Kinetic theory of flowing, magnetized plasma

R. D. Hazeltine
Institute for Fusion Studies and Department of Physics, The University of Texas at Austin, Austin, Texas 78712-0262

F. L. Hinton
General Atomics, P.O. Box 85608, San Diego, California 92186-5608

(Received 11 June 2005; accepted 31 August 2005; published online 5 October 2005)

The drift-kinetic equation for a rapidly flowing magnetized plasma is derived, allowing for arbitrary anisotropy of the distribution function. © *2005 American Institute of Physics*.
[DOI: 10.1063/1.2082007]

I. INTRODUCTION

This work derives a drift-kinetic equation[1,2,14]—that is, a kinetic equation for guiding-center motion—for a magnetized plasma whose flow velocity is comparable to its thermal speed. Such rapidly flowing plasmas play an increasingly important role in several areas of plasma physics research. For example, they occur in a number of astrophysical phenomena, such as galactic jets;[3] they play a key role in various laboratory plasmas, such as the "centrifugal confinement" device;[4] and they underly studies of novel plasma equilibria, such as Ref. 5. Finally, rapid plasma flow can have dramatic effects on plasma confinement in tokamak devices.[6]

Most previous studies of rapidly flowing, magnetized plasmas are based on fluid models. However, in many applications the collision frequency is relatively small, allowing significant departure from Maxwellian distribution functions and requiring kinetic analysis. There are a number of kinetic treatments,[7–9,11,12] but these are not sufficiently general to treat some plasmas of recent interest. In particular the previous studies assume the lowest-order distribution function to be Maxwellian and thus isotropic in velocity space. This assumption simplifies the kinetic equation and its derivation, but it may not apply to the low collisionality regimes associated with astrophysical jets, to laboratory plasmas in nontoroidal geometry, or to toroidal plasmas that are strongly driven. The main objective of this work is to remove the isotropy assumption and derive a drift-kinetic that is fully general with regard both to isotropy and plasma geometry.

In deriving the new result, we also attempt to improve upon previous literature in other ways. First, we have tried to make the derivation as transparent and systematic as possible, and to express the result in a convenient form. Second, we demonstrate its invariance to velocity-coordinate rotations in the plane normal to the magnetic field, by expressing all the drift-kinetic coefficients in term of the local magnetic field and its gradients. Third, we relate the new result to the well-known kinetic equation[13] used in "kinetic magnetohydrodynamics (MHD)," showing in the process that both new and previous results conserve phase space. More generally we include detailed comparisons of our result to previous kinetic descriptions, including Refs. 7, 11, and 13.

Our work is similar in some respects to Ref. 10, in which the authors carried out a derivation of the gyrokinetic equation, in a small gyroradius expansion limit. They assumed a flow which did not need to be small compared with the thermal speed, but they assumed that this large flow was perpendicular to the magnetic field. That is, any parallel component of the flow was assumed to be much smaller than the thermal speed. In contrast, we have assumed that the flow has an arbitrary parallel component, which need not be small. Our results are expected to be more useful in the treatment of magnetically confined plasmas, where the perpendicular and parallel flow components may be comparable. Also, the general discussion of phase-space conservation given in Ref. 10 applies directly to the zeroth-order terms in our drift-kinetic equation, but it is not clear how to apply it to the first-order (drift) terms, because the operator Λ defined in Sec. III appears twice in Eqs. (10) and (11). We have included a direct proof of phase-space conservation for the first-order terms, in Sec. VI, which is independent of the discussion in Ref. 10.

The following section presents our notational and geometrical conventions. The heart of the analysis is displayed in Sec. III, and the results given in Sec. IV. A discussion of the zero-gyroradius limit and its relation to MHD is presented in Sec. V. Phase-space conservation through first order in gyroradius is demonstrated in Sec. VI. Our conclusions are summarized in Sec. VII, which displays a relatively self-contained statement of the drift-kinetic equation.

II. GEOMETRY AND NOTATION

A. Kinetic equation and small gyroradius

The starting point for our analysis is the kinetic equation for the distribution function $f(\boldsymbol{x},\boldsymbol{v},t)$ of charged particles in a quasi-neutral, magnetized plasma. To conveniently treat rapidly moving plasma, this equation is written in a noninertial reference frame which approximates the rest frame of the distribution. Thus, while the first velocity moment of f will be relatively small, we allow the frame velocity U to be comparable to the thermal speed v_t associated with f. In this case "inertial" forces play an important role.

The kinetic equation in the moving frame has the well-known form[11]

$$\frac{\partial f}{\partial t} + (\boldsymbol{v}+\boldsymbol{U})\cdot\nabla f + \left(\frac{e}{m}\boldsymbol{v}\times\boldsymbol{B}+\boldsymbol{F}-\boldsymbol{v}\cdot\nabla\boldsymbol{U}\right)\cdot\frac{\partial f}{\partial \boldsymbol{v}} = C(f), \tag{1}$$

where e is the particle charge, m its mass, $\boldsymbol{B}$ is the magnetic field, C a collision operator and the force

$$\boldsymbol{F} = \frac{e}{m}(\boldsymbol{E}+\boldsymbol{U}\times\boldsymbol{B}) - \frac{\partial \boldsymbol{U}}{\partial t} - \boldsymbol{U}\cdot\nabla\boldsymbol{U} \tag{2}$$

includes the inertial forces along with the Lorentz force associated with plasma flow. The *drift*-kinetic equation applies to a magnetized plasma, in which the thermal gyroradius

$$\rho = \frac{v_t}{\Omega},$$

where $\Omega = eB/m$ is the gyrofrequency, is small compared to a gradient scale length L. The corresponding gyroradius parameter is denoted by

$$\delta \equiv \frac{\rho}{L}.$$

It is then convenient to introduce unit vectors $(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{b})$, with $\boldsymbol{b}\equiv\boldsymbol{B}/B=\boldsymbol{e}_1\times\boldsymbol{e}_2$, and to express the velocity coordinate as

$$\boldsymbol{v} = \boldsymbol{u}+\boldsymbol{s},$$

where $\boldsymbol{u}$ is the parallel velocity,

$$\boldsymbol{u} = \boldsymbol{b}\boldsymbol{b}\cdot\boldsymbol{v}$$

and $\boldsymbol{s}$ is the perpendicular velocity,

$$\boldsymbol{s} = s\hat{s} = s(\boldsymbol{e}_1\cos\alpha + \boldsymbol{e}_2\sin\alpha).$$

Here α is the gyrophase angle. It is helpful to decompose the distribution function into its gyrophase average, $\bar{f}$, and the residual $\tilde{f}$:

$$f(\boldsymbol{x},u,s,\alpha,t) = \bar{f}(\boldsymbol{x},u,s,t) + \tilde{f}(\boldsymbol{x},u,s,\alpha,t),$$

where

$$\bar{f} \equiv \langle f\rangle \equiv \oint \frac{d\alpha}{2\pi} f.$$

Rapid gyration implies that $\tilde{f}$ is relatively small:[7]

$$\bar{f}/\tilde{f} \sim \delta. \tag{3}$$

Returning to the kinetic equation, Eq. (1), we note that the last term on the left-hand side,

$$\frac{e}{m}(\boldsymbol{v}\times\boldsymbol{B})\cdot\frac{\partial f}{\partial \boldsymbol{v}} = -\Omega\frac{\partial f}{\partial \alpha},$$

is measured by the gyrofrequency. Since this term acts exclusively on $\tilde{f}$, our basic ordering, Eq. (3), will hold as long as the other terms on the right-hand side correspond to slower motions. For a precise statement of this condition, we introduce the operator

$$\Lambda = (\boldsymbol{v}+\boldsymbol{U})\cdot\nabla + (\boldsymbol{F}-\boldsymbol{v}\cdot\nabla\boldsymbol{U})\cdot\frac{\partial}{\partial \boldsymbol{v}}. \tag{4}$$

This operator will be expressed in the variables $(\boldsymbol{x},u,s,\alpha)$ in Sec. III. It is clear that Eq. (3) will describe evolution on time scales long compared to the gyroperiod, provided the collision frequency is smaller than the gyrofrequency and provided that

$$\Lambda(f) \sim \delta\Omega f. \tag{5}$$

Most of the terms in Λ obviously satisfy Eq. (5). The exceptions are two terms contained in the force $\boldsymbol{F}$:

$$\Lambda = \frac{e}{m}(\boldsymbol{E}+\boldsymbol{U}\times\boldsymbol{B})\cdot\frac{\partial}{\partial \boldsymbol{v}} + \mathcal{O}(\delta).$$

These terms will contradict Eq. (5) unless they nearly cancel—unless $\boldsymbol{E}$ balances the large $\boldsymbol{U}\times\boldsymbol{B}$ force. We therefore assume

$$\boldsymbol{E}+\boldsymbol{U}\times\boldsymbol{B} = \boldsymbol{b}E_\parallel, \tag{6}$$

where

$$\frac{e}{m}E_\parallel \sim \delta\Omega v_t. \tag{7}$$

Then Eq. (2) becomes

$$\boldsymbol{F} = \frac{e}{m}\boldsymbol{b}E_\parallel - \frac{\partial \boldsymbol{U}}{\partial t} - \boldsymbol{U}\cdot\nabla\boldsymbol{U}. \tag{8}$$

Furthermore, the velocity $\boldsymbol{U}$ must have the form

$$\boldsymbol{U} = \boldsymbol{b}U_\parallel + \boldsymbol{V}_E, \tag{9}$$

where

$$\boldsymbol{V}_E = \frac{(\boldsymbol{E}\times\boldsymbol{b})}{B}$$

is the familiar $E\times B$ drift, and the parallel flow $U_\parallel$ is arbitrary.

The drift-kinetic equation is the gyrophase-average of Eq. (1):

$$\frac{\partial \bar{f}}{\partial t} + \langle\Lambda(\bar{f}+\tilde{f})\rangle = C(\bar{f}). \tag{10}$$

Here $\tilde{f}$ is to be expressed in terms of $\bar{f}$, using the first-order equation

$$\Omega\frac{\partial \tilde{f}}{\partial \alpha} = \Lambda(\bar{f}) - \langle\Lambda(\bar{f})\rangle. \tag{11}$$

While this recursive procedure[14] evidently omits some second-order terms in $\tilde{f}$,[15] it captures in the drift-kinetic equation all first-order contributions, along with second-order terms coming from the first-order part of $\tilde{f}$. Equations including all second-order terms, with more specialized assumptions concerning plasma flow or collisionality, have also been derived.[15,16]

B. Geometry

Because the magnetic field is not uniform, the unit vectors $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{b})$ depend upon position. To describe this variation we will sometimes use the curvature coefficients R_{ij} defined by

$$\boldsymbol{e}_i \cdot (\nabla \boldsymbol{e}_j) \cdot \boldsymbol{e}_i = \frac{1}{R_{ij}}$$

and the torsion coefficients T_{ij} defined by

$$\boldsymbol{e}_k \cdot (\nabla \boldsymbol{e}_j) \cdot \boldsymbol{e}_i = \frac{1}{T_{ij}} = -\frac{1}{T_{ji}}.$$

Here k is the unique index unequal to i or j. Notice that we use the notation

$$\boldsymbol{e}_i \cdot (\nabla \boldsymbol{V}) \cdot \boldsymbol{e}_j \equiv [(\boldsymbol{e}_i \cdot \nabla)\boldsymbol{V}] \cdot \boldsymbol{e}_j.$$

We often abbreviate

$$\boldsymbol{b} \cdot \nabla = \nabla_{\parallel}.$$

The curvatures and torsions essentially measure components of the symmetric tensor $\boldsymbol{\mathcal{B}} \equiv \nabla \boldsymbol{b} + (\nabla \boldsymbol{b})^T$, with components

$$\mathcal{B}_{ij} \equiv \boldsymbol{e}_i \cdot (\nabla \boldsymbol{b}) \cdot \boldsymbol{e}_j + \boldsymbol{e}_j \cdot (\nabla \boldsymbol{b}) \cdot \boldsymbol{e}_i. \tag{12}$$

In particular,

$$\mathcal{B}_{12} = \frac{1}{T_{13}} + \frac{1}{T_{23}} \tag{13}$$

and

$$\frac{1}{2}(\mathcal{B}_{11} - \mathcal{B}_{22}) = \frac{1}{R_{13}} - \frac{1}{R_{23}}. \tag{14}$$

Our drift-kinetic equation will apply only for relatively slow temporal change,

$$\frac{\partial}{\partial t} \sim \delta^2 \Omega. \tag{15}$$

In that case temporal variation of the unit vectors is too weak to affect Eq. (10), and we treat the set $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{b})$ as constant in time.

This is the most relevant assumption about the time dependence for many applications, e.g., plasma transport. However, we temporarily relax this assumption, allowing faster time dependence, in Sec. V, where we discuss the zero gyroradius limit and demonstrate phase-space conservation for this limit.

C. Consistency condition

It follows from Eq. (15) that the electric field is primarily electrostatic:

$$\boldsymbol{E} \approx -\nabla \Phi,$$

where Φ is the potential. In this case a large perpendicular flow, $U_\perp \sim v_t$, can occur only if

$$\left(\frac{e}{T}\right)|\nabla \Phi| \sim \frac{1}{\rho}. \tag{16}$$

It is not obvious that the ordering Eq. (16) is consistent with the small-gyroradius assumption. We consider that consistency here.

We call the plasma magnetized if its parameters vary slowly on the scale of the gyroradius. The electrostatic potential must be counted as one of these parameters; in particular, its variation is typically linked to that of the plasma density.[11] But what does it mean to say that the potential varies slowly? The obvious condition

$$\rho|\nabla \Phi| \ll |\Phi| (?)$$

is not gauge-invariant and therefore meaningless.

A gauge-invariant measure involves the difference $\Delta\Phi$ between the potential's maximum and its minimum in the system. If $e|\Delta\Phi|/T$ is comparable to or less than unity, and if other plasma parameters vary sufficiently slowly, then the plasma is magnetized, and its velocity is small: $U_\perp \sim \delta v_t$. The present study emphasizes the case of rapid flow, $U_\perp \sim U_\parallel \sim v_t$. For electrostatic $E \times B$ motion on this scale we must allow

$$\frac{e|\Delta\Phi|}{T} \sim \frac{1}{\delta} \tag{17}$$

while at the same time imposing the gauge-invariant magnetization requirement,

$$\rho|\nabla \Phi| \sim \delta |\Delta\Phi|. \tag{18}$$

It is easily seen that Eqs. (17) and (18) combine to reproduce the rapid flow condition, Eq. (16). Hence they express the consistency conditions for an electrostatic $E \times B$ flow to approach thermal speed, in a magnetized plasma.

III. ANALYSIS

A. The operator Λ

To understand the structure of the operator Λ, we first isolate its averaged part, writing

$$\Lambda = \bar{\Lambda} + \widetilde{\Lambda},$$

where

$$\bar{\Lambda} = \dot{\boldsymbol{x}} \cdot \nabla \dot{u} \frac{\partial}{\partial u} + \dot{s} \frac{\partial}{\partial s}.$$

Here $\dot{\boldsymbol{x}} = \boldsymbol{U} + \boldsymbol{b}u$ and

$$\dot{u} = \frac{s^2}{2} \nabla \cdot \boldsymbol{b} + \boldsymbol{F} \cdot \boldsymbol{b} - u(\nabla_\parallel \boldsymbol{U}) \cdot \boldsymbol{b}, \tag{19}$$

$$\dot{s} = -\frac{us}{2} \nabla \cdot \boldsymbol{b} + \frac{s}{2}[(\nabla_\parallel \boldsymbol{U}) \cdot \boldsymbol{b} - \nabla \cdot \boldsymbol{U}]. \tag{20}$$

The residual $\widetilde{\Lambda}$ is expressed in terms of a scalar, J, a vector $\boldsymbol{K}$ and a second-rank tensor, $\boldsymbol{L}$:

$$\widetilde{\Lambda} = J\frac{\partial}{\partial\alpha} + \boldsymbol{s}\cdot\boldsymbol{K} - \widetilde{(\boldsymbol{ss})}:\boldsymbol{L}, \tag{21}$$

where

$$\widetilde{(\boldsymbol{ss})} = \frac{s^2}{2}[\cos 2\alpha(\boldsymbol{e}_1\boldsymbol{e}_1 - \boldsymbol{e}_2\boldsymbol{e}_2) + \sin 2\alpha(\boldsymbol{e}_1\boldsymbol{e}_2 + \boldsymbol{e}_2\boldsymbol{e}_1)]$$

and

$$J = (\boldsymbol{U} + \boldsymbol{b}u + \boldsymbol{s})\cdot\nabla\alpha + (\boldsymbol{F} - u\nabla_\parallel\boldsymbol{U} - \boldsymbol{s}\cdot\nabla\boldsymbol{U})\cdot\frac{\boldsymbol{b}\times\boldsymbol{s}}{s^2}, \tag{22}$$

$$\boldsymbol{K} = \nabla + (\boldsymbol{U} + \boldsymbol{b}u)\cdot\nabla\boldsymbol{b}\left(\frac{\partial}{\partial u} - \frac{u}{s}\frac{\partial}{\partial s}\right) - (\nabla\boldsymbol{U})\cdot\boldsymbol{b}\frac{\partial}{\partial u} + (\boldsymbol{F} - u\nabla_\parallel\boldsymbol{U})\frac{1}{s}\frac{\partial}{\partial s}, \tag{23}$$

and

$$L_{ij} = \frac{1}{2s}\left[(\mathcal{W}_{ij} + u\mathcal{B}_{ij})\frac{\partial}{\partial s} - s\mathcal{B}_{ij}\frac{\partial}{\partial u}\right]. \tag{24}$$

Here $\mathcal{B}_{ij}$ is defined by Eq. (12) and, similarly,

$$\mathcal{W}_{ij} \equiv \boldsymbol{e}_i\cdot(\nabla\boldsymbol{U})\cdot\boldsymbol{e}_j + \boldsymbol{e}_j\cdot(\nabla\boldsymbol{U})\cdot\boldsymbol{e}_i.$$

All gradients are performed at fixed u, s, α except $\nabla\alpha$, which is performed at fixed $\boldsymbol{v}\equiv\boldsymbol{b}u+\boldsymbol{s}$. Now the drift-kinetic equation can be written as

$$\frac{\partial\bar{f}}{\partial t} + \overline{\Lambda f} + \left\langle J\frac{\partial\tilde{f}}{\partial\alpha}\right\rangle + \langle\boldsymbol{s}\cdot\boldsymbol{K}(\tilde{f})\rangle - \langle\boldsymbol{ss}:\boldsymbol{L}(\tilde{f})\rangle = C(\bar{f}). \tag{25}$$

Notice that we use parentheses to indicate the operand of each operator.

We refer to the last three terms on the left-hand side of Eq. (26) as the scalar, vector, and tensor terms, in the obvious order. We evaluate these terms after finding the distribution $\tilde{f}$.

B. Gyro-varying distribution

The function $\tilde{f}$ is obtained from Eq. (11), or,

$$\Omega\frac{\partial\tilde{f}}{\partial\alpha} = \widetilde{\Lambda}(\bar{f}) = \boldsymbol{s}\cdot\boldsymbol{K}(\bar{f}) - \widetilde{(\boldsymbol{ss})}:\boldsymbol{L}(\bar{f}).$$

Here we used Eq. (21). The integration constant is determined by requiring $\langle\tilde{f}\rangle=0$. Thus we find

$$\tilde{f} = -\boldsymbol{\rho}\cdot\boldsymbol{K}(\bar{f}) - \frac{s^2}{4\Omega}\hat{\eta}:\boldsymbol{L}(\bar{f}), \tag{26}$$

where

$$\boldsymbol{\rho} = \frac{s}{\Omega}(-\boldsymbol{e}_1\sin\alpha + \boldsymbol{e}_2\cos\alpha) \tag{27}$$

and

$$\hat{\eta} = \sin 2\alpha(\boldsymbol{e}_1\boldsymbol{e}_1 - \boldsymbol{e}_2\boldsymbol{e}_2) - \cos 2\alpha(\boldsymbol{e}_1\boldsymbol{e}_2 + \boldsymbol{e}_2\boldsymbol{e}_1). \tag{28}$$

The expression, Eq. (26), agrees with the previously published version.[7]

C. Scalar term

Straightforward calculation shows that

$$\nabla\alpha = (\boldsymbol{e}_1\cdot\nabla)\boldsymbol{e}_2 + \boldsymbol{e}_1\times\nabla\times\boldsymbol{e}_2 + \frac{u}{s}[\cos\alpha(\boldsymbol{b}\cdot\nabla\boldsymbol{e}_2 + \boldsymbol{b}\times\nabla\times\boldsymbol{e}_2) - \sin\alpha(\boldsymbol{b}\cdot\nabla\boldsymbol{e}_1 + \boldsymbol{b}\times\nabla\times\boldsymbol{e}_1)]. \tag{29}$$

We use this result to compute

$$\left\langle J\frac{\partial\tilde{f}}{\partial\alpha}\right\rangle = \boldsymbol{Z}\cdot\boldsymbol{K}(\bar{f}) + \frac{s^2}{16\Omega}(\mathcal{B},\mathcal{W})\frac{\partial\bar{f}}{\partial u} \tag{30}$$

in terms of the abbreviations

$$(\mathcal{B},\mathcal{W}) \equiv \mathcal{B}_{12}(\mathcal{W}_{11} - \mathcal{W}_{22}) - \mathcal{W}_{12}(\mathcal{B}_{11} - \mathcal{B}_{22}) \tag{31}$$

and

$$\boldsymbol{Z} \equiv -\frac{s^2}{2\Omega}\left(\frac{\boldsymbol{e}_2}{R_{21}} - \frac{\boldsymbol{e}_1}{R_{12}}\right) - \frac{\boldsymbol{b}}{2\Omega}\times(\boldsymbol{F} - u^2\boldsymbol{\kappa} - u\boldsymbol{\nu}), \tag{32}$$

where $\boldsymbol{\kappa}=\boldsymbol{b}\cdot\nabla\boldsymbol{b}$ and

$$\boldsymbol{\nu} \equiv \boldsymbol{U}\cdot\nabla\boldsymbol{b} + \boldsymbol{b}\cdot\nabla\boldsymbol{U}. \tag{33}$$

Note that the s derivatives have canceled in the second term on the right-hand side of Eq. (30).

The terms in Eq. (31) involving Z will combine with terms from the vector contribution, computed next, to give the correct guiding-center drift; the remaining terms combine with identical terms from the tensor contribution to give the correct $\dot{u}$.

D. Vector term

Next consider

$$\langle\boldsymbol{s}\cdot\boldsymbol{K}(\tilde{f})\rangle = -\langle s_i K_i(\rho_j K_j(\bar{f}))\rangle = -\langle s_i K_i(\rho_j)\rangle K_j(\bar{f}) - \langle s_i\rho_j\rangle K_i(K_j)\bar{f}. \tag{34}$$

It is convenient to write

$$\boldsymbol{K} = \nabla + \boldsymbol{K}^u\frac{\partial}{\partial u} + \boldsymbol{K}^s\frac{\partial}{\partial s}$$

with

$$\boldsymbol{K}^u = (\boldsymbol{U} + \boldsymbol{b}u)\cdot\nabla\boldsymbol{b} - (\nabla\boldsymbol{U})\cdot\boldsymbol{b}, \tag{35}$$

$$\boldsymbol{K}^s = s^{-1}[\boldsymbol{F} - u^2\boldsymbol{\kappa} - u\boldsymbol{\nu}]. \tag{36}$$

Then one finds that

$$\langle\boldsymbol{s}\cdot\boldsymbol{K}(\tilde{f})\rangle = \boldsymbol{v}_{*D}\cdot\nabla\bar{f} + K^{*u}\frac{\partial\bar{f}}{\partial u} + K^{*s}\frac{\partial\bar{f}}{\partial s}, \tag{37}$$

where

$$\boldsymbol{v}_{*D} \equiv \boldsymbol{u}_D + \frac{s^2}{2\Omega}\left[\boldsymbol{b}\times\nabla\log B + \left(\frac{\boldsymbol{e}_2}{R_{21}} - \frac{\boldsymbol{e}_1}{R_{12}}\right)\right] - \frac{\boldsymbol{b}}{2\Omega}(\boldsymbol{F} - u^2\boldsymbol{\kappa} - u\boldsymbol{\nu}) \quad (38)$$

and

$$K^{*u} = \boldsymbol{v}_{*D}\cdot\boldsymbol{K}^u - \frac{s^2}{2\Omega}\boldsymbol{b}\cdot(\nabla\times\boldsymbol{K}^u - \boldsymbol{\kappa}\times\boldsymbol{K}^u), \quad (39)$$

$$K^{*s} = \boldsymbol{v}_{*D}\cdot\boldsymbol{K}^s - \frac{s^2}{2\Omega}\boldsymbol{b}\cdot[\nabla\times\boldsymbol{K}^s - s^{-1}\boldsymbol{K}^u\times(u\boldsymbol{\kappa}+\boldsymbol{b}\cdot\boldsymbol{\mathcal{W}})]. \quad (40)$$

In Eq. (38)

$$\boldsymbol{u}_D \equiv \frac{s^2}{2\Omega}\boldsymbol{b}\boldsymbol{b}\cdot\nabla\times\boldsymbol{b}$$

is the usual parallel drift.

E. Tensor term

Because the tensor operator $\boldsymbol{L}$ acts only on u and s, and because the α-varying part of $\boldsymbol{ss}$ contains only $\cos 2\alpha$ and $\sin 2\alpha$ terms, we have

$$-\langle\boldsymbol{ss}:\boldsymbol{L}(\tilde{f})\rangle = \frac{s^2}{4\Omega}\langle\hat{s}_i\hat{s}_j\hat{\eta}_{kl}\rangle L_{ij}(s^2L_{kl}\bar{f}). \quad (41)$$

Here we have noted that the tensor $\hat{E}\equiv 4\langle\hat{s}_i\hat{s}_j\hat{\eta}_{kl}\rangle$ is antisymmetric under $(i,j)\leftrightarrow(k,l)$. It is given by

$$\hat{E} = (1211) - (1222) + (2111) - (2122) - (1112) - (1121) + (2212) + (2221),$$

where, for example, $(1211)\equiv\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_1\boldsymbol{e}_1$.

The only nonvanishing contribution to Eq. (41) comes from

$$L_{ij}(s^2) = \mathcal{W}_{ij} + u\mathcal{B}_{ij}.$$

Since, furthermore,

$$L_{ij}(s^2)L_{kl} = -\frac{1}{2}\mathcal{W}_{ij}\mathcal{B}_{kl},$$

we have

$$-\langle\boldsymbol{ss}:\boldsymbol{L}(\tilde{f})\rangle = -\frac{s^2}{32\Omega}\hat{E}_{ijkl}\mathcal{W}_{ij}\mathcal{B}_{kl}$$

or

$$-\langle\boldsymbol{ss}:\boldsymbol{L}(\tilde{f})\rangle = \frac{s^2}{16\Omega}(\mathcal{B},\mathcal{W})\frac{\partial\bar{f}}{\partial u}. \quad (42)$$

IV. DRIFT-KINETIC COEFFICIENTS

A. Drift velocity and accelerations

We use the notation

$$\langle\widetilde{\Lambda}(\bar{f})\rangle = \boldsymbol{v}_D\cdot\nabla\bar{f} + \dot{u}_D\frac{\partial\bar{f}}{\partial u} + \dot{s}_D\frac{\partial\bar{f}}{\partial s}.$$

The coefficients are found by combining Eqs. (30), (37), and (42). We find that

(1) The guiding-center velocity $\boldsymbol{v}_D$ has contributions from Eqs. (30) and (37), and is given by $\boldsymbol{v}_D=\boldsymbol{v}_{*D}+\boldsymbol{Z}$ or

$$\boldsymbol{v}_D \equiv \boldsymbol{u}_D + \frac{s^2}{2\Omega}\boldsymbol{b}\times\nabla\log B - \frac{\boldsymbol{b}}{\Omega}\times(\boldsymbol{F} - u^2\boldsymbol{\kappa} - u\boldsymbol{\nu}). \quad (43)$$

Note that the $E\times B$ drift is contained in $\boldsymbol{U}$, allowing it to become large; this drift is therefore missing from $\boldsymbol{v}_D$.

(2) The guiding-center parallel acceleration has contributions from all three parts of $\widetilde{\Lambda}$, and is given by

$$\dot{u}_D = \boldsymbol{v}_D\cdot\boldsymbol{K}^u - \frac{s^2}{2\Omega}\boldsymbol{b}\cdot(\nabla\times\boldsymbol{K}^u - \boldsymbol{\kappa}\times\boldsymbol{K}^u) + \frac{s^2}{8\Omega}(\mathcal{B},\mathcal{W}). \quad (44)$$

Here $\boldsymbol{K}^u$ is defined by Eq. (35).

(3) The guiding-center perpendicular acceleration is given by

$$\dot{s}_D = \boldsymbol{v}_D\cdot\boldsymbol{K}^s - \frac{s^2}{2\Omega}\boldsymbol{b}\cdot[\nabla\times\boldsymbol{K}^s - s^{-1}\boldsymbol{K}^u\times(u\boldsymbol{\kappa}+\boldsymbol{b}\cdot\boldsymbol{\mathcal{W}})]. \quad (45)$$

The above expression for $\langle\widetilde{\Lambda}(\bar{f})\rangle$ is to be used in Eq. (10), along with the expressions given at the beginning of Sec. III, to form the complete drift-kinetic equation in the variables (u,s).

B. Guiding-center energy change

The particle kinetic energy is defined by $w=v^2/2=(u^2+s^2)/2$. We compute $\dot{w}_D=u\dot{u}_D+s\dot{s}_D$ from Eqs. (44) and (45), noticing that

$$u\boldsymbol{K}^u + s\boldsymbol{K}^s = \boldsymbol{F} - u\boldsymbol{b}\cdot\boldsymbol{\mathcal{W}}. \quad (46)$$

We also introduce the conventional[17] measure of velocity shear:

$$\boldsymbol{W} \equiv \tfrac{1}{2}\left[\nabla\boldsymbol{U} + (\nabla\boldsymbol{U})^T - \tfrac{2}{3}\boldsymbol{I}\nabla\cdot\boldsymbol{U}\right].$$

Thus

$$\boldsymbol{\mathcal{W}} = \tfrac{2}{3}\boldsymbol{I}\nabla\cdot\boldsymbol{U} + 2\boldsymbol{W} \quad (47)$$

and we find

$$\dot{w}_D = \boldsymbol{v}_D\cdot(\boldsymbol{F} - 2u\boldsymbol{b}\cdot\boldsymbol{W}) - \frac{s^2}{2\Omega}\boldsymbol{b}\cdot\nabla\times(\boldsymbol{F} - 2u\boldsymbol{b}\cdot\boldsymbol{W}) + \frac{s^2}{\Omega}\boldsymbol{b}\cdot\boldsymbol{K}^u\times(\boldsymbol{b}\cdot\boldsymbol{W}) + \frac{us^2}{4\Omega}(\mathcal{B},W). \quad (48)$$

To make Eq. (48) more explicit, we first observe that

$$(\mathcal{B},W) = \left(\frac{1}{T_{13}} + \frac{1}{T_{23}}\right)(W_{11} - W_{22}) - 2W_{12}\left(\frac{1}{R_{13}} - \frac{1}{R_{23}}\right).$$

We also note that

$$\boldsymbol{b}\cdot\boldsymbol{K}^u\times(\boldsymbol{b}\cdot\boldsymbol{W})=\boldsymbol{b}\times\boldsymbol{\nu}\cdot(\boldsymbol{b}\cdot\boldsymbol{W})+(\boldsymbol{b}\times u\boldsymbol{\kappa})\cdot(\boldsymbol{b}\cdot\boldsymbol{W})$$

where the last term can be written explicitly as

$$(\boldsymbol{b}\times u\boldsymbol{\kappa})\cdot(\boldsymbol{b}\cdot\boldsymbol{W})=u\left(\frac{W_{13}}{R_{32}}-\frac{W_{23}}{R_{31}}\right).$$

Combining these observations we find

$$\dot{w}_D=\dot{w}_1+\dot{w}_2+\dot{w}_3+\dot{w}_4 \tag{49}$$

with

$$\begin{aligned}
\dot{w}_1&=\boldsymbol{v}_D\cdot(\boldsymbol{F}-2u\boldsymbol{b}\cdot\boldsymbol{W})-\frac{s^2}{2\Omega}\boldsymbol{b}\cdot\nabla\times(\boldsymbol{F}-2u\boldsymbol{b}\cdot\boldsymbol{W}),\\
\dot{w}_2&=\frac{2\mu B}{\Omega}\left[u\left(\frac{W_{13}}{R_{32}}-\frac{W_{23}}{R_{31}}\right)+\boldsymbol{b}\times\boldsymbol{\nu}\cdot(\boldsymbol{b}\cdot\boldsymbol{W})\right],\\
\dot{w}_3&=\frac{u\mu B}{2\Omega}\left(\frac{1}{T_{13}}+\frac{1}{T_{23}}\right)(W_{11}-W_{22}),\\
\dot{w}_4&=\frac{u\mu B}{\Omega}W_{12}\left(\frac{1}{R_{23}}-\frac{1}{R_{13}}\right).
\end{aligned} \tag{50}$$

Notice that the curl in Eq. (50) is performed at fixed u. This expression for $\dot{w}_D$ is easily compared to previous literature,[7] with which it agrees. An alternative, somewhat simpler expression is derived presently and given by Eq. (64).

C. Guiding-center magnetic moment change

The rate of change of magnetic moment $\mu\equiv s^2/(2B)$ is

$$\dot{\mu}=B^{-1}(s\dot{s}-\mu\dot{\boldsymbol{x}}\cdot\nabla B).$$

We show in Sec. V that the lowest-order terms in this expression, coming from $\dot{\bar{s}}$ and $\dot{\bar{\boldsymbol{x}}}=\boldsymbol{U}+\boldsymbol{b}u$, nearly cancel, essentially because of Eq. (9). Here we consider only the higher-order terms,

$$\dot{\mu}_D\equiv B^{-1}(s\dot{s}_D-\mu\boldsymbol{v}_D\cdot\nabla B) \tag{51}$$

without using Eq. (9). From Eq. (45) we have

$$B\dot{\mu}_D=\boldsymbol{v}_D\cdot(s\boldsymbol{K}^s-\mu\nabla B)-\frac{\mu B}{\Omega}\boldsymbol{b}\cdot[\nabla\times s\boldsymbol{K}^s-\boldsymbol{K}^u\times(u\boldsymbol{\kappa}+\boldsymbol{b}\cdot\boldsymbol{\mathcal{W}})]. \tag{52}$$

Recalling Eq. (36) and taking advantage of several cancellations, we find that Eq. (52) can be written as

$$\dot{\mu}_D=\dot{\mu}_1+\dot{\mu}_2+\dot{\mu}_3 \tag{53}$$

with

$$B\dot{\mu}_1=u_D(F_\parallel-u\boldsymbol{b}\cdot\nabla_\parallel-\mu\nabla_\parallel B), \tag{54}$$

$$B\dot{\mu}_2=-\frac{\mu B}{\Omega}\boldsymbol{b}\cdot\nabla\times(\boldsymbol{F}-u^2\boldsymbol{\kappa}-u\boldsymbol{\nu}), \tag{55}$$

$$B\dot{\mu}_3=\frac{\mu B}{\Omega}\boldsymbol{b}\cdot[u(\boldsymbol{\nu}-4\boldsymbol{b}\cdot\boldsymbol{W})\times\boldsymbol{\kappa}+2\boldsymbol{\nu}\times\boldsymbol{b}\cdot\boldsymbol{W}]. \tag{56}$$

In Eq. (55), as in Eq. (50), the curl is performed at fixed u.

D. Summary

The drift-kinetic equation can be written as

$$\frac{\partial f}{\partial t}+(\boldsymbol{u}+\boldsymbol{U}+\boldsymbol{v}_D)\cdot\nabla f+\dot{w}\frac{\partial f}{\partial w}+\dot{\mu}\frac{\partial f}{\partial\mu}=C(f), \tag{57}$$

where we suppress the overbar on f for simplicity, and where

$$\dot{w}=\dot{\bar{w}}+\dot{w}_D, \tag{58}$$

$$\dot{\mu}=\dot{\bar{\mu}}+\dot{\mu}_D. \tag{59}$$

Here $\dot{w}_D$ and $\dot{\mu}_D$ are given by Eqs. (49) and (53), respectively, while the leading terms are

$$\dot{\bar{w}}=u\dot{\bar{u}}+s\dot{\bar{s}}, \tag{60}$$

$$\dot{\bar{\mu}}=B^{-1}s\dot{\bar{s}}-\mu(\boldsymbol{b}u+\boldsymbol{U})\cdot\nabla\log B. \tag{61}$$

These barred terms are the only terms that survive in the limit of vanishing gyroradius, $\delta\to 0$. They are considered in more detail in Sec. V.

E. Symmetry

Since the choice of the unit vectors $\boldsymbol{e}_1(\boldsymbol{x})$ and $\boldsymbol{e}_2(\boldsymbol{x})$ is constrained only by $\boldsymbol{e}_1\times\boldsymbol{e}_2=\boldsymbol{b}$, the kinetic equation must be invariant under rotation of the coordinate system about the local field $\boldsymbol{b}(\boldsymbol{x})$. (Such rotation is equivalent to redefining the gyrophase at each point, and is closely analogous to local gauge symmetry in quantum field theory.) A simple way to verify this "gyrosymmetry" is to begin with Eq. (48), in which every term is clearly symmetric except the last, involving $(\mathcal{B},W)$. However, this quantity can be written as

$$(\boldsymbol{\mathcal{B}},W)=\tfrac{1}{2}[(\boldsymbol{\mathcal{B}}:\boldsymbol{\tau})(W:\boldsymbol{\sigma})-(\boldsymbol{\mathcal{B}}:\boldsymbol{\sigma})(W:\boldsymbol{\tau})],$$

where $\boldsymbol{\sigma}=\boldsymbol{e}_1\boldsymbol{e}_1-\boldsymbol{e}_2\boldsymbol{e}_2$ and $\boldsymbol{\tau}=\boldsymbol{e}_1\boldsymbol{e}_2+\boldsymbol{e}_2\boldsymbol{e}_1$. Therefore $(\mathcal{B},W)$ is symmetric if the tensor $\boldsymbol{\tau}\boldsymbol{\sigma}-\boldsymbol{\sigma}\boldsymbol{\tau}$ is gyrosymmetric; but this gyrosymmetry is easily demonstrated by direct substitution, $\alpha\to\alpha+\varphi$.

A *manifestly* symmetric version of the kinetic equation is obtained by expressing Eq. (57) in terms of $\boldsymbol{b}$ alone, without reference to $\boldsymbol{e}_1$ or $\boldsymbol{e}_2$. This version is in fact relatively simple, so we derive it here.

The only term in Eq. (57) that involves the $\boldsymbol{e}$'s is the coefficient of $u\mu B/\Omega$ in $\dot{w}_D$; we denote this quantity by

$$X\equiv 2\left(\frac{W_{13}}{R_{32}}-\frac{W_{23}}{R_{31}}\right)+\frac{1}{2}\left(\frac{1}{T_{13}}+\frac{1}{T_{23}}\right)(W_{11}-W_{22})+W_{12}\left(\frac{1}{R_{23}}-\frac{1}{R_{13}}\right).$$

It can be seen that

$$\frac{1}{R_{3i}}=-\kappa_i,$$

where $\boldsymbol{\kappa}\equiv\boldsymbol{b}\cdot\nabla\boldsymbol{b}$ is the magnetic field curvature. Combining this result with Eqs. (13) and (14) we obtain

$$X = 2(\kappa_1 W_{23} - \kappa_2 W_{13}) + \tfrac{1}{2}[(\mathcal{B}_{12}(W_{11} - W_{22}) - W_{12}(\mathcal{B}_{11} - \mathcal{B}_{22})\ ]. \tag{62}$$

Next consider the quantity

$$Y \equiv b_i \epsilon_{ijk} W_{jm} \mathcal{B}_{mk},$$

where ϵ_{ijk} is the unit antisymmetric tensor. Because W_{jm} and $\mathcal{B}_{mk}$ are components of second-rank tensors, Y is manifestly gyrosymmetric (in variant under rotation in the transverse plane). Straightforward expansion reveals that

$$Y = \mathcal{B}_{12}(W_{11} - W_{22}) - W_{12}(\mathcal{B}_{11} - \mathcal{B}_{22}) + \kappa_2 W_{13} - \kappa_1 W_{23}.$$

Thus

$$X = \tfrac{1}{2}Y - \tfrac{5}{2}(\kappa_2 W_{13} - \kappa_1 W_{23})$$

or

$$X = \tfrac{1}{2} b_i e_{ijk} W_{jm} \mathcal{B}_{mk} - \tfrac{5}{2} \boldsymbol{b} \cdot \boldsymbol{\kappa} \times (\boldsymbol{W} \cdot \boldsymbol{b}). \tag{63}$$

We conclude that Eq. (49) can be expressed in gyrosymmetric form as

$$\dot{w}_D = \dot{w}_1 + \frac{2\mu B}{\Omega} \boldsymbol{b} \times \boldsymbol{\nu} \cdot (\boldsymbol{W} \cdot \boldsymbol{b}) + \frac{u\mu B}{\Omega} X, \tag{64}$$

where $\dot{w}_1$ is given by Eq. (50) and X by Eq. (63).

V. ZERO GYRORADIUS LIMIT

A. Magnetic moment change

We compute the lowest-order magnetic moment variation $\dot{\bar{\mu}}$ from Eqs. (61) and (20). After noting that

$$\nabla \cdot \boldsymbol{b} = -\nabla_\parallel \log B$$

we find that

$$\dot{\bar{\mu}} = \mu[(\nabla_\parallel \boldsymbol{U}) \cdot \boldsymbol{b} - \nabla \cdot \boldsymbol{U} - \boldsymbol{U} \cdot \nabla \log B].$$

The identities

$$(\nabla_\parallel \boldsymbol{U}) \cdot \boldsymbol{b} = \nabla_\parallel U_\parallel - \boldsymbol{\kappa} \cdot \boldsymbol{U}, \tag{65}$$

$$\nabla \cdot \boldsymbol{U} = \nabla_\parallel U_\parallel + \nabla \cdot \boldsymbol{V}_E - U_\parallel \nabla_\parallel \log B, \tag{66}$$

where $\boldsymbol{\kappa} = \nabla_\parallel \boldsymbol{b}$ is the magnetic curvature, then provide

$$\dot{\bar{\mu}} = -\mu(\nabla \cdot \boldsymbol{V}_E + \boldsymbol{\kappa} \cdot \boldsymbol{V}_E + \boldsymbol{V}_E \cdot \nabla \log B). \tag{67}$$

We next use Faraday's law to compute

$$\nabla \cdot \boldsymbol{V}_E = -\boldsymbol{V}_E \cdot \nabla \log B - \boldsymbol{\kappa} \cdot \boldsymbol{V}_E - \frac{E_\parallel}{B} \boldsymbol{b} \cdot \nabla \times \boldsymbol{b}. \tag{68}$$

Here we consistently neglect the term $(\mu/B)\partial B/\partial t$; however, had we kept this term it would precisely cancel with the time-derivative term omitted from Eq. (61). Thus Eq. (67) reduces to

$$\dot{\bar{\mu}} = \mu\left(\frac{E_\parallel}{B}\right) \boldsymbol{b} \cdot \nabla \times \boldsymbol{b}. \tag{69}$$

This quantity is $\mathcal{O}(\delta)$, in view of Eq. (7): there is no zeroth-order magnetic moment change.

B. Energy change

We compute the lowest-order energy change from Eqs. (19), (20), and (60):

$$\dot{\bar{w}} = u\boldsymbol{b} \cdot (\boldsymbol{F} - u\nabla_\parallel \boldsymbol{U}) + \mu B[(\nabla_\parallel \boldsymbol{U}) \cdot \boldsymbol{b} - \nabla \cdot \boldsymbol{U}]. \tag{70}$$

The last term in this expression brings in a first-order contribution, given by the last term in Eq. (68); all the other terms are nominally $\mathcal{O}(1)$. The small term is consistently omitted from the lowest-order energy change, which we denote by $\dot{\bar{w}}_0$. Using Eqs. (65), (66), and (68), we find that

$$\dot{\bar{w}}_0 = u\left[\frac{e}{m}E_\parallel - \boldsymbol{U} \cdot (\nabla \boldsymbol{U}) \cdot \boldsymbol{b}\right] + \mu B \boldsymbol{U} \cdot \nabla \log B - u^2(\nabla_\parallel \boldsymbol{U}) \cdot \boldsymbol{b}. \tag{71}$$

The lowest-order drift-kinetic equation is therefore

$$(\boldsymbol{b}u + \boldsymbol{U}) \cdot \nabla f_0 + \dot{\bar{w}}_0 \frac{\partial f_0}{\partial w} = C(f_0), \tag{72}$$

where f_0 is the lowest-order distribution and $\dot{\bar{w}}_0$ is given by Eq. (71).

C. Phase-space conservation

In this subsection we relax the time-ordering Eq. (15), allowing temporal variation in zeroth order. Provided that Eq. (70) is used for $\dot{\bar{w}}$, the only change in the lowest-order kinetic equation is the addition of an obvious term to Eq. (72):

$$\frac{\partial f}{\partial t} + (\boldsymbol{b}u + \boldsymbol{U}) \cdot \nabla f + \dot{\bar{w}} \frac{\partial f}{\partial w} = 0. \tag{73}$$

We also omit collisions, for simplicity, and suppress 0 subscripts. The kinetic equation, Eq. (73), is essentially the same as that of Rosenbluth and Rostoker (RR):[13] it is part of the foundation of "kinetic MHD." Thus one can view the present analysis as an extension of the RR theory to include terms of first order in the gyroradius.

We let $\boldsymbol{\xi} = (\xi^1, \ldots, \xi^6)$ denote arbitrary curvilinear coordinates for the six-dimensional phase space; in our case $\boldsymbol{\xi} = (\boldsymbol{x}, \mu, w, \alpha)$. Thus the total number of particles is

$$N \equiv \int d^3x d^3v f = \int d^6\xi J f,$$

where J denotes the Jacobian for the six-dimensional coordinate transformation $(\boldsymbol{x}, \boldsymbol{v}) \to \boldsymbol{\xi}$. Any kinetic equation expresses the conservation of N; if each ξ^i changes according to some specified $\dot{\xi}^i(\boldsymbol{\xi})$, then the natural, particle-conserving form of the kinetic equation is

$$\frac{1}{J}\frac{\partial (Jf)}{\partial t} + \frac{1}{J}\frac{\partial}{\partial \xi^i}(\dot{\xi}^i J f) = 0. \tag{74}$$

The kinetic equation, Eq. (72), has a different form:

$$\frac{\partial f}{\partial t} + \dot{\xi}^i \frac{\partial f}{\partial \xi^i} = 0. \tag{75}$$

However, if the rates of change $\dot{\xi}^i$ satisfy the phase-space conservation law,

$$\frac{\partial J}{\partial t} + \frac{\partial}{\partial \xi^i}(J\dot{\xi}^i) = 0, \tag{76}$$

then the two forms of the kinetic equation, Eqs. (74) and (75), are equivalent and N is manifestly conserved.

To study the conservation properties of the RR equation, Eq. (73), we first note that the variables

$$\boldsymbol{\xi} = (\boldsymbol{x}, \mu, w, \alpha)$$

have the Jacobian

$$J = \frac{B}{u}$$

and the dynamical laws

$$\dot{\boldsymbol{x}} = \boldsymbol{b}u + \boldsymbol{U}, \quad \dot{\mu} = 0, \quad \dot{\alpha} = 0,$$
$$\dot{w} = -u^2 \nabla_\parallel \boldsymbol{U} \cdot \boldsymbol{b} + \mu B[(\nabla_\parallel \boldsymbol{U}) \cdot \boldsymbol{b} - \nabla \cdot \boldsymbol{U}]. \tag{77}$$

In Eq. (77) we omit a term proportional to u, since it cannot contribute to Eq. (76).

Hence Eq. (76) takes the form

$$\frac{\partial}{\partial t}\left(\frac{B}{u}\right) + \nabla \cdot \left(\frac{B}{u}\boldsymbol{U}\right) + \frac{\partial}{\partial w}\left(\frac{B}{u}\dot{w}\right) = 0. \tag{78}$$

Note that the velocity term $(B/u)u\boldsymbol{b} = \boldsymbol{B}$ is obviously divergence free.

To verify Eq. (78), we first use

$$u = \sqrt{2(w - \mu B)}$$

to compute

$$\frac{\partial}{\partial t}\left(\frac{B}{u}\right) + \nabla \cdot \left(\frac{B}{u}\boldsymbol{U}\right) = \frac{u^2 + \mu B}{u^3}\frac{dB}{dt} + \frac{B}{u}\nabla \cdot \boldsymbol{U},$$

where

$$\frac{dB}{dt} \equiv \frac{\partial B}{\partial t} + \boldsymbol{U} \cdot \nabla B.$$

Faraday's law, using Eq. (9) for $\boldsymbol{U}$, can be used to show that

$$\frac{dB}{dt} = B[(\nabla_\parallel \boldsymbol{U}) \cdot \boldsymbol{b} - \nabla \cdot \boldsymbol{U}] + \mathcal{O}(\delta)$$

and therefore that

$$\frac{\partial}{dt}\left(\frac{B}{u}\right) + \nabla \cdot \left(\frac{B}{u}\boldsymbol{U}\right) = \frac{u^2 + \mu B}{u^3} B[(\nabla_\parallel \boldsymbol{U}) \cdot \boldsymbol{b} - \nabla \cdot \boldsymbol{U}] + \frac{B}{u}\nabla \cdot \boldsymbol{U}. \tag{79}$$

Finally, we use the identity $\partial u/\partial w = 1/u$ to find that the last term in Eq. (78) precisely cancels the right-hand side of Eq. (79).

Hence the RR equation conserves phase space. We restore the collision operator to express the equation as

$$\frac{\partial}{\partial t}\left(\frac{Bf}{u}\right) + \nabla \cdot \left[\frac{Bf}{u}(\boldsymbol{b}u + \boldsymbol{U})\right] + \frac{\partial}{\partial w}\left(\frac{Bf}{u}\dot{w}\right) = \frac{B}{u}C(f). \tag{80}$$

This form is especially convenient for computing moments of the RR equation.

D. Lowest-order distribution: Maxwellian case

We consider the lowest-order theory in the special case of a Maxwellian f_0, as must occur in toroidal geometry without external sources. This case has been studied previously;[7,11] our purpose here is to see how the familiar results are obtained from the present formulation.

When

$$f_0 = (2\pi T)^{-3/2} n(\boldsymbol{x}) e^{-mw/T(\boldsymbol{x})}$$

we find that Eq. (72) implies

$$(\boldsymbol{b}u + \boldsymbol{U}) \cdot \left(\nabla \log n - \frac{3}{2}\nabla \log T + \frac{mw}{T}\nabla \log T\right) + \frac{mu}{T}\left(\frac{e}{m}\nabla_\parallel \Phi + \boldsymbol{U} \cdot \nabla \boldsymbol{U} \cdot \boldsymbol{b}\right) - \frac{m\mu B}{T}\boldsymbol{U} \cdot \nabla \log B + \frac{mu^2}{T}\boldsymbol{b} \cdot \nabla \boldsymbol{U} \cdot \boldsymbol{b} = 0.$$

Noting that this relation must hold for all μ and w, we infer that

$$\nabla_\parallel \log T = 0,$$

$$\boldsymbol{U} \cdot \left(\nabla \log n - \tfrac{3}{2}\nabla \log T\right) = 0,$$

$$\boldsymbol{U} \cdot \nabla \log T - \boldsymbol{U} \cdot \nabla \log B = 0,$$

$$\nabla_\parallel \log n + \frac{e}{T}\nabla_\parallel \Phi + \frac{m}{T}\boldsymbol{U} \cdot \nabla \boldsymbol{U} \cdot \boldsymbol{b} = 0,$$

$$\boldsymbol{b} \cdot \nabla \boldsymbol{U} \cdot \boldsymbol{b} + \tfrac{1}{2}\boldsymbol{U} \cdot \nabla \log B = 0.$$

If we next specialize to axisymmetric (tokamak) geometry, then these relations quickly yield the familiar[11] rigid-body toroidal rotation, with density and potential variation modified by centrifugal force.

VI. PHASE-SPACE CONSERVATION

We now demonstrate phase-space conservation for our drift-kinetic equation including the first-order (drift) terms. This is simplest using the variables (u, s), for which the Jacobian is s. The zeroth-order coefficients are given by $\dot{\bar{\boldsymbol{x}}} = \boldsymbol{U} + \boldsymbol{b}u$ and Eqs. (19) and (20) for $\dot{\bar{u}}$ and $\dot{\bar{s}}$. We find easily

$$\nabla \cdot \dot{\bar{\boldsymbol{x}}} + \frac{\partial \dot{\bar{u}}}{\partial u} + \frac{1}{s}\frac{\partial}{\partial s}(s\dot{\bar{s}}) = 0. \tag{81}$$

The first-order coefficients are given by Eqs. (43), (44), and (45). We write Eq. (43) in the form

$$\boldsymbol{v}_D = \boldsymbol{u}_D + \frac{\boldsymbol{b}}{\Omega} \times \left(\frac{s^2}{2B} \nabla B - s\boldsymbol{K}^s \right), \tag{82}$$

where $s\boldsymbol{K}^s$ is given by Eq. (36). Using

$$\boldsymbol{b} \cdot \nabla(\boldsymbol{b} \cdot \nabla \times \boldsymbol{b}) + (\nabla \cdot \boldsymbol{b})\boldsymbol{b} \cdot \nabla \times \boldsymbol{b} = \boldsymbol{b} \cdot \nabla \times \boldsymbol{\kappa}, \tag{83}$$

we have

$$\nabla \cdot \boldsymbol{u}_D = \frac{s^2}{2\Omega}\left[\boldsymbol{b} \cdot \nabla \times \boldsymbol{\kappa} - (\boldsymbol{b} \cdot \nabla \times \boldsymbol{b}) \frac{\boldsymbol{b} \cdot \nabla B}{B} \right]. \tag{84}$$

Also,

$$\nabla \cdot (\boldsymbol{v}_D - \boldsymbol{u}_D) = \frac{s^2}{2\Omega}(\nabla \times \boldsymbol{b}) \cdot \nabla \log B - \frac{1}{\Omega}\left(\nabla \times \boldsymbol{b} + \frac{\boldsymbol{b} \times \nabla B}{B} \right) \cdot (s\boldsymbol{K}^s) + \frac{1}{\Omega}\boldsymbol{b} \cdot \nabla \times (s\boldsymbol{K}^s). \tag{85}$$

Then, using

$$\nabla \times \boldsymbol{b} = \boldsymbol{b}(\boldsymbol{b} \cdot \nabla \times \boldsymbol{b}) + \boldsymbol{b} \times \boldsymbol{\kappa}, \tag{86}$$

we have

$$\nabla \cdot \boldsymbol{v}_D = \frac{s^2}{2\Omega}\left(\boldsymbol{b} \cdot \nabla \times \boldsymbol{\kappa} + \frac{\boldsymbol{b} \times \boldsymbol{\kappa} \cdot \nabla B}{B} \right) - \frac{1}{\Omega}\left(\nabla \times \boldsymbol{b} + \frac{\boldsymbol{b} \times \nabla B}{B} \right) \cdot (s\boldsymbol{K}^s) + \frac{1}{\Omega}\boldsymbol{b} \cdot \nabla \times (s\boldsymbol{K}^s). \tag{87}$$

Also, using Eq. (44) for $\dot{u}_D$ and, from Eq. (35), $\partial \boldsymbol{K}^u / \partial u = \boldsymbol{\kappa}$, we have

$$\frac{\partial \dot{u}_D}{\partial u} = \frac{\boldsymbol{b}}{\Omega} \times (2u\boldsymbol{\kappa} + \boldsymbol{\nu}) \cdot \boldsymbol{K}^u + \boldsymbol{\kappa} \cdot \boldsymbol{v}_D - \frac{s^2}{2\Omega}\boldsymbol{b} \cdot \nabla \times \boldsymbol{\kappa}. \tag{88}$$

And, finally, from Eq. (45) for $\dot{s}_D$, noting that $\partial / \partial s (s\boldsymbol{K}^s) = 0$, we find

$$\frac{1}{s}\frac{\partial}{\partial s}(s\dot{s}_D) = \frac{1}{\Omega}\left[\boldsymbol{b}\boldsymbol{b} \cdot \nabla \times \boldsymbol{b} + \frac{\boldsymbol{b} \times \nabla B}{B} \right] \cdot (s\boldsymbol{K}^s) - \frac{1}{\Omega}\boldsymbol{b} \cdot \nabla \times (s\boldsymbol{K}^s) - \frac{1}{\Omega}\boldsymbol{b} \times (u\boldsymbol{\kappa} + \boldsymbol{b} \cdot \mathcal{W}) \cdot \boldsymbol{K}^u. \tag{89}$$

By adding these together, using Eq. (82) for $\boldsymbol{v}_D$ and

$$\boldsymbol{K}^u = u\boldsymbol{\kappa} + \boldsymbol{\nu} - \boldsymbol{b} \cdot \mathcal{W}, \tag{90}$$

we obtain

$$\nabla \cdot \boldsymbol{v}_D + \frac{\partial \dot{u}_D}{\partial u} + \frac{1}{s}\frac{\partial}{\partial s}(s\dot{s}_D) = 0. \tag{91}$$

Finally, by adding the zeroth-order and first-order terms, we have

$$\nabla \cdot (\boldsymbol{U} + \boldsymbol{b}u + \boldsymbol{v}_D) + \frac{\partial}{\partial u}(\dot{\bar{u}} + \dot{u}_D) + \frac{1}{s}\frac{\partial}{\partial s}[s(\dot{\bar{s}} + \dot{s}_D)] = 0 \tag{92}$$

which expresses phase-space conservation through first order in δ.

VII. CONCLUSIONS

The main result of this paper is the drift-kinetic equation, Eq. (57), where the coefficients are given in Sec. IV. For convenience we recapitulate the key results here:

$$\frac{\partial f}{\partial t} + (\boldsymbol{u} + \boldsymbol{U} + \boldsymbol{v}_D) \cdot \nabla f + \dot{w}\frac{\partial f}{\partial w} + \dot{\mu}\frac{\partial f}{\partial \mu} = C(f),$$

where the kinetic energy and magnetic moment are defined by $w = v^2/2$ and $\mu = s^2/(2B)$, and where

$$\boldsymbol{v}_D = \boldsymbol{u}_D + \frac{s^2}{2\Omega}\boldsymbol{b} \times \nabla \log B - \frac{\boldsymbol{b}}{\Omega} \times (\boldsymbol{F} - u^2\boldsymbol{\kappa} - u\boldsymbol{\nu}),$$

$$\dot{w} = u\left[\frac{e}{m}E_\parallel - \boldsymbol{U} \cdot (\nabla \boldsymbol{U}) \cdot \boldsymbol{b} \right] + \mu B\boldsymbol{U} \cdot \nabla \log B - u^2(\nabla_\parallel \boldsymbol{U}) \cdot \boldsymbol{b} + \frac{2\mu B}{\Omega}\boldsymbol{b} \times \boldsymbol{\nu} \cdot (\boldsymbol{b} \cdot \boldsymbol{W}) + \frac{u\mu B}{2\Omega}b_i e_{ijk} W_{jm} \mathcal{B}_{mk} - \frac{5}{2}\boldsymbol{b} \cdot \boldsymbol{\kappa} \times (\boldsymbol{b} \cdot \boldsymbol{W}) + \mu E_\parallel \boldsymbol{b} \cdot \nabla \times \boldsymbol{b},$$

$$\dot{\mu} = \mu\left(\frac{E_\parallel}{B} \right)\boldsymbol{b} \cdot \nabla \times \boldsymbol{b} + \boldsymbol{u}_D \cdot (\boldsymbol{F} - u\nabla_\parallel \boldsymbol{U} - \mu \nabla B) - \frac{\mu B}{\Omega}\boldsymbol{b} \cdot \nabla \times (\boldsymbol{F} - u^2\boldsymbol{\kappa} - u\boldsymbol{\nu}) + \frac{\mu B}{\Omega}\boldsymbol{b} \cdot [u(\boldsymbol{\nu} - 4\boldsymbol{b} \cdot \boldsymbol{W}) \times \boldsymbol{\kappa} + 2\boldsymbol{\nu} \times (\boldsymbol{b} \cdot \boldsymbol{W})].$$

Here the curl is performed at fixed u, and we use the abbreviations

$$\boldsymbol{u}_D = \frac{\mu B}{\Omega}\boldsymbol{b}\boldsymbol{b} \cdot \nabla \times \boldsymbol{b},$$

$$\boldsymbol{F} = \frac{e}{m}\boldsymbol{b}E_\parallel - \frac{\partial \boldsymbol{U}}{\partial t} - \boldsymbol{U} \cdot \nabla \boldsymbol{U},$$

$$\boldsymbol{\nu} = \boldsymbol{U} \cdot \nabla \boldsymbol{b} + \boldsymbol{b} \cdot \nabla \boldsymbol{U},$$

$$\boldsymbol{\mathcal{B}} = \boldsymbol{\nabla b} + (\boldsymbol{\nabla b})^T,$$

$$\boldsymbol{W} = \tfrac{1}{2}\left[\nabla \boldsymbol{U} + (\nabla \boldsymbol{U})^T - \tfrac{2}{3}\boldsymbol{I}\nabla \cdot \boldsymbol{U} \right]$$

and $\boldsymbol{\kappa} = \boldsymbol{b} \cdot \nabla \boldsymbol{b}$ is the magnetic curvature. We remind the reader that the $E \times B$ drift is missing from $\boldsymbol{v}_D$ because it appears in $\boldsymbol{U}$; the point is that this drift is allowed to become large as part of $\boldsymbol{U}$, while $\boldsymbol{v}_D \sim \delta$. We also point out that the first-order term in $\dot{\bar{w}}$, omitted in the lowest-order Eq. (71), has now been restored; it appears as the last term in $\dot{w}$, above.

Our drift-kinetic equation includes all terms of first order in the gyroradius (as well as certain second-order terms), while allowing the lowest-order flow velocity $\boldsymbol{U} = \boldsymbol{b}U_\parallel + \boldsymbol{V}_E$ to be comparable to the thermal speed. There is no assumption about the form of the guiding-center distribution f. We have demonstrated phase-space conservation for our drift-kinetic equation.

ACKNOWLEDGMENTS

We wish to thank H. Vernon Wong for allowing us to see his work prior to publication. One of us (R.D.H.) wishes to thank Ker-Chung Shaing for helpful discussions concerning Ref. 7. We also wish to thank Ronald E. Waltz for helpful comments on the manuscript.

This work was supported in part by DOE Grants No. DE-FG02-04ER54742 and No. DE-FG03-95ER54309.

[1] E. A. Freiman, Phys. Fluids **13**, 490 (1970).
[2] P. H. Rutherford and E. A. Frieman, Phys. Fluids **11**, 569 (1968).
[3] M. C. Begelman, R. D. Blandford, and M. J. Rees, Rev. Mod. Phys. **56**, 255 (1984).
[4] R. F. Ellis, A. B. Hassam, S. Messer, and B. R. Osborn, Phys. Plasmas **8**, 2057 (2001).
[5] S. M. Mahajan and Z. Yoshida, Phys. Rev. Lett. **81**, 4863 (1998).
[6] K. H. Burrell, Phys. Plasmas **4**, 1499 (1997).
[7] R. D. Hazeltine and A. A. Ware, Plasma Phys. **20**, 673 (1978).
[8] L. Cheung, Phys. Rev. A **13**, 1209 (1976).
[9] L. Cheung and W. Horton, Jr., Ann. Phys. (N.Y.) **81**, 201 (1970).
[10] I. B. Bernstein and P. J. Catto, Phys. Fluids **28**, 1342 (1985).
[11] F. L. Hinton and S. K. Wong, Phys. Fluids **28**, 3082 (1985).
[12] R. D. Hazeltine, D. Dobrott, and T. S. Wang, Phys. Fluids **18**, 1778 (1976).
[13] M. N. Rosenbluth and N. Rostoker, Phys. Fluids **2**, 23 (1959).
[14] R. D. Hazeltine, Plasma Phys. **15**, 77 (1973).
[15] A. N. Simakov and P. J. Catto, Phys. Plasmas **12**, 012105 (2005).
[16] H. Vernon Wong (to be published).
[17] S. I. Braginskii, "Transport processes in a plasma," *Reviews of Plasma Physics* (Consultants Bureau, New York, 1965), Vol. 1.