehat{\mathsf{b}}_0 \,\right] \boldsymbol{\cdot} d\overline{\mathbf{X}} \\ &\quad+\; \frac{mc}{e}\, \overline{\mu}\, d\overline{\zeta} \;-\; \overline{w}\, dt \\ &\equiv\; \overline{\Gamma}_0 \;+\; \epsilon_\delta\, \overline{\Gamma}_1, \end{aligned} \tag{155}$$

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

$$\delta\mathbf{A}_{\rm gy} \;\equiv\; \alpha\,\langle \delta\mathbf{A}_{\perp{\rm gc}}\rangle \;+\; \beta\,\langle \delta A_{\|{\rm gc}}\rangle\, \widehat{\mathsf{b}}_0. \tag{156}$$

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

gyrocenter Model	α	β	$\overline{p}_\|$
Hamiltonian	0	0	canonical
Symplectic	1	1	kinetic
$\perp -$ Symplectic	1	0	canonical
$\| -$ Symplectic	0	1	kinetic

The Hamiltonian gyrocenter model ($\alpha = 0 = \beta$) and the symplectic gyrocenter model ($\alpha = 1 = \beta$) are presented by Brizard (1989a), while the parallel-symplectic gyrocenter model ($\beta = 1, \alpha = 0$) is 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 $\mathcal{J} = m^2\, B_\|^*$, where

$$B_\|^* \;\equiv\; B_{0\|}^* \;+\; \epsilon_\delta\left(\alpha\,\langle \delta B_{\|{\rm gc}}\rangle\right), \tag{157}$$

while the general form for the gyrocenter Poisson bracket is

$$\begin{aligned} \{\mathcal{F},\, \mathcal{G}\} \;&=\; \frac{e}{mc}\left(\frac{\partial\mathcal{F}}{\partial\zeta}\,\frac{\partial\mathcal{G}}{\partial\mu^*} \;-\; \frac{\partial\mathcal{F}}{\partial\mu^*}\,\frac{\partial\mathcal{G}}{\partial\zeta}\right) \\ &\quad+\; \frac{\mathbf{B}^*}{B_\|^*}\boldsymbol{\cdot}\left(\nabla^*\mathcal{F}\,\frac{\partial\mathcal{G}}{\partial p_\|} \;-\; \frac{\partial\mathcal{F}}{\partial p_\|}\,\nabla^*\mathcal{G}\right) \\ &\quad-\; \frac{c\widehat{\mathsf{b}}_0}{e\,B_\|^*}\boldsymbol{\cdot}\nabla^*\mathcal{F}\times\nabla^*\mathcal{G} \\ &\quad+\; \left(\frac{\partial\mathcal{F}}{\partial w}\,\frac{\partial\mathcal{G}}{\partial t} \;-\; \frac{\partial\mathcal{F}}{\partial t}\,\frac{\partial\mathcal{G}}{\partial w}\right), \end{aligned} \tag{158}$$

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

$$\begin{aligned} \nabla^*\mathcal{F} \;&\equiv\; \nabla\mathcal{F} \;-\; \epsilon_\delta\,\frac{e}{c}\left(\frac{\partial\delta\mathbf{A}_{\rm gy}}{\partial t}\,\frac{\partial\mathcal{F}}{\partial w} \;-\; \frac{\Omega}{B}\,\frac{\partial\delta\mathbf{A}_{\rm gy}}{\partial\mu}\,\frac{\partial\mathcal{F}}{\partial\zeta}\right), \\ \frac{\partial\mathcal{F}}{\partial\mu^*} \;&\equiv\; \frac{\partial\mathcal{F}}{\partial\mu} \;-\; \epsilon_\delta^2\left(\frac{e}{c}\,\frac{\partial\delta\mathbf{A}_{\rm gy}}{\partial t}\boldsymbol{\cdot}\frac{\partial\delta\mathbf{A}_{\rm gy}}{\partial\mu}\times\widehat{\mathsf{b}}_0\right)\frac{\partial\mathcal{F}}{\partial w}. \end{aligned}$$

We recover the guiding-center Poisson bracket (149) from Eq. (158) with the Hamiltonian gyrocenter model ($\delta\mathbf{A}_{\rm gy} = 0$).

The nonlinear gyrocenter Hamilton's equations are

$$\begin{aligned} \dot{\overline{\mathbf{X}}} \;&=\; \frac{c\widehat{\mathsf{b}}_0}{eB_\|^*}\times\left(\nabla\overline{H} \;+\; \epsilon_\delta\,\frac{e}{c}\,\frac{\partial\delta\mathbf{A}_{\rm gy}}{\partial t}\right) \\ &\quad+\; \frac{\partial\overline{H}}{\partial\overline{p}_\|}\,\frac{\mathbf{B}^*}{B_\|^*}, \end{aligned} \tag{159}$$

$$\dot{\overline{p}}_\| \;=\; -\,\frac{\mathbf{B}^*}{B_\|^*}\boldsymbol{\cdot}\left(\nabla\overline{H} \;+\; \epsilon_\delta\,\frac{e}{c}\,\frac{\partial\delta\mathbf{A}_{\rm gy}}{\partial t}\right), \tag{160}$$

where the gyrocenter Hamiltonian $\overline{H} = \overline{H}_0 + \epsilon_\delta\,\overline{H}_1 + \epsilon_\delta^2\,\overline{H}_2$ is derived in the next sections. Here, we note that the gyrocenter Hamilton's equations (159)-(160) satisfy the gyrocenter Liouville theorem

$$0 \;=\; \frac{\partial B_\|^*}{\partial t} \;+\; \nabla\boldsymbol{\cdot}\left(B_\|^*\,\dot{\overline{\mathbf{X}}}\right) \;+\; \frac{\partial}{\partial\overline{p}_\|}\left(B_\|^*\,\dot{\overline{p}}_\|\right). \tag{161}$$

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. (103), with $\overline{\Gamma}_1 \equiv (e/c)\,\delta\mathbf{A}_{\rm gy}\boldsymbol{\cdot}\mathsf{d}\overline{\mathbf{X}}$, the first-order generating vector field for the gyrocenter phase-space transformation is

$$\begin{aligned} G_1^a \;&=\; \{S_1,\; \mathcal{Z}^a\}_0 \;+\; \frac{e}{c}\,\delta\mathbf{A}_{\rm gc}\boldsymbol{\cdot}\{\mathbf{X}+\boldsymbol{\rho},\; \mathcal{Z}^a\}_0 \\ &\quad-\; \frac{e}{c}\,\delta\mathbf{A}_{\rm gy}\boldsymbol{\cdot}\{\mathbf{X},\; \mathcal{Z}^a\}_0\,, \end{aligned} \tag{162}$$

or its components can be explicitly given as

$$G_1^{\mathbf{X}} \;=\; -\,\widehat{\mathsf{b}}_0\,\frac{\partial S_1}{\partial p_\|} \;-\; \frac{c\widehat{\mathsf{b}}_0}{eB_0}\times\nabla S_1$$

$$+ \left(\delta\mathbf{A}_{\rm gc} \;-\; \alpha\,\langle\delta\mathbf{A}_{\perp{\rm gc}}\rangle \right) \times \frac{\widehat{\mathsf{b}}_0}{B_0}\;, \qquad (163)$$

$$G_1^{p_\|} \;=\; \widehat{\mathsf{b}}_0\cdot\nabla S_1 \;+\; \frac{e}{c}\Big(\delta A_{\|{\rm gc}} \;-\; \beta\,\langle\delta A_{\|{\rm gc}}\rangle\Big)\,, \qquad (164)$$

$$G_1^{\mu} \;=\; \frac{e}{mc}\left(\frac{e}{c}\,\delta\mathbf{A}_{\perp{\rm gc}}\cdot\frac{\partial\boldsymbol{\rho}}{\partial\zeta} \;+\; \frac{\partial S_1}{\partial\zeta}\right), \qquad (165)$$

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

$$G_1^{w} \;=\; -\,\frac{\partial S_1}{\partial t}, \qquad (167)$$

where effects due to background magnetic field nonuniformity are omitted for clarity. Here, we note that the gyrocenter parallel momentum is expanded as

$$\overline{p}_\| \;=\; p_\| \;+\; \epsilon_\delta\,\frac{e}{c}\left(\delta A_{\|{\rm gc}} \;-\; \beta\,\langle\delta A_{\|{\rm gc}}\rangle\right) + \cdots, \qquad (168)$$

which shows that the gyrocenter parallel momentum $\overline{p}_\|$ 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 (108) as

$$\overline{H}_1 \;\equiv\; e\,\delta\psi_{\rm gc} \;-\; \{S_1,\;\mathcal{H}_0\}_0,$$

where the effective first-order potential is defined as

$$\delta\psi_{\rm gc} \;\equiv\; \delta\phi_{\rm gc} \;-\; \delta\mathbf{A}_{\rm gc}\cdot\frac{\mathbf{v}}{c} \;+\; \beta\;\frac{v_\|}{c}\,\langle\delta A_{\|{\rm gc}}\rangle. \qquad (169)$$

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

$$\overline{H}_1 \;\equiv\; e\,\langle\delta\psi_{\rm gc}\rangle \;=\; e\left\langle \delta\phi_{\rm gc} \;-\; \frac{\mathbf{v}_\perp}{c}\cdot\delta\mathbf{A}_{\perp{\rm gc}}\right\rangle \;-\; \frac{ev_\|}{c}\,(1-\beta)\,\langle\delta A_{\|{\rm gc}}\rangle, \qquad (170)$$

while the solution for the scalar field S_1 is

$$S_1 \;=\; \frac{e}{\Omega_0}\int \delta\widetilde{\psi}_{\rm gc}\,d\zeta \;\equiv\; \frac{e}{\Omega_0}\,\delta\widetilde{\Psi}_{\rm gc}, \qquad (171)$$

where $\delta\widetilde{\psi}_{\rm gc} \equiv \delta\psi_{\rm gc} - \langle\delta\psi_{\rm gc}\rangle$ denotes the gyroangle-dependent part of the first-order effective potential (169).

While the (linear) first-order gyrocenter Hamiltonian (170) is sufficient for applications of 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 $\overline{H}_2$ for two important reasons. First, the second-order gyrocenter Hamiltonian $\overline{H}_2$ is needed in order to obtain the important polarization and magnetization effects which, within the variational formulation of self-consistent gyrokinetic Vlasov-Maxwell theory presented here, have variational definitions expressed in terms of the partial derivatives $\partial\overline{H}_2/\partial\mathbf{E}_1$ and $\partial\overline{H}_2/\partial\mathbf{B}_1$, respectively (see Section III). Second, once polarization and magnetization effects are included in the gyrokinetic Maxwell equations, the second-order gyrocenter Hamiltonian $\overline{H}_2$ must be kept in the gyrokinetic Vlasov Lagrangian density in order to obtain an exact energy conservation law (derived by Noether method) for the gyrokinetic Vlasov-Maxwell equations.

Hence, the general expression for the second-order gyrocenter Hamiltonian is obtained from Eq. (114) as

$$\begin{aligned}\overline{H}_2 \;=\;& -\,\frac{e^2}{2\Omega_0}\,\left\langle\left\{\delta\widetilde{\Psi}_{\rm gc},\;\delta\widetilde{\psi}_{\rm gc}\right\}_0\right\rangle \\ &+\;\frac{e^2}{2mc^2}\left(\left\langle|\delta\mathbf{A}_{\rm gc}|^2\right\rangle \;-\;\beta\,\langle\delta A_{\|{\rm gc}}\rangle^2\right) \\ &+\;\alpha\,\langle\delta\mathbf{A}_{\perp{\rm gc}}\rangle\cdot\frac{\widehat{\mathsf{b}}_0}{B_0}\times\nabla\overline{H}_1, \end{aligned} \qquad (172)$$

where the first term (denoted $\overline{K}_2$) describes low-frequency ponderomotive effects associated with the elimination of the fast gyro-motion time scale while the remaining terms explicitly involve magnetic perturbations and the choice of gyrocenter-model parameters (α,β).

D. Nonlinear Gyrokinetic Vlasov Equation

Once the linear gyrocenter Hamiltonian (170) and the nonlinear gyrocenter Hamiltonian (172) is obtained, it is a simple step to derive the corresponding nonlinear gyrokinetic Vlasov equation for the gyrocenter Vlasov distribution $\overline{F}$:

$$0 \;=\; \frac{\partial\overline{F}}{\partial t} \;+\; \left\{\overline{F},\;\overline{H}_{\rm gy}\right\}, \qquad (173)$$

where the nonlinear gyrocenter Hamiltonian is $\overline{H}_{\rm gy} = \overline{H}_{\rm gc} + e\,\delta\Psi_{\rm gy}$. Up to second order in the amplitude parameter ϵ_δ, the extended phase-space gyrocenter Hamiltonian is, therefore, expressed as

$$\begin{aligned}\overline{\mathcal{H}} \;=\;& \overline{\mathcal{H}}_0 \;+\; \epsilon_\delta\;e\langle\delta\psi_{\rm gc}\rangle \;+\; \frac{\epsilon_\delta^2\,e^2}{2\,mc^2}\,\left\langle|\delta\mathbf{A}_{\rm gc}|^2\right\rangle \\ &-\;\frac{\epsilon_\delta^2\,e^2}{2\,\Omega_0}\,\left\langle\left\{\delta\widetilde{\Psi}_{\rm gc},\;\delta\widetilde{\psi}_{\rm gc}\right\}\right\rangle, \end{aligned} \qquad (174)$$

where $\overline{\mathcal{H}}_0 = \overline{p}_\|^2/2m + \overline{\mu}\,B_0 - \overline{w}$ denotes the unperturbed extended guiding-center Hamiltonian. In order to simplify our presentation, we, henceforth, adopt the Hamiltonian gyrocenter model (with $\alpha = 0 = \beta$) for the remainder of this Section.

We have, thus, 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 include an electromagnetic field response to the gyrocenter Hamiltonian dynamics, a set of low-frequency Maxwell's 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's equations and the exact gyrokinetic energy invariant, we investigate the connection between the particle Vlasov distribution f and the gyrocenter Vlasov distribution $\overline{F}$.

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

$$f \;=\; \mathsf{T}_{\rm gc}F \;=\; e^{-\boldsymbol{\rho}\cdot\nabla}\; F \tag{175}$$

The pull-back transformation from the gyrocenter Vlasov distribution $\overline{F}$ to the guiding-center Vlasov distribution F, on the other hand, is expressed as

$$\begin{aligned} F \;=\; \mathsf{T}_{\rm gy}\overline{F} \;=\; \overline{F} \;+\; \epsilon_\delta\left\{S_1,\;\overline{F}\right\} \\ +\;\epsilon_\delta\,\frac{e}{c}\,\delta\mathbf{A}_{\rm gc}\cdot\left\{\mathbf{X}+\boldsymbol{\rho},\;\overline{F}\right\}. \end{aligned} \tag{176}$$

We point that no information is lost in transforming the Vlasov equation in particle phase space to the gyrokinetic Vlasov equation in gyrocenter phase space since

$$\begin{aligned} \frac{df}{dt} \;&=\; \frac{d}{dt}\left(\mathsf{T}_{\rm gc}\,\mathsf{T}_{\rm gy}\,\overline{F}\right) \;=\; \mathsf{T}_{\rm gc}\left[\left(\mathsf{T}_{\rm gc}^{-1}\frac{d}{dt}\mathsf{T}_{\rm gc}\right)\mathsf{T}_{\rm gy}\,\overline{F}\right] \\ &=\; \mathsf{T}_{\rm gc}\left[\frac{d_{\rm gc}}{dt}\left(\mathsf{T}_{\rm gy}\,\overline{F}\right)\right] \\ &=\; \mathsf{T}_{\rm gc}\,\mathsf{T}_{\rm gy}\left[\left(\mathsf{T}_{\rm gy}^{-1}\frac{d_{\rm gc}}{dt}\mathsf{T}_{\rm gy}\right)\overline{F}\right] \\ &=\; \mathsf{T}_{\rm gc}\,\mathsf{T}_{\rm gy}\left(\frac{d_{\rm gy}\overline{F}}{dt}\right), \end{aligned} \tag{177}$$

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

Next, we introduce the guiding-center Poisson bracket associated with the coordinates $(\mathbf{X}, \mathcal{E}, \mu, \zeta)$:

$$\begin{aligned} \{F,\,G\} \;=\; \Omega\left[\frac{\partial F}{\partial\zeta}\left(\frac{\partial G}{\partial\mathcal{E}} + \frac{1}{B}\frac{\partial G}{\partial\mu}\right)\right. \\ \left.-\;\left(\frac{\partial F}{\partial\mathcal{E}} + \frac{1}{B}\frac{\partial F}{\partial\mu}\right)\frac{\partial G}{\partial\zeta}\right] \\ +\;\mathbf{v}_{\rm gc}\cdot\left(\nabla F\,\frac{\partial G}{\partial\mathcal{E}} \;-\; \frac{\partial F}{\partial\mathcal{E}}\,\nabla G\right) \\ -\;\frac{c\widehat{\mathsf{b}}}{eB}\cdot\nabla F\times\nabla G, \end{aligned} \tag{178}$$

where $\mathbf{v}_{\rm gc} = v_\parallel\,\widehat{\mathsf{b}}$ in the absence of magnetic-field nonuniformity. Hence, by combining the guiding-center and gyrocenter pull-backs, we find the pull-back transformation from the gyrocenter Vlasov distribution $\overline{F}$ and the particle Vlasov distribution f:

$$\begin{aligned} f \;=\; e^{-\boldsymbol{\rho}\cdot\nabla}\left[\overline{F} \;-\; e\,\langle\delta\psi_{\rm gc}\rangle\;\left(\frac{\partial\overline{F}}{\partial\mathcal{E}} + \frac{1}{B}\frac{\partial\overline{F}}{\partial\mu}\right)\right] \\ +\;e\delta\phi\,\frac{\partial\overline{F}}{\partial\mathcal{E}} \;+\; \frac{e}{B}\left(\delta\phi - \frac{v_\parallel}{c}\,\delta A_\parallel\right)\frac{\partial\overline{F}}{\partial\mu} \\ +\;\delta\mathbf{A}\times\frac{\widehat{\mathsf{b}}}{B}\cdot\nabla\overline{F}. \end{aligned} \tag{179}$$

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 $\overline{F}$ and the nonadiabatic component of the perturbed particle Vlasov distribution.

Lastly, by comparing the pull-back decomposition (179) with the Frieman-Chen decomposition (14), we obtain a relation between the first-order correction $\overline{F}_1$ to the gyrocenter distribution $\overline{F} = \overline{F}_0 + \epsilon_\delta\,\overline{F}_1$ and the nonadiabatic part $\overline{G}_1$:

$$\overline{F}_1 \;\equiv\; \overline{G}_1 \;+\; e\,\langle\delta\psi_{\rm gc}\rangle\,\frac{\partial\overline{F}_0}{\partial\mathcal{E}} \tag{180}$$

Substituting this relation into the nonlinear gyrokinetic Vlasov equation (173), with the gyrocenter Hamiltonian truncated at first order $\overline{H}_{\rm gy} = \overline{H}_{\rm gc} + \epsilon_\delta\,e\,\langle\delta\psi_{\rm gc}\rangle$:

$$0 \;=\; \frac{d_{\rm gy}\overline{F}}{dt} \;\equiv\; \frac{d_{\rm gc}\overline{F}}{dt} \;+\; \epsilon_\delta\;e\left\{\overline{F},\;\langle\delta\psi_{\rm gc}\rangle\right\},$$

we obtain

$$\begin{aligned} 0 \;&=\; \frac{d_{\rm gy}}{dt}\left[\overline{F}_0 \;+\; \epsilon_\delta\left(\overline{G}_1 \;+\; e\,\langle\delta\psi_{\rm gc}\rangle\,\frac{\partial\overline{F}_0}{\partial\mathcal{E}}\right)\right] \\ &=\; \epsilon_\delta\left\{\overline{F}_0,\;e\,\langle\delta\psi_{\rm gc}\rangle\right\} \;+\; \epsilon_\delta\,\frac{d_{\rm gc}\overline{G}_1}{dt} \\ &\quad+\;\epsilon_\delta^2\left\{\overline{G}_1,\;e\,\langle\delta\psi_{\rm gc}\rangle\right\} \;+\; \epsilon_\delta\,e\,\frac{d_{\rm gc}\langle\delta\psi_{\rm gc}\rangle}{dt}\,\frac{\partial\overline{F}_0}{\partial\mathcal{E}}. \end{aligned} \tag{181}$$

Using the guiding-center Poisson bracket (178), we find

$$\begin{aligned} \left\{\overline{F}_0,\;e\,\langle\delta\psi_{\rm gc}\rangle\right\} \;=\; \frac{c\widehat{\mathsf{b}}}{B}\times\nabla\langle\delta\psi_{\rm gc}\rangle\cdot\nabla\overline{F}_0 \\ -\;\left(e\mathbf{v}_{\rm gc}\cdot\nabla\langle\delta\psi_{\rm gc}\rangle\right)\frac{\partial\overline{F}_0}{\partial\mathcal{E}} \end{aligned}$$

and the nonlinear gyrokinetic Vlasov equation (181) becomes the Frieman-Chen nonlinear gyrokinetic Vlasov:

$$\frac{d_{\rm gc}\overline{G}_1}{dt} \;=\; -\left(e\,\frac{\partial\langle\delta\psi_{\rm gc}\rangle}{\partial t}\,\frac{\partial\overline{F}_0}{\partial\mathcal{E}} \;+\; \frac{c\widehat{\mathsf{b}}}{B}\times\nabla\langle\delta\psi_{\rm gc}\rangle\cdot\nabla\overline{F}_0\right)$$

$$- \; \frac{c\widehat{\mathsf{b}}}{B} \times \nabla\langle\delta\psi_{\rm gc}\rangle \cdot \nabla\overline{G}_1, \qquad (182)$$

where higher-order terms (e.g., $\epsilon_{\rm B}\epsilon_\delta^2$) were omitted. Hence, the Frieman-Chen nonlinear gyrokinetic Vlasov equation is contained in the nonlinear gyrokinetic Vlasov equation (173). Furthermore, while the adiabatic/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, in fact, is that it performs a partial solution of the Vlasov equation associated with the fast-time gyro-motion dynamics.

VI. GYROKINETIC VARIATIONAL FORMULATION

After having derived various expressions for the nonlinear gyrocenter Hamiltonian and its associated nonlinear gyrokinetic Vlasov equation, we now turn our attention to deriving 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, we also wish to derive the exact energy conservation law these nonlinear gyrokinetic equations satisfy. These two tasks are simultaneously performed in this Section by using a variational formulation for the nonlinear gyrokinetic Vlasov-Maxwell equations (Brizard, 2000b).

A. Reduced Variational Principle

Before proceeding to the variational formulation formulation for the nonlinear gyrokinetic Vlasov-Maxwell equations, we briefly review the variational formulation of the reduced Vlasov-Maxwell equations (118) and (126)-(127). Here, we assume that the transformation associated with the dynamical reduction is motivated by electromagnetic-field perturbations generated by the first-order perturbed four-potential $A_1^\mu = (\phi_1, \mathbf{A}_1)$; the symbol "$\delta$" is, henceforth, uniquely reserved to denote a functional variation in the remainder of this Section.

We begin with the reduced plasma variational principle (Brizard, 2000a)

$$\begin{aligned} 0 \;&=\; \delta\int\overline{\mathcal{L}}\; d^4x \qquad (183) \\ &\equiv\; \delta\int\left[\;\frac{1}{16\pi}\,\mathsf{F}_{\mu\nu}\,\mathsf{F}^{\nu\mu} \;+\; \overline{\mathcal{L}}_{\rm V}(\overline{\mathcal{F}}, A_{1\mu}; \mathsf{F}_{1\mu\nu})\;\right] d^4x, \end{aligned}$$

where the first term represents the electromagnetic action functional (where $\mathsf{F} = \mathsf{F}_0 + \epsilon_\delta\,\mathsf{F}_1$ denotes the total electromagnetic tensor) and the second term represents the reduced plasma dynamics, with the reduced Vlasov Lagrangian density defined as (Brizard, 2000a,b)

$$\overline{\mathcal{L}}_{\rm V} \;\equiv\; -\sum\int d^4\overline{p}\;\overline{\mathcal{F}}\;\overline{\mathcal{H}}(\overline{\mathcal{Z}}, A_{1\mu}; \mathsf{F}_{1\mu\nu}), \qquad (184)$$

where $\overline{\mathcal{H}}$ denotes the reduced Hamiltonian and $\overline{\mathcal{F}}$ denotes the reduced Vlasov distribution defined in Eqs. (94) and (117), respectively. Note that the reduced particle Lagrangian density (184) depends not only on the perturbed electromagnetic four-potential $A_1^\mu = (\phi_1, \mathbf{A}_1)$, through the standard interaction Lagrangian, but also depends on the electromagnetic tensor $\mathsf{F}_{1\mu\nu} = \partial_\mu A_{1\nu} - \partial_\nu A_{1\mu}$ as a result of the dynamical reduction introduced by the extended phase-space transformation (84).

The reduced variational principle (183) considers *Eulerian* variations for the extended reduced Vlasov distribution $\overline{\mathcal{F}}$ defined in terms of the *gauge* function $\overline{\mathcal{S}}(\overline{\mathcal{Z}})$ and the extended Poisson bracket (78) as

$$\delta\overline{\mathcal{F}} \;\equiv\; \{\overline{\mathcal{S}},\, \overline{\mathcal{F}}\}_\epsilon, \qquad (185)$$

where $\{\;,\;\}_\epsilon$ denotes the transformed Poisson bracket.

Under Eulerian variations of the electromagnetic potentials δA^μ and the Eulerian variation (185) of the extended Vlasov distribution $\overline{\mathcal{F}}$, the variation of the reduced Lagrangian density (183) can be expressed as

$$\begin{aligned} \delta\overline{\mathcal{L}} \;&=\; -\sum\int d^4\overline{p}\;\overline{\mathcal{S}}\;\{\overline{\mathcal{F}},\,\overline{\mathcal{H}}\}_{\mathsf{z}} \qquad (186) \\ &+\; \delta A_{1\mu}\left[\;\frac{\partial}{\partial x^\nu}\left(\frac{\partial\overline{\mathcal{L}}}{\partial\mathsf{F}_{1\nu\mu}}\right) \;-\; \sum\int d^4\overline{p}\;\overline{\mathcal{F}}\;\frac{\partial\overline{\mathcal{H}}}{\partial A_{1\mu}}\;\right] \\ &+\; \frac{\partial}{\partial x^\nu}\left(\sum d^4\overline{p}\;\overline{\mathcal{S}}\,\overline{\mathcal{F}}\;\frac{\partial\overline{\mathcal{H}}}{\partial\overline{p}_\nu} \;+\; \delta A_{1\mu}\;\frac{\partial\overline{\mathcal{L}}}{\partial\mathsf{F}_{1\mu\nu}}\right). \end{aligned}$$

Stationarity of the reduced action functional (183) with respect to arbitrary virtual phase-space displacements generated by $\overline{\mathcal{S}}$ yields the reduced Vlasov equation (119) in extended phase space. Next, stationarity with respect to arbitrary variations δA_ν yields the reduced Maxwell equations (126)-(127) expressed as Euler-Lagrange equations:

$$\frac{\partial}{\partial x^\mu}\left(\frac{\partial\overline{\mathcal{L}}}{\partial\mathsf{F}_{1\mu\nu}}\right) \;=\; \frac{\partial\overline{\mathcal{L}}}{\partial A_{1\nu}}, \qquad (187)$$

where the reduced four-current density $\overline{J}^\nu \equiv (c\overline{\rho},\; \overline{\mathbf{J}})$ is defined as

$$\overline{J}^\nu \;\equiv\; c\,\frac{\partial\overline{\mathcal{L}}_{\rm V}}{\partial A_{1\nu}} \;=\; -\sum\int d^4\overline{p}\;\overline{\mathcal{F}}\;\left(c\,\frac{\partial\overline{\mathcal{H}}}{\partial A_{1\nu}}\right). \qquad (188)$$

It is convenient to introduce the antisymmetric *macroscopic* electromagnetic tensor (Boghosian, 1987)

$$\begin{aligned} \mathsf{M}^{\mu\nu} \;&\equiv\; -\,\frac{\partial\overline{\mathcal{L}}}{\partial\mathsf{F}_{1\mu\nu}} \\ &=\; \frac{1}{4\pi}\,\mathsf{F}^{\mu\nu} \;+\; \sum\int d^4\overline{p}\;\overline{\mathcal{F}}\;\frac{\partial\overline{\mathcal{H}}}{\partial\mathsf{F}_{1\mu\nu}}, \qquad (189) \end{aligned}$$

with components $\mathsf{M}^{0i} = D^i/4\pi$ and $\mathsf{M}^{ij} = \epsilon^{ijk}\, H_k/4\pi$, so that the variational definitions of the polarization and magnetization vectors are

$$
\begin{aligned}
(\mathbf{P}_\epsilon,\; \mathbf{M}_\epsilon) \;&\equiv\; \left(\frac{\partial\overline{\mathcal{L}}_{\rm V}}{\partial\mathbf{E}_1},\; \frac{\partial\overline{\mathcal{L}}_{\rm V}}{\partial\mathbf{B}_1}\right) \\
&=\; -\sum \int d^4\overline{p}\;\overline{\mathcal{F}}\; \left(\frac{\partial\overline{\mathcal{H}}}{\partial\mathbf{E}_1},\, \frac{\partial\overline{\mathcal{H}}}{\partial\mathbf{B}_1}\right). \qquad (190)
\end{aligned}
$$

Note that since only the second-order reduced Hamiltonian $\overline{H}_2$ depends on the perturbed electromagnetic field $(\mathbf{E}_1, \mathbf{B}_1)$, the reduced polarization and magnetization vectors (190) are first-order expressions themselves.

Lastly, using these definitions, the macroscopic Maxwell equations (187) become

$$
\frac{\partial\mathsf{M}^{\mu\nu}}{\partial x^\mu} \;=\; -\,\frac{1}{c}\,\overline{J}^\nu. \qquad (191)
$$

and we note that the reduced polarization-magnetization four-current density $K_\epsilon^\mu \equiv (c\,\rho_{\rm pol}, \mathbf{J}_{\rm pol} + \mathbf{J}_{\rm mag})$ can be covariantly defined in terms of the reduced Lagrangian $\overline{\mathcal{L}}_{\rm V}$ as

$$
K_\epsilon^\mu \;\equiv\; -\,c\,\frac{\partial}{\partial x^\nu}\left(\frac{\partial\overline{\mathcal{L}}_{\rm V}}{\partial\mathsf{F}_{1\nu\mu}}\right), \qquad (192)
$$

which is manifestly space-time divergenceless ($\partial_\mu K_\epsilon^\mu \equiv 0$) as a result of the antisymmetry of the electromagnetic tensor $\mathsf{F}_{1\mu\nu}$.

B. Nonlinear Gyrokinetic Vlasov-Maxwell Equations

We now derive the nonlinear self-consistent gyrokinetic Vlasov-Maxwell equations (using the Hamiltonian gyrocenter model) from a reduced variational principle, which 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

$$
\begin{aligned}
\mathcal{A}_{\rm gy} \;&=\; -\int d^8\mathcal{Z}\;\mathcal{F}(\mathcal{Z})\;\mathcal{H}(\mathcal{Z}; A_{1\mu}, \mathsf{F}_{1\mu\nu}) \\
&\quad+\; \int \frac{d^4x}{8\pi}\;\left(|\nabla\Phi|^2 \;-\; |\mathbf{B}|^2\right), \qquad (193)
\end{aligned}
$$

where $\mathcal{H}$ denotes the nonlinear gyrocenter Hamiltonian (174) and we, henceforth, use the notation

$$
\Phi \;\equiv\; \epsilon\,\phi_1 \quad \text{and} \quad \mathbf{B} \;\equiv\; \mathbf{B}_0 \;+\; \epsilon\,\nabla\times\mathbf{A}_1,
$$

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 wherever 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 (193) 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 nonlinear gyrokinetic ordering (5)-(6).

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

$$
\begin{aligned}
\delta\mathcal{A}_{\rm gy} \;&=\; \int \frac{d^4x}{4\pi}\Big(\;\epsilon\,\nabla\delta\phi_1\cdot\nabla\Phi \;-\; \epsilon\,\nabla\times\delta\mathbf{A}_1\cdot\mathbf{B}\;\Big) \\
&\quad-\; \int d^8\mathcal{Z}\Big[\;\delta\mathcal{F}(\mathcal{Z})\;\mathcal{H} \\
&\quad+\; \mathcal{F}(\mathcal{Z})\int d^3x\left(\delta A_{1\mu}(\mathbf{x})\;\frac{\delta H}{\delta A_{1\mu}(\mathbf{x})}\right)\Big]. \qquad (194)
\end{aligned}
$$

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

$$
\delta\mathcal{F} \;\equiv\; \{\mathcal{S},\; \mathcal{F}\}, \qquad (195)
$$

where $\{\;,\;\}$ is the extended guiding-center Poisson bracket (149). The functional derivatives $\delta H/\delta A_{1\mu}(\mathbf{x})$ in Eq. (194), on the other hand, are evaluated using the gyrocenter Hamiltonian (174) (to second order in ϵ) as

$$
\frac{\delta H}{\delta A_{1\mu}(\mathbf{x})} \;\equiv\; -\,\epsilon\,e\left\langle \mathsf{T}_{\rm gy}^{-1}\left(\frac{v^\mu}{c}\,\delta_{\rm gc}^3\right)\right\rangle, \qquad (196)
$$

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

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

After re-arranging terms and integrating by parts, the variation (194) becomes

$$
\begin{aligned}
\delta\mathcal{A}_{\rm gy} \;&=\; \int d^4x\,(\partial\cdot\mathcal{J}_{\rm gy}) - \int d^8\mathcal{Z}\,\mathcal{S}\,\{\mathcal{F},\;\mathcal{H}\} \qquad (197) \\
&-\; \int d^4x\Big\{\epsilon\,\delta\phi_1\left[\;\frac{\nabla^2\Phi}{4\pi} \;+\; e\int d^6Z\;F\;\left\langle\mathsf{T}_{\rm gy}^{-1}\delta_{\rm gc}^3\right\rangle\;\right] \\
&+\; \epsilon\,\delta\mathbf{A}_1\cdot\left[\frac{\nabla\times\mathbf{B}}{4\pi} - e\int d^6Z\;F\;\left\langle\mathsf{T}_{\rm gy}^{-1}\left(\frac{\mathbf{v}}{c}\,\delta_{\rm gc}^3\right)\right\rangle\right]\Big\},
\end{aligned}
$$

where the first term on the right side of Eq. (197) involves the exact space-time divergence

$$
\begin{aligned}
\partial\cdot\mathcal{J}_{\rm gy} \;&\equiv\; \frac{\partial}{\partial x^\mu}\left(\int d^4p\;\mathcal{S}\;\mathcal{F}\;\dot{X}^\mu\right) \\
&\quad+\; \nabla\cdot\left(\epsilon\,\frac{\delta\phi_1}{4\pi}\,\nabla\Phi \;-\; \epsilon\,\frac{\delta\mathbf{A}_1}{4\pi}\times\mathbf{B}\right), \qquad (198)
\end{aligned}
$$

where $\dot{X}^\mu \equiv \{X^\mu,\, \mathcal{H}\}$ denotes the lowest-order gyrocenter four-velocity. Since Eq. (198) is an exact space-time divergence, it does not contribute to the reduced variational principle $\delta\mathcal{A}_{\rm gy} \equiv 0$.

By requiring that the action functional $\mathcal{A}_{\rm gy}$ be stationary with respect to arbitrary variations $\mathcal{S}$ and δA_1^μ (which vanish on the integration boundaries), we find the nonlinear gyrokinetic Vlasov equation

$$0 \;=\; \{\mathcal{F},\, \mathcal{H}\}, \tag{199}$$

and the gyrokinetic Maxwell equations: the gyrokinetic Poisson equation

$$\begin{aligned}\nabla^2\Phi(\mathbf{x}) \;&=\; -\,4\pi\, e \int d^6Z\; F\; \left\langle \mathsf{T}_{\rm gy}^{-1}\; \delta^3_{\rm gc}\right\rangle \\ &\equiv\; -\,4\pi\, e \int d^3p\; \left\langle e^{-\boldsymbol{\rho}\cdot\nabla}\,(\mathsf{T}_{\rm gy}F)\right\rangle, \end{aligned}\tag{200}$$

and the gyrokinetic Ampère equation

$$\begin{aligned}\nabla\times\mathbf{B}(\mathbf{x}) \;&=\; \frac{4\pi\, e}{c}\int d^6Z\; F(Z)\; \left\langle \mathsf{T}_{\rm gy}^{-1}\left(\mathbf{v}\,\delta^3_{\rm gc}\right)\right\rangle \\ &\equiv\; \frac{4\pi\, e}{c}\int d^3p\; \left\langle e^{-\boldsymbol{\rho}\cdot\nabla}\,(\mathbf{v}\;\mathsf{T}_{\rm gy}F)\right\rangle, \end{aligned}\tag{201}$$

which are actually valid for all gyrocenter models discussed in Sec. V.B. If we now integrate the extended gyrokinetic Vlasov equation $\{\mathcal{F},\, \mathcal{H}\}_{\mathcal{Z}} = 0$ over the energy coordinate w, we obtain the standard nonlinear gyrokinetic Vlasov equation written explicitly as

$$\begin{aligned}&\frac{\partial F}{\partial t} \;+\; \left(\frac{\mathbf{B}^*}{B_\parallel^*}\,\frac{\partial H}{\partial p_\parallel} \;+\; \frac{c\widehat{\mathsf{b}}}{eB_\parallel^*}\times\nabla H\right)\cdot\nabla F \\ &-\;\left(\frac{\mathbf{B}^*}{B_\parallel^*}\cdot\nabla H\right)\frac{\partial F}{\partial p_\parallel} \;=\; 0. \end{aligned}\tag{202}$$

The nonlinear equations (200), (201), and (202), with the gyrocenter Hamiltonian (174), are the self-consistent nonlinear gyrokinetic Vlasov-Maxwell equations in general magnetic field geometry (Brizard, 1989a).

C. Gyrokinetic Energy Conservation Law

We now apply the Noether method on the gyrokinetic action functional (193) to derive an exact gyrokinetic energy conservation law. By substituting Eqs. (199), (200), and (201) into Eq. (197), the variational equation $\delta\mathcal{A}_{\rm gy} \equiv \int \delta\mathcal{L}_{\rm gy}\, d^4x$ yields the Noether equation

$$\delta\mathcal{L}_{\rm gy} \;\equiv\; \partial\cdot\mathcal{J}_{\rm gy}. \tag{203}$$

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

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

$$\left.\begin{aligned}\mathcal{S} \;&=\; -\,w\;\delta t \\ \delta\phi_1 \;&=\; -\delta t\;\partial_t\phi_1 \\ \delta\mathbf{A}_1 \;&=\; -\delta t\;\partial_t\mathbf{A}_1 \;\equiv\; c\,\delta t\;(\mathbf{E}_1 + \nabla\phi_1) \\ \delta\mathcal{L}_{\rm gy} \;&=\; -\delta t\;\partial_t\mathcal{L}_{\rm gy}\end{aligned}\right\}. \tag{204}$$

In Eq. (204), the gyrokinetic Vlasov-Maxwell Lagrangian density is $\mathcal{L}_{\rm gy} = (|\nabla\Phi|^2 - |\mathbf{B}|^2)/8\pi$ after the physical constraint $\mathcal{H} = 0$ is imposed in the space-time integrand of the reduced action functional (193).

By combining Eq. (204) with Eqs. (198) and (203), we obtain, after rearranging and cancelling some terms (Brizard, 2000b), the *local* gyrokinetic energy conservation law:

$$\frac{\partial\mathcal{E}_{\rm gy}}{\partial t} \;+\; \nabla\cdot\mathbf{S}_{\rm gy} \;=\; 0, \tag{205}$$

where the gyrokinetic energy density is

$$\begin{aligned}\mathcal{E}_{\rm gy} \;&=\; \int d^3p\; F\;\left(H \;-\; e\left\langle \mathsf{T}_{\rm gy}^{-1}\Phi_{\rm gc}\right\rangle\right) \\ &+\; \frac{1}{8\pi}\left(|\nabla\Phi|^2 \;+\; |\mathbf{B}|^2\right), \end{aligned}\tag{206}$$

while the gyrokinetic energy density flux is

$$\begin{aligned}\mathbf{S}_{\rm gy} \;&=\; \int d^3p\; F\left(H\,\dot{\mathbf{X}} \;-\; e\left\langle \mathsf{T}_{\rm gy}^{-1}\;\mathbf{v}\Phi_{\rm gc}\right\rangle\right) \\ &+\; \frac{\epsilon}{4\pi}\left(\,c\,\mathbf{E}_1\times\mathbf{B} \;-\; \Phi\;\nabla\frac{\partial\phi_1}{\partial t}\,\right). \end{aligned}\tag{207}$$

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

$$\begin{aligned}E \;&=\; \int\frac{d^3x}{8\pi}\left(|\nabla\Phi|^2 \;+\; |\mathbf{B}|^2\right) \\ &+\; \int d^6Z\; F\;\left(H \;-\; e\left\langle \mathsf{T}_{\rm gy}^{-1}\Phi_{\rm gc}\right\rangle\right). \end{aligned}\tag{208}$$

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 (Brizard, 1989a; Hahm *et al.*, 1988) gyrokinetic equations.

VII. SUMMARY

The foundations of modern nonlinear gyrokinetic theory are based on three important pillars: (1) a gyrokinetic Vlasov equation written in terms of a Hamiltonian with quadratic low-frequency ponderomotive-like terms;

(2) a set of gyrokinetic Maxwell equations written in terms of the gyrocenter Vlasov distribution that contain low-frequency polarization and magnetization terms (derived from the quadratic nonlinearities in the Hamiltonian); and (3) an exact energy conservation law for the gyrokinetic Vlasov-Maxwell equations that includes all the relevant linear and nonlinear coupling terms.

These three pillars were emphasized in Section III, where simplified forms of the nonlinear gyrokinetic equations were presented for the cases of electrostatic fluctuations, shear-Alfvenic 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.

Through the use of Lie-transform perturbation methods on extended particle phase space, we showed 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. 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

The authors would like to 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 would like to thank P. H. Diamond, J. A. Krommes, and W. M. Tang for crucial advice on various applications, and we wish to acknowledge 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. Lastly, 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, X. Garbet, Ph. Ghendrih, G. W. Hammett, F. L. Hinton, W. Horton, Y. Idomura, S. I. Itoh, K. Itoh, Y. Kishimoto, J. N. Leboeuf, N. Mattor, E. Mazzucato, H. Naitou, Y. Nishimura, C. Oberman, M. Ottaviani, W. Park, S. E. Parker, F. W. Perkins, 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. Zweben for travel support.

This work was support by the U. S. Department of Energy Contracts No. DE-AC03-76SFOO098 (AJB) and DE-AC02-CHO-3073 (TSH) and partly by the U. S. Department of Energy SciDAC Center for Gyrokinetic Particle Simulation of Turbulent Transport in Burning Plasmas (TSH).

APPENDIX A: MATHEMATICAL PRIMER

This Appendix presents a brief summary of the differential geometric foundations of Lie-transform perturbation methods. We also present a brief discussion of general magnetic geometry in terms of magnetic coordinates.

1. Exterior Differential Calculus

Differential k-forms (Flanders, 1989)

$$\boldsymbol{\omega}_k \;=\; \frac{1}{k!}\,\omega_{i_1 i_2 \ldots i_k}\,\mathsf{d}z^{i_1}\;\wedge\;\mathsf{d}z^{i_2}\;\wedge\;\cdots\;\wedge\;\mathsf{d}z^{i_k}$$

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

Note that the exterior derivative $\mathsf{d}\boldsymbol{\omega}_k$ of a differential k-form (or k-form for short) $\boldsymbol{\omega}_k$ is a $(k+1)$-form. Here, we are interested in the exterior derivatives of scalar fields f (defined as 0-forms) and 1-forms $\Gamma = \Gamma_a\,\mathsf{d}z^a$. First, the exterior derivative of a 0-form f is defined as

$$\mathsf{d}f \;\equiv\; \partial_a f\,\mathsf{d}z^a, \tag{A1}$$

and, thus, $\mathsf{d}f$ is a differential 1-form; note that its components are the components of ∇f. Next, the exterior derivative of a 1-form Γ is a 2-form:

$$\mathsf{d}\Gamma \;\equiv\; \mathsf{d}\Gamma_b\;\wedge\;\mathsf{d}z^b \;=\; \partial_a\Gamma_b\,\mathsf{d}z^a\;\wedge\;\mathsf{d}z^b,$$

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

$$\begin{aligned}\mathsf{d}\Gamma \;&=\; \frac{1}{2}\,(\partial_a\Gamma_b \;-\; \partial_b\Gamma_a)\,\mathsf{d}z^a\;\wedge\;\mathsf{d}z^b \\ &\equiv\; \frac{1}{2}\,\omega_{ab}\,\mathsf{d}z^a\;\wedge\;\mathsf{d}z^b,\end{aligned} \tag{A2}$$

where $\omega_{ab} = -\,\omega_{ba}$ denotes the antisymmetric components of the 2-form $\boldsymbol{\omega} \equiv \mathsf{d}\Gamma$.

An important difference between the exterior derivative d and the standard derivative d comes from the property that $\mathsf{d}^2\boldsymbol{\omega}_k = \mathsf{d}(\mathsf{d}\boldsymbol{\omega}_k) \equiv 0$ for any k-form $\boldsymbol{\omega}_k$. Indeed, for a 0-form, we find

$$\mathsf{d}^2 f \;=\; \partial^2_{ab} f\;\mathsf{d}z^a\;\wedge\;\mathsf{d}z^b \;=\; 0,$$

since $\partial^2_{ab} f$ is symmetric with respect to interchange $a \leftrightarrow b$, and for a 1-form:

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

A k-form $\boldsymbol{\omega}_k$ is said to be *closed* if its exterior derivative is $\mathsf{d}\boldsymbol{\omega}_k \equiv 0$, while a k-form $\boldsymbol{\omega}_k$ is said to be *exact* if it can be written in terms of a (k-1)-form Γ_{k-1} as $\boldsymbol{\omega}_k \equiv \mathsf{d}\Gamma_{k-1}$. 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 $\mathcal{J}$:

$$\boldsymbol{\Omega} \;\equiv\; \mathcal{J}(\mathbf{u})\,\mathsf{d}u^1 \;\wedge\; \mathsf{d}u^2 \;\wedge\; \mathsf{d}u^3$$

is a closed 3-form since $\mathsf{d}\boldsymbol{\Omega} \equiv 0$. Hence, according to the converse of Poincaré's Lemma, there exists a 2-form $\boldsymbol{\sigma}$ such that $\boldsymbol{\Omega} \equiv \mathsf{d}\boldsymbol{\sigma}$, where

$$\boldsymbol{\sigma} \;\equiv\; \frac{1}{2}\,\epsilon_{ijk}\,\sigma^k(\mathbf{u})\,\mathsf{d}u^i \;\wedge\; \mathsf{d}u^j$$

defines the infinitesimal area 2-form, with the Jacobian defined as $\mathcal{J} \equiv \partial\sigma^i(\mathbf{u})/\partial u^i$.

We now introduce the inner-product operation involving a vector field $\mathbf{v}$ and a k-form $\boldsymbol{\omega}_k$, denoted in the present work as $\mathbf{v}\cdot\boldsymbol{\omega}_k$, which creates a $(k-1)$-form. For example, for a 1-form, it is defined as $\mathbf{v}\cdot\Gamma = v^a\,\Gamma_a$ while for a 2-form, it is defined as

$$\mathbf{v}\cdot\boldsymbol{\omega} \;\equiv\; \frac{1}{2}\,\left(v^a\,\omega_{ab}\,\mathsf{d}z^b \;-\; \omega_{ab}\,v^b\,\mathsf{d}z^a\right) \;=\; v^a\,\omega_{ab}\,\mathsf{d}z^b.$$

Note that $\mathsf{d}(\mathbf{v}\cdot\boldsymbol{\Omega}) = \mathcal{J}^{-1}\partial_a(\mathcal{J}\,v^a)\,\boldsymbol{\Omega} \equiv (\nabla\cdot\mathbf{v})\,\boldsymbol{\Omega}$.

The Lie derivative $\pounds_{\mathbf{v}}$ along the vector field $\mathbf{v}$ of a k-form $\boldsymbol{\omega}_k$ is defined in terms of the Homotopy formula (Abraham and Marsden, 1978)

$$\pounds_{\mathbf{v}}\boldsymbol{\omega}_k \;\equiv\; \mathbf{v}\cdot\mathsf{d}\boldsymbol{\omega}_k \;+\; \mathsf{d}\left(\mathbf{v}\cdot\boldsymbol{\omega}_k\right). \tag{A3}$$

Here, we see that the Lie derivative of a k-form is itself a k-form. For example, the Lie derivative of a 0-form f along the vector field $\mathbf{v}$ is $\pounds_{\mathbf{v}} f \equiv \mathbf{v}\cdot\mathsf{d}f = v^a\,\partial_a f$ (i.e., the directional derivative $\mathbf{v}\cdot\nabla f$), while the Lie derivative of a 1-form $\Gamma = \Gamma_a\,\mathsf{d}z^a$ is

$$\pounds_{\mathbf{v}}\Gamma \;=\; \left[v^a\,\omega_{ab} \;+\; \partial_b\left(\mathbf{v}\cdot\Gamma\right)\right]\,\mathsf{d}z^b.$$

Note that the Lie derivative satisfies the Leibnitz property $\pounds_{\mathbf{v}}(f\,g) = (\pounds_{\mathbf{v}} f)\,g + f\,(\pounds_{\mathbf{v}} g)$ and $\pounds_{\mathbf{v}}$ commutes with d. For example, we consider the four-dimensional electromagnetic 1-form $\mathsf{A} = A_\mu\,\mathsf{d}x^\mu$ and its Lie derivative $\pounds_{\delta x}$ along the four-vector $\delta x^\mu = (c\delta t, \delta\mathbf{x})$:

$$\pounds_{\delta x}\mathsf{A} \;=\; \delta x^\mu\,\mathsf{F}_{\mu\nu}\,\mathsf{d}x^\nu \;+\; \mathsf{d}\left(\delta x^\mu\,A_\mu\right) \;\equiv\; -\,\delta\mathsf{A},$$

which defines the perturbed electromagnetic four-potential $\delta A^\mu = (\delta\phi, \delta\mathbf{A})$.

Next, we define the Lie transform generated by the vector field $\mathbf{v}$ as $\mathsf{T} \equiv \exp\pounds_{\mathbf{v}}$ such that T is distributive $\mathsf{T}(f\,g) \equiv (\mathsf{T}f)\,(\mathsf{T}g)$ and T commutes with d. For example, the pull-back operator $\mathsf{T}_\epsilon = \exp(\epsilon\,\pounds_{\boldsymbol{\xi}})$ associated with the nonuniform space transformation

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

which is generated by the vector field $\epsilon\,\boldsymbol{\xi}$, yields the identity

$$\begin{aligned} f(\mathbf{x}) \;&=\; F(\mathbf{X}) \;=\; F\left(\mathbf{x} \;+\; \epsilon\,\boldsymbol{\xi} \;+\; \frac{\epsilon^2}{2}\,\boldsymbol{\xi}\cdot\nabla\boldsymbol{\xi} \;+\; \cdots\right) \\ &=\; F(\mathbf{x}) \;+\; \epsilon\,\boldsymbol{\xi}\cdot\nabla F(\mathbf{x}) \;+\; \frac{\epsilon^2}{2}\,\boldsymbol{\xi}\cdot\nabla(\boldsymbol{\xi}\cdot\nabla F) \;+\; \cdots \\ &=\; \exp\left(\epsilon\,\boldsymbol{\xi}\cdot\nabla\right)\,F(\mathbf{x}) \;\equiv\; \exp\left(\epsilon\,\pounds_{\boldsymbol{\xi}}\right)\,F(\mathbf{x}). \end{aligned} \tag{A4}$$

Hence, we see that 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.

2. General Magnetic Field Geometry

A general magnetic field is written in divergenceless form (at least locally) as

$$\mathbf{B} \;\equiv\; \nabla\alpha\times\nabla\beta, \tag{A5}$$

where α and β are called Euler potentials (Stern, 1970). Here, each magnetic field line is labeled by α and β (since $\mathbf{B}\cdot\nabla\alpha = 0 = \mathbf{B}\cdot\nabla\beta$), and the magnetic vector potential $\mathbf{A}$ (with $\mathbf{B} \equiv \nabla\times\mathbf{A}$) can be written as

$$\mathbf{A} \;\equiv\; \frac{1}{2}\,\left(\alpha\,\nabla\beta \;-\; \beta\,\nabla\alpha\right) \;+\; \nabla\gamma, \tag{A6}$$

where the gauge function γ (which may be multi-valued) is involved in the definition of magnetic helicity $\mathbf{A}\cdot\mathbf{B} = \mathbf{B}\cdot\nabla\gamma$. Since the gauge term $\nabla\gamma$ does not play a role in what follows, however, we henceforth omit it.

Another useful representation for the magnetic field is the covariant (Clebsch) representation

$$\mathbf{B} \;\equiv\; \sum_i\,\lambda_i\,\nabla\chi^i, \tag{A7}$$

where (λ_i, χ^i) are Clebsch potentials (Seliger and Whitham, 1968); here, the index i goes from 1 to 3 (at most). Note that these potentials must still satisfy the condition $\nabla\cdot\mathbf{B} = 0$. The Clebsch representation is useful when an explicit expression for $\nabla\times\mathbf{B} \equiv \sum_i\,\nabla\lambda_i\times\nabla\chi^i$ is known. For example, if $\nabla\times\mathbf{B} \equiv 0$, the magnetic field $\mathbf{B}$ can then simply be written as $\mathbf{B} \equiv \nabla\chi$.

The magnetic field representations (A5) and (A7) allow the introduction of the magnetic *coordinates* $\Psi^i \equiv$

(α, β, s), where α and β are the Euler potentials for $\mathbf{B}$ and s is the spatial coordinate along a magnetic field line:

$$\frac{\partial \mathbf{x}}{\partial s} \;\equiv\; \widehat{\mathbf{b}} \;\equiv\; \frac{\mathbf{B}}{B}. \tag{A8}$$

Using the notation $\mathbf{y} \equiv (\alpha, \beta)$ for coordinates in the space of field-line labels (i.e., each magnetic field line is represented as a point in $\mathbf{y}$-space), the magnetic vector potential (A6) can also be written as

$$\mathbf{A} \;\equiv\; \frac{1}{2} \sum_{a,b} \eta_{ab}\, y^a\, \nabla y^b, \tag{A9}$$

where η_{ab} is anti-symmetric in its indices (with $\eta_{12} = +1 = -\eta_{21}$).

To define a magnetic geometry, we require the complete sets of contravariant basis vectors $(\partial \mathbf{x}/\partial \Psi^i)$ and covariant basis vectors $(\nabla \Psi^i)$. Since the vectors ∇y^a are given in Eq. (A9) and $\partial \mathbf{x}/\partial s$ is given by Eq. (A8), we only need expressions for ∇s and $\partial \mathbf{x}/\partial y^a$. In deriving these expressions, we use the orthogonality relations

$$\nabla \Psi^i \boldsymbol{\cdot} \frac{\partial \mathbf{x}}{\partial \Psi^j} \;=\; \delta^i_j, \tag{A10}$$

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

$$\nabla s \;\equiv\; \widehat{\mathbf{b}} \;-\; \sum_a \mathcal{R}_a\, \nabla y^a \tag{A11}$$

where

$$\mathcal{R}_a \;\equiv\; \widehat{\mathbf{b}} \boldsymbol{\cdot} \frac{\partial \mathbf{x}}{\partial y^a} \;=\; \sum_i \frac{\lambda_i}{B} \frac{\partial \chi^i}{\partial y^a}, \tag{A12}$$

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

$$\frac{\partial \mathbf{x}}{\partial y^a} \;\equiv\; \mathcal{R}_a\, \widehat{\mathbf{b}} \;+\; \sum_b \eta_{ab}\, \nabla y^b \times \frac{\mathbf{B}}{B^2}. \tag{A13}$$

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 (A10).

Next, we construct the parallel gradient operator $\partial_\| \equiv \widehat{\mathbf{b}} \boldsymbol{\cdot} \nabla = \partial/\partial s$ and the perpendicular gradient operator $\nabla_\perp \equiv -\,\widehat{\mathbf{b}} \times (\widehat{\mathbf{b}} \times \nabla)$, which only has y^a-components:

$$\frac{\partial \mathbf{x}}{\partial y^a} \boldsymbol{\cdot} \nabla_\perp \;\equiv\; \frac{\partial}{\partial y^a} \;-\; \mathcal{R}_a \frac{\partial}{\partial s} \;\equiv\; \partial_{\perp a}.$$

Hence, the gradient operator can be expressed as $\nabla \equiv \widehat{\mathbf{b}}\, \partial_\| + \nabla y^a\, \partial_{\perp a}$. Lastly, the Jacobian for the transformation $\mathbf{x} \rightarrow (\alpha, \beta, s)$ is

$$\frac{\partial \mathbf{x}}{\partial \alpha} \times \frac{\partial \mathbf{x}}{\partial \beta} \boldsymbol{\cdot} \frac{\partial \mathbf{x}}{\partial s} \;\equiv\; B^{-1} \;\equiv\; (\nabla \alpha \times \nabla \beta \boldsymbol{\cdot} \nabla s)^{-1}, \tag{A14}$$

so that $d^3x \;\equiv\; B^{-1}\; d^2y\, ds$ is the infinitesimal volume element in magnetic coordinates.

APPENDIX B: UNPERTURBED GUIDING-CENTER HAMILTONIAN DYNAMICS

1. Guiding-center Phase-space Transformation

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

$$\begin{aligned} \Gamma_{\rm gc} \;=\; & \left[\, \epsilon^{-1}\, \frac{e}{c}\, \mathbf{A} \;+\; p_\|\, \widehat{\mathsf{b}} \;-\; \epsilon \left(\frac{mc}{e}\right) \mu\, \mathbf{R}^* \,\right] \boldsymbol{\cdot} \mathsf{d}\mathbf{X} \\ & +\; \epsilon \left(\frac{mc}{e}\right) \mu\, \mathsf{d}\zeta \;-\; H_{\rm gc}\, \mathsf{d}t, \end{aligned} \tag{B1}$$

where $\epsilon \equiv \epsilon_{\rm B}$ denotes 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

$$H_{\rm gc} \;=\; \frac{p_\|^2}{2m} \;+\; \mu\, B. \tag{B2}$$

The guiding-center phase-space Lagrangian (B1) and guiding-center Hamiltonian (B2) were originally derived by Littlejohn (1983) by Lie-transform methods in the form of asymptotic expansions $Z^\alpha_{\rm gc} = Z^\alpha_0 + \epsilon\, G^\alpha_1 + \cdots$, where the components of the first-order generating vector field are

$$G_1^{\mathbf{x}} \;=\; -\,\boldsymbol{\rho}_0 \;=\; -\,\frac{mc}{e} \sqrt{\frac{2\,\mu}{mB}}\; \widehat{\rho}, \tag{B3}$$

$$\begin{aligned} G_1^{p_\|} \;=\; & (mc/e)\, \mu \left(\mathsf{a}_1 : \nabla \widehat{\mathsf{b}} \;+\; \widehat{\mathsf{b}} \boldsymbol{\cdot} \nabla \times \widehat{\mathsf{b}}\right) \\ & -\; p_\|\, \boldsymbol{\rho}_0 \boldsymbol{\cdot} \left(\widehat{\mathsf{b}} \boldsymbol{\cdot} \nabla \widehat{\mathsf{b}}\right), \end{aligned} \tag{B4}$$

$$\begin{aligned} G_1^{\mu} \;=\; & \boldsymbol{\rho}_0 \boldsymbol{\cdot} \left(\mu\, \nabla \ln B \;+\; \frac{m v_\|^2}{B}\, \widehat{\mathsf{b}} \boldsymbol{\cdot} \nabla \widehat{\mathsf{b}}\right) \\ & -\; \mu\, \frac{v_\|}{\Omega} \left(\mathsf{a}_1 : \nabla \widehat{\mathsf{b}} \;+\; \widehat{\mathsf{b}} \boldsymbol{\cdot} \nabla \times \widehat{\mathsf{b}}\right), \end{aligned} \tag{B5}$$

$$\begin{aligned} G_1^{\zeta} \;=\; & -\,\boldsymbol{\rho}_0 \boldsymbol{\cdot} \mathbf{R} \;+\; \frac{\partial \boldsymbol{\rho}_0}{\partial \zeta} \boldsymbol{\cdot} \nabla \ln B \;+\; \frac{v_\|}{\Omega}\, \mathsf{a}_2 : \nabla \widehat{\mathsf{b}} \\ & +\; \frac{m v_\|^2}{2\,\mu B} \left(\widehat{\mathsf{b}} \boldsymbol{\cdot} \nabla \widehat{\mathsf{b}} \boldsymbol{\cdot} \frac{\partial \boldsymbol{\rho}_0}{\partial \zeta}\right). \end{aligned} \tag{B6}$$

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

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

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

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

$$\nabla^* \;\equiv\; \nabla \;+\; \mathbf{R}^*\, \frac{\partial}{\partial \zeta}, \tag{B7}$$

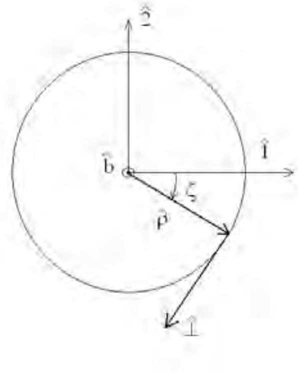

FIG. 6 Fixed unit vectors $(\widehat{1}, \widehat{2}, \widehat{\mathsf{b}})$ and rotating unit vectors $(\widehat{\perp}, \widehat{\rho}, \widehat{\mathsf{b}})$. Gyrogauge invariance involves an arbitrary rotation of the perpendicular unit vectors $\widehat{1}$ and $\widehat{2}$ about the parallel unit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

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

$$
\begin{aligned}
\mathsf{a}_1 &= -\frac{1}{2}\left(\widehat{\rho}\,\widehat{\perp} + \widehat{\perp}\,\widehat{\rho}\right) \;=\; \frac{\partial \mathsf{a}_2}{\partial\zeta}, \\
\mathsf{a}_2 &= \frac{1}{4}\left(\widehat{\perp}\,\widehat{\perp} - \widehat{\rho}\,\widehat{\rho}\right) \;=\; -\frac{1}{4}\,\frac{\partial \mathsf{a}_1}{\partial\zeta}.
\end{aligned}
$$

We also note that the guiding-center kinetic energy $\mathcal{E} = p_{\|}^2/2m + \mu\,B$ is identical to the particle kinetic energy (to first order) since

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

Note that 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. Hence, by introducing the gyrogauge transformation $\zeta' = \zeta + \chi(\mathbf{X})$, the perpendicular unit vectors $(\widehat{1}, \widehat{2})$ are transformed as $\widehat{1}' = \widehat{1}\,\cos\chi + \widehat{2}\,\sin\chi$ and $\widehat{2}' = -\,\widehat{1}\,\sin\chi + \widehat{2}\,\cos\chi$ 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 term $\mathsf{d}\zeta - \mathbf{R}\cdot\mathsf{d}\mathbf{X}$.

2. Guiding-center Hamiltonian Dynamics

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

$$
\mathbf{B}^* \;\equiv\; \mathbf{B} \;+\; \epsilon\,\left(\frac{c p_{\|}}{e}\right)\,\nabla\times\widehat{\mathsf{b}},
$$

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 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 $(\mathbf{X}, p_{\|}, \mu, \zeta)$ as

$$
\begin{aligned}
\{F,\, G\}_{\rm gc} \;=\; & \epsilon^{-1}\,\frac{e}{mc}\,\left(\frac{\partial F}{\partial\zeta}\,\frac{\partial G}{\partial\mu} \;-\; \frac{\partial F}{\partial\mu}\,\frac{\partial G}{\partial\zeta}\right) \\
& +\; \frac{\mathbf{B}^*}{B_{\|}^*}\cdot\left(\nabla^* F\,\frac{\partial G}{\partial p_{\|}} \;-\; \frac{\partial F}{\partial p_{\|}}\,\nabla^* G\right) \\
& -\; \epsilon\,\frac{c\widehat{\mathsf{b}}}{eB_{\|}^*}\cdot\nabla^* F\times\nabla^* G,
\end{aligned} \tag{B8}
$$

where ∇^* is the gradient operator (B7) and the ϵ-ordering clearly separates the fast gyro-motion 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 are, therefore, given in terms of the guiding-center Poisson bracket (B8) and the guiding-center Hamiltonian (B2) as $\dot{Z}^\alpha = \{Z^\alpha,\; H_{\rm gc}\}_{\rm gc}$:

$$
\begin{aligned}
\dot{\mathbf{X}} \;&=\; v_{\|}\,\frac{\mathbf{B}^*}{B_{\|}^*} \;+\; \epsilon\,\frac{c\widehat{\mathsf{b}}}{eB_{\|}^*}\times\mu\,\nabla B \;\equiv\; \mathbf{v}_{\rm gc} \qquad \text{(B9)} \\
&=\; v_{\|}\,\widehat{\mathsf{b}} \;+\; \epsilon\,\frac{c\widehat{\mathsf{b}}}{eB_{\|}^*}\times\left(\mu\,\nabla B \;+\; m v_{\|}^2\,\widehat{\mathsf{b}}\cdot\nabla\widehat{\mathsf{b}}\right),
\end{aligned}
$$

$$
\dot{p}_{\|} \;=\; -\,\frac{\mathbf{B}^*}{B_{\|}^*}\cdot\mu\,\nabla B, \tag{B10}
$$

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

$$
\dot{\zeta} \;=\; \epsilon^{-1}\,\Omega \;+\; v_{\|}\,\widehat{\mathsf{b}}\cdot\mathbf{R}^* \;+\; \mathcal{O}(\epsilon). \tag{B11}
$$

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

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

from which we derive the guiding-center Liouville theorem

$$
\nabla\cdot\left(B_{\|}^*\,\frac{\mathbf{X}}{dt}\right) \;+\; \frac{\partial}{\partial p_{\|}}\left(B_{\|}^*\,\frac{dp_{\|}}{dt}\right) \;=\; 0.
$$

3. Guiding-center Pull-back Transformation

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

$$
f \;=\; F \;-\; \boldsymbol{\rho}_0\cdot\nabla F \;+\; \epsilon\,G_1^{p_{\|}}\,\frac{\partial F}{\partial p_{\|}} \;+\; \epsilon\,G_1^{\mu}\,\frac{\partial F}{\partial\mu}, \tag{B12}
$$

where the second term on the right side is also ordered at ϵ. Here, $f = f(\mathbf{x}, p_{0\|}, \mu_0, \zeta_0)$ is a function of the particle position $\mathbf{x}$, its kinetic momentum parallel to the magnetic field $p_{0\|} = \mathbf{p}\cdot\widehat{\mathsf{b}}$, the lowest-order magnetic moment $\mu_0 = |\mathbf{p}_\perp|^2/2mB$ and gyration angle $\zeta_0 = \arctan[(-\mathbf{p}\cdot\widehat{1})/(-\mathbf{p}\cdot\widehat{2})]$. Hence, the Vlasov equation in particle phase space $(\mathbf{x}, p_{0\|}, \mu_0, \zeta_0)$ is expressed as

$$0 \;=\; \frac{df}{dt} \;\equiv\; \frac{\partial f}{\partial t} \;+\; \mathbf{v}\cdot\nabla f \;+\; \dot{z}_0^i\,\frac{\partial f}{\partial z_0^i}, \tag{B13}$$

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

$$\dot{p}_{0\|} \;=\; -\,\epsilon\left(\mu_0\,\widehat{\mathsf{b}}\cdot\nabla B \;+\; \Omega\,\frac{\partial G_1^{p_\|}}{\partial\zeta}\right), \tag{B14}$$

$$\dot{\mu}_0 \;=\; -\,\epsilon\,\Omega\,\frac{\partial G_1^{\mu}}{\partial\zeta}, \tag{B15}$$

$$\begin{aligned}\dot{\zeta}_0 \;&=\; (\Omega - \epsilon\,\boldsymbol{\rho}_0\cdot\nabla\Omega) \\ &\quad+\;\epsilon\left(v_\|\,\widehat{\mathsf{b}}\cdot\mathbf{R}^* \;-\; \Omega\,\frac{\partial G_1^{\zeta}}{\partial\zeta}\right).\end{aligned} \tag{B16}$$

We now show that Eq. (B12) is a solution of the Vlasov equation (B13) provided the guiding-center Vlasov distribution F satisfies the guiding-center Vlasov equation

$$0 \;=\; \frac{d_{\rm gc}F}{dt} \;\equiv\; \frac{\partial F}{\partial t} \;+\; \mathbf{v}_{\rm gc}\cdot\nabla F \;+\; \dot{p}_\|\,\frac{\partial F}{\partial p_\|}, \tag{B17}$$

where $\partial F/\partial\zeta \equiv 0$ and $\dot{\mu} \equiv 0$; here, we use $\mathbf{v}_{\rm gc} = v_\|\,\widehat{\mathsf{b}}$ and $\dot{p}_\| = -\,\epsilon\,\mu\,\widehat{\mathsf{b}}\cdot\nabla B$. First, we write

$$\begin{aligned}\frac{df}{dt} \;&=\; \frac{\partial F}{\partial t} \;+\; \epsilon\left(\mathbf{v} \;-\; \Omega\,\frac{\partial\boldsymbol{\rho}_0}{\partial\zeta}\right)\cdot\nabla F \\ &\quad+\;\left(\dot{p}_{0\|} \;+\; \epsilon\,\Omega\,\frac{\partial G_1^{p_\|}}{\partial\zeta}\right)\frac{\partial F}{\partial p_\|} \\ &\quad+\;\left(\dot{\mu}_0 \;+\; \epsilon\,\Omega\,\frac{\partial G_1^{\mu}}{\partial\zeta}\right)\frac{\partial F}{\partial\mu},\end{aligned} \tag{B18}$$

where we have used the fact that F is independent of the gyroangle ζ. By inserting definitions (B14)-(B15) for $(\dot{p}_{0\|}, \dot{\mu}_0)$, we readily find $df/dt = d_{\rm gc}F/dt$, so that the particle Vlasov equation (B13) is satisfied if the guiding-center Vlasov equation (B17) is satisfied.

We conclude, therefore, that the pull-back operator $\mathsf{T}_{\rm gc}$ 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 (B12) 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 $J \equiv J_{\rm b}$) can be constructed. The resulting *bounce-averaged guiding-center* dynamics takes place in a reduced two-dimensional phase space with spatial coordinates (y^1, y^2), where each coordinate y^a (with $a = 1$ or 2) satisfies the condition $\mathbf{B}\cdot\nabla y^a = 0$; the coordinates (y^1, y^2) are known as magnetic field line *labels* and were described in Sec. A.2. Bounce-averaged guiding-center dynamics in static magnetic fields has also been shown to possess a canonical Hamiltonian structure (Littlejohn, 1982b).

First, we begin with the unperturbed guiding-center phase-space Lagrangian (B1) written in magnetic coordinates as

$$\begin{aligned}\Gamma_0 \;&=\; \left(\epsilon_{\rm d}^{-1}\,\frac{e}{2c}\,\eta_{ab}\,y^a + p_\|\,\mathcal{R}_b\right)dy^b \;+\; p_\|\,ds \\ &\quad-\;\left(\mu\,B \;+\; \frac{p_\|^2}{2m}\right)dt \\ &\equiv\; \mathcal{F}_b\,dy^b \;+\; p_\|\,ds \;-\; H_0\,dt,\end{aligned} \tag{B19}$$

where the gyro-motion dynamics has been removed and $\epsilon_{\rm d} \ll 1$ 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 $\epsilon_{\rm d}$-ordering, the fast guiding-center motion is described by the quasi-periodic bounce motion:

$$\dot{s} \;=\; v_\| \quad\text{and}\quad \dot{v}_\| \;=\; -\,(\mu/m)\,\partial_\| B, \tag{B20}$$

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. Here, the action-angle coordinates (J, Θ) associated with periodic bounce motion have the following lowest-order expressions: for the bounce action $J \equiv J_{\rm b} = J_0 + \cdots$, we find (Littlejohn, 1982b; Northrop, 1963)

$$\begin{aligned}J_0(\mathcal{E}, \mu; \mathbf{y}) \;&\equiv\; \frac{1}{2\pi}\oint p_\|(s, \mathcal{E}, \mu; \mathbf{y})\,ds \\ &=\; \frac{1}{\pi}\int_{s_0}^{s_1}\sqrt{2m\,[\mathcal{E} - \mu\,B(s; \mathbf{y})]}\,ds,\end{aligned} \tag{B21}$$

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

$$\Theta_0(s, \mathcal{E}, \mu; \mathbf{y}) \;\equiv\; \pi \;\pm\; \omega_{\rm b}\int_{s_0}^{s}\frac{ds'}{\sqrt{\frac{2}{m}\,[\mathcal{E} - \mu\,B(s'; \mathbf{y})]}}, \tag{B22}$$

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

Eq. (B21) as

$$\omega_{\rm b}(\mathbf{y};\mathcal{E},\mu) \;\equiv\; \left(\frac{\partial J}{\partial\mathcal{E}}\right)^{-1} \;=\; 2\pi\left(\oint\frac{ds}{v_\|}\right)^{-1}. \quad \text{(B23)}$$

We can now proceed to perform the substitution $(s,p_\|) \to (J,\Theta)$ in the guiding-center phase-space Lagrangian (B19). First, we note that the transformation $(s,\,p_\|) \to (J,\Theta) \equiv \mathbf{u}$ is canonical since $\mathsf{d}p_\|\wedge\mathsf{d}s \equiv \mathsf{d}J\wedge\mathsf{d}\Theta$. In the guiding-center phase-space Lagrangian (B19), the differential $\mathsf{d}s$ becomes $\mathsf{d}s = \partial_\alpha s\;\mathsf{d}u^\alpha$ and, thus, we have

$$\begin{aligned}\Gamma_0 \;\equiv\; & \left(\frac{q}{2c\epsilon_{\rm d}}\;\eta_{ab}\,y^a + p_\|\,\mathcal{R}_b\right)\mathsf{d}y^b \\ & +\;\left(p_\|\,\frac{\partial s}{\partial u^\alpha}\right)\mathsf{d}u^\alpha \;-\; H_0(\mathbf{y},\mathbf{u})\,\mathsf{d}t,\end{aligned} \quad \text{(B24)}$$

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 (B24). Because of its dependence on the field-line labels $\mathbf{y}$, the bounce action (B21) is not conserved at order $\epsilon_{\rm d}$ [i.e., $dJ/dt = \mathcal{O}(\epsilon_{\rm d})$]. To remove the bounce-angle dependence in the guiding-center phase-space Lagrangian (B24) and construct an asymptotic expansion for the bounce-action adiabatic invariant, we proceed by performing an infinitesimal transformation $(\mathbf{y},\mathbf{u}) \to (\overline{\mathbf{y}},\overline{\mathbf{u}})$, where the relation between the guiding-center coordinates $(\mathbf{y},\mathbf{u})$ and the bounce-guiding-center coordinates $(\overline{\mathbf{y}},\overline{\mathbf{u}})$ is given in terms of the asymptotic expansions

$$\left.\begin{aligned}\overline{y}^a \;&\equiv\; y^a \;+\; \epsilon_{\rm d}\,G_1^a \;+\;\cdots \\ \overline{u}^\alpha \;&\equiv\; u^\alpha \;+\; \epsilon_{\rm d}\,G_1^\alpha \;+\;\cdots\end{aligned}\right\}, \quad \text{(B25)}$$

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

$$G_1^a \;=\; -\,\eta^{ab}\;\frac{c}{e}\left(\frac{\partial S_1}{\partial\overline{y}^b} \;+\; p_\|\,\mathcal{R}_b\right), \quad \text{(B26)}$$

where the gauge function $S_1(\overline{\mathbf{y}},\overline{\mathbf{u}})$ is defined from the relation

$$\frac{\partial S_1}{\partial\overline{u}^\beta} \;\equiv\; -\,\frac{\eta_{\alpha\beta}}{2}\,\overline{u}^\alpha \;-\; p_\|(\overline{\mathbf{y}},\overline{\mathbf{u}})\,\frac{\partial s(\overline{\mathbf{u}})}{\partial\overline{u}^\beta}, \quad \text{(B27)}$$

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

The purpose of this transformation is, thus, to remove the bounce-angle dependence at all orders in $\epsilon_{\rm d}$. Hence, the unperturbed bounce-averaged guiding-center (or *bounce-center*) phase-space Lagrangian becomes

$$\overline{\Gamma}_0 \;\equiv\; \epsilon_{\rm d}^{-1}\,\frac{e}{2c}\,\eta_{ab}\,\overline{y}^a\,\mathsf{d}\overline{y}^b \;+\; \overline{J}\,\mathsf{d}\overline{\Theta} \;-\; \overline{H}_0(\overline{\mathbf{y}},\overline{J};\epsilon_{\rm d})\,\mathsf{d}t, \quad \text{(B28)}$$

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

$$\overline{H}_0 \;\equiv\; H_0 \;-\; \frac{\epsilon_{\rm d}}{2}\left(\omega_{\rm b}\,\eta_{ab}\left\langle G_1^a\,\frac{\partial G_1^b}{\partial\overline{\Theta}}\right\rangle_{\rm b}\right), \quad \text{(B29)}$$

where $\langle\;\rangle_{\rm b}$ denotes averaging with respect to $\overline{\Theta}$. 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 $(\overline{\mathbf{y}},\overline{\mathbf{u}})$ as

$$\{\mathcal{F},\,\mathcal{G}\} \;=\; \frac{\partial\mathcal{F}}{\partial\overline{u}^\alpha}\,\eta^{\alpha\beta}\,\frac{\partial\mathcal{G}}{\partial\overline{u}^\beta} \;+\; \epsilon_{\rm d}\,\frac{c}{e}\,\frac{\partial\mathcal{F}}{\partial\overline{y}^a}\,\eta^{ab}\,\frac{\partial\mathcal{G}}{\partial\overline{y}^b}, \quad \text{(B30)}$$

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.

Lastly, we note that the bounce-guiding-center position $\overline{y}^a$ is the (bounce-motion) time-averaged position of the guiding-center position y^a, i.e., $\overline{y}^a \equiv \langle y^a\rangle_{\rm b}$, and thus

$$\Lambda_{\rm b}^a \;\equiv\; y^a \;-\; \langle y^a\rangle_{\rm b} \;=\; -\,\epsilon_{\rm d}\,G_1^a(\overline{\mathbf{y}},\overline{\mathbf{u}}) \quad \text{(B31)}$$

represents the bounce-angle dependent *bounce* radius.

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. In 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 (124) 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 $(\mathsf{G}_1,\mathsf{G}_2,\cdots)$ by Eq. (125), so that we obtain

$$\begin{aligned}\|v^\mu\| \;=\; & \int d^4\overline{p}\;\left(\mathsf{T}_\epsilon^{-1}v^\mu\right)\;\overline{\mathcal{F}} \\ & -\;\nabla\cdot\left[\int d^4\overline{p}\;\boldsymbol{\rho}_\epsilon\;\left(\mathsf{T}_\epsilon^{-1}v^\mu\right)\;\overline{\mathcal{F}} \;+\;\cdots\right],\end{aligned} \quad \text{(C1)}$$

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). Next, we derive the push-forward representations for the four-current $J^\mu = (c\rho,\mathbf{J}) \equiv \sum e\,\|v^\mu\|$. First, we derive the push-forward expression for the charge density (130), where $\overline{\rho} \equiv \sum e\int d^4\overline{p}\,\overline{\mathcal{F}}$ denotes the reduced charge density and the polarization vector is defined as

$$\mathbf{P}_\epsilon \;\equiv\; \sum\;e\;\int d^4\overline{p}\;\boldsymbol{\rho}_\epsilon\,\overline{\mathcal{F}}, \quad \text{(C2)}$$

where $e\boldsymbol{\rho}_\epsilon$ denotes the electric-dipole moment associated with the charge separation induced by the phase-space transformation.

Next, we derive the push-forward expression for the current density (131), where the push-forward of the particle velocity $\mathbf{v} = d\mathbf{x}/dt$ (using the Lagrangian representation)

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

is expressed in terms of the reduced total time derivative d_ϵ/dt. Here, $\overline{\mathbf{v}} \equiv d_\epsilon\overline{\mathbf{x}}/dt$ denotes the reduced velocity and we may replace the term $d_\epsilon\boldsymbol{\rho}_\epsilon/dt$ in Eq. (C3) by using the following identity based on the expression (C2) for the reduced polarization vector:

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

where the reduced Vlasov equation (93) was used and integration by parts was performed. The push-forward representation of the current density is, therefore, expressed as

$$\begin{aligned}\mathbf{J} &= \overline{\mathbf{J}} \;+\; \frac{\partial\mathbf{P}_\epsilon}{\partial t} \;+\; \nabla\times\left[\sum e\int d^4\overline{p}\left(\boldsymbol{\rho}_\epsilon\times\frac{d_\epsilon\overline{\mathbf{x}}}{dt}\right)\overline{\mathcal{F}}\right]\\ &\quad -\;\nabla\cdot\left[\sum e\int d^4\overline{p}\left(\boldsymbol{\rho}_\epsilon\,\frac{d_\epsilon\boldsymbol{\rho}_\epsilon}{dt}\right)\overline{\mathcal{F}}\right],\end{aligned} \tag{C5}$$

where $\overline{\mathbf{J}} \equiv \sum e\int d^4\overline{p}\,(d_\epsilon\overline{\mathbf{x}}/dt)\,\overline{\mathcal{F}}$ denotes the reduced current density, $\mathbf{J}_{\rm pol} \equiv \partial\mathbf{P}_\epsilon/\partial t$ denotes the polarization current, the third term represents the moving-electric-dipole contribution to the magnetization vector (Jackson, 1975), and the last term represents the intrinsic magnetization vector (see below).

For example, the oscillation-center transformation (discussed in Sec. IV.F) introduces nonlinear (ponderomotive) polarization and magnetization effects as follows. First, the oscillation-center polarization vector is expressed in terms of the eikonal-averaged (denoted by an overbar) displacement vector

$$\overline{\boldsymbol{\rho}}_\epsilon \;=\; \epsilon_\delta^2\,\mathbf{k}\times\left(i\widetilde{\boldsymbol{\xi}}\times\widetilde{\boldsymbol{\xi}}^*\right) \;\equiv\; \epsilon_\delta^2\,\overline{\boldsymbol{\rho}}_2,$$

so that the reduced polarization vector (C2) becomes the oscillation-center polarization vector

$$\mathbf{P}_{\rm osc} \;=\; \epsilon_\delta^2\sum\int d^4\overline{p}\,\overline{\boldsymbol{\pi}}_2\,\overline{\mathcal{F}}, \tag{C6}$$

where $\overline{\boldsymbol{\pi}}_2 \equiv e\,\overline{\boldsymbol{\rho}}_2$ denotes the second-order ponderomotive electric-dipole moment. Next, the oscillation-center magnetization is expressed in terms of the eikonal-averaged expressions

$$\overline{(\boldsymbol{\rho}_\epsilon\times\dot{\mathbf{x}}_\epsilon)} \;=\; \overline{\boldsymbol{\rho}}_\epsilon\times\dot{\overline{\mathbf{x}}}_\epsilon \;=\; \epsilon_\delta^2\,\overline{\boldsymbol{\rho}}_2\times\mathbf{v} \;+\;\cdots$$

and

$$\overline{\left(\boldsymbol{\rho}_\epsilon\,\frac{d_\epsilon\boldsymbol{\rho}_\epsilon}{dt}\right)} \;=\; -\,\epsilon_\delta^2\left[\,i\,\omega'\left(\widetilde{\boldsymbol{\xi}}^*\,\widetilde{\boldsymbol{\xi}} \;-\; \widetilde{\boldsymbol{\xi}}\,\widetilde{\boldsymbol{\xi}}^*\right)\right] \;+\;\cdots$$

so that Eq. (C5) becomes

$$\mathbf{J} \;=\; \overline{\mathbf{J}} \;+\; \frac{\partial\mathbf{P}_{\rm osc}}{\partial t} \;+\; c\,\nabla\times\mathbf{M}_{\rm osc},$$

where the oscillation-center magnetization vector

$$\mathbf{M}_{\rm osc} \;=\; \epsilon_\delta^2\sum\int d^4\overline{p}\left(\overline{\boldsymbol{\mu}}_2 \;+\; \overline{\boldsymbol{\pi}}_2\times\frac{\mathbf{v}}{c}\right)\overline{\mathcal{F}}, \tag{C7}$$

where $\overline{\boldsymbol{\mu}}_2 \equiv \omega'\,(i\,\widetilde{\boldsymbol{\xi}}\times\widetilde{\boldsymbol{\xi}}^*)$ denotes the second-order intrinsic magnetic-dipole moment.

2. Push-forward Representation of Gyrocenter Fluid Moments

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

$$\boldsymbol{\rho}_{\rm gy} \;\equiv\; \boldsymbol{\rho}_{\rm gc} \;-\; \epsilon_\delta\,\mathbf{G}_1^* \;+\;\cdots, \tag{C8}$$

where $\boldsymbol{\rho}_{\rm gc}$ denotes the gyroangle-dependent gyroradius and the effective first-order vector field $\mathbf{G}_1^*$ is defined as [see Eqs. (163), (165), and (166)]

$$\begin{aligned}\mathbf{G}_1^* &= G_1^{\mathbf{x}} \;+\; G_1^\mu\,\frac{\partial\overline{\boldsymbol{\rho}}_0}{\partial\mu} \;+\; G_1^\zeta\,\frac{\partial\overline{\boldsymbol{\rho}}_0}{\partial\zeta}\\ &= \{S_1,\;\overline{\mathbf{X}}+\overline{\boldsymbol{\rho}}_0\}_0 \;+\; \alpha\,\frac{\widehat{\mathsf{b}}_0}{B_0}\times\langle\delta\mathbf{A}_{\perp{\rm gc}}\rangle.\end{aligned} \tag{C9}$$

We begin with the push-forward expression for the fluid density

$$n \;=\; \overline{N} \;-\; \nabla\cdot\left(\int d^4\overline{p}\,\langle\boldsymbol{\rho}_{\rm gy}\rangle\,\overline{\mathcal{F}}\right), \tag{C10}$$

where n denotes the *particle* fluid density while $\overline{N}$ denotes the *gyrocenter* fluid density. Here, the gyroangle-averaged gyrocenter displacement (C8) is expressed as

$$\langle\boldsymbol{\rho}_{\rm gy}\rangle \;=\; -\,\frac{e}{B_0}\,\frac{\partial}{\partial\mu}\left\langle\delta\widetilde{\psi}_{\rm gc}\,\boldsymbol{\rho}_0\right\rangle \;-\; \alpha\,\frac{\widehat{\mathsf{b}}_0}{B_0}\times\langle\delta\mathbf{A}_{\perp{\rm gc}}\rangle. \tag{C11}$$

The gyrocenter electric-dipole moment is now defined in the fluid limit as

$$\begin{aligned}\boldsymbol{\pi}_{\rm gy} \;\equiv\; e\,\langle\boldsymbol{\rho}_{\rm gy}\rangle_{\rm fluid} &= -\,\frac{mc^2}{B_0^2}\left(\nabla_\perp\delta\phi \;-\; \frac{u_\parallel}{c}\,\nabla_\perp\delta A_\parallel\right)\\ &\quad +\;(1-\alpha)\,\frac{\widehat{\mathsf{b}}_0}{B_0}\times\delta\mathbf{A}_\perp,\end{aligned} \tag{C12}$$

where $u_\parallel$ denotes the parallel drift-fluid velocity. Hence, we see that the perturbed density $\delta\rho \equiv -\nabla\cdot(n_0\,\boldsymbol{\pi}_{\rm gy})$ is expressed as

$$\delta\rho \;=\; \nabla\cdot\left[\,\frac{mc^2 n_0}{B_0^2}\;\left(\nabla_\perp\delta\phi \;-\; \frac{u_\parallel}{c}\,\widehat{\mathsf{b}}_0\times\delta\mathbf{B}_\perp\right)\,\right] \;-\; (1-\alpha)\,\nabla\cdot\left(en_0\,\frac{\widehat{\mathsf{b}}_0}{B_0}\times\delta\mathbf{A}_\perp\right), \tag{C13}$$

where the first term denotes the standard polarization density due to low-frequency electrostatic fluctuations, the second term denotes the covariant correction of the polarization density due to magnetic fluctuations perpendicular to the background magnetic field (Hahm *et al.*, 1988), and the third term is a term that appears only in the Hamiltonian and $\|$-symplectic versions (where $\alpha = 0$) of gyrokinetic theory. Here, for a uniform background plasma, the last term becomes $-\,en_0\;\delta B_\parallel/B_0$, which vanishes when summation of particle species is done and quasi-neutrality is assumed. Extension of the gyrokinetic polarization density to gyro-bounce-kinetics leads to the neoclassical polarization density (Brizard, 2000c; Fong and Hahm, 1999) as discussed in Appendix E.

Lastly, we note that, in the absence of electromagnetic fluctuations (i.e., considering the guiding-center transformation alone), the difference between the perpendicular particle flux and the perpendicular guiding-center flux, expressed in terms of Eq. (C1) with $\mathbf{v}_\perp$, is given as

$$\nabla\cdot\left(e\int d^3p\;\langle\boldsymbol{\rho}\,\mathbf{v}_\perp\rangle\;F\right) \;=\; \nabla\times\left(\int d^3p\;\mu B_0\;\frac{cF\,\widehat{\mathsf{b}}_0}{B_0}\right) \;\equiv\; \nabla\times\left(p_\perp\,\frac{c\widehat{\mathsf{b}}_0}{B_0}\right),$$

where $-\,p_\perp\widehat{\mathsf{b}}_0/B_0$ denotes the parallel magnetization vector (and we have ignored background magnetic field nonuniformity). Hence, we see that the push-forward and variational methods yield identical expressions for the polarization and magnetization effects appearing in reduced Maxwell's equations.

APPENDIX D: DIRECT PROOF OF GYROKINETIC ENERGY CONSERVATION

In this Appendix, we present a direct proof of the gyrokinetic energy conservation law (205). First, we express the time derivative of Eq. (206) in four separate terms:

$$\frac{\partial\mathcal{E}}{\partial t} \;\equiv\; \frac{\partial\mathcal{E}_I}{\partial t} \;+\; \frac{\partial\mathcal{E}_{II}}{\partial t} \;+\; \frac{\partial\mathcal{E}_{III}}{\partial t} \;+\; \frac{\partial\mathcal{E}_{IV}}{\partial t} \tag{D1}$$

where (with $\epsilon \equiv \epsilon_\delta$ and $\mathsf{T}_\epsilon \equiv \mathsf{T}_{\rm gy}$)

$$\begin{aligned}
\frac{\partial\mathcal{E}_I}{\partial t} &= \int \frac{\partial F}{\partial t}\;H \\
\frac{\partial\mathcal{E}_{II}}{\partial t} &= \int F\left(\frac{\partial H}{\partial t} \;-\; \epsilon\left\langle\frac{\partial}{\partial t}\left(\mathsf{T}_\epsilon^{-1}e\,\phi_{1\rm gc}\right)\right\rangle\right) \\
\frac{\partial\mathcal{E}_{III}}{\partial t} &= \frac{\epsilon^2}{4\pi}\,\nabla\phi_1\cdot\nabla\frac{\partial\phi_1}{\partial t} \;-\; \epsilon\int\frac{\partial F}{\partial t}\;\langle\mathsf{T}_\epsilon^{-1}e\,\phi_{1\rm gc}\rangle \\
\frac{\partial\mathcal{E}_{IV}}{\partial t} &= \epsilon\,\frac{\mathbf{B}}{4\pi}\cdot\nabla\times\frac{\partial\mathbf{A}_1}{\partial t}
\end{aligned}$$

where $\mathbf{B} = \mathbf{B}_0 + \epsilon\,\nabla\times\mathbf{A}_1$ and $\int(...)$ denotes $\int d^3p(...)$. The partial time derivatives of the gyrocenter pull-back T_ϵ and the gyrocenter push-forward T_ϵ^{-1} operators are defined in terms of an arbitrary function G as

$$\frac{\partial}{\partial t}\left(\mathsf{T}_\epsilon^{\pm1}G\right) \;=\; \mathsf{T}_\epsilon^{\pm1}\left(\frac{\partial G}{\partial t}\right) \;\pm\;\epsilon\,\left\{\frac{\partial S_1}{\partial t},\;G\right\} \;\pm\;\frac{\epsilon\,e}{c}\,\frac{\partial\mathbf{A}_{1\rm gc}}{\partial t}\cdot\{\mathbf{X}+\boldsymbol{\rho},\;G\} \;+\;\cdots, \tag{D2}$$

where $\{\;,\;\} \equiv \{\;,\;\}_{\rm gc}$.

The first term can be written as

$$\frac{\partial\mathcal{E}_I}{\partial t} \;=\; -\,\nabla\cdot\left(\int\;\dot{\mathbf{X}}\;F\,H\right), \tag{D3}$$

where $\dot{\mathbf{X}} \equiv \{\mathbf{X},\;H\}$ denotes the gyrocenter velocity. For the second term, we express the partial time derivative of the gyrocenter Hamiltonian (174) as

$$\begin{aligned}
\frac{\partial H}{\partial t} &= \epsilon\,e\,\left\langle\frac{\partial\psi_{1\rm gc}}{\partial t}\right\rangle \;+\; \frac{\epsilon^2e^2}{mc^2}\,\left\langle\mathbf{A}_{1\rm gc}\cdot\frac{\partial\mathbf{A}_{1\rm gc}}{\partial t}\right\rangle \;-\;\epsilon^2 e\,\left\langle\left\{S_1,\;\frac{\partial\psi_{1\rm gc}}{\partial t}\right\}\right\rangle \\
&\equiv \epsilon\,\left\langle\mathsf{T}_\epsilon^{-1}\,\left(e\,\frac{\partial\psi_{1\rm gc}}{\partial t}\right)\right\rangle
\end{aligned}$$

and we use the operator formula (D2) for $G = e\,\phi_{1\rm gc}$:

$$e\,\left\langle\frac{\partial}{\partial t}\left(\mathsf{T}_\epsilon^{-1}\phi_{1\rm gc}\right)\right\rangle \;=\; e\,\left\langle\mathsf{T}_\epsilon^{-1}\left(\frac{\partial\phi_{1\rm gc}}{\partial t}\right)\right\rangle \;-\;\left\langle\left\{\frac{\partial S_1}{\partial t},\;e\,\phi_{1\rm gc}\right\}\right\rangle$$

to obtain

$$\frac{\partial\mathcal{E}_{II}}{\partial t} \;=\; \epsilon\,\int\;F\left[\,\epsilon\,\left\langle\left\{\frac{\partial S_1}{\partial t},\;e\,\phi_{1\rm gc}\right\}\right\rangle \;-\;\left\langle\mathsf{T}_\epsilon^{-1}\left(e\frac{\mathbf{v}}{c}\cdot\frac{\partial\mathbf{A}_{1\rm gc}}{\partial t}\right)\right\rangle\,\right], \tag{D4}$$

where we used the identity $\{\mathbf{X}+\boldsymbol{\rho},\;\phi_{\rm gc}\} \equiv 0$.

Next, we consider the third (electrostatic-energy) term

$$\frac{\partial\mathcal{E}_{III}}{\partial t} \;=\; \nabla\cdot\left(\frac{\epsilon^2\,\phi_1}{4\pi}\,\nabla\frac{\partial\phi_1}{\partial t}\right) + \epsilon\,\int\;\left[\,e\,\phi_{1\rm gc}\,\frac{\partial(\mathsf{T}_\epsilon F)}{\partial t} \;-\;\frac{\partial F}{\partial t}\;\langle\mathsf{T}_\epsilon^{-1}e\,\phi_{1\rm gc}\rangle\,\right]$$

where we have inserted the partial time derivative of Poisson's equation (200). By using the identity (D2) for the pull-back operator (with $G \equiv F$)

$$\int\;e\,\phi_{1\rm gc}\,\left[\frac{\partial}{\partial t}\left(\mathsf{T}_\epsilon F\right)\right] \;=\; \int\;\frac{\partial F}{\partial t}\;\langle\mathsf{T}_\epsilon^{-1}e\,\phi_{1\rm gc}\rangle \;-\;\epsilon\,\int\;F\;\left\langle\left\{\frac{\partial S_1}{\partial t},\;e\,\phi_{1\rm gc}\right\}\right\rangle,$$

the electrostatic energy term becomes

$$\frac{\partial \mathcal{E}_{III}}{\partial t} \;=\; \nabla\cdot\left(\epsilon^2\,\frac{\phi_1}{4\pi}\,\nabla\frac{\partial\phi_1}{\partial t}\right) \;-\; \epsilon^2\int F\,\left\langle\left\{\frac{\partial S_1}{\partial t},\;e\,\phi_{1\mathrm{gc}}\right\}\right\rangle \quad \text{(D5)}$$

Lastly, we consider the fourth (magnetic-energy) term

$$\frac{\partial \mathcal{E}_{IV}}{\partial t} \;=\; \epsilon\,\nabla\cdot\left(\frac{\partial \mathbf{A}_1}{\partial t}\times\frac{\mathbf{B}}{4\pi}\right) \;+\; \epsilon\int F\,\left\langle \mathsf{T}_\epsilon^{-1}\left(\frac{e\mathbf{v}}{c}\cdot\frac{\partial \mathbf{A}_{1\mathrm{gc}}}{\partial t}\right)\right\rangle$$

where

$$\begin{aligned}\nabla\cdot\left(\frac{\partial \mathbf{A}_1}{\partial t}\times\frac{\mathbf{B}}{4\pi}\right) \;&=\; -\,\epsilon\,\nabla\cdot\left[(\mathbf{E}_1+\nabla\phi_1)\times\frac{c\mathbf{B}}{4\pi}\right] \\ &=\; -\,\epsilon\,\nabla\cdot\left(\frac{c\mathbf{E}_1}{4\pi}\times\mathbf{B}\;-\;\frac{c\,\phi_1}{4\pi}\,\nabla\times\mathbf{B}\right) \\ &=\; -\,\epsilon\,\nabla\cdot\left(\frac{c}{4\pi}\,\mathbf{E}_1\times\mathbf{B}\right) \\ &\quad +\;\epsilon\,\nabla\cdot\left[\int F\,\left\langle\mathsf{T}_\epsilon^{-1}\left(e\mathbf{v}\,\phi_{1\mathrm{gc}}\right)\right\rangle\right]\end{aligned}$$

Hence, the magnetic energy term becomes

$$\begin{aligned}\frac{\partial \mathcal{E}_{IV}}{\partial t} \;&=\; -\,\epsilon\,\nabla\cdot\left[\,\frac{c}{4\pi}\,\mathbf{E}_1\times\mathbf{B} \;-\;\int F\,\left\langle\mathsf{T}_\epsilon^{-1}\left(e\mathbf{v}\,\phi_{1\mathrm{gc}}\right)\right\rangle\right] \\ &\quad +\;\epsilon\int F\,\left\langle \mathsf{T}_\epsilon^{-1}\left(\frac{e\mathbf{v}}{c}\cdot\frac{\partial \mathbf{A}_{1\mathrm{gc}}}{\partial t}\right)\right\rangle \qquad \text{(D6)}\end{aligned}$$

By adding the four terms (D3)-(D6), we obtain

$$\frac{\partial \mathcal{E}}{\partial t} \;=\; -\,\nabla\cdot\left[\,\epsilon\,\frac{c}{4\pi}\,\mathbf{E}_1\times\mathbf{B}\;-\;\epsilon^2\,\frac{\phi_1}{4\pi}\,\nabla\frac{\partial\phi_1}{\partial t} \;+\;\int F\,\left(\dot{\mathbf{X}}\,H\;-\;\epsilon\,\left\langle\mathsf{T}_\epsilon^{-1}\left(e\mathbf{v}\,\phi_{1\mathrm{gc}}\right)\right\rangle\right)\right]$$

and we recover the exact gyrokinetic energy conservation law (205).

APPENDIX E: EXTENSIONS OF NONLINEAR GYROKINETIC EQUATIONS

In this Appendix, we present two important extensions of nonlinear gyrokinetic theory. First, we discuss the extension of the nonlinear gyrokinetic equations presented in the text by introducing the effects of an inhomogeneous equilibrium electric field. Here, two new ordering parameters must 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, we stated that the various expansion parameters appearing in nonlinear gyrokinetic theory originate from different physical reasons, and that the standard nonlinear gyrokinetic ordering is not a unique ordering. In this Appendix, we present an example where a further ordering consideration is necessary. This example not only demonstrates the flexibility and the power of 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 last decade has demontrated that a new parameter regime characterized by a strong shear in $E\times B$ flow, a steep pressure gradient, and a low fluctuation level can be reproduced routinely. This motivates a further improvement of the standard nonlinear gyrokinetic ordering.

The analytical nonlinear theories of the $E\times B$ shear decorrelation of turbulence (Biglari *et al.*, 1990; Shaing *et al.*, 1990) and of transition dynamics (Carreras *et al.*, 1994; Diamond *et al.*, 1994) in cylindrical geometry have demonstrated a possible important role of the $E\times B$ shear in L (low) mode to H (high) mode transition (Burrell, 1997; Wagner *et al.*, 1982). 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 (Burrell, 1997; Mazzucato *et al.*, 1996; Synakowski *et al.*, 1997) and has been utilized in the analytical threshold calculation for the transport bifurcation (Diamond *et al.*, 1997).

While there have been significant progress in both shear-flow physics (see, for example, 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 further improved 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 (Artun and Tang, 1994; Bernstein and Catto, 1985; Brizard, 1995; Hahm, 1992), 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 flow or the diamagnetic flow plays a dominant role. 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 (Artun and Tang, 1994; Brizard, 1995; Hahm, 1992) because one can formally treat the guiding-center motion part separately from the equilibrium mass-flow issue, which is related to the determination of ion distribution function from neoclassical theory.

A general formulation can be pursued with $u_E/v_{\rm th} \sim 1$, in addition to the standard gyrokinetic ordering $\omega/\Omega \sim e\delta\phi/T_{\rm i} \sim \rho_{\rm i}k_\parallel \sim \epsilon$ and $k_\perp\rho_{\rm i} \sim 1$. Here, u_E is the equilibrium $E \times B$ velocity. [12] Only the electrostatic fluctuations are considered in this Appendix.

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

$$\Gamma_0 \equiv \left(\frac{e}{c}\mathbf{A} + m\mathbf{u}_E + mv_\parallel\widehat{\mathsf{b}}\right)\cdot \mathsf{d}\mathbf{X} + \frac{\mu B}{\Omega}\mathsf{d}\zeta - H_0\mathsf{d}t, \quad \text{(E1)}$$

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

$$H_0 = e\Phi \;+\; \mu\,(B + B_{\rm E}) \;+\; \frac{m}{2}\left(v_\parallel^2 + |\mathbf{u}_{\rm E}|^2\right), \quad \text{(E2)}$$

where

$$\begin{aligned} B_{\rm E} \;&\equiv\; \frac{\mathbf{B}}{2\Omega}\cdot\nabla\times\mathbf{u}_{\rm E} \qquad \text{(E3)} \\ &=\; \frac{c}{2}\left[\,\nabla\cdot\left(\frac{\nabla\Phi}{\Omega}\right) \;+\; \frac{\nabla\Phi}{\Omega}\cdot\left(\widehat{\mathsf{b}}\cdot\nabla\widehat{\mathsf{b}}\right)\,\right] \end{aligned}$$

describes the finite-Larmor-orbit-average reduction of the equilibrium potential (Brizard, 1995). While the term $\mu B_{\rm E}$ in the guiding-center Hamiltonian (E2) might be smaller than $m|\mathbf{u}_{\rm E}|^2/2$, 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 in Hahm (1996). Perturbation analysis up to the second order is required for energy conservation up to $O(\epsilon_\delta^2)$ in the formulation in terms of total distribution function (Brizard, 1989a; Dubin *et al.*, 1983; Hahm, 1988). The total phase-space Lagrangian is given up to the second order by

$$\begin{aligned} \overline{\Gamma} \;=\; & \left(\frac{e}{c}\mathbf{A} + m\mathbf{u}_E + m\overline{v}_\parallel\widehat{\mathsf{b}}\right)\cdot\mathsf{d}\overline{\mathbf{X}} + \frac{\overline{\mu}B}{\Omega}\,\mathsf{d}\overline{\zeta} \\ & -\;(H_0 \;+\; e\,\delta\Psi_{\rm gy})\,\mathsf{d}t, \qquad \text{(E4)} \end{aligned}$$

where the effective gyrocenter perturbation potential is

$$\delta\Psi_{\rm gy} \;\equiv\; \langle\delta\phi_{\rm gc}\rangle \;-\; \frac{e}{2B}\,\frac{\partial}{\partial\overline{\mu}}\left\langle\delta\widetilde{\phi}_{\rm gc}^2\right\rangle.$$

The corresponding Euler-Lagrange equation is

$$-\frac{e\mathbf{B}^*}{c}\times\frac{d\overline{\mathbf{X}}}{dt} - m\widehat{\mathsf{b}}\,\frac{d\overline{v}_\parallel}{dt} \;=\; \overline{\nabla}(H_0 + e\,\delta\Psi_{\rm gy}), \qquad \text{(E5)}$$

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

$$\begin{aligned} \frac{d\overline{\mathbf{X}}}{dt} \;=\; & \overline{v}_\parallel\,\frac{\mathbf{B}^*}{B_\parallel^*} \;+\; \frac{c\widehat{\mathsf{b}}}{eB_\parallel^*}\times\Big[\,e\,\overline{\nabla}(\Phi + \delta\Psi_{\rm gy}) \\ & +\overline{\mu}\,\overline{\nabla}(B + B_{\rm E}) \;+\; \frac{m}{2}\overline{\nabla}|\mathbf{u}_{\rm E}|^2\Big], \qquad \text{(E6)} \end{aligned}$$

and

$$\begin{aligned} \frac{d\overline{v}_\parallel}{dt} \;=\; & -\frac{\mathbf{B}^*}{mB_\parallel^*}\cdot\Big[\,e\,\overline{\nabla}(\Phi + \delta\Psi_{\rm gy}) \\ & +\overline{\mu}\,\overline{\nabla}(B + B_{\rm E}) \;+\; \frac{m}{2}\overline{\nabla}|\mathbf{u}_{\rm E}|^2\Big] \qquad \text{(E7)} \end{aligned}$$

Although Eqs. (E6) and (E7) are mathematically concise, those can be written in the following form which is closer to the result of previous works in terms of the mass flow (Artun and Tang, 1994; Brizard, 1995).

$$\begin{aligned} \frac{d\overline{\mathbf{X}}}{dt} = & \mathbf{u}_E + \overline{v}_\parallel\widehat{\mathsf{b}} + \frac{c\widehat{\mathsf{b}}}{eB_\parallel^*}\times\Big[e\overline{\nabla}\delta\Psi_{\rm gy} + \overline{\mu}\overline{\nabla}(B + B_{\rm E}) \\ & +m(\mathbf{u}_E + \overline{v}_\parallel\widehat{\mathsf{b}})\cdot\overline{\nabla}(\mathbf{u}_E + \overline{v}_\parallel\widehat{\mathsf{b}})\Big], \qquad \text{(E8)} \end{aligned}$$

and

$$\begin{aligned} \frac{d\overline{v}_\parallel}{dt} = & -\frac{\mathbf{B}^{*(0)}}{mB_\parallel^{*(0)}}\cdot\Big[e\overline{\nabla}(\Phi_1 + \delta\Psi_{\rm gy}) + \overline{\mu}\,\overline{\nabla}(B + B_{\rm E}^{(0)}) \\ & +m(\mathbf{u}_E^{(0)} + \overline{v}_\parallel\widehat{\mathsf{b}})\cdot\overline{\nabla}(\mathbf{u}_E^{(0)} + \overline{v}_\parallel\widehat{\mathsf{b}})\Big]. \qquad \text{(E9)} \end{aligned}$$

Here, $\mathbf{u}_E^{(0)} \equiv \widehat{\mathsf{b}}\times\nabla\Phi^{(0)}/B$, $\mathbf{B}^{*(0)} \equiv \mathbf{B} + \frac{m}{e}\nabla\times(\mathbf{u}_E^{(0)} + v_\parallel\widehat{\mathsf{b}})$, and $B_\parallel^{*(0)} \equiv \widehat{\mathsf{b}}\cdot\mathbf{B}^{*(0)}$. Although Eq. (E8) is valid for an arbitrary Φ, Eq.(E9) can be only obtained from Eq. (E7) via a perturbative analysis (Brizard, 1995). The equilibrium electrostatic potential, in general, consists of two

[12] A tokamak-specific ordering, $B_\theta/B \simeq q\,r/R \ll 1$, with further subsidiary ordering, simplifies the formulation for applications. This exemplifies the nonuniqueness of the standard nonlinear gyrokinetic ordering; details can be found in Hahm (1996) and Hahm *et al.* (2004b).

parts $\Phi \equiv \Phi_0 + \Phi_1$. In most cases, Φ can be approximated by a flux function $\Phi_0(\psi)$ satisfying $\widehat{\mathsf{b}} \cdot \nabla \Phi_0 = 0$. The poloidal-angle-dependent $\Phi_1(\psi, \theta)$ can be produced, for instance, by the centrifugal-force-driven charge separation in strongly rotating plasmas (Connor *et al.*, 1987; Hinton and Wong, 1985). According to the ordering in this section, $\Phi_0 = \mathcal{O}(\epsilon_{\rm E}^{-1})$ and $\Phi_1 = O(1)$. Theory of $E \times B$ flow shear suppression of turbulence has been also extended to include the poloidal-angle-dependent $\Phi_1(\psi, \theta)$ (Hahm and Burrell, 1996) exhibiting the tensor nature of shearing process by large convective cells (Diamond *et al.*, 2005).

With Eqs. (E8) and (E9), one can write explicitly the gyrokinetic Vlasov equation for the gyrocenter distribution function $\overline{F}(\overline{\mathbf{X}}, \overline{\mu}, \overline{v}_\|, t)$,

$$\frac{\partial \overline{F}}{\partial t} + \frac{d\overline{\mathbf{X}}}{dt} \boldsymbol{\cdot} \overline{\nabla} \overline{F} + \frac{d\overline{v}_\|}{dt} \frac{\partial \overline{F}}{\partial \overline{v}_\|} \;=\; 0. \tag{E10}$$

Here, we note that $d\overline{\mu}/dt \equiv 0$ and $\partial F/\partial \overline{\zeta} \equiv 0$ have been used. The accompanying gyrokinetic Poisson's equation expressed in terms of the gyrocenter distribution function $\overline{F}(\overline{\mathbf{X}}, \overline{\mu}, \overline{v}_\|, t)$ is (Hahm, 1996):

$$\nabla^2 (\Phi + \delta\phi) \;=\; 4\pi e \, \left(n_{\rm e} \;-\; \overline{N}_{\rm i} \right), \tag{E11}$$

where the ion gyrofluid density

$$\overline{N}_{\rm i} \;\equiv\; \int d^3\overline{p} \left\langle e^{-\boldsymbol{\rho} \boldsymbol{\cdot} \nabla} \left(F + \frac{e\delta\widetilde{\phi}_{\rm gc}}{B} \frac{\partial F}{\partial \overline{\mu}} \right) \right\rangle$$

includes the ion polarization density and the electron density $n_{\rm e}$ can be obtained from the drift-kinetic equation ($\boldsymbol{\rho}_{\rm e} \boldsymbol{\cdot} \nabla \rightarrow 0$), for instance. The invariant energy for Eqs. (E10) and (E11) is obtained by transforming the energy constant of the original Vlasov-Poisson system as described in Appendix D,

$$\begin{aligned} E \;=\; & \int d^6\overline{Z}\,\overline{F}_{\rm i} \left[\overline{\mu}\,(B + B_{\rm E}) + \frac{m}{2}(|\mathbf{u}_E|^2 + \overline{v}_\|^2) \right] \\ & + \int d^6 z \, f_{\rm e} \, \left(\frac{m_{\rm e}}{2}\, v^2 \right) + \int \frac{d^3 x}{8\pi} \, |\mathbf{E}|^2 \\ & + \frac{e^2}{2B} \int d^6\overline{Z}\,\overline{F}_{\rm i} \left(\frac{\partial}{\partial \overline{\mu}} \left\langle \delta\widetilde{\phi}_{\rm gc}^2 \right\rangle \right), \end{aligned} \tag{E12}$$

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.

On a related subject, for extension of nonlinear gyrokinetic formulations to edge turbulence, a different ordering is desirable due to high relative fluctuation amplitude in L-mode plasmas and strong $E \times B$ flow shear in H-mode plasmas (Hahm *et al.*, 2004b). We note that, via a rigorous derivation, additional terms (other than the intuitively obvious, radially-dependent Doppler-shift-like term) appear in the gyrocenter equations of motion. Som of these terms are kept in some comprehensive gyrokinetic stability analysis addressing the $E \times B$ shear effects (Peeters and Strintzi, 2004; Rewoldt *et al.*, 1998).

For nonlinear gyrokinetic simulations of turbulence, on the other hand, much of emphasis in the last decade has been concentrated on the study of zonal flows which are spontaneously generated by turbulence. The self-generated zonal flows are radially localized ($k_r L_{\rm F} \gg 1$), axisymmetric ($k_\varphi = 0$), and mainly poloidal $E \times B$ flows.

There have been early indications from a fluid simulation (Hasegawa and Wakatani, 1985) and nonlinear gyrofluid simulations in 90's (Beer, 1995; Dorland, 1993; Hammett *et al.*, 1993; Waltz *et al.*, 1994) that self-generated zonal flows can be important in drift wave turbulence. Based on nonlinear gyrokinetic simulations (Dimits *et al.*, 2000; Lin *et al.*, 1998) 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 that the plasma microturbulence problem can be referred to as "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), 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, Conway *et al.* (2005); McKee *et al.* (2003)).

One important effect of zonal flows on drift wave turbulence is the shearing of turbulent eddys. While the shearing due to mean $E \times B$ flow is well understood and pedagogical explanations are available, the complex spatio-temporal 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 has been first characterized via the "effective shearing rate" (Hahm *et al.*, 1999). Later, this trend was confirmed in particular turbulence models (Kim, 2005; Kim *et al.*, 2004). This is also the reason why the Geodesic Acoustic Mode (GAM) (Winsor *et al.*, 1968), with $\omega_{GAM} \sim v_{\rm th}/R$, does not reduce the ambient turbulence significantly for typical core parameters (Angelino *et al.*, 2006; Miyato *et al.*, 2004). At the edge, sharp pressure gradients make the diamagnetic drift frequency at the relevant long-wavelength closer to the GAM frequency, i.e., $\omega_*/\omega_{\rm GAM} \sim (k_\theta R)\,\rho_{\rm i}/l_{\rm p} \sim 1$. Therefore, the GAM can possibly affect the edge ambient turbulence (Hallatschek and Biskamp, 2001; Scott, 2003). The second reason is the chaotic pattern of the zonal flows. Due to this the shearing due to zonal flows is better characterized by a random diffraction derived from statistical approaches (Diamond *et al.*, 2005) rather than the coherent stretching which is applicable to the shearing due to mean $E \times B$ shear.

2. Bounce-center-kinetic Vlasov Equation

As a second example of the extension of the nonlinear gyrokinetic formalism presented in the text, we construct nonlinear Hamilton equations for charged particles in the presence of low-frequency electromagnetic fluctuations with characteristic mode frequency ω such that

$$\omega_{\rm d}\,,\,\omega \;\ll\; \omega_{\rm b} \;\ll\; \Omega, \tag{E13}$$

where $\omega_{\rm b}$ and $\omega_{\rm d}$ denote the bounce and drift frequencies of a trapped guiding-center particle. This new time-scale ordering, thus, 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 $J = J_{\rm b}$.

In deriving these reduced equations, we ignore finite-Larmor-radius effects associated with the electromagnetic field perturbations (i.e., we take the long-wavelength limit $k_\perp^2\rho_{\rm i}^2 \ll 1$), and we refrain from ordering the perpendicular and parallel wavenumbers since $k_\|/k_\perp$ 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 $\overline{J}$. Here, the electromagnetic field fluctuations are represented by: the perturbed scalar potential $\delta\phi$, the parallel component of the perturbed vector potential $\delta A_\|$ ($\equiv \widehat{\mathsf{b}}_0\cdot\delta\mathbf{A}$), and the parallel component of the perturbed magnetic field $\delta B_\|$ ($\equiv \widehat{\mathsf{b}}_0\cdot\nabla\times\delta\mathbf{A}$). We shall assume that the characteristic mode frequency ω is much smaller than the bounce frequency $\omega_{\rm b}$, i.e.,

$$\frac{\omega}{\omega_{\rm b}} \;\sim\; \epsilon_\omega, \tag{E14}$$

where ϵ_ω is a small ordering parameter; we henceforth set $\epsilon_{\rm d}$ equal to one for clarity.

The perturbed guiding-center phase-space Lagrangian can be written as

$$\overline{\Gamma} \;=\; \overline{\Gamma}_0 \;+\; \epsilon_\delta\,\overline{\Gamma}_1, \tag{E15}$$

where the first-order guiding-center phase-space Lagrangian is (Brizard, 1989a)

$$\overline{\Gamma}_1 \;\equiv\; \frac{e}{c}\left(\delta A_{\|{\rm bc}}\,\frac{\partial s}{\partial\overline{u}^\alpha}\right)d\overline{u}^\alpha, \tag{E16}$$

and the first-order guiding-center Hamiltonian is

$$\overline{H}_1 \;\equiv\; e\,\delta\phi_{\rm gc} \;+\; \overline{\mu}\,\delta B_{\|{\rm bc}}. \tag{E17}$$

Here, dependence on the fast bounce-angle $\overline{\Theta}$ is reintroduced in $\overline{\Gamma}_n$ ($n \geq 1$) because the perturbation fields $(\delta\phi_{\rm gc}, \delta A_{\|{\rm bc}}, \delta B_{\|{\rm bc}})$ depend on $\overline{\Theta}$ through $s(\overline{\mathbf{u}}) \equiv s(\mathbf{u})$ (to lowest order in $\epsilon_{\rm d}$) and $\mathbf{y} \equiv \overline{\mathbf{y}} + \mathbf{\Lambda}_{\rm b}$. For example, the perturbed scalar potential $\delta\phi_{\rm gc}(\overline{\mathbf{y}},\overline{\mathbf{u}})$ is defined as

$$\delta\phi_{\rm gc}(\overline{\mathbf{y}},\overline{\mathbf{u}};t) \;\equiv\; \delta\phi\left(\overline{\mathbf{y}} + \mathbf{\Lambda}_{\rm b},\, s(\overline{\mathbf{u}});t\right). \tag{E18}$$

In what follows, we make no assumptions about the orderings of the parallel and perpendicular wavenumbers. In Eq. (E17), the gyrocenter magnetic moment

$$\overline{\mu} \;\equiv\; \mu \;+\; \epsilon_\delta\left[\frac{e\boldsymbol{\rho}}{B_0}\cdot\overline{\nabla}\left(\delta\phi_{\rm gc} \;-\; \frac{v_\|}{c}\,\delta A_{\|{\rm bc}}\right) - \mu\,\frac{\delta B_{\|{\rm bc}}}{B_0}\right] \;+\; \cdots \tag{E19}$$

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 $\rho^2k_\perp^2 \ll 1$) is

$$\overline{H}_2 \;\equiv\; -\,\frac{mc^2}{2B_0^2}\left|\overline{\nabla}_\perp\left(\delta\phi_{\rm gc} \;-\; \frac{v_\|}{c}\,\delta A_{\|{\rm bc}}\right)\right|^2 \;-\; e\delta\mathbf{A}_{\perp{\rm bc}}\cdot\frac{\widehat{\mathsf{b}}_0}{B_0}\times\overline{\nabla}_\perp\left(\delta\phi_{\rm gc} \;-\; \frac{v_\|}{c}\,\delta A_{\|{\rm bc}}\right). \tag{E20}$$

The new bounce-gyrocenter phase-space Lagrangian is chosen to be of the form

$$\widehat{\Gamma} \;\equiv\; \frac{e}{2c}\,\eta_{ab}\,\widehat{y}^a\,d\widehat{y}^b \;+\; \widehat{J}\,d\widehat{\Theta} \;-\; \widehat{w}\,dt, \tag{E21}$$

i.e., all the electromagnetic perturbation effects have been transfered to the bounce-gyrocenter Hamiltonian

$$\widehat{H}(\widehat{\mathbf{y}},t;\widehat{J}) \;\equiv\; \widehat{H}_0 \;+\; \epsilon_\delta\,\widehat{H}_1 \;+\; \epsilon_\delta^2\,\widehat{H}_2, \tag{E22}$$

where the second-order bounce-center Hamiltonian contains low-frequency ponderomotive terms associated with the asymptotic decoupling of the bounce-motion time scale. Here, the first-order bounce-center Hamiltonian is

$$\widehat{H}_1 \;\equiv\; e\,\langle\delta\psi_{\rm bc}\rangle_{\rm b} \;=\; \left\langle e\left(\delta\phi_{\rm gc} \;-\; \frac{v_\|}{c}\,\delta A_{\|{\rm bc}}\right)\right\rangle_{\rm b} \;+\; \overline{\mu}\,\left\langle\delta B_{\|{\rm bc}}\right\rangle_{\rm b}, \tag{E23}$$

where the bounce-angle averaging with respect to $\widehat{\Theta}$ is denoted $\langle\ \rangle_{\rm b}$. The second-order bounce-center Hamiltonian, on the other hand, is

$$\widehat{H}_2 \;\equiv\; \langle\overline{H}_2\rangle_{\rm b} \;+\; \frac{e^2}{2mc^2}\left\langle(\delta A_{\|{\rm bc}})^2\right\rangle_{\rm b} \;-\; \frac{e^2}{2\omega_{\rm b}}\left\langle\left\{\delta\widetilde{\Psi}_{\rm bc},\; \delta\widetilde{\psi}_{\rm bc}\right\}_{\rm bc}\right\rangle_{\rm b}, \tag{E24}$$

where $\delta\widetilde{\Psi}_{\rm bc} \equiv \int \delta\widetilde{\psi}_{\rm bc}\, d\widehat{\Theta}$ and $\{\ ,\ \}_{\rm bc}$ denotes the unperturbed bounce-center Poisson bracket.

The nonlinear bounce-gyrocenter Hamiltonian is, therefore, expressed as

$$\begin{aligned}\widehat{H} \;\equiv\; & \widehat{H}_0 \;+\; \epsilon_\delta\left\langle e\left(\delta\phi_{\rm gc} \;-\; \frac{v_\|}{c}\,\delta A_{\|{\rm bc}}\right) \;+\; \overline{\mu}\,\delta B_{\|{\rm bc}}\right\rangle_{\rm b} \\ & +\; \epsilon_\delta^2\left[\,\langle\overline{H}_2\rangle_{\rm b} \;+\; \frac{e^2}{2mc^2}\left\langle\left(\delta A_{\|{\rm bc}}\right)^2\right\rangle_{\rm b} \right. \\ & \left. -\; \frac{e^2}{2\omega_{\rm b}}\left\langle\left\{\delta\widetilde{\Psi}_{\rm bc},\; \delta\widetilde{\psi}_{\rm bc}\right\}_{\rm bc}\right\rangle_{\rm b}\,\right].\end{aligned} \tag{E25}$$

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 ($\langle\overline{H}_2\rangle_{\rm b}$) and neoclassical ($\langle\{\delta\widetilde{\Psi}_{\rm bc},\, \delta\widetilde{\psi}_{\rm bc}\}\rangle_{\rm b}$) terms. The bounce-center phase-space transformation $(\overline{\mathbf{y}}, \overline{\mathbf{u}}) \rightarrow (\widehat{\mathbf{y}}, \widehat{\mathbf{u}})$ is defined up to first order in ϵ_δ as

$$\left.\begin{aligned}\widehat{y}^a \;&=\; \overline{y}^a \;+\; \epsilon_\delta\;\left\{\overline{S}_1,\; \overline{y}^a\right\}_{\rm bc} \\ \widehat{u}^\alpha \;&=\; \overline{u}^\alpha \;+\; \epsilon_\delta\;\left\{\overline{S}_1,\; \overline{u}^\alpha\right\}_{\rm bc} \\ &\quad +\; \epsilon_\delta(e/c)\,\delta A_{\|{\rm bc}}\;\left\{\overline{s},\; \overline{u}^\alpha\right\}_{\rm bc}\end{aligned}\right\}, \tag{E26}$$

where $\overline{S}_1 \equiv e\,\delta\widetilde{\Psi}_{\rm bc}/\omega_{\rm b}$. Lastly, the neoclassical polarization density can be defined in terms of the push-forward expression $\rho_{\rm pol} \equiv -\sum e\,\nabla\cdot\|\boldsymbol{\rho}_\epsilon\|$, where $\|\ \|$ denotes a momentum integration over the bounce-center distribution function and

$$\rho_\epsilon^a \;\equiv\; -\,\epsilon_\delta\;\left\{\overline{S}_1,\; \overline{y}^a \;+\; \Lambda^a\right\}_{\rm bc}.$$

By definition, the bounce-center moment $\|\rho_\epsilon^a\|$ involves a bounce-angle average and, thus, to lowest order in the bounce-kinetic ordering, we find

$$\langle\rho_\epsilon^a\rangle \;=\; \epsilon_\delta\,\frac{e}{\omega_{\rm b}}\,\frac{\partial}{\partial\widehat{J}}\left\langle\delta\widetilde{\psi}_{\rm bc}\;\Lambda^a\right\rangle.$$

We, therefore, 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., 2006, Plasma Phys. Control. Fusion, (to appear).

Antonsen, T. M., and B. Lane, 1980, Phys. Fluids **23**, 1205-1214.

Arnold, V. I., 1989, *Mathematical Methods of Classical Mechanics, 2nd ed.* (Springer, New York).

Artun, M., and W. M. Tang, 1994, Phys. Plasmas **1**, 2682-2692.

Beer, M. A., 1995, Ph. D. Thesis (Princeton University).

Bernstein, I. B., and P. J. Catto, 1985, Phys. Fluids **28**, 1342-1353.

Biglari, H., P. H. Diamond, and P. W. Terry, 1990, Phys. Fluids B **2**, 1-4.

Boghosian, B. M., 1987, Ph. D. thesis (University of California, Davis); see also arXiv:physics/0307148.

Briguglio, S., L. Chen, J. Q. Dong, G. Fogaccia, R. A. Santoro, G. Vlad, and F. Zonca, 2000, Nucl. Fusion **40**, 701-706.

Briguglio, S., F. Zonca, and G. Vlad, 1998, Phys. Plasmas **5**, 3287-3301.

Brizard, A. J., 1989a, J. Plasma Phys. **41**, 541-559.

Brizard, A. J., 1989b, Phys. Fluids B **1**, 1381-1384.

Brizard, A. J., 1990, Ph. D. Thesis (Princeton University).

Brizard, A. J., 1992, Phys. Fluids B **4**, 1213-1228.

Brizard, A. J., 1994a, Phys. Plasmas **1**, 2460-2472.

Brizard, A. J., 1994a, Phys. Plasmas **1**, 2473-2479.

Brizard, A. J., 1995, Phys. Plasmas **2**, 459-471.

Brizard, A. J., 2000a, Phys. Rev. Lett. **84**, 5768-5771.

Brizard, A. J., 2000b, Phys. Plasmas **7**, 4816-4822.

Brizard, A. J., 2000c, Phys. Plasmas **7**, 3238-3246.

Brizard, A. J., 2001, Phys. Lett. A **291**, 146-149.

Brizard, A. J., 2004, Phys. Plasmas **11**, 4429-4438.

Brizard, A. J., 2005a, J. Plasma Phys. **71**, 225-236.

Brizard, A. J., 2005b, Phys. Plasmas **12**, 092302:1-11.

Brizard, A. J., and A. A. Chan, 1999, Phys. Plasmas **6**, 4548-4558.

Burrell, K. H., 1997, Phys. Plasmas **4**, 1499-1518.

Carreras, B. A., D. E. Newman, P. H. Diamond, and Y. M. Liang, 1994, Phys. Plasmas **1**, 4014-4021.

Cary, J. R., 1977, J. Math. Phys. **18**, 2432-2435.

Cary, J. R., and A. N. Kaufman, 1981, Phys. Fluids **24**, 1238-1250.

Cary, J. R., and R. G. Littlejohn, 1983, Ann. Phys. (NY) **151**, 1-34.

Catto, P. J., W. M. Tang, and D. E. Baldwin, 1981, Plasma Phys. **23**, 639-650.

Cendra, H., D. D. Holm, M. J. W. Doyle, and J. E. Marsden, 1998, J. Math. Phys. **39**, 3138-3157.

Chang, C. S., and F. L. Hinton, 1982, Phys. Fluids **25**, 1493-1494.

Chen, L., 1994, Phys. Plasmas **1**, 1519-1522.

Chen, L., 1999, J. Geophys. Res. **104**, 2421-2428.

Chen, L., and A. Hasegawa, 1994, J. Geophys. Res. **99**, 179-182.

Chen, L., Z. Lin, and R. B. White, 2000, Phys. Plasmas **7**, 3129-3132.

Chen, L., Z. Lin, R. B. White, and F. Zonca, 2001, Nucl. Fusion **41**, 747-753.

Chen, L., R. B. White, and F. Zonca, 2004, Phys. Rev. Lett. **92**, 075004.

Chen, L., F. Zonca, and Z. Lin, 2005, Plasma Phys. Control. Fusion **47**, B71-B81.

Chen, Y., and S. E Parker, 2001, Phys. Plasmas **8**, 2095-2100.

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**, 1121-1127.

Connor, J. W., S. C. Cowley, R. J. Hastie, and L. R. Pan, 1987, Plasma Phys. Control. Fusion **29**, 919-931.

Connor, J. W., and H. R. Wilson, 1994, Plasma Phys. Control. Fusion **36**, 719-795.

Connor, J. W., and H. R. Wilson, 2000, Plasma Phys. Control. Fusion **42**, R1-R74.

Conway, G. D., B. D. Scott, J. Schirmer, M. Reich, A. Kendl, and the ASDEX Upgrade Team, 2005, Plasma Phys. Control. Fusion **47**, 1165-1185.

Dannert, T., and F. Jenko, 2005, Phys. Plasmas **12**, 072309:1-8.

Depret, G., X. Garbet, P. Bertrand, and A. Ghizzo, 2000, Plasma Phys. Control. Fusion **42**, 949-971.

Diamond, P. H., and T. S. Hahm, 1995, Phys. Plasmas **2**, 3640-3649.

Diamond, P. H., S. I. Itoh, K. Itoh, and T. S. Hahm, 2005, Plasma Phys. Control. Fusion **47**, R35-R161.

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**, 1472-1475.

Diamond, P. H., Y. M. Liang, B. A. Carreras, and P. W. Terry, 1994, Phys. Rev. Lett. **72**, 2565-2568.

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**, 969-983.

Dimits, A. M., and B. I. Cohen, 1994, Phys. Rev. E **49**, 709-721.

Dimits, A. M., L. L. Lodestro, and D. H. E. Dubin, 1992, Phys. Fluids B **4**, 274-277.

Dimits, A. M., T. J. Williams, J. A. Byers, and B. I. Cohen, 1996, Phys. Rev. Lett. **77**, 71-74.

Dorland, W., 1993, Ph. D. Thesis (Princeton University).

Dorland, W., F. Jenko, M. Kotschenreuther, and B. N. Rogers, 2000, Phys. Rev. Lett. **85**, 55795582.

Dubin, D. H. E., J. A. Krommes, C. Oberman, and W. W. Lee, 1983, Phys. Fluids **26**, 3524-3535.

Endler, M., H. Niedermeyer, L. Giannone, E. Kolzhauer, A. Rudyj, G. Theimer, and N. Tsois, 1995, Nucl. Fusion **35**, 1307-1339.

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**, 2637-2648.

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**, 3736-3739.

Fong, B. H., and T. S. Hahm, 1999, Phys. Plasmas **6**, 188-199.

Frieman, E. A., and L. Chen, 1982, Phys. Fluids **25**, 502-508.

Frisch, U., 1995, *Turbulence* (Cambridge University Press, Cambridge).

Gang, F. Y., and P. H. Diamond, 1990, Phys. Fluids B **2**, 2976-2985.

Gang, F. Y., P. H. Diamond, and M. N. Rosenbluth, 1991, Phys. Fluids B **3**, 68-86.

Garbet, X., L. Laurent, A. Samain, and J. Chinardet, 1994, Nucl. Fusion **34**, 963-974.

Garbet, X., and R. E. Waltz, 1998, Phys. Plasmas **5**, 2836-2845.

Goldstein, H., C. Poole, and J. Safko, 2002, *Classical Mechanics, 3rd ed.* (Addison-Wesley, San Francisco).

Gurcan, O. D., P. H. Diamond, T. S. Hahm, and Z. Lin, 2005, Phys. Plasmas **12**, 032303.

Gurcan, O. D., P. H. Diamond, and T. S. Hahm, 2006, Phys. Plasmas **(to appear)**.

Hahm, T. S., 1988, Phys. Fluids **31**, 2670-2673.

Hahm, T. S., 1992, Phys. Fluids B **4**, 2801-2806.

Hahm, T. S., 1994, Phys. Plasmas **1**, 2940-2944.

Hahm, T. S., 1996, Phys. Plasmas **3**, 4658-4664.

Hahm, T. S., 2002, Plasma Phys. Control. Fusion **44**, A87-A101.

Hahm, T. S., M. A. Beer, Z. Lin, G. W. Hammett, W. W. Lee, and W. M. Tang, 1999, Phys. Plasmas **6**, 922-926.

Hahm, T. S., and K. H. Burrell, 1995, Phys. Plasmas **2**, 1648-1651.

Hahm, T. S., and K. H. Burrell, 1996, Plasma Phys. Control. Fusion **38**, 1427-1431.

Hahm, T. S., K. H. Burrell, Z. Lin, R. Nazikian, and E. J. Synakowski, 2000, Plasma Phys. Control. Fusion **42**, A205-A210.

Hahm, T. S., P. H. Diamond, Z. Lin, K. Itoh, and S. I. Itoh, 2004a, Plasma Phys. Control. Fusion **46**, A323-A333.

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**, 1940-1948.

Hahm, T. S., Z. Lin, P. H. Diamond, G. Rewoldt, W. X. Wang, S. Ethier, O. Gurcan, W. W. Lee, and W. M. Tang, 2004b, *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**, 1815-1821.

Hahm, T. S., and W. M. Tang, 1991, Phys. Fluids B **3**, 989-999.

Hahm, T. S., and W. M. Tang, 1996, Phys. Plasmas **3**, 242-247.

Hallatschek, K., and D. Biskamp, , 2001, Phys. Rev. Lett. **86**, 1223-1226.

Hammett, G. W., M. A. Beer,W. Dorland, S. C. Cowley, and S. A. Smith, 1993, Plasma Phys. Control. Fusion **35**, 973-985.

Hasegawa, A., and M. Wakatani, 1987, Phys. Rev. Lett. **59**, 1581-1584.

Hastie, R. J., J. B. Taylor, and F. A. Haas, 1967, Ann. Phys. **41**, 302-338.

Hatori, T., and H. Washimi, 1981, Phys. Rev. Lett. **46**, 240-243.

Hatzky, R., T. M. Tran, A. Koenies, R. Kleiber, and S. J. Allfrey, 2002, Phys. Plasmas **9**, 898-912.

Hinton, F. L., and R. D. Hazeltine, 1976, Rev. Mod. Phys. **48**, 239-308.

Hinton, F. L., and S. K. Wong, 1985, Phys. Fluids **28**, 3082-3098.

Hinton, F. L., J. Kim, Y. B. Kim, A. Brizard, and K. H. Burrell, 1994, Phys. Rev. Lett. **72**, 1216-1219.

Horton, W., 1999, Rev. Mod. Phys. **71**, 735-778.

Horton, W., P. Zhu, G. T. Hoang, T. Aniel, M. Ottaviani, and X. Garbet, 2000, Phys. Plasmas **7**, 1494-1510.

Howes, G. G., S. C. Cowley, W. Dorland, G. W. Hammett, E. Quataert, and A. A. Schekochihin, 2006, Astrophys. J. , (submitted).

Idomura, Y., M. Wakatani, and S. Tokuda, 2000, Phys. Plasmas **7**, 3551-3566.

Itoh, S. I., and K. Itoh, 2001, Plasma Phys. Control. Fusion **43**, 1055-1102.

Itoh, K., S. I. Itoh, T. S. Hahm, and P. H. Diamond, 2005, J. Phys. Soc. Japan **74**, 2001-2006.

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**, 1904-1910.

Jenko, F., and B. D. Scott, 1999, Phys. Plasmas **6**, 2705-2713.

Kim, E. J., 2005, Phys. Plasmas **12**, 090902:1-4.

Kim, E. J., P. H. Diamond, and T. S. Hahm, 2004, Phys. Plasmas **11**, 4554-4558.

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**, 961-968.

Kim, E. J., C. Holland, and P. H. Diamond, 2003, Phys. Rev. Lett. **91**, 075003:1-4.

Kishimoto, Y., S. Tokuda, and K. Sakamoto, 1995, Phys. Plasmas **2**, 1316-1325.

Krommes, J. A., 1993a, Phys. Rev. Lett. **70**, 3067-3070.

Krommes, J. A., 1993b, Phys. Fluids B **5**, 2405-2411.

Krommes, J. A., 2002, Phys. Rep. **360**, 1-352.

Krommes, J. A., and C. B. Kim, 1988, Phys. Fluids **31**, 869-886.

Krommes, J. A., W. W. Lee, and C. Oberman, 1986, Phys. Fluids **29**, 2421-2425.

Krommes, J. A., C. Oberman, and R. G Kleva, 1983, J. Plasma Phys. **30**, 11-56.

Kruskal, M., 1962, J. Math. Phys. **3**, 806-828.

Lashmore-Davies, C. N., and R. O. Dendy, 1989, Phys. Fluids B **1**, 1565-1577.

Lee, W. W., 1983, Phys. Fluids **26**, 556-562.

Lee, W. W., 1987, J. Comput. Phys. **72**, 243-269.

Lee, W. W., J. A. Krommes, C. R. Oberman, and R. A. Smith, 1984, Phys. Fluids **27**, 2652-2663.

Lee, W. W., J. L. V. Lewandowski, T. S. Hahm, and Z. Lin, 2001, Phys. Plasmas **8**, 4435-4440.

Lee, W. W., and W. M. Tang, 1988, Phys. Fluids **31**, 612-624.

Lee, X. S., J. R. Myra, and P. J. Catto, 1983, Phys. Fluids **26**, 223-229.

Lichtenberg, A. J., and M. A. Lieberman, 1984, *Regular and Stochastic Motion* (Springer, New York).

Liewer, P. C., 1985, Nucl. Fusion **25**, 543-621.

Lin, Y., X. Wang, Z. Lin, and L. Chen, 2005, Plasma Phys. Control. Fusion **47**, 657-669.

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., T. S. Hahm, W. W. Lee, W. M. Tang, and R. B. White, 1998, Science **281**, 1835-1837.

Lin, Z., T. S. Hahm, W. W. Lee, W. M. Tang, and P. H. Diamond, 1999, Phys. Rev. Lett. **83**, 3645-3648.

Lin, Z., and T. S. Hahm, 2004, Phys. Plasmas **11**, 1099-1108.

Littlejohn, R. G., 1981, Phys. Fluids **24**, 1730-1749.

Littlejohn, R. G., 1982a, J. Math. Phys. **23**, 742-747.

Littlejohn, R. G., 1982b, Physica Scr. **T2/I**, 119-125.

Littlejohn, R. G., 1983, J. Plasma Phys. **29**, 111-125.

Littlejohn, R. G., 1988, Phys. Rev. A **38**, 6034-6045.

Low, F. E., 1958, Proc. R. Soc. London A **248**, 282-287.

Marsden, J. E., and A. Weinstein, 1982, Physica D **4**, 394-406.

Matsumoto, T., H. Naitou, S. Tokuda, and Y. Kishimoto, 2005, Nucl. Fusion **45**, 1264-1270.

Matsumoto, T., S. Tokuda, Y. Kishimoto, and H. Naitou, 2003, Phys. Plasmas **10**, 195-203.

Mattor, N., 1992, Phys. Fluids B **4**, 3952-3961.

Mattor, N., and P. H. Diamond, 1989, Phys. Fluids B **1**, 1980-1992.

Mazzucato, E., 1982, Phys. Rev. Lett. **48**, 1828-1830.

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. Synakowski, W. M. Tang, G. Taylor, and M. C. Zarnstorff, 1996, Phys. Rev. Lett. **77**, 3145-3148.

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**, 1712-1719.

Mishchenko, A., R. Hatzky, and A. Könies, 2004, Phys. Plasmas **11**, 5480-5486.

Miyato, N., Y. Kishimoto, and J. Li, 2004, Phys. Plasmas **11**, 5557-5564.

Naitou, H., K. Tsuda, W. W. Lee, and R. D. Sydora, 1995, Phys. Plasmas **2**, 4257-4268.

Naulin, V., J. Nycander, and J. J. Rasmussen, 1998, Phys. Rev. Lett. **81**, 4148-4151.

Naulin, V., A. H. Neilsen, and J. J. Rasmussen, 2005, Phys. Plasmas **12**, 122306.

Newman, D. E., B. A. Carreras, P. H. Diamond, and T. S. Hahm, 1996, Phys. Plasmas **3**, 1858-1866.

Northrop, T. G., 1963, *Adiabatic Motion of Charged Particles* (Wiley, New York).

Northrop, T. G., and E. Teller, 1960, Phys. Rev. **117**, 215225.

Ono, M., M. G. Bell, R. E. Bell, *et al.*, 2003, Plasma Phys. Control. Fusion **45**, A335-A350.

Otsuka, F., and T. Hada, 2003, Space Sci. Rev. **107**, 499-502.

Park, W., S. E. Parker, H. Biglari, *et al.*, 1992, Phys. Fluids **4**, 2033-2037.

Parker, S. E., Y. Chen, W. Wan, B. I. Cohen, and W. M. Nevins, 2004, Phys. Plasmas **11**, 2594-2599.

Parker, S. E., and W. W. Lee, 1993, Phys. Fluids B **5**, 77-86.

Parker, S. E., W. W. Lee, and R. A. Santoro, 1993, Phys. Rev. Lett. **71**, 2042-2045.

Peeters, A. G., and D. Strintzi, 2004, Phys. Plasmas **11**, 3748-3751.

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**, 1962-1969.

Qin, H., W. M. Tang, W. W. Lee, and G. Rewoldt, 1999, Phys. Plasmas **6**, 1575-1588.

Qin, H., W. M. Tang, and W. W. Lee, 2000, Phys. Plasmas **7**, 4433-4445.

Rechester, A. B., and M. N. Rosenbluth, 1978, Phys. Rev. Lett. **40**, 38-41.

Rewoldt, G., 1998, Phys. Plasmas **5**, 1815-1821.

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. Control. Fusion **47**, B323-B336.

Rosenbluth, M. N., R. D. Hazeltine, and F. L. Hinton, 1972, Phys. Fluids **15**, 116-140.

Rosenbluth, M. N., and F. L. Hinton, 1998, Phys. Rev. Lett. **80**, 724-727.
Rutherford, P. H., and E. A. Frieman, 1968, Phys. Fluids **11**, 569-585.
Sagdeev, R. Z., and A. A. Galeev, 1969, *Nonlinear Plasma Theory* (Benjamin, New York).
Santoro, R .A., and L. Chen, 1996, Phys. Plasmas **6**, 2349-2352.
Sarazin, Y., and Ph. Ghendrih, 1998, Phys. Plasmas **5**, 4214-4228.
Sarazin, Y., V. Grandgirard, E. Fleurence, X. Garbet, P. Ghendrih, P. Bertrand, and G. Depret, 2005, Plasma Phys. Control. Fusion **47**, 1817-1839.
Schulz, M., and L. J. Lanzerotti, 1974, *Particle Diffusion in the Radiation Belts* (Springer, New York).
Scott, B. D., 2003, Phys. Lett. A **320**, 53-62.
Scott, B. D., 2005, New J. Phys. **7**, 92:1-45.
Seliger, R. L., and G. B. Whitham, 1968, Proc. R. Soc. London Ser. A **305**, 1-25.
Shaing, K. C., E. C. Crume, and W. A. Houlberg, 1990, Phys. Fluids B **2**, 1492-1498.
Similon, P. L., 1985, Phys. Lett. A **112**, 33-37.
Similon, P. L., and P. H. Diamond, 1984, Phys. Fluids **27**, 916-924.
Smith, R. A., J. A. Krommes, and W. W. Lee, 1985, Phys. Fluids **28**, 1069-1076.
Stern, D. P., 1970, Am. J. Phys. **38**, 494-501.
Strintzi, D., B. D. Scott, and A. J. Brizard, 2005, Phys. Plasmas **12**, 052517.
Sugama,H., 2000, Phys. Plasmas **7**, 466-480.
Sugama,H., M. Okamoto, W. Horton, and M. Wakatani 1996, Phys. Plasmas **3**, 2379-2394.
Sydora, R. D., 2001, Phys. Plasmas **8**, 1929-1934.
Sydora, R. D., 1990, Phys. Fluids B**2**, 1455-1460.
Sydora, R. D., V. K. Decyk, and J. M. Dawson, 1996, Plasma Phys. Control. Fusion **12**, A281-A294.
Sydora, R. D., T. S. Hahm, W. W. Lee, and J. M. Dawson, 1990, Phys. Rev. Lett. **64**, 2015-2018.
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**, 2101-2106.
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**, 1736-1744.
Tang, W. M., 1978, Nucl. Fusion **18**, 1089-1160.
Tang, W. M., 2002, Phys. Plasmas **9**, 1856-1872.
Tang, W. M., and V. S Chan, 2005, Plasma Phys. Control. Fusion **47**, R1-R34.
Tang, W. M., J. W. Connor, and R. J. Hastie, 1980, Nucl. Fusion **20**, 1439-1453.
Taylor, J. B., 1967, Phys. Fluids **10**, 1357-1359.
Terry, P. W., 2000, Rev. Mod. Phys. **72**, 109-165.
Terry, P. W., P. H. Diamond, and T. S. Hahm, 1986, Phys. Rev. Lett. **57**, 1899-1902.
Terry, P. W., P. H. Diamond, and T. S. Hahm, 1988, Phys. Rev. Lett. **60**, 966.
Thoul, A. A., P. L. Similon, and R. N. Sudan, 1987, Phys. Rev. Lett. **59**, 1448-1451.
Thoul, A. A., P. L. Similon, and R. N. Sudan, 1988, Phys. Rev. Lett. **60**, 967.
Tsai, S. T., J. W. Van Dam, and L. Chen, 1984, Plasma Phys. Control. Fusion **26**, 907-914.
Todo, Y., H. L. Berk, and B. N. Breizman, 2003, Phys. Plasmas **10**, 2888-2902.
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, 2004a, Nucl. Fusion **44**, 172-180.
Villard, L., P. Angelino, A. Bottino, S. J. Allfrey, R. Hatzky, Y. Idomura, O. Sauter, M. Brunetti, and T. M. Tran, 2004b, Plasma Phys. Control. Fusion **46**, B51-B62.
Vlad, G., F. Zonca, and S. Briguglio, 1999, La Rivista del Nuovo Cimento **22**, 1-97.
Wagner, F., G. Becker, K. Behringer, *et al.*, 1982, Phys. Rev. Lett. **49**, 1408-1412.
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**, 1938-1946.
Waltz, R. E., G. D. Kerbel, and J. Milovich, 1994, Phys. Plasmas **1**, 2229-2244.
Watanabe,T. H., H. Sugama, 2006, Nucl. Fusion **46**, 24-32.
Winsor, N., J. L. Johnson, and J. M. Dawson, 1968, Phys. Fluids **11**, 2448-2450.
Wong, K. L., N. Bretz, T. S. Hahm, and E. J. Synakowski, 1997, Phys. Lett. A **236**, 339-344.
Wooton, 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**, 2879-2903.
Yagi, M., S. I. Itoh, A. Fukuyama, and M. Azumi, 1995, Phys. Plasmas **2**, 4140-4148.
Ye, H., and P. J. Morrison, 1992, Phys. Fluids B **4**, 771-777.
Zonca, F., S. Briguglio, L. Chen, S. Dettrick, G. Fogaccia, D. Testa, and G. Vlad, 2002, Phys. Plasmas **9**, 4939-4956.
Zonca, F., S. Briguglio, L. Chen, G. Fogaccia, and G. Vlad, 2005, Nucl. Fusion **45**, 477-484.
Zonca, F., R. B. White, and L. Chen, 2004, Phys. Plasmas **11**, 2488-2496.
Zweben, S. J., and S. S. Medley, 1989, Phys. Fluids B **1**, 2058-2065.

TABLE I Applications of Nonlinear Gyrokinetic Equations

Instability	**Nonlinear Theory**	**Nonlinear Simulation**
Drift (Universal or Dissipative) Instability	Frieman and Chen (1982)	Lee (1983)
	Smith *et al.* (1985)	Lee *et al.* (1984)
	Hahm (1992)	Parker and Lee (1993)
Ion Temperature Gradient (ITG) Mode	Mattor and Diamond (1989)	Lee and Tang (1988)
	Hahm and Tang (1990)	Sydora *et al.* (1990)
	Mattor (1992)	Parker *et al.* (1993)
ITG Turbulence with Zonal Flows	Rosenbluth and Hinton (1998)	Dimits *et al.* (1996)
	Chen *et al.* (2000)	Lin *et al.* (1999)
		Refs. in Diamond *et al.* (2005)
ITG Turbulence with Velocity-space		Hatzky *et al.* (2002)
Nonlinearity addressing Energy Conservation		Villard *et al.* (2004a)
Trapped Electron Mode	Similon and Diamond (1984)	Sydora (1990)
	Gang *et al.* (1991)	Chen and Parker (2001)
	Hahm and Tang (1991)	Ernst *et al.* (2004)
		Dannert and Jenko (2005)
Trapped-Ion Mode	Hahm and Tang (1996)	Depret *et al.* (2000)
Electron-Temperature-Gradient (ETG) Mode	Kim *et al.* (2003)	Jenko *et al.* (2000)
	Chen *et al.* (2005)	Idomura *et al.* (2000)
		Dorland *et al.* (2000)
		Lin *et al.* (2005)
Interchange turbulence		Sarazin *et al.* (2005)
(Kinetic) Shear Alfvén Wave	Frieman and Chen (1982)	Lee *et al.* (2001)
	Hahm *et al.* (1988)	Parker *et al.* (2004)
Drift-Alfvén Turbulence	Briguglio *et al.* (2000)	Briguglio *et al.* (1998)
	Chen *et al.* (2001)	Jenko and Scott (1999)
		Chen *et al.* (2001)
Tearing & Internal Kink Instability		Naitou *et al.* (1995)
		Matsumoto *et al.* (2005, 2003)
Micro-tearing & Drift-tearing Mode		Sydora (2001)
		Parker *et al.* (2004)
Energetic Particle driven MHD Instabilities	Chen (1994)	Park *et al.* (1992)
	Vlad *et al.* (1999)	Santoro and Chen (1996)
	Zonca *et al.* (2005)	Zonca *et al.* (2002)
		Todo *et al.* (2003)
Geo-magnetic Pulsation	Chen and Hasegawa (1994)	
Whistler Lower-hybrid Instability		Lin and Wang *et al.* (2005)

External Distribution

Plasma Research Laboratory, Australian National University, Australia
Professor I.R. Jones, Flinders University, Australia
Professor João Canalle, Instituto de Fisica DEQ/IF - UERJ, Brazil
Mr. Gerson O. Ludwig, Instituto Nacional de Pesquisas, Brazil
Dr. P.H. Sakanaka, Instituto Fisica, Brazil
The Librarian, Culham Science Center, England
Mrs. S.A. Hutchinson, JET Library, England
Professor M.N. Bussac, Ecole Polytechnique, France
Librarian, Max-Planck-Institut für Plasmaphysik, Germany
Jolan Moldvai, Reports Library, Hungarian Academy of Sciences, Central Research Institute for Physics, Hungary
Dr. P. Kaw, Institute for Plasma Research, India
Ms. P.J. Pathak, Librarian, Institute for Plasma Research, India
Dr. Pandji Triadyaksa, Fakultas MIPA Universitas Diponegoro, Indonesia
Professor Sami Cuperman, Plasma Physics Group, Tel Aviv University, Israel
Ms. Clelia De Palo, Associazione EURATOM-ENEA, Italy
Dr. G. Grosso, Instituto di Fisica del Plasma, Italy
Librarian, Naka Fusion Research Establishment, JAERI, Japan
Library, Laboratory for Complex Energy Processes, Institute for Advanced Study, Kyoto University, Japan
Research Information Center, National Institute for Fusion Science, Japan
Professor Toshitaka Idehara, Director, Research Center for Development of Far-Infrared Region, Fukui University, Japan
Dr. O. Mitarai, Kyushu Tokai University, Japan
Mr. Adefila Olumide, Ilorin, Kwara State, Nigeria
Dr. Jiangang Li, Institute of Plasma Physics, Chinese Academy of Sciences, People's Republic of China
Professor Yuping Huo, School of Physical Science and Technology, People's Republic of China
Library, Academia Sinica, Institute of Plasma Physics, People's Republic of China
Librarian, Institute of Physics, Chinese Academy of Sciences, People's Republic of China
Dr. S. Mirnov, TRINITI, Troitsk, Russian Federation, Russia
Dr. V.S. Strelkov, Kurchatov Institute, Russian Federation, Russia
Kazi Firoz, UPJS, Kosice, Slovakia
Professor Peter Lukac, Katedra Fyziky Plazmy MFF UK, Mlynska dolina F-2, Komenskeho Univerzita, SK-842 15 Bratislava, Slovakia
Dr. G.S. Lee, Korea Basic Science Institute, South Korea
Dr. Rasulkhozha S. Sharafiddinov, Theoretical Physics Division, Insitute of Nuclear Physics, Uzbekistan
Institute for Plasma Research, University of Maryland, USA
Librarian, Fusion Energy Division, Oak Ridge National Laboratory, USA
Librarian, Institute of Fusion Studies, University of Texas, USA
Librarian, Magnetic Fusion Program, Lawrence Livermore National Laboratory, USA
Library, General Atomics, USA
Plasma Physics Group, Fusion Energy Research Program, University of California at San Diego, USA
Plasma Physics Library, Columbia University, USA
Alkesh Punjabi, Center for Fusion Research and Training, Hampton University, USA
Dr. W.M. Stacey, Fusion Research Center, Georgia Institute of Technology, USA
Director, Research Division, OFES, Washington, D.C. 20585-1290

The Princeton Plasma Physics Laboratory is operated
by Princeton University under contract
with the U.S. Department of Energy.

Information Services
Princeton Plasma Physics Laboratory
P.O. Box 451
Princeton, NJ 08543

Phone: 609-243-2750
Fax: 609-243-2751
e-mail: pppl_info@pppl.gov
Internet Address: http://www.pppl.gov