PPPL Technical Report #4945

Report of the Study Group GK2 on Momentum Transport in Gyrokinetics*

J. A. Krommes† and G. W. Hammett‡

September 30, 2013

Abstract

This is the report of a study group that was asked to review recent work on the issue of turbulent momentum transport in gyrokinetic theory, forms of gyrokinetics that lead to a momentum conservation law, the order of accuracy needed to adequately predict momentum transport, and related topics. Very subtle issues are involved. Agreement is found with recent work that has argued that a very high order of accuracy (beyond what has been calculated or implemented) would be required in the low-flow gyro-Bohm ordering regime if one were to try a standard gyrokinetic approach using the gyrokinetic Poisson equation to determine the long-wavelength potential on the transport time scale. That work presents an alternate approach that directly employs a momentum conservation law in order to reduce the order of accuracy required. Another method is to couple a gyrokinetic code on the short turbulence time scale with a transport code for the longer transport time scale. Other regimes are also of interest and lead to somewhat simpler, though still challenging, equations. While progress has been made and there are some definite conclusions that can be stated at the present time, there are still important questions that are unsolved and momentum transport remains an area of ongoing research. The report includes extensive tutorial material that provides some of the background needed for an understanding of recent papers on gyrokinetics.

*This version omits certain calculations, such as the derivation of the gyrocenter momentum conservation law from the particle moment equations, and in general should be considered to be a work in progress. It is made available in the hope that it may be useful, but with the understanding that it is not a fully completed body of work. Someday an updated version of this report may be issued with additional calculations included.

†E-mail: krommes@princeton.edu

‡E-mail: hammett@princeton.edu

Contents

[1] Some of the material of this section has been distilled into the Brief Communication by Krommes (2013).

1 Prologue

This report deals with ordering issues relating to gyrokinetic momentum transport. It reflects our understanding of the literature through mid-September, 2013. We believe that the main conclusions here are well founded, as will be explained further. However, there are still some open questions and there is ongoing work on various aspects of momentum transport. Specifically, we are aware of research by Brizard, Calvo, Parra, Scott, and Sugama (and there are likely others) that is either nearly completed or soon to be published; we have not seen written details. Therefore, the present document should properly be considered to be a status report on our own understanding of the issues associated with an actively evolving field rather than a review of a fully mature research topic.

1.1 A few words about emphasis and balance

As explained thoroughly in Sec. 2.1, this report focuses on issues raised by Parra & Catto in a series of papers dating from about the time of Parra's PhD dissertation (Parra, 2009). Although we make extensive reference to those papers, it does not seem necessary to provide a complete review of their methodology, which has been thoroughly and clearly explained in the original publications. (Some extended quotations from those are provided in Appendix L.) Instead, we attempt to provide a slightly different perspective and offer some previously unpublished insights. Thus we have chosen to conduct much of our discussion in the framework of the momentum conservation law for gyrocenters that was originally derived by Scott & Smirnov (2010) and which we consider to be an

important contribution to the field. Our discussion goes beyond that of Scott & Smirnov and thus fills a gap in the literature. We do not mean to imply that proceeding in this way is 'better' than the route taken by Parra & Catto, nor do we wish to diminish in any way the seminal and extensive contributions of Parra and his collaborators. But as is frequently the case in advanced physics research, it is illuminating to establish consistency between alternate approaches, and that is what we have tried to do.

1.2 Reading guide

In order to make this report useful to both beginners and experts, it contains a wealth of material, including extensive sections on technical details of recent gyrokinetic (GK) work related to momentum transport as well as tutorial sections that provide some of the background needed to understand the recent papers. Much of that can be skipped on an initial reading. For an quick overview, one can read Secs. 1.3 (p. 6) and 1.4 (p. 8), which serve as an executive summary; read through Secs. 2.1 (p. 9) and 2.2 (p. 12) for an introduction and an explanation of the assumptions; then skip to Sec. 5.1.2 (p. 51) to see the ordering arguments of Parra & Catto presented in a simple way that demonstrates the need to keep high-order corrections to the Hamiltonian if one is relying on the gyrokinetic Poisson equation to determine the long-wavelength potential on transport time scales. Aficionados of gyrokinetic field theory can peruse Secs. 5.3 (p. 55) and Sec. 5.4 (p. 58), where the ordering arguments are presented specifically from the point of the gyrokinetic momentum conservation law originally derived by Scott & Smirnov. A discussion of the outstanding questions, and an overall summary, are given in Sec. 6 (p. 61). Many technical details are relegated to the appendices, some of which are pedagogical.

1.3 Charge letter

The GK2 Study Group of the PPPL Theory Department was commissioned by Riccardo Betti in January, 2011.[2] In this section the content of the original charge letter to the Study Group, including background paragraph and specific questions, is reproduced verbatim. That letter makes it clear that the goal of this Study Group was to focus on one specific technical question (related to the accuracy of gyrokinetics for the description of momentum transport). It is not within the purview of this report to address the larger universe of interesting and important issues relating to the applications of gyrokinetics to magnetic confinement fusion or other related physics areas such as space plasma physics, or even to consider other momentum-related issues that do not satisfy the assumptions described below.

[2] An early version of this report was circulated among a few experts in October, 2011. The present report is a major expansion of the original version; the main conclusions have not changed.

Verbatim copy of the charge letter:

TITLE: Ordering requirements in the gyrokinetic equations for simulations of turbulence on transport time scales

MEMBERS: G. Hammett, J. Krommes (Chair)

BACKGROUND: It has been recently stated in several articles that the physics models in current global gyrokinetic codes are not accurate enough to correctly predict evolution of the momentum profile via solving the quasineutrality equation over long (i.e. transport) time scales. This conclusion was based upon the observation that the particle drift velocities apparently need to be computed to high accuracy (up to third order in the normalized Larmor radius, ρ/a) to correctly predict the radial electric field and, therefore, the transport of momentum over long time scales. Since flows can have profound effects on regulating the turbulence level, such high accuracy would be required for that approach to the gyrokinetic simulation of turbulence. In light of the difficulties required in deriving a third-order accurate model and in maintaining the required numerical accuracy, existing gyrokinetic models were deemed inadequate for simulating turbulence on long time scales. There are very subtle issues here, and this conclusion has been challenged by other experts. An alternate view is that second-order gyrokinetics is sufficient to correctly predict the radial electric field and to simulate turbulence on long time scales. The issue to be addressed and resolved by the GK2 Study Group concerns the minimum ordering required in the gyrokinetic equations for correctly predicting turbulent transport on long time scales.

QUESTIONS TO BE ADDRESSED:

[1] Are the gyrokinetic equations used in existing codes (including `GTS`, `GYRO`, and `GTC`) adequate to simulate the evolution of turbulence over transport time scales?

[2] Are second- and/or third-order corrections in the normalized Larmor radius required for (a) the gyrokinetic equation; (b) the gyrokinetic Maxwell equations?

[3] If the answer to [2] is positive, is the vorticity equation proposed by Parra & Catto a possible fix to extend the validity of gyrokinetic codes to long time scales?

[4] If second-order gyrokinetics is sufficient, what needs to be done to fully implement this into codes with sufficient numerical accuracy, either directly or through separate gyrokinetic transport equations for the long time scale?

1.4 Summary

Here are concise answers to the previous questions. In the body of the report, we will provide introductory discussion, background information on some of the techniques, and a summary of our understanding of the present state of affairs. Most details and advanced pedagogical material are relegated to appendices.

In addition to the questions identified in the Charge, a more basic question must be answered first:

[0] Are the physics contents of the particle and the gyrocenter conservation laws equivalent or different?

The physics contents are equivalent. The mathematical representations differ.

With question [0] answered, a major source of uncertainty is removed and one can focus on the original questions:

[1] Are the gyrokinetic equations used in existing codes (including `GTS`, `GYRO`, and `GTC`) adequate to simulate the evolution of turbulence over transport time scales?

No existing gyrokinetic code contains all of the terms that are required for completely consistent simulations of momentum transport in the low-flow gyro-Bohm regime on transport time scales. However, efforts in that direction are being made in the context of coupling turbulence-time-scale gyrokinetic simulations with long-time transport solvers.

It should be noted that most present codes are adequate (perhaps with some extension to fully incorporate second-order Hamiltonian effects into the fluxes) to study turbulent momentum fluxes on the shorter turbulence time scale in the high-flow regime, such as if there is strong beam injection. Present codes are also adequate to study the particle and heat fluxes on the shorter turbulence time scale.

[2] Are second- and/or third-order corrections in the normalized Larmor radius required for (a) the gyrokinetic equation; (b) the gyrokinetic Maxwell equations?

In principle, under a particular set of assumptions relevant to a specific physics regime (see Sec. 2.1), third-order corrections are required for the gyrokinetic equation with a global full-F approach, while second-order corrections are necessary in the gyrokinetic Maxwell equations. If a fluid momentum equation is adjoined to the gyrokinetic equation, it is possible to work to just second order in the kinetic equation (by using a moment method described by Parra & Catto).

[3] If the answer to [2] is positive, is the vorticity equation proposed by Parra & Catto a possible fix to extend the validity of gyrokinetic codes to long time scales?

Yes.

[4] If second-order gyrokinetics is sufficient, what needs to be done to fully implement this into codes with sufficient numerical accuracy, either directly or through separate gyrokinetic transport equations for the long time scale?

Second-order gyrokinetics should be sufficient if it is coupled directly to a momentum transport equation, or if one is in a parameter regime that relaxes the low-flow gyro-Bohm ordering. Second-order effects could be added in principle to existing codes (along with adding the momentum equation). The polarization term in the gyrokinetic Poisson equation should be calculated in a particular way to be consistent. Calculating the effects of small second-order drifts accurately when they are combined with larger first-order drifts in the same equation may require higher numerical accuracy than is typical at present. Analysis of the accuracy requirements for the several types of gyrokinetic codes (which use various PIC and continuum discretization and time-integration methods that possess various conservation properties) should be carried out. The accuracy requirements might be more manageable near the plasma edge where the local $\epsilon \doteq \rho/L$ (the ratio of the gyroradius ρ to the local gradient scale length L) is not as small and there is less separation between the turbulence and profile time-evolution scales.

2 Introduction

2.1 Motivations, and brief statement of the problem

We consider magnetically confined, toroidal fusion devices. The physics of the radial electric field and plasma rotation are important for a number of reasons, including MHD stability and the saturation of microturbulence. Plasmas have been observed to rotate intrinsically, i.e., spin up in the absence of external torques, initial mean flow, or flow gradients; such rotation can result from the presence of turbulence with density or temperature gradients and is a subject of great current interest.

An important distinction must be made between *axisymmetric* and *nonaxisymmetric* devices. The external magnetic-field configuration of an ideal tokamak is axisymmetric (invariant under a toroidal rotation), whereas tokamaks with magnetic field ripple (or externally imposed 3D magnetic perturbations) and stellarators are nonaxisymmetric. In the presence of axisymmetry, the conservation of toroidal canonical angular momentum provides an important constraint and is the source of the ordering subtleties discussed below. Practically, it is possible that those subtleties may in some cases be overwhelmed by the consequences of nonaxisymmetry. However, considerable further research remains to be done on this point, and in any event it is not the purpose of this report to discuss all possible physics regimes. The original concerns related to axisymmetric devices, and we focus on those here. *Unless noted otherwise, toroidal axisymmetry of the background magnetic-field configuration is assumed*

from here on. Also, we shall consider only electrostatic turbulence ($\boldsymbol{E} = -\boldsymbol{\nabla}\phi$) for simplicity. Electromagnetic effects have been considered by Scott & Smirnov (2010).

Low-frequency turbulence with $\omega \ll \omega_{ci}$ is generally addressed with the gyrokinetic formalism, and elaborate computer codes have been constructed to solve the gyrokinetic equations (coupled to the gyrokinetic Maxwell equations). Traditionally those codes have been used to study the physics of turbulence saturation, so they have been generally run through only a few turbulence autocorrelation times. However, modern interest is turning to physics on the *transport time scale* on which profiles develop in response to the neoclassical and turbulence transport processes. Several approaches can be envisaged:

1. Simulations on the turbulence time scale can be used to calculate turbulent fluxes. Those can then be used in a macroscopic transport solver that evolves the profiles.

2. A gyrokinetic simulation can be integrated directly to transport time scales.

In either of these approaches, it is of interest to understand whether the analytical formulation of the gyrokinetic system is complete enough to capture the relevant physics. The answer is unclear *a priori* because practical implementations of gyrokinetics are inevitably approximate, involving truncations of asymptotic expansions. As we will see, momentum fluxes can be very small in some cases, raising the possibility that relevant terms may have been omitted in many or all of the extant gyrokinetic codes.

This concern was first articulated clearly in the PhD dissertation of Parra (2009) [see also the earlier paper of Parra & Catto (2008)], which was primarily concerned with method 2 above, and has been developed in many subsequent papers by Parra & Catto and Parra & Calvo (see Appendix L for some excerpts from the key papers). Those authors have argued strongly that current gyrokinetic codes are inadequate for a complete treatment of momentum transport on the transport time scale. (However, present codes could be used to calculate momentum fluxes on the shorter turbulence time scale in the presence of strong flows driven by beam injection.) Papers by Scott & Smirnov (2010) and Idomura (2012) contain statements that have been interpreted by some as counterarguments.[3] The purpose of the present report is to provide background for, and to describe the current state of this issue. In brief, we find that the principal conclusion of Parra & Catto is correct regarding the need for higher-

[3]Both papers are excellent and contain interesting, important results; however, neither goes far enough to disprove the specific points that Parra & Catto have made, as we will demonstrate later. Further concerns have been expressed by B. Scott (private communication, 2013). But Parra & Calvo (private communication, 2013) have evidently calculated all of the terms to the necessary order; they have concluded that it is not possible to prove that the new higher-order terms do not contribute to momentum transport in general, and they speculate that there is always some regime in which they will matter. The present authors have not seen the details underlying either of the last two private communications.

order terms for momentum transport, given the particular set of assumptions they considered. Those are

1. low-flow ordering;
2. gyro-Bohm scaling;
3. toroidal axisymmetry;
4. up–down symmetry.

It is very important to stress that these assumptions are not relevant to all possible physics regimes of interest or all possible machine designs. For example, they may not hold near the plasma edge where eddy sizes are not very small compared to gradients and gyro-Bohm scaling might not hold, or with large shear flows such as may occur with beam injection or near transport barriers. Therefore, extant codes (perhaps with some extensions) remain useful for studies of a variety of important physical processes.

The problem can be boiled down to the following questions, which constitute the charge of the Study Group:

[1] Are the gyrokinetic equations used in existing codes (including `GTS`, `GYRO`, and `GTC`) adequate to simulate the evolution of turbulence over transport time scales?

[2] Are second- and/or third-order corrections in the normalized Larmor radius required for (a) the gyrokinetic equation; (b) the gyrokinetic Maxwell equations?

[3] If the answer to [2] is positive, is the vorticity equation proposed by Parra & Catto a possible fix to extend the validity of gyrokinetic codes to long time scales?

[4] If second-order gyrokinetics is sufficient, what needs to be done to fully implement this into codes with sufficient numerical accuracy, either directly or through separate gyrokinetic transport equations for the long time scale?

One can note that these questions do not, in fact, make direct reference to the fundamental assumptions listed above. The questions are open-ended and can be asked for various physics regimes. Answering them in all generality would require nothing less than a discussion of the utility of gyrokinetic simulations for all possible physical regimes of toroidal devices. Here, however, we shall interpret the questions narrowly and address them only within the context of the basic assumptions of toroidally axisymmetric devices with gyro-Bohm scaling in the low-flow regime.

In fact, an even more fundamental issue underlies the discussion. We will address questions 1–4 from the point of view of momentum conservation.[4] Now

[4] For a more general approach, see Calvo & Parra (2012).

momentum conservation laws can be formulated in the phase spaces of either particles or gyrocenters, and the mathematical forms of those laws differ depending on which is chosen. There has been some concern that although a gyrocenter momentum conservation may exist, its physics content may differ from the particle conservation law; if so, the gyrokinetic conservation law would be irrelevant to the momentum transport of real particles. Thus, the most basic question is

[0] Are the physics contents of the particle and the gyrocenter conservation laws equivalent or different?

We will argue that they are, in fact, equivalent. The important implication is that manipulations can be done in either particle or gyrocenter phase space. If they are done consistently, one will be led to identical physics conclusions, although they will have different mathematical representations. As a simple example, it is well known that while the diamagnetic velocity emerges from the particle-based momentum equation, it is absent (explicitly) from the gyrokinetic equation. However, both equations can correctly represent low-frequency physics.

2.2 Momentum flux in the low-flow ordering

In the subsequent discussion, it is crucial to understand the size of the momentum flux that is expected. We shall give two different but consistent arguments.

Unless otherwise indicated, by 'flux' we mean statistically averaged, time averaged, or simply mean flux Π. This is the quantity that is measured experimentally and is the one that is relevant for macroscopic considerations of confinement and rotation-related issues. In a turbulent system, there are also random fluxes $\widetilde{\Pi}$, e.g., momentum flows due to randomly directly $\boldsymbol{E}\times\boldsymbol{B}$ flows. One has $\Pi = \langle\widetilde{\Pi}\rangle$, where the angular brackets denote a flux-surface average as well as a statistical ensemble average. It is important that such averages are assumed to vary only on a macroscopic spatial scale L. (If necessary, we will think of L as the minor radius a; however, generally we will not distinguish between the macroscopic length scales, such as a or the major radius R, or equilibrium scale lengths such as $L_n \doteq |\nabla \ln\langle n\rangle|^{-1}$.) Of course, knowledge of the mean says nothing about the variance of the fluctuations. Typically the probability density function (PDF) of those fluxes is broad.[5] It is quite possible for the ratio of mean to variance to be asymptotically small, and that is in fact the scenario we shall describe.

2.2.1 Gyro-Bohm scaling and momentum flux

For turbulent transport (we will mention collisional transport in the next paragraph), one can easily estimate[6] the mean flux as follows. Consider a hypothet-

[5] B. Scott (private communication, 2013) has performed measurements of those PDFs in simulations.

[6] One or another version of this estimate can be found in many of the papers of Parra & Catto.

ical local conservation law for the fluid momentum density (which is predominantly that of the ions) in a slab model:

$$m_i n_i \frac{\partial \langle u \rangle(x,t)}{\partial t} = -\frac{\partial \Pi}{\partial x}. \tag{1}$$

Here m_i is the ion mass, n_i is the ion density, u is the toroidal velocity, and x is a radial coordinate. For ordering purposes, we will assume that the mean flow is diffusive:

$$-\frac{\partial \Pi}{\partial x} = D \frac{\partial^2 \langle m_i n_i u \rangle}{\partial x^2}. \tag{2}$$

We do not mean to imply that $\langle u \rangle$ literally obeys a diffusion equation. Later we will refer to pinch and residual-stress terms; the diffusion form is just an artifice for establishing a basic scaling. Now the dimensions of Π (being a flux of momentum) are $[\Pi] = [(mnu)u]$. Therefore, it is natural to normalize Π to $m_i n_i c_{\rm s}^2$, where $c_{\rm s}$ is the sound speed. Then

$$\frac{\Pi}{m_i n_i c_{\rm s}^2} \sim \frac{L}{c_{\rm s}^2} \frac{\partial \langle u \rangle}{\partial t} \sim \left(\frac{L}{c_{\rm s}^2} \right) \left(\frac{D}{L^2} \right) \langle u \rangle = \frac{D}{L c_{\rm s}} \left(\frac{\langle u \rangle}{c_{\rm s}} \right). \tag{3}$$

One can consider either a sonic ('high-flow') or a subsonic ('low-flow') ordering; we follow Parra & Catto in concentrating here on the low-flow regime.[7] By definition, in the low-flow ordering one has $\langle u \rangle / c_{\rm s} = O(\epsilon)$, where ϵ is the basic gyrokinetic expansion parameter $\epsilon \doteq \rho_{\rm s}/L$. [Here $\rho_{\rm s}$ is the sound radius: $\rho_{\rm s} \doteq c_{\rm s}/\omega_{ci}$. The ordering makes $\langle u \rangle = O(V_*)$, where $V_* \doteq \rho_{\rm s} c_{\rm s}/L_n = \epsilon c_{\rm s}$ is the diamagnetic speed and L_n is the density scale length.] A general theory should be able to handle the gyro-Bohm regime (which is what is expected in the small-ρ_* limit of local turbulence and is also what is found in neoclassical theory when

[7] Important physics regimes, such as might occur with strong beam injection or near the edge or scrape-off layer, cannot be treated in the low-flow approximation. We study that regime because it is the most difficult and the one treated in much of Parra's work.

It is not yet clear how intrinsic rotation in tokamaks scales. One can show that the ensemble-averaged intrinsic momentum flux vanishes in high-flow gyrokinetics in standard core regimes (Parra et al., 2011b), so it is necessary to keep higher-order terms from the low-flow regime in order to explain intrinsic rotation observations from a core turbulence mechanism. Some proposed theories involve edge mechanisms that drive an intrinsic momentum flux to the wall (deGrassie et al., 2009; Stoltzfus-Dueck, 2012), and some of those mechanisms might force the core flows into a high-flow regime. Some empirical scalings (Rice et al., 2007) suggest that intrinsic flows can be in a high-flow regime, while other studies of rotation data find that the core–edge difference is consistent with low-flow gyro-Bohm theoretical scalings (Parra et al., 2012). In particular, observations of a reversal in the toroidal flow in the tokamak core region indicate the need to keep high-order terms from the low-flow regime (Parra et al., 2012).

When there is strong rotation from beam injection, it is clearer that one can focus on the simpler high-flow regime in the plasma core. A thorough treatment of many aspects of the high-flow regime has been given by Abel et al. (2013) [see also Sugama & Horton (1998)].

$\nu/\omega_{ci} \sim \epsilon$), so we also assume gyro-Bohm transport scaling[8]:

$$D_{\mathrm{gB}} = \underbrace{\left(\frac{\rho_{\mathrm{s}}}{L}\right)}_{\epsilon} D_B, \tag{4}$$

where the Bohm diffusion coefficient is defined to be

$$D_B \doteq \frac{cT_e}{eB} = \rho_{\mathrm{s}} c_{\mathrm{s}}. \tag{5}$$

Then

$$\frac{\Pi}{m_i n_i c_{\mathrm{s}}^2} \sim \underbrace{\left(\frac{\rho_{\mathrm{s}}}{L}\right)^2}_{\epsilon^2} \underbrace{\left(\frac{\langle u \rangle}{c_{\mathrm{s}}}\right)}_{\epsilon} = O(\epsilon^3). \tag{6}$$

Thus, for gyro-Bohm scaling in the low-flow ordering, the flux of linear momentum density is $O(\epsilon^3)$. That scaling also holds for the flux of angular momentum density if it is normalized appropriately. Angular momentum contains an extra factor of the major radius R. Technically, this follows from the covariant toroidal (φ) component (see Appendix A for a review of generalized coordinates) of $\boldsymbol{u}$:

$$u_\varphi \doteq \boldsymbol{u} \cdot \boldsymbol{e}_\varphi = R\boldsymbol{u} \cdot \widehat{\boldsymbol{\varphi}}. \tag{7}$$

Therefore, one has

$$\frac{\Pi_\varphi}{m_i n_i c_{\mathrm{s}}^2 R} = O(\epsilon^3). \tag{8}$$

In the presence of collisions, this estimate is modified. Parra & Catto (2008) showed that the collision-driven viscosity is $O\big(\epsilon^2(\nu/\omega_{ci})\big)$. For a typical maximal ordering, one can take $\nu/\omega_{ci} = O(\epsilon)$; then both the collisional and turbulent contributions are $O(\epsilon^3)$.

The $O(\epsilon^3)$ scaling is the key result that must be kept in mind during the subsequent discussion. Athough it may not yet be obvious, one way of interpreting it is to say that the physics of an axisymmetric[9] torus is *intrinsically ambipolar through second order* (Parra & Catto, 2009c; Sugama et al., 2011; Calvo & Parra, 2012, and references therein), so one must work to third order to determine the long-wavelength (flux-surface-averaged) part of the radial electric field (which is intimately related to toroidal rotation). This was already known for neoclassical theory, and the discussion suggests that it holds for turbulence as well. Here the result has been (heuristically) argued by looking just at the rotation. However, Calvo & Parra (2012) have discussed the problem in depth from the point of view of the complete gyrokinetic equation; the same conclusion is reached. We shall refer to that important paper again; it is essential reading for a complete understanding of the subject.

[8] Gyro-Bohm scaling, corresponding to local, microturbulent transport, is the worst case. In regimes where the transport scales more adversely than gyro-Bohm, i.e., $D \sim \epsilon^n D_B$ with $n < 1$, then the highest, third-order terms may not be needed, though second-order terms still will be.

[9] In the present discussion, the restriction to axisymmetry is hidden in the assumption of gyro-Bohm scaling.

2.2.2 Symmetry breaking and momentum flux

We now rederive the ϵ^3 scaling from a different argument that makes strong use of symmetry considerations. Consider the Reynolds stresses associated with microturbulence. In a slab geometry, a standard contribution to momentum flux is the Reynolds stress[10] $\Pi = (mn)_i \langle \delta V_x \, \delta V_y \rangle$, where $\delta \boldsymbol{V}$ is the fluctuation in the turbulent $\boldsymbol{E}\times\boldsymbol{B}$ velocity. For characteristic microturbulence that is saturated at the mixing-length level, those $\boldsymbol{E}\times\boldsymbol{B}$ velocities are at the diamagnetic level $V_* = \epsilon c_\mathrm{s}$. Nominally, therefore, $\Pi \equiv \Pi^{[2]} = O(\epsilon^2)$. (Here the bracketed superscript indicates the *apparent* order of the term.) However, in the absence of a preferred direction, one expects that δV_x and δV_y should be uncorrelated, in which case $\Pi^{[2]}$ would vanish.[11]

This argument fails in the presence of a sheared mean flow, which breaks the symmetry of the microscopic turbulence. One expects that the resulting flux should be proportional to the size of the symmetry-breaking effect:

$$\Pi \sim \epsilon^2 \times (\text{size of the symmetry breaking}). \tag{9}$$

In the low-flow regime, one has by assumption that $\langle u \rangle = O(\epsilon)$. Therefore, one expects for symmetry breaking by a sheared mean flow that

$$\Pi \sim \epsilon^2 \times \epsilon = \epsilon^3, \tag{10}$$

which agrees with the estimate in Sec. 2.2.1. (This argument is amplified in Appendix K.) Working backwards, we see that by invoking symmetry breaking by a mean diamagnetic-level flow we have *derived* gyro-Bohm scaling rather than assumed it from the outset.

In making this argument, it is crucial that the assumption of uncorrelated $\boldsymbol{E}\times\boldsymbol{B}$ flows actually holds for flow-free turbulence. That has been argued by two sets of authors: Sugama et al. (2011), and Parra et al. (2011b). They identify transformations T that preserve the form of the local gyrokinetic equation (in the absence of a mean flow or flow shear). That transformation includes invariance under reflection through the midplane of the poloidal field configuration, i.e., up–down symmetry. They then argue that since $\mathrm{T}\langle \delta V_x \, \delta V_y \rangle = -\langle \delta V_x \, \delta V_y \rangle$, the correlation function must vanish.[12] Again, that symmetry is broken by a mean flow or mean flow shear.

[10] In toroidal coordinates, the proper (contravariant and covariant) components are $\Pi = (mn)_i \langle \delta V^\psi \, \delta V_\varphi \rangle$; see Eq. (54).

[11] This is a key observation to which we shall return multiple times in this report; see, for example, Sec. 5.3. If it were the case that the basic Reynolds stress were in fact $O(\epsilon^2)$, one would not need to worry about possible contributions from higher-order terms.

[12] There is a subtlety that one might worry about, which is the possible breakdown of the ergodic hypothesis that equates an ensemble average with an appropriate average of a single realization over space and time. In principle, even if the ensemble average of the flux-surface-averaged $\delta V_x \delta V_y$ vanishes, it is conceivable that the equivalent space- and time-averaged flux could have a large nonzero value that would violate the ordering assumptions. Of course the instantaneous $\delta V_x \delta V_y$ at a single space point in a particular realization will not vanish in general and can be large. Sometimes it is assumed that the flux-surface average alone eliminates large fluctuations and behaves like an ensemble average. However, that is frequently not true in practice, so when processing numerical or experimental data one does

This symmetry argument thus implies an ϵ^3 scaling of a diffusive term ($\Pi \propto \partial_x \langle u \rangle$) and a pinch term ($\Pi \propto \langle u \rangle$) in the momentum balance. It is important to also consider residual stress, which by definition is not proportional to mean flow at all. However, the estimate (9) is still expected to hold. Various mechanisms may act. For example, Idomura (2012) has performed simulations in which the symmetry breaking arises from shear in the background density and temperature profiles, i.e., from the radial variation of the diamagnetic velocity V_*. (That leads to radial shear in the poloidal phase velocities of ITG/drift waves.) In a local gyro-Bohm limit, such effects vanish. If one includes them perturbatively, they enter as a correction of the order of ρ_s/L_*, where $L_*^{-1} \doteq |\partial_r \ln V_*|$. If one takes $L_* \sim L$, one finds that the symmetry breaking again scales with ϵ, corresponding to momentum fluxes of $O(\epsilon^3)$ or an intrinsic rotation of the order of ϵc_s. (See Appendix K for further discussion of a possible mechanism that might give a stronger symmetry breaking.)

This is a scaling argument; it does not speak to the absolute size of the effect. The distinction is important. For example, Idomura (2012) suggested that because his simulation result $U_\varphi/v_{ti} \sim 5\%$ "is an order of magnitude larger than" the low-flow gyro-Bohm estimate of $U_\varphi/v_{ti} \sim \rho_*$, the residual stress from the lower-order terms is much larger than third order. However, the estimate $U_\varphi/v_{ti} \sim \rho_*$ is only meant to indicate the *scaling* with ρ_*; there may be coefficients multiplying it (like qR/r) that may make the precise value significantly larger. One would need to do a ρ_* scan to determine the scaling (as Idomura acknowledges), although that is numerically very expensive. (It might be that this value of ρ_{*L} is not small enough to be deeply in the gyro-Bohm regime, so it would not contradict the scaling estimates we have given.) Furthermore, if one uses the more relevant local $\rho_{*L} = \rho/L_T \approx \frac{1}{42}$ instead of $\rho_* = \rho/a \approx \frac{1}{150}$, then one finds that the simulation U_φ/v_{ti} is actually about 2.1 ρ_{*L}, not very big. Finally, although Idomura's simulation methodology was formulated and executed exquisitely carefully, his analysis of the effects of higher-order terms was based on a limiting form of the third-order Hamiltonian H_3 that, as he ac-

further space and time averaging, where the averages are over widths in radius and time that are long compared to the turbulent decorrelation length and time but short compared to the equilibrium space and time scales; this selects out just the components of the momentum flux that contribute to the equilibrium-scale flow evolution. (Similarly, the particle and energy fluxes may have large fluctuations at any radius and instant of time, but transport theory is concerned with particle and energy fluxes that have been averaged over a radial width and time period sufficiently wide that fast zonal fluctuations have been eliminated and one can focus on the longer-time equilibrium profile evolution.) But while space-time averages and ensemble averages are clearly distinct in principle, so could lead to different predictions, they are equivalent under the ergodic hypothesis, and we are not aware of any demonstration that this assumption is breaking down.

Another possibility is that the surface-averaged $\delta V_x \delta V_y$ has a large positive or negative value for an extended time, and only after very long time averaging does it average to zero. In fact, there are observations (van Heijst & Clercx, 2009) in bounded 2D hydrodynamics that the fluid randomly spins up in one direction or the other (this spin-up is related to the inverse cascade in 2D), then occasionally flips direction spontaneously. We are not aware of this occurring in the equilibrium-scale toroidal rotation in gyrokinetic simulations. While one might worry about inverse cascades from the toroidal component of $\boldsymbol{E} \times \boldsymbol{B}$, the normal viscosity acting on the parallel flow probably prevents this.

knowledged, is valid only in the long-wavelength limit (Mishchenko & Brizard, 2011). For that case, we show analytically in Sec. J.2.1 that the contributions from H_3 are expected to be small. Studies of cases with $k_\perp \rho_i \sim 1$ remain to be done.

This concludes our introductory discussion of the basic scaling and symmetry arguments. A more detailed analysis of the terms in the gyrokinetic momentum balance is given in Appendix J; the results are consistent with the simple estimates given above.

2.3 Momentum conservation and axisymmetry

Except for a few remarks in Sec. 6, we consider only axisymmetric toroidal devices. By definition, that axisymmetry is with respect to the *background* magnetic field arising from the external coils and equilibrium plasma currents. The turbulent fluctuations need not be axisymmetric (that is, the toroidal mode number n can be nonzero).

The virtue of the restriction to axisymmetry is that clean results about the form of momentum conservation laws can be obtained from Noether symmetry arguments (reviewed in Appendix H), and that was first done for the gyrokinetic system in a seminal paper by Scott & Smirnov (2010). [Their results will be discussed at length in Secs. 5.3 and Appendix J. A subsequent paper by Brizard & Tronko (2011) gave a cleaner and more complete derivation.] But that possibility also leads to a subsidiary concern: Even if a conservation law for gyrokinetic toroidal angular momentum can be formulated, is it clear that it is the 'right' one? Perhaps one obtains proper results only by working with conservation laws for the actual particles, not the gyrocenters. Indeed, in the course of the last several years, statements to that effect have sometimes been made. Those assertions transcend the application to momentum conservation and call into question the fundamental justifications of the gyrokinetic formalism. Therefore, in Sec. 3 we briefly review the history and nature of gyrokinetics. We will conclude that one can legitimately focus on the content of the gyrokinetic conservation laws.

2.4 Global versus local, and full-F versus δF formulations

Gyrokinetic simulations can be formulated in terms of either global or local ('two-scale') approaches. 'Global' means that the physics of (at least a fair fraction of) the entire cross section is treated, including profile variations. 'Local' means that just an thin annulus or flux tube is treated, assuming constant profile gradients.

A distinction that is in principle independent of global vs local is between a 'full-F'[13] and a δF formulation. Full-F simply means that one solves the gyrokinetic equation for the entire distribution function F. No assumption is made about a Maxwellian or quasi-Maxwellian character of F; its form follows

[13] In this report we use F for the gyrocenter distribution and f for the particle distribution. The methods in question refer to solutions of the gyrokinetic equation, which evolves F.

from the GK dynamics in the presence of imposed boundary conditions and sources. The δF method is most precisely defined[14] by writing the complete gyrokinetic distribution function F as $F = F_0 + \delta F$, where F_0 is a known distribution function. The method can serve as a way of focusing on just the fluctuating components of the turbulence, since F_0 does not vary significantly on the turbulence time scale. [This was used in the Frieman–Chen gyrokinetics, as is often done in Hasegawa–Mima and other fluid-plasma-turbulence studies, and it can be employed in both continuum and particle-in-cell (PIC) codes.] Specifically in PIC simulation, the application of numerical techniques to just δF provides a way of reducing Monte Carlo sampling noise.[15] Parra's original concern was with the accuracy of full-F approaches integrated to transport time scales, and for most of the following discussion it is legitimate to assume that we are speaking of full-F.

The original two-scale δF approach focuses on the asymptotic $\rho_* \to 0$ gyro-Bohm limit. While that is an important limit to study, the modern full-F approach relaxes this assumption and allows one to consider nonlocal effects such as turbulence spreading and the transition between gyro-Bohm and Bohm scaling at finite ρ_*, which may become important even in large devices near transport barriers or the edge region. [Some later δF codes like `GYRO` or `GENE` were formulated to have a global option to study some nonlocal effects, and a more general nonlocal iterative derivation has been presented by Parra & Catto (2008).] Full-F is the natural way of implementing a global simulation, especially in the presence of regions close to the edge where fluctuations are large and there is no intrinsic separation between a background and fluctuating distribution function. However, it is clear that if no terms are dropped, equations formulated in terms of δF carry exactly the same content as do those describing full-F. Thus, it is in principle possible to perform a global simulation using a δF method, as is done in several PIC codes as a noise-reduction method, and as is done in some continuum codes like `GYRO` and `GENE`, which have options for running in a nonlocal mode as well as a local mode. However, what is sometimes meant by a δF method, even when run nonlocally, is one that solves only for a time short compared to the transport time scale, so that the background F_0 does not evolve much, in order to focus on calculating the turbulence spectrum. Thus one can run a δF simulation on the turbulence time scale in order to calculate fluxes, then use those fluxes in a transport solver in order to advance the profiles on the longer transport time scale. Some further remarks about δF formulations are given in Sec. 6.

2.5 Why has there been a controversy?

In advance of the detailed discussion to follow, let us assert that it is a firm result that the mean momentum flux is $O(\epsilon^3)$, given the stated assumptions. The forms of the momentum conservation laws for both particles and gyrocenters are also

[14] The paper by Hu & Krommes (1994) contains both introductory and advanced discussion of this point.

[15] For a review and further references, see Krommes (2007).

unambiguous. Why then has the topic of higher-order terms in gyrokinetics been so controversial, and why has it been so difficult to get everyone to agree? There are a variety of reasons.

First, note that although some small terms are omitted from conventional approximations, that does not mean that conventional codes cannot calculate some fluxes of the required order; they do. Both Idomura (2012) and Scott (private communication, 2013) have obtained mean momentum fluxes from simulations that include just first-order drifts and first-order polarization effects in Poisson's equation. (As discussed below, the latter effects are associated with a second-order Hamiltonian H_2.) The question is whether the omitted terms give rise to fluxes that are comparable to those that are already calculated.[16]

Let us explain this further. Normally one would think that in any asymptotic expansion the lowest-order nonvanishing terms should be sufficient. It is not too hard to argue that one must keep effects related to the second-order Hamiltonian H_2. The question is whether one also needs to consider the third-order Hamiltonian H_3; intuitively, most people dismissed that as being small and negligible. Indeed, H_3 itself is much smaller than H_2. The problem is that, as explained in Sec. 2.2.2, there are certain near-cancellations in an up–down symmetric tokamak (Sugama et al., 2011; Parra et al., 2011b; Peeters et al., 2011) that occur in the low-flow regime when calculating the flux-surface-averaged momentum flux (the perpendicular Reynolds-stress part of which is related to H_2 at lowest order); those make the resulting momentum flux smaller (by one factor of ϵ) than one would at first expect. This is why it is possible for H_3-related effects to compete with and cause just as much momentum flux as those associated with H_2, even though H_3 itself is $O(\epsilon)$ relative to H_2.

Note that these cancellations can only occur in axisymmetric tokamaks (and perhaps quasisymmetric stellarators). In general nonsymmetric devices, the radial electric field will be determined by ambipolarity constraints at lower order (Sugama et al., 2011); then it is not necessary to calculate H_3. Indeed, one of the complaints about the focus on higher-order terms has been that even if small terms are important for the axisymmetric problem, realistic devices are not perfectly axisymmetric. The worry is that if ripple or other asymmetries are sufficiently large, the deviations from axisymmetry may dominate small axisymmetric corrections. *A priori*, this is a legitimate concern, and further research is called for to identify regimes where asymmetry may be important. However, in practice the magnetic ripple $\delta B/B$ in many tokamaks is very small over most of the minor radius[17] and can likely be neglected, although there

[16] In addition to his simulations, Scott (private communication, 2013) has performed analysis that indicates to him that higher-order effects are negligible. We have seen some information about his arguments, and we agree with certain particular points. However, as of the date of the writing of this report we have not been able to obtain enough details for us to be able to discuss them in depth or to change our general conclusions about the ordering issues that are described in this report.

[17] For example, the relative magnetic field ripple $\delta B/B$ at the edge of the Joint European Torus (JET) is normally only 0.08% and drops rapidly by two orders of magnitude as one moves towards the magnetic axis. However, JET experiments (Nave et al., 2010) find that if the edge ripple is increased to 1.5% then it can have an effect on rotation that is comparable

may be some regimes near the plasma edge where the situation may be more complicated. Nevertheless, it remains useful to consider the consequences of axisymmetry. The logic and form of the gyrokinetic formalism are subtle; if one does not master them in a simple setting, there is no guarantee that errors will not be made in other situations as well.

In the gyrokinetic formalism, physics embodied in the Hamiltonian shows up in two places: through the drifts in the gyrokinetic equation, and in the form of the polarization effect in the gyrokinetic Poisson equation (which determines the electrostatic potential ϕ and thus the radial electric field E_r). A point of possible confusion is that it is not H itself that enters into the Poisson equation but rather terms derived from H by functional differentiation with respect to ϕ. (We remind the reader that we consider only electrostatic effects in this report.) That differentiation reduces the order by one, so use of a third-order H implies that for the electric polarization one must consider effects of no higher than second order. Although H_3 may be extremely small in practice and intuitively may seem to be negligible, second-order effects are not so small. [Another way of stating this point is that one must consider second-order terms in the 'pullback' transformation (Appendix D) that relates gyrocenter and particle coordinates (Brizard, 2000). Most codes only implement first-order terms in that transformation.]

The functional derivative $\delta H/\delta\phi$ also appears in the formal definition of the perpendicular Reynolds stresses, as we will show later in detail. Because it turns out that the derivative gets multiplied by another factor of $O(\epsilon)$ [see, for example, Eq. (93b)], H_3 can be shown to give rise to stresses that are nominally[18] $O(\epsilon^3)$. While that is very small, it is no smaller than the effects that arise from the symmetry breaking of $\Pi^{[2]}$, and it is the order that is expected for momentum fluxes in the gyro-Bohm regime, as indicated by the elementary arguments in Sec. 2.2.1. This is the basic logic that argues for the necessity of considering higher-order terms.

Yet another reason why it has been difficult to get everyone to agree is that there are two very different approaches to deriving nonlinear gyrokinetic equations. The more conventional original approach, initiated with the work of Frieman & Chen (1982), used a two-scale δF asymptotic expansion of the Vlasov equation, expanding the distribution function into large-scale equilibrium and small-scale fluctuating components. The more recent full-F reformulation uses Hamiltonian and/or Lagrangian methods and does not directly expand the distribution function. These two approaches use very different terminology and (superficially) different mathematical-physics techniques, and it is rare to find someone who understands both approaches well. However, there are some

in magnitude to the normal intrinsic rotation. The edge ripple for ITER is 0.5% to 1.2% (depending on the arrangement of ferritic inserts) and also drops rapidly towards the center, so its effect near the edge probably needs to be considered along with other possible mechanisms for intrinsic rotation.

[18] "Nominally of $O(\epsilon^3)$" means that the apparent size of the term is $O(\epsilon^3)$, although it could in principle be smaller, perhaps due to some unanticipated cancellation. In Sec. J.2.2 we note that some such terms are indeed smaller, but also that no argument has yet been advanced to prove that all contributions from H_3 are smaller than $O(\epsilon^3)$, and we doubt that is true.

exceptions such as Parra and, earlier, Sugama, who wrote seminal papers using both the two-scale expansion method [with a 1998 paper (Sugama & Horton, 1998) extending to the high-flow ordering and systematically deriving the next-order gyrokinetic transport equations] and the Hamiltonian method [with a 2000 paper (Sugama, 2000) that introduced the very useful approach of Lagrangian field theory]. It is often useful to have two independent approaches to a problem, as they provide different insights and useful cross checks against each other.

The more conventional two-scale asymptotic expansion technique has its roots in the linear gyrokinetics developed in the 1960's and was used in the invention of nonlinear gyrokinetics in the early 1980's. The transition to a modern formalism that involves a single 'full-F' gyrocenter distribution function began with the recursive derivation by Lee (1983) and the reformulation of Lee's work into a Hamiltonian representation of gyrokinetics by Dubin, Krommes, Oberman, & Lee (Dubin et al., 1983), which in turn was founded on earlier seminal work by Littlejohn. It has been thoroughly developed since then by such people as (in historical order) Hahm, Brizard, Qin, Sugama, and Scott. It uses a very different, more geometrical language (e.g., symplectic forms) and methodology (e.g., Lie transforms and variational principles). Catto and then Parra & Catto initially used the original two-scale expansion technique and furthermore discussed some of their results in terms of the language of neoclassical theory, where they have an extensive background. The technical language and body of known results can differ significantly between various physics subspecialties, so cross communication between the two approaches has been difficult. However, recently Parra and coworker Calvo have begun doing calculations with the modern Lagrangian methods. That has helped to provide a common ground for discussion.

In the end, however, the principle reason for past and, to some extent, continuing confusion is not differences in methodology. Rather, it is simply that gyrokinetics is subtle, both physically and mathematically, and the problem of momentum conservation is very difficult, involving the calculation and interpretation of nonlinear terms through several orders. There is no reason to expect that this should be trivial. Also, we should emphasize again that the original controversy referred to a specific set of assumptions. If those are violated, the conclusions must be revisited, although we do not attempt that here.

2.6 The utility of modern gyrokinetics

To repeat, it is useful to consider difficult problems in light of multiple approaches, and gyrokinetics is no exception. Certainly work based on the traditional two-scale approach to gyrokinetics has been very useful. That procedure uses techniques that are more familiar to many physicists and provides a down-to-earth way of looking at the physics that has led to many valuable insights, including much of the pioneering work in gyrokinetics. However, 'modern' gyrokinetics (Lagrangian methods, noncanonical variables, Lie transforms, *etc.*) can also be useful both conceptually and practically. The history of physics contains numerous examples of the importance of concise reformulations, which

are useful not only for enhancing physical understanding but also for ease of technical manipulations. Indeed, the very notion of gyrokinetics represents a radical change of point of view in which the particle is replaced by the gyrocenter as a fundamental entity. Specifically with regard to the modern approach[19]:

- Whereas the kinetic equation of Frieman & Chen (1982) buries the polarization effect in an obscure way,[20] the modern formulation brings it to the fore in the gyrokinetic Poisson equation, enabling one to make explicit connections to classical discussions of free and bound charge in electromagnetism. This is related to the specific meaning of a gyrocenter, which is given a precise[21] definition by the modern methods.

- The ability to use both noncanonical variables and Hamiltonian methods simultaneously is a huge technical convenience. Coordinate-independent formulations of Hamiltonian mechanics have been developed by such people as Arnold (1978). Fundamentally, this is possible because Newtonian mechanics (in its Hamiltonian formulation) preserves symplectic structure (Appendix E).

- Although one can work out perturbation expansions using various methodologies and eventually get to the correct answer, Lie transformations provide the most succinct (and, in principle, nonperturbative) approach to the choice of appropriate gyrocenter variables; their justification lies in the theory of flows on phase space and in other results from differential geometry, including the important concept of the Lie derivative (Sec. D.3).

- No matter what the approach, if the concept of gyrocenter is to be introduced at all one must face up to the relation between the particle distribution f and the gyrocenter distribution F. Although this can be obtained and discussed in various ways and is a simple and straightforward concept, in the mathematical language of modern differential geometry it is represented elegantly and concisely by the pullback transformation T^*: $f = \mathrm{T}^* F$ (Appendix D). The field of differential geometry provides a natural stage and the technical results useful to discuss pullbacks and other tools used in modern gyrokinetics. Physically, the pullback[22] gives precise meaning to the concept of polarization charge (Sec. 3.1.3).

[19] Some prior background is required in order to appreciate all of the nuances of these bulleted items. The material is developed later in the report, but is mentioned here in order to set the stage.

[20] For more explanation, see footnote 30 on p. 30.

[21] Actually, the transformation between particle and gyrocenter phase-space coordinates is not uniquely specified; some freedom exists that can be exploited for various purposes of technical convenience. This can lead to some subtleties; see, for example, footnote 54 on p. 56.

[22] Brizard (private communication, 2013) has stressed that polarization (and magnetization) are actually more closely related to the pushforward transformation T_*. As used for the practical manipulations in this report, the pushforward T_* is related to the pullback T^* by $\mathrm{T}_* = (\mathrm{T}^*)^{-1}$. For discussion, see the last paragraph of footnote 64 on p. 76.

- By preserving the Hamiltonian/Lagrangian structure of the dynamics and making approximations only to the Hamiltonian or Lagrangian, one ensures that key conservation properties are maintained exactly at each order in the expansion. This is not to imply that previous two-scale δF theories did not satisfy the proper conservation laws in the appropriate limit. In fact, they do, although one must go to next order in the expansion, namely to the transport equations, to obtain a complete set of conservation laws (Sugama & Horton, 1998; Abel et al., 2013). The Hamiltonian approaches ensure that conservation laws hold for full-F global simulations as well.

While this report cannot be a self-contained primer on differential geometry and related methods, we have provided a number of appendices that explain some of the key ideas that are referred to in the body of the paper. We hope those enhance the value of the report for newcomers to the field.

3 Review of gyrokinetic history and concepts

Gyrokinetics is built crucially on the adiabatic invariance of the magnetic moment μ. We do not discuss the history of adiabatic invariants here, but note an important paper by Taylor (1967) in which he demonstrated the adiabatic invariance of a modified magnetic moment even in the presence of a short-wavelength electrostatic perturbation. Calculations of linear gyrokinetic physics were made early on by such authors as Taylor & Hastie (1968) and Rutherford & Frieman (1968). A transformative paper by Catto (1978) "avoids the substantial mathematical complications inherent in [those] prior treatments by introducing the transformation to the guiding center variables and performing the guiding center gyrophase average before specifying the magnetic coordinates to be employed"; all methodology that we now know as gyrokinetics uses that approach. Nonlinear gyrokinetics dates from the early 1980's, the seminal paper being that of Frieman & Chen (1982). Thirty years later, the subject has evolved substantially with the advent of such modern techniques as noncanonical variables, Lie transforms, field-theoretic formulations, *etc.* In principle, none of those are necessary in order to address the basic ordering issues related to momentum transport, and their relative unfamiliarity may obscure rather than enlighten the subject; indeed, confusion has arisen because different groups have approached the problem in different ways and used different language. Since one of the purposes of this report is to describe the connections between the various methodologies, it is necessary to briefly review the fundamental concepts that will be referred to later. We do that in this section. Much, although not all, of the material is elementary; some more advanced topics are discussed in the appendices. We devote Sec. 4 to an exposition of momentum conservation in its various guises. With the background of Secs. 3 and 4 in hand, we are then prepared to discuss and answer in Sec. 5 the questions posed in the Charge. The reader interested only in the answers to the Charge questions and our discussion of the current state of affairs can skip Secs. 3 and 4 altogether and continue with Sec. 5 (p. 50).

A number of review articles develop the subject of gyrokinetics from different points of view. Garbet et al. (2010) focused on gyrokinetic simulations, whereas the articles of Krommes (2010) and Krommes (2012) are both more theoretical and aimed at audiences of non-plasma physicists. The technical methodology that underlies modern gyrokinetics was reviewed in detail by Brizard & Hahm (2007). Finally, many details of gyrokinetics as applied to rotating plasmas were described by Abel et al. (2013).

3.1 Fundamental concepts

Some elementary treatments of gyrokinetics approach it from the point of view of an average over one gyroperiod of a particle's motion. That provides useful intuition regarding the fact that a gyrocenter moves under the influence of an *effective* or gyroaveraged electric field, but disguises other important features of the formalism; because the charges and currents in Maxwell's equations are those of the real particles, gyrophase-*dependent* information is required in some places, so one must consider the deviation from the average as well. Note that a gyroperiod average is not well defined in general, as the projection of a particle's position onto a locally perpendicular plane does not close in the presence of magnetic inhomogeneity or electric fields. That suggests that most formalisms that exploit the rapid gyration of a particle will be inherently perturbative, since the averaging operation must be defined with respect to the lowest-order motion. Indeed, modern gyrokinetics proceeds by implementing a variable transformation that is developed perturbatively (asymptotically). Of course, a mere change of variables cannot change the physics content of the original Vlasov–Poisson system. In order to focus on just the low-frequency fluctuations of interest, a gyrokinetic closure must be made as well, analogous to the statistical closure approximations familiar in turbulence theory (Krommes, 2002); we will discuss it in Sec. 3.2.

3.1.1 The magnetic moment

A useful set of lowest-order variables is obtained from an action–angle transformation applied to circular motion. The lowest-order action is defined by[23]

$$J_0 \doteq \frac{1}{2\pi} \oint \boldsymbol{p} \cdot d\boldsymbol{q} \tag{11}$$

over one period of the lowest-order motion The canonical momentum for a charged particle is (Goldstein, 1951)

$$\boldsymbol{p} = m\boldsymbol{v} + \frac{q}{c}\boldsymbol{A}(\boldsymbol{x}). \tag{12}$$

[23] Here q is used as both a particle charge (scalar) and canonical coordinate (vector).

First consider the kinetic part:

$$J_{0,K} = \frac{1}{2\pi} m \oint \boldsymbol{v} \cdot d\boldsymbol{x} \tag{13a}$$

$$= \frac{1}{2\pi} m \oint v_\perp^2 \, dt \tag{13b}$$

$$= \frac{1}{2\pi} m v_\perp^2 \left(\frac{2\pi}{|\omega_c|} \right) \tag{13c}$$

$$= m v_\perp^2 / |\omega_c|. \tag{13d}$$

For the vector-potential part, write $\boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{A}(\boldsymbol{X} + \boldsymbol{\rho}) \approx \boldsymbol{A}(\boldsymbol{X}) + \boldsymbol{\rho} \cdot \boldsymbol{\nabla} \boldsymbol{A}(\boldsymbol{X})$. The lowest-order term vanishes after integration over the closed path. It is easy to show that the first-order piece is $-\frac{1}{2}$ of the kinetic part. Thus, the lowest-order action is

$$J_0 \doteq J_{0,K} + J_{0,A} = \frac{1}{2} m v_\perp^2 / |\omega_c|. \tag{14}$$

This quantity is closely related to the magnetic moment μ' associated with the circulating current $\boldsymbol{j} = q\delta(\boldsymbol{x} - \widetilde{\boldsymbol{x}})\boldsymbol{v}$, where $\widetilde{\boldsymbol{x}}$ describes circular motion involving the gyroradius vector $\boldsymbol{\rho} \doteq \widehat{\boldsymbol{b}} \times \boldsymbol{v}/\omega_c$:

$$\mu' \doteq \frac{1}{2} c^{-1} \widehat{\boldsymbol{b}} \cdot \int d\boldsymbol{x} \, \boldsymbol{x} \times \boldsymbol{j} = \frac{1}{2} c^{-1} q \boldsymbol{\rho} \times \boldsymbol{v} = -\frac{1}{2} q c^{-1} v_\perp^2 / \omega_c = -\frac{m v_\perp^2}{2B}. \tag{15}$$

The last result is charge-independent. Conventionally the minus sign is dropped and one defines the (lowest-order) 'magnetic moment' as $\mu \doteq m v_\perp^2 / 2B$. (Of course, the sign matters when one is discussing the plasma magnetization.) Technically it is often more convenient to use a quantity whose dimensions (indicated by brackets) are literally those of an action, i.e., [action] = [energy] / [frequency], so an alternate definition is

$$\mu \doteq \frac{1}{2} m v_\perp^2 / |\omega_c|. \tag{16}$$

This is, of course, just J_0, but the notation μ seems to be more common. In this report, we shall use the definition (16), but since we will be considering ions we will drop the absolute value.

There is some ambiguity about whether ω_c should be evaluated at the particle position $\boldsymbol{x}$ or the gyrocenter position $\boldsymbol{X}$. The previous derivation defines μ as a property of the gyrocenter. However, in the systematic treatments it is often simpler to begin with the particle variables, then to deduce the transformation $\boldsymbol{X} \approx \boldsymbol{x} - \boldsymbol{\rho}$ by proceeding through first order in $\epsilon \doteq \rho/L$. Thus, the complete set of lowest-order gyrocenter variables (sometimes called the guiding-center variables) can be taken to be

$$\boldsymbol{z} \doteq \{\boldsymbol{x}, U, \mu, \zeta\}, \tag{17}$$

where $U \doteq \boldsymbol{v} \cdot \widehat{\boldsymbol{b}} \equiv v_\parallel$, $\mu \doteq \frac{1}{2} m v_\perp^2 / \omega_c(\boldsymbol{x})$, and ζ is the gyrophase angle. ζ is defined to increase for clockwise (ion) motion, so $\zeta = \tan^{-1}(-v_y/v_x)$; see Fig. 1. The perpendicular velocity vector can be written in terms of a rotating unit vector $\widehat{\boldsymbol{c}}$ according to $\boldsymbol{v}_\perp = v_\perp \widehat{\boldsymbol{c}}(\zeta)$; then one defines $\widehat{\boldsymbol{a}} \doteq \widehat{\boldsymbol{b}} \times \widehat{\boldsymbol{c}}$ such that $\boldsymbol{\rho} = \rho \widehat{\boldsymbol{a}}$, where $\rho \doteq v_\perp / \omega_c$.

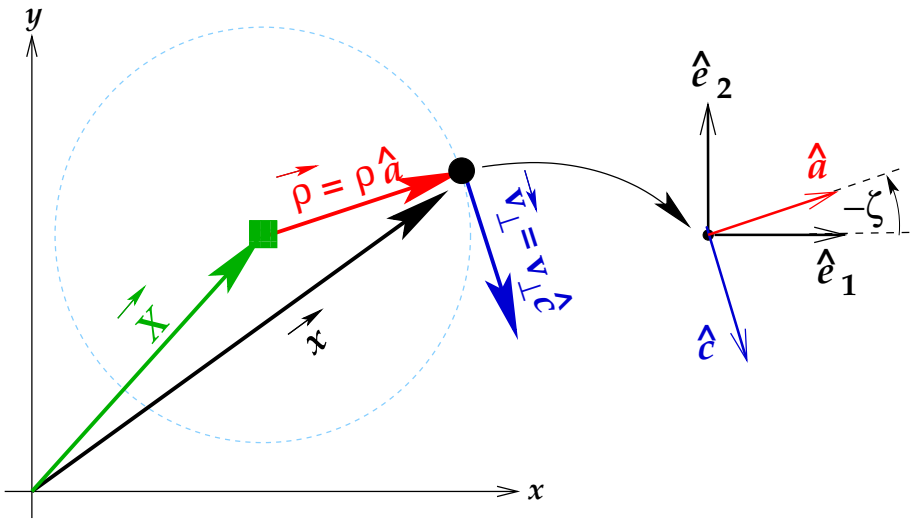

Figure 1: Illustration of the lowest-order gyrocenter coordinates. The magnetic field is $\boldsymbol{B} = B\widehat{\boldsymbol{b}}$; at the position $\boldsymbol{x}$ of the particle (black dot), $\boldsymbol{B}$ is in the z direction. The particle velocity is $U\widehat{\boldsymbol{b}} + v_\perp \widehat{\boldsymbol{c}}$; if $\boldsymbol{B}$ were constant, the particle would circle a gyrocenter at position $\boldsymbol{X} \doteq \boldsymbol{x} - \boldsymbol{\rho}$ (green square) with angular velocity $\dot{\zeta} = \omega_c$, where the gyroradius vector is $\boldsymbol{\rho} \doteq \widehat{\boldsymbol{b}} \times \boldsymbol{v}_\perp / \omega_c = \rho \widehat{\boldsymbol{a}}$ ($\rho \doteq v_\perp / \omega_c$). Instead of resolving vectors onto the orthonormal triad $(\widehat{\boldsymbol{a}}, \widehat{\boldsymbol{b}}, \widehat{\boldsymbol{c}})$, it is frequently convenient to use a local Cartesian system $(\widehat{\boldsymbol{e}}_1, \widehat{\boldsymbol{e}}_2, \widehat{\boldsymbol{b}})$ centered on the magnetic field line at the position of the particle. The lowest-order magnetic moment is $\mu \doteq \frac{1}{2} m v_\perp^2 / \omega_c$, and the lowest-order gyrocenter coordinates are $\{\boldsymbol{X}, U, \mu, \zeta\}$. This figure was originally published as Fig. 1 of the review article of Krommes (2012). Subsequently, it appeared on Sheldon's whiteboard in the January 31, 2013 episode of the TV series The Big Bang Theory entitled "The Cooper/Kripke Inversion."

By construction, μ is conserved for pure circular motion (which occurs only for $B = \text{const}$ and $\boldsymbol{E} = \boldsymbol{0}$). More profoundly, general results from the theory of adiabatic invariants suggest that a quantity $\overline{\mu} \doteq \mu + \cdots$ will be conserved through all orders[24] if quantities seen by the gyrocenter change slowly in time. It is convenient to use $\overline{\mu}$ rather than μ as a variable because then $\overline{\mu}$ will just

[24] "Conserved through all orders" does not mean that $\overline{\mu}$ is conserved exactly. Resonances between gyromotion and high harmonics of bounce motion (related to the second adiabatic invariant) can destroy $\overline{\mu}$ conservation, as discussed in some detail by Dubin & Krommes (1982). Mathematically, the issue is closely related to the fact that the asymptotic expansion of the nonzero function $\exp(-1/\epsilon)$ vanishes.

appear as a parameter in the gyrocenter distribution function $\overline{F}$; in the kinetic equation for $\overline{F}$, no derivative with respect to $\overline{\mu}$ will appear. A basic part of the gyrokinetic program is thus to construct the transformation $\overline{\boldsymbol{z}} = \mathrm{T}\boldsymbol{z}$ such that $\overline{\mu} = \mathrm{T}\mu$ is one of the new variables.[25] That is conventionally done in terms of Lie transforms, discussed in the next section. A simple example of the construction of $\overline{\mu}$ is given in Appendix F.

As indicated in Fig. 1, the particle motion is usually resolved onto a set of local orthonormal unit vectors $\widehat{\boldsymbol{e}}_1$, $\widehat{\boldsymbol{e}}_2$, and $\widehat{\boldsymbol{b}}$. Concern has been raised that in some magnetic field configurations $\widehat{\boldsymbol{e}}_1$ and $\widehat{\boldsymbol{e}}_2$ may not exist (Sugiyama, 2008), which calls into question the asymptotic construction of $\overline{\mu}$ and the proof of its adiabatic invariance. The issue was formally discussed in a published exchange between Krommes (2009) and Sugiyama (2009), but a consensus was not reached. However, Burby & Qin (2012) have recently given an authoritative and solid mathematical analysis that has laid the matter to rest. The conclusion is that the unit vectors always exist in a torus. In some other configurations they may not exist (although for reasons different than those considered by Sugiyama); however, the proof of asymptotic adiabatic invariance remains intact.

3.1.2 Lie transforms

Newton's laws of motion derive from a Hamiltonian, and that fundamentally geometric property (Arnold, 1978) is not lost under the gyrokinetic change of variables. Thus not only does one derive a new set of variables $\overline{\boldsymbol{z}}$, one derives a *gyrocenter Hamiltonian* from which the collisionless gyrocenter equations of motion can be obtained from standard Poisson-bracket relations (see Appendix E). Those equations of motion are encapsulated in the *gyrokinetic equation* for the gyrocenter distribution $F(\overline{\boldsymbol{z}})$.

In traditional Hamiltonian theory, changes of variables are *canonical* (i.e., the Poisson brackets between the coordinates q^i and momenta p_j obey $\{q^i, p_j\} = \delta^i_j$) and determined by various kinds of *mixed-variable generating functions*. However, the use of those functions becomes highly awkward already at second order in perturbation theory because of the need to untangle the implicit definitions of the new variables in terms of the old ones. Modern theory employs two major technical advances: (i) one need not work with canonical variables (see Appendix E); (ii) *Lie transformations* provide explicit formulas for the new variables in terms of the old ones. Because the Lie transformation can be written in a covariant fashion, it can be applied to both canonical and noncanonical variables.

Essentially, Lie transforms do for perturbation theory what propagators do for temporal evolution. The solution of the nonlinear system of ODEs $\dot{\boldsymbol{x}} = \boldsymbol{V}(\boldsymbol{x}, t)$ can be represented as $x^i(t) = \mathrm{U}(\boldsymbol{x}, t)x^i$; the nonlinear operator $\mathrm{U}(\boldsymbol{x}, t) \equiv$

[25] As a side note, it is in principle possible to represent all of the gyrocenter physics in terms of the lowest-order variables $\boldsymbol{z}$ instead of $\overline{\boldsymbol{z}}$. One reason that is not usually desirable is that then the gyrokinetic equation would contain a ∂_μ term since the lowest-order μ is not conserved.

$U(t)$ propagates the initial conditions forward to their final value. Analogously, Lie transforms propagate variables in a perturbation parameter ϵ, which replaces time t. Proceeding analogously to the construction of $\mathrm{U}(t)$, one finds that the basic Lie transformation operator is $\mathrm{T}(\epsilon) = \exp[w(\epsilon)]$. Here $w(\epsilon)$ is a vector field (a differential operator; see Sec. A.6): $w \doteq w^\nu \partial_\nu$. In perturbation theory, w could be expanded in powers of ϵ. More commonly, however, a procedure advocated by Dragt & Finn (1976) is followed, wherein T is constructed as the compound transformation

$$\mathrm{T} = \ldots e^{\mathcal{L}_3} e^{\mathcal{L}_2} e^{\mathcal{L}_1}, \tag{18}$$

where $\mathcal{L}_n \equiv \epsilon^n w_n$. The w_n are determined order by order in such a way that $\overline{\mu}$ is conserved.

It must be clearly understood that $\overline{\mu}$ conservation is a physical property of the particle motion, not of any particular mathematical representation. Thus, although the gyrokinetic equation is particularly clean in that no $\partial_{\overline{\mu}}$ appears, the consequences of $\overline{\mu}$ conservation permeate representations of the physics in terms of particle variables as well, including the real-space momentum conservation law.

3.1.3 Polarization

For purely circular motion, gyrocenters move only in the parallel direction. More generally, they also drift slowly across the magnetic field with the effective electric and magnetic drifts. The formalism will focus on the gyrocenters as the basic entities. From the point of view of macroscopic electrodynamics, gyrocenters therefore behave as *free charge*, and the average deviation of the gyrating particles from the gyrocenters can be considered to be *bound charge* (bound to the gyrocenters). It is therefore useful to replace Poisson's original equation $\boldsymbol{\nabla} \cdot \boldsymbol{E} = 4\pi\rho$ with the macroscopic equation $\boldsymbol{\nabla} \cdot \boldsymbol{D} = 4\pi\rho^{\mathrm{free}} \equiv 4\pi\rho^G$, where $\boldsymbol{D} = \boldsymbol{E} + 4\pi\boldsymbol{\mathcal{P}}$ and $\boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{P}} = -\rho^{\mathrm{bound}} \equiv -\rho^{\mathrm{pol}}$. Here pol stands for polarization. This notation is used because the bound charge can polarize (slightly move away from the gyrocenter) under various effects, such as in response to a slowly time-varying electric field. Under the assumption of *quasineutrality*, which will be justified in the next paragraph, Poisson's equation therefore reduces to a balance between the free gyrocenter charge and the polarization charge due to the gyrating particles: $0 \approx \rho^G + \rho^{\mathrm{pol}}$. Because the electron gyroradius is very small, the polarization effect is negligible for electrons and is due entirely to the ions.

Many of the difficulties and conceptual issues associated with modern gyrokinetics are associated with the proper calculation (and approximation) of ion polarization through several orders in ϵ. However, it is easy to hand-wave the lowest-order effect. Take as given the result that in slowly varying fields the

polarization or inertial drift of an ion fluid is[26]

$$\boldsymbol{u}^{\rm pol} = \frac{1}{\omega_{ci}} \frac{\partial}{\partial t} \left(\frac{c}{B} \boldsymbol{E}_\perp \right). \tag{19}$$

That drift leads to a *polarization charge density* $\rho^{\rm pol}$ according to the continuity equation $\partial_t \rho^{\rm pol} + \boldsymbol{\nabla} \cdot (\boldsymbol{u}^{\rm pol} \rho^{\rm pol}) = 0$. Upon linearizing this around a background state with no field and removing the partial time derivatives that appear in both terms, one finds that

$$\rho_i^{\rm pol} = \boldsymbol{\nabla} \cdot \left(\frac{1}{4\pi} \chi_\perp \boldsymbol{\nabla}_\perp \phi \right). \tag{20}$$

Here $\chi_\perp$ is the perpendicular dielectric permittivity[27]

$$\chi_\perp \doteq \frac{\rho_{\rm s}^2}{\lambda_{De}^2} = \frac{\omega_{pi}^2}{\omega_{ci}^2}. \tag{21}$$

Many plasmas have $\chi_\perp \gg 1$,which leads to some very useful properties (Krommes et al., 1986). When that is satisfied,[28] one has $4\pi \boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{P}} \gg \boldsymbol{\nabla} \cdot \boldsymbol{E}$, i.e., $\boldsymbol{D} \approx 4\pi \boldsymbol{\mathcal{P}}$, or $-\rho^{\rm pol} \approx \rho^G$; this is the *quasineutrality* approximation $\rho^G + \rho^{\rm pol} \approx 0$ that will be used in the remainder of the discussion.[29]

[26] For a single particle, this follows from a simple iteration of Newton's second law of motion; nonlinear corrections are omitted here. [Some early discussions of polarization drift are given by Spitzer (1962) and Chandrasekhar (1960).] For a statistical distribution of ions, this is correct only in the cold-ion limit $T_i \to 0$, i.e., in the absence of finite-Larmor-radius (FLR) effects.

[27] Yet another way of writing the permittivity is $\chi_\perp = c^2/c_A^2$, where $c_A \doteq [B^2/4\pi(nm)_i]^{1/2}$ is the Alfvén speed. However, since we only discuss electrostatics in this report, we will have no need to introduce the Alfvén speed.

[28] Most fusion and many other plasmas satisfy $\chi_\perp \gg 1$ (a typical value is $\chi_\perp \sim 10^3$), so quasineutrality is usually a good approximation on scales large compared to the tiny Debye length. However, since the non-quasineutral corrections to the quasineutrality equation are of order $\chi_\perp^{-1} \sim 10^{-3}$, if one were to try to calculate momentum transport in the low-flow, gyro-Bohm regime using the standard approach, which requires keeping corrections to the dominant polarization that are of a factor of $\epsilon \sim \rho/L$ smaller, then one might need to relax the quasineutrality assumption as well, unless ϵ is significantly larger than $\chi_\perp^{-1}$.

[29] A frequently voiced paradox relating to the quasineutrality condition raises the question, "If the particle charge density ρ vanishes, doesn't one conclude from Poisson's equation that there can be no electric field?" The answer is no, as can be seen from consideration of the simple equation

$$\epsilon x = 1 - x, \tag{22}$$

where ϵ is small. The solution is $x \approx 1$. However, from that approximate solution, which annihilates the right-hand side of Eq. (22), one cannot conclude that the resulting statement $\epsilon x \approx 0$ leads to $x = 0$. From $\epsilon x \approx 0$, which is an approximation, not an equality, one can conclude only that x is the indeterminate ratio of two small terms, so it is not a useful way of solving for x. In plasma physics, x is analogous to the potential ϕ, which frequently appears as one piece of the charge density (for example, when electrons have approximately adiabatic response). Thus, when $\nabla^2 \phi$ is small the potential is dominantly determined by setting the right-hand side of Poisson's equation to zero. Corrections to that solution can be determined by iteration. In gyrokinetics, the ϵ used in this footnote is not the gyrokinetic expansion parameter, but rather $\chi_\perp^{-1}$, reflecting the dominance of ion polarization.

3.2 Derivation of the gyrokinetic–Poisson system

The early works on gyrokinetics, including the nonlinear gyrokinetic equation of Frieman & Chen (1982),[30] represented polarization in an awkward form that disguised the significance of the gyrocenters and the interpretation in terms of free and bound charge; as observed by Frieman & Chen, polarization was "contained, not very obviously, within the finite Larmor radius corrections in [the Bessel function] $J_0 \ldots$." Here $J_0 \equiv J_0(k_\perp v_\perp/\omega_c)$; it represents the reduction of the effective potential felt by the gyrocenter due to gyroaveraging of the motion of a particle with nonzero gyroradius. The appearance of J_0 in connection with a discussion of polarization is confusing because *polarization is fundamentally not a finite-Larmor-radius effect*. Expressions (19), (20), and (21) do not involve the ion temperature; they remain nonzero for cold ions, for which $J_{0,i} = 1$. Basic polarization is extracted from the Frieman–Chen formulation as a limiting

[30] The kinetic equation used by Frieman & Chen, as well as many subsequent authors, is written not in terms of the gyrocenter PDF but rather for the quantity h in the decomposition of the fluctuating Vlasov distribution

$$\delta f_s = -\frac{q_s \delta\phi}{T_s} F_{0s} + h_s, \tag{23}$$

where F_0 is the background distribution and h_s has no dependence on gyroangle. That is, h represents the non-Boltzmann response. The derivation of the equation for h is described in great detail by Abel et al. (2013). They discuss (p. 71) how the h equation contains polarization as follows:

> "Unfortunately, due to our choice of decomposition ... for δf_s, [polarization] is not apparent from [the quasineutrality condition
>
> $$\sum_s \frac{Z_s^2 e^2 n_s \delta\phi}{T_s} = \sum_s Z_s e \int d^3\boldsymbol{w}\, \langle h_s \rangle_{\boldsymbol{r}}. \tag{24}$$
>
> Here $\boldsymbol{r}$ and $\boldsymbol{w}$ are the particle position and velocity, respectively.] If, instead of using h_s, we work with $g_s = \langle \delta f_s \rangle_{\boldsymbol{R}} = h_s - Z_s e F_{0s} \langle \delta\varphi \rangle_{\boldsymbol{R}}/T_s$ [Lee (1983); $\boldsymbol{R}$ is the gyrocenter position], then the quasineutrality condition becomes
>
> $$\sum_s \frac{Z_s^2 e^2 n_s (1 - \Gamma_{0s}) \delta\phi}{T_s} = \sum_s Z_s e \int d^3\boldsymbol{w}\, \langle g_s \rangle_{\boldsymbol{r}}, \tag{25}$$
>
> where Γ_{0s} is an operator defined by $\Gamma_{0s}\delta\phi = (1/n_s) \int d^3\boldsymbol{w}\, \langle \langle \delta\phi \rangle_{\boldsymbol{R}} \rangle_{\boldsymbol{r}} F_{0s}$. In this formulation, the left-hand side of this equation is precisely the density of polarization charge (i.e. that density which accumulates because $\boldsymbol{\nabla} \cdot \boldsymbol{V}_{\rm pol} \neq 0$). Indeed, if $k_\perp \rho_i$ is small, then [for $\Gamma_e = 1$ and $\Gamma_i \equiv \Gamma$] $1 - \Gamma_0 \approx (1/2)\rho_i^2 \nabla_\perp^2$ and this correspondence becomes obvious (the quasineutrality condition now looks like Poisson's equation for $\delta\phi$, with an enhanced permittivity).
>
> "The difference between the two formulations [in terms of either h or the gyrocenter distribution F] is purely interpretative; we interpret [the h equation] and the field equations as describing the dynamics of physically extended rings of charge moving in a vacuum, whereas the approach emphasising the polarisation drift interprets gyrokinetics as describing a gas of point-particle-like gyrocenters (with distribution function g_s) in a polarisable vacuum, with the above form of the quasineutrality condition playing the role of Poisson's equation. For more discussion of this second interpretation see [Krommes (2012)] and references therein."

procedure in which the smallness of the FLR corrections to $J_{0,i} \approx 1$ is cancelled by dividing by small ion temperature in the limit as $T_i \to 0$. (Of course, there are true FLR corrections to that basic limit.) The formalism of Frieman & Chen focuses on rings of particle charge, not gyrocenters, as discussed in more detail in footnote 30 on p. 30.

An important advance was made by Lee (1983), who produced a gyrokinetic equation in a more conventional form. He was initially motivated by practical considerations relating to the goal of formulating a particle-in-cell gyrokinetic simulation scheme. To that end, he needed a kinetic equation in characteristic form, and he also required an efficient way of calculating the potential under the quasineutrality constraint. If that constraint is written in terms of the particle charge density ρ, the equation $\rho = 0$ defines the potential ϕ implicitly and is difficult to handle numerically. But if the polarization charge density is split off and balanced with the gyrocenter charge density, one obtains a Poisson equation for ϕ that is much easier to solve. Practical considerations aside, Lee's formulation amounts (as was only fully appreciated later) to a shift in focus to the dynamics of a gyrocenter, so it yields to the interpretation in terms of free and bound charge given above — a powerful argument in its favor. By displaying the polarization effect explicitly in Poisson's equation as he did, it becomes clear how it survives in the cold-ion limit in which all FLR effects are neglected and all J_0's approach 1. Although there may be certain technical advantages to the Frieman–Chen form of the gyrokinetic equation, and it is used in various extant codes, the physics of the situation is best brought out by the gyrocenter representation. Furthermore, the gyrocenter interpretation emerges naturally from the modern Hamiltonian formalism, as is described below.

Dubin et al. (1983) reconsidered Lee's (collisionless) work from the point of view of Hamiltonian dynamics. They used noncanonical variables and Lie transforms, and they worked with an extended (8D) phase space. In order to focus on the self-consistent determination of the potential ϕ, they considered an electrostatic slab model with constant magnetic field. Let $\widetilde{F}(\overline{\boldsymbol{X}}, \overline{U}, \overline{\mu}, \overline{\zeta}, t)$ be the gyrocenter distribution function; the tilde denotes dependence on the gyrophase angle $\overline{\zeta}$.[31] Dubin *et al.* found the kinetic equation[32]

$$\frac{\partial \widetilde{F}}{\partial t} + \overline{U}\widehat{\boldsymbol{b}} \cdot \overline{\boldsymbol{\nabla}}\widetilde{F} + \overline{\boldsymbol{V}}_E[\psi] \cdot \overline{\boldsymbol{\nabla}}\widetilde{F} + \frac{q}{m}\overline{E}_\parallel[\psi]\frac{\partial \widetilde{F}}{\partial \overline{U}} - \omega_c \frac{\partial \widetilde{F}}{\partial \overline{\zeta}} = 0. \tag{26}$$

Here $\boldsymbol{E}[\phi] \doteq -\overline{\boldsymbol{\nabla}}\phi$ and $\overline{\boldsymbol{E}}[\phi] \equiv \boldsymbol{E}[\overline{\phi}]$, $\overline{\phi}$ denoting the gyroaveraged potential. The quantity ψ is the effective potential (the gyrocenter Hamiltonian divided by charge q), which Dubin *et al.* calculated through second order.

One must take care with the gyrokinetic Poisson equation. Poisson's equation is written in terms of the particle PDF f, but the GKE evolves the different

[31] In our previous discussion we merely wrote F for the quantity here called $\widetilde{F}$. Here we need to be more precise.

[32] The gyrokinetic equation including gyrophase dependence is not written in this specific form in the Dubin paper, but cf. the unnumbered equation between their Eqs. (16) and (17). Regarding unnumbered equations, see Mermin (1990, p. 69, Rule 1).

quantity $\widetilde{F}$. As discussed in Appendix D, the necessary change of variables that relates the particle and gyrocenter PDFs is represented succinctly as[33] $f = \mathrm{T}^* \widetilde{F}$, where T^* (an operator induced[34] from T) is called the *pullback transformation*.[35] One then finds that the particle charge density is[36]

$$\rho(\boldsymbol{x}, t) = \sum_s (\overline{n} q)_s \int \mathcal{J}\, d\overline{\boldsymbol{z}}\, \delta(\boldsymbol{x} - \overline{\boldsymbol{X}}) \mathrm{T}^* \widetilde{F}(\overline{\boldsymbol{z}}, t). \tag{27}$$

Here $\mathcal{J}$ is the Jacobian $\partial(\boldsymbol{x})/\partial(\overline{\boldsymbol{z}})$. The delta function was inserted in order that a complete integration could be performed over all six phase-space variables.

It is important to appreciate that $\boldsymbol{x}$ is a dummy variable in this equation. Equation (27) determines the charge density (and thus the potential) at any point in physical space. Therefore, there is no conceptual difficulty in using a gyrokinetic simulation to determine a laboratory-coordinate-system-based electric field.

Equation (27) is not yet the usual gyrokinetic Poisson equation because it involves the gyrophase-dependent PDF $\widetilde{F} \doteq \overline{F} + \Delta F$, where $\overline{F}$ does not depend on gyrophase but ΔF does. Indeed, if all one does is introduces a change of variables in the kinetic equation ($\overline{z} = \mathrm{T} z$, where for the present discussion we allow for the possibility that T includes effects beyond all orders in the gyrokinetic expansion parameter ϵ) and represents Poisson's equation using $f = \mathrm{T}^* \widetilde{F}$, *no physics has been lost*; the formalism still contains high-frequency physics and uses a 6D kinetic equation. In order to obtain a reduced formalism that is efficient for fluctuations with $\omega \ll \omega_{ci}$, a *gyrokinetic closure*[37] must be made.

[33] There are notational difficulties with the formula $f = \mathrm{T}^* \widetilde{F}$. The Vlasov distribution function is typically written as $f(\boldsymbol{x}, \boldsymbol{v}, t)$. To proceed with the transformation theory, one needs to write f in terms of lowest-order gyrocenter variables $\boldsymbol{z}$. In terms of those variables, the distribution has a different functional form, so one needs a new symbol. It would make some sense to write $f(\boldsymbol{x}, \boldsymbol{v}, t) = F(\boldsymbol{z})$ (the particle PDF expressed in a different set of variables), then to transform to $\overline{F}(\overline{\boldsymbol{z}})$; however, F is used in different ways by different authors, and uppercase may imply a gyrocenter PDF, so use of that symbol may be confusing. Therefore, we will just use f to stand generically for the particle PDF; it should be clear from the context on which variables that depends.

[34] If $T = \exp(w)$, where w is a vector field, then $\mathrm{T}^* = \exp(L_w)$, where L_w is the Lie derivative (Sec. D.3) in the direction of w.

[35] We hope that the reader is not intimidated by this technical term. We use the notation T^* and the word pullback merely as a shorthand that indicates that the particle distribution f and the gyrocenter distribution $\widetilde{F}$ have different functional forms; they express the same physics, but in different variables.

[36] In the work of Dubin *et al.* and in other places, the delta function that appears in Eq. (27) is instead written as $\delta(\boldsymbol{x} - \overline{\boldsymbol{X}} - \overline{\boldsymbol{\rho}})$ because a preliminary transformation to lowest-order gyrocenter variables was made (following early work of Littlejohn). Brizard (1990) has shown that the preliminary transformation is unnecessary; it can be obtained as part of the Lie transformation T. It is formally easier to carry all of the gyroradius corrections in T, so we shall not make the preliminary transformation.

[37] The phrase 'gyrokinetic closure' is used here in the same sense as 'statistical closure' is used in turbulence theory. In the latter (Krommes, 2002, and references therein), ensemble averages are used to reduce the enormous amount of information contained in the primitive amplitude equation; the price one pays is that an approximation must be made in order to close the theory in terms of a finite number of moments. Gyrokinetics is analogous, with

That means, first, constructing T through only a few orders in ϵ and expanding out resonance functions of the form $(\omega - n\omega_{ci})^{-1}$ ($|n| \geq 1$) by treating ω as small; and second, reducing the 6D phyics to 5D. That is easily possible by integrating the kinetic equation over gyrophase because T is constructed to make the characteristic velocities and accelerations independent of gyrophase; thus one arrives at a conventional gyrokinetic equation for the reduced, 5D distribution $\overline{F} \doteq \int_0^{2\pi} d\overline{\zeta}\, \widetilde{F}$. However, one must still face up to the presence of the gyrophase-dependent quantity ΔF in Poisson's equation. Traditionally ΔF has been dropped altogether. (This does *not* mean that all information about particle gyration is ignored; one still has gyroaveraged electric fields, and the differences between particle and gyrocenter positions show up as the polarization effect in Poisson's equation.) In collisionless theory, that can be justified by noting that one can find gyrophase-independent solutions of the kinetic equation (Dubin et al., 1983). This is not possible in the presence of collisions. Brizard (2004) pointed out that collisions drive a gyrophase-dependent correction $\Delta F = O(\nu/\omega_{ci})\delta F$, where δF is the deviation from a Maxwellian distribution [of course, this is born out by the collisional calculations of Parra & Catto (2008)]. Given the other assumptions already made in obtaining low-frequency gyrokinetics, this correction may be calculated perturbatively. Dropping it altogether is consistent when only first-order polarization terms are retained (see later discussion), but one may need to retain it when higher-order effects are of concern.[38] Indeed, the desire to develop a unified theory that includes both collisional and turbulence effects on equal footing was a key motivation for the work of Parra & Catto (2008) and Catto et al. (2008). For now, we follow the collisionless calculation of Dubin et al. (1983) by dropping ΔF and thus writing the gyrokinetic Poisson equation as

$$\rho(\boldsymbol{x}, t) = \sum_{\overline{s}} (\overline{n}q)_{\overline{s}} \int \mathcal{J}\, d\overline{\boldsymbol{z}}\, \delta(\boldsymbol{x} - \overline{\boldsymbol{X}}) \mathrm{T}^* \overline{F}(\overline{\boldsymbol{z}}, t). \tag{28}$$

We now have a closed gyrokinetic–Poisson system that determines $\overline{F}$ and the potential ϕ.

One can proceed from Eq. (28) to develop a formally exact expression for the polarization; this will be useful in later manipulations that lead to the gyrocenter momentum conservation law (see Sec. I.2). Following Brizard, we write Eq. (28)

gyrophase average replacing ensemble average and the goal of closure being a closed system that determines the gyrocenter PDF $\overline{F}$. In turbulence theory, closures sometimes have negative connotations because in situations of strong turbulence they can entail errors of order unity and/or be qualitatively wrong. However, in situations involving small autocorrelation time statistical closures can become asymptotically exact, and that situation is more analogous to gyrokinetics, which exploits an asymptotic expansion in ϵ.

[38] In other words, while the lowest-order collisional effects have been included in gyrokinetics in various ways [see, for example, Dimits & Cohen (1994), Sugama & Horton (1998), Abel et al. (2013), Li & Ernst (2011), and references therein], if one pursues momentum transport in the low-flow gyro-Bohm regime by including second-order terms in the polarization, then one may also need to keep higher-order collisional corrections related to the gyrophase dependence of F.

as

$$\rho(\boldsymbol{x},t) = \sum_{\overline{s}} (\overline{n}q)_{\overline{s}} \int \mathcal{J}\, d\overline{\boldsymbol{z}}\, [(\mathrm{T}^*)^{-1}\delta(\boldsymbol{x}-\overline{\boldsymbol{X}})]\overline{F}(\overline{\boldsymbol{z}},t). \tag{29a}$$

$$= \sum_{\overline{s}} (\overline{n}q)_{\overline{s}} \int d\overline{\boldsymbol{x}}\, \mathcal{J}\, \delta(\boldsymbol{x}-\overline{\boldsymbol{X}}-\overline{\boldsymbol{\rho}}_\epsilon)\overline{F}, \tag{29b}$$

where[39] $\boldsymbol{\rho}_\epsilon \doteq (\mathrm{T}^*)^{-1}\boldsymbol{X} - \boldsymbol{X}$. The delta function can be expanded in small $\overline{\boldsymbol{\rho}}_\epsilon$:

$$\delta(\boldsymbol{x}-\overline{\boldsymbol{X}}-\overline{\boldsymbol{\rho}}_\epsilon) = \delta(\boldsymbol{x}-\overline{\boldsymbol{X}}) - \boldsymbol{\nabla}\cdot[\overline{\boldsymbol{\rho}}_\epsilon\delta(\boldsymbol{x}-\overline{\boldsymbol{X}})] + \frac{1}{2}\boldsymbol{\nabla}\boldsymbol{\nabla} : [\overline{\boldsymbol{\rho}}_\epsilon\,\overline{\boldsymbol{\rho}}_\epsilon\delta(\boldsymbol{x}-\overline{\boldsymbol{X}})] + \cdots . \tag{30}$$

The zeroth-order term gives rise to the gyrocenter charge density

$$\rho^G \doteq \sum_{\overline{s}} (\overline{n}q)_{\overline{s}} \int d\overline{\boldsymbol{p}}\, \overline{F} \tag{31}$$

(the Jacobian is subsumed in the formal $d\boldsymbol{p}$ momentum integration), and the remaining terms can be interpreted as $-\boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{P}}$, where the polarization vector is

$$\boldsymbol{\mathcal{P}} \doteq \sum_{\overline{s}} (\overline{n}q)_{\overline{s}} \int d\overline{\boldsymbol{p}}\, \langle\overline{\boldsymbol{\rho}}_\epsilon\rangle\overline{F} - \frac{1}{2}\boldsymbol{\nabla}\cdot\left(\sum_{\overline{s}} (\overline{n}q)_{\overline{s}} \int d\overline{\boldsymbol{p}}\, \langle\overline{\boldsymbol{\rho}}_\epsilon\,\overline{\boldsymbol{\rho}}_\epsilon\rangle\overline{F}\right) + \cdots . \tag{32}$$

(Here the angle brackets denote an average over gyrophase $\overline{\zeta}$; that is legitimate because no other quantities depend on $\overline{\zeta}$ after the gyrokinetic closure has been made.)

Note that some of these correction terms are present even in the absence of an electric field. The familiar polarization (20) vanishes with the potential. That agrees with the common intuition that macroscopic polarization exists only in the presence of an electric field. However, to the extent that there are pressure-related or other contributions to polarization that persist even in the absence of potential, that intuition must be incomplete. Since generations of plasma physicists have learned the physics of polarization from Jackson's well-known treatise on electrodynamics, it is worth quoting from his discussion; we will find that he actually covered the general situation! Quoting specifically from Jackson's first edition (Jackson, 1962, p. 108),

> In the absence of external fields, atoms or molecules may or may not have electric dipole moments, but if they do, the moments are randomly oriented. In the presence of a field, the atoms become polarized (or their permanent moments tend to align with the field) and possess on the average a dipole moment. These dipole moments can contribute to the averaged charge density… *In the absence of a field there is no average polarization.*[a]

[a] Except for *electrets*, which have a permanent electric polarization.

[39] Here the pushforward operator $(\mathrm{T}^*)^{-1}$ is understood to act on the Cartesian components of $\boldsymbol{X}$. For further discussion of this point, see the last paragraph of footnote 64 on p. 76.

(italics added). Note his crucial footnote,[40] which allows that some materials known as electrets (see Wikipedia:Electret) can have a permanent polarization. In this sense, *the magnetized plasma at nonzero temperature is an electret.*

3.3 Symplectic dynamics and the 1-form method

The methodology used by Dubin et al. (1983) relied on Darboux's theorem[41] to justify and implement the construction of semi-canonical variables; in particular, the magnetic moment $\overline{\mu}$ and the gyrophase $\overline{\zeta}$ were constructed to be canonically conjugate. That method did not fully exploit the symplectic nature of the dynamics or the technical simplifications that follow from the method of differential forms. As explained by Littlejohn (1983) and Cary & Littlejohn (1983), it is most natural to represent the dynamics in terms of the fundamental Poincaré–Cartan (differential) 1-form (for background information on differential forms, see Appendix C) $\gamma \doteq \boldsymbol{p}\cdot\boldsymbol{dq} - H\,\boldsymbol{d}t$, in terms of which Hamilton's variational principle is written as $\delta\int\gamma = 0$. Because the value of the particle action is indifferent to the particular variables used in its evaluation, γ may be written in terms of any variables whatsoever; then the equations of motion for the new variables follow as the new Euler–Lagrange equations.

There are several ways of representing this theory covariantly. First let us follow Cary & Littlejohn (1983) and arrange the phase-space coordinates for a single particle according to

$$z^i \doteq \{q^1, q^2, q^3, p_1, p_2, p_3\} \quad (i = 1, \ldots, 6), \tag{33}$$

then introduce the extended phase-space coordinates[42]

$$z^\mu = \{t, z^i\} \quad (\mu = 0, \ldots, 6). \tag{34}$$

Then the (covariant) coordinates of the canonical 1-form are

$$\gamma_\mu = \{-H, p_1, p_2, p_3, 0, 0, 0\} \tag{35}$$

and one has the manifestly covariant form

$$\boldsymbol{\gamma} = \gamma_\mu dz^\mu. \tag{36}$$

It is shown in Appendix C that the Euler–Lagrange equations are

$$\omega_{\mu\nu}\frac{dz^\nu}{d\lambda} = 0, \tag{37}$$

[40] This footnote seems to have vanished from subsequent editions.

[41] Darboux's theorem (in symplectic geometry) states very loosely that in a local region of a symplectic manifold (Appendix E) one can construct a change of variables that brings the symplectic form ω to its canonical form. Dubin *et al.* followed Littlejohn (1979) in finding a *semi-canonical* form involving the term $\overline{\mu}\,d\overline{\zeta}$.

[42] Phase-space coordinates are superscripted (contravariant); the coordinates of forms are subscripted (covariant). See Appendix A.

where λ parametrizes the trajectories (λ could be the time t) and

$$\omega_{\mu\nu} \doteq \partial_\mu \gamma_\nu - \partial_\nu \gamma_\mu. \tag{38}$$

The $\omega_{\mu\nu}$ are the coefficients of the differential 2-form $\omega \doteq d\gamma$ (see Sec. C.3).

Another, yet more elegant way of proceeding is to work in an 8D extended phase space possessing space-time coordinates $\{t, x^1, x^2, x^3\}$ and momenta $\{H, p_1, p_2, p_3\}$. Trajectories are then parametrized by an auxiliary time-like parameter τ (which from the extended equations of motion turns out to be just the time t). That is the approach used in Brizard's variational formulation (see Appendix G).

3.4 Gyrokinetic field theory

To the extent that one formulates the gyrokinetic–Poisson (or Maxwell) system in terms of some truncation of the pullback transformation, one is faced with the fundamental uncertainty of how to ensure that the truncations are consistent between the gyrokinetic equation and the Poisson equation. 'Consistent' means that one should be able to demonstrate exact conservation laws for the truncated system. Although one may be able to find such laws by experimentation [for several important examples, see Dubin et al. (1983)], the trial-and-error approach is fundamentally unsatisfactory. A major advance was achieved when Sugama (2000) and Brizard (2000) formulated gyrokinetics as a Lagrangian field theory. In this method, both the collisionless gyrokinetic equation and *a consistent set of* gyrokinetic Maxwell equations are obtained by appropriate variational derivatives of a gyrokinetic action *functional* $\mathcal{A}[\overline{F}, \varphi]$ (not to be confused with the single-particle action). Provided that approximations are made only on $\mathcal{A}$, one can use powerful Noether methods (reviewed in Appendix H) to both prove that exact conservation laws exist in the presence of particular symmetries and simply obtain the form of those conservation laws. For present purposes, the two most important examples are conservation of energy (for a time-stationary background) and conservation of toroidal angular momentum (for a background state that is independent of toroidal angle φ). We will be particularly concerned with the latter law, which was first derived by Scott & Smirnov (2010) and is discussed in detail in Sec. 5.3 and Appendixes I and J. Scott has argued strongly that modern gyrokinetics should be *defined* by its representation as a field theory.

An important criticism of the variational approach is that it naturally generates *collisionless* gyrokinetics; a way of embedding collisional effects directly into the action is not necessarily apparent. There has been some limited progress on methodologies that include dissipation: Kaufman (1984) and Morrison (1984) have discussed *dissipative brackets*, various kinds of statistical closures (Krommes, 2002) can be used to relate dissipation-free and dissipative systems, and Burby et al. (2013) have recently made progress on the closely related topic of Hamiltonian Langevin equations. It is by no means excluded that collisional and turbulence effects can be ultimately unified by means of a

dissipative variational formalism. However, that is an advanced topic that we do not pursue further here; it is not ready for prime time.

4 Form and interpretation of the gyrokinetic momentum conservation law

In this section, we discuss various representations of the conservation law for toroidal angular momentum, both for particles (Sec. 4.1) and for gyrocenters (Sec. 4.2). We give a derivation of the gyrocenter conservation law for a simplified situation. Then we sketch the derivation of the complete conservation law and present its final form. Armed with this background (which can be skipped on a first reading), we address in Sec. 5 the questions posed in the charge to the study group.

Because of axisymmetry, one focuses on the conservation law for toroidal angular momentum. Within the context of low-frequency physics, that law is derived most simply from the gyrokinetic equation, as shown by Scott & Smirnov (2010) and Brizard & Tronko (2011). However, there has been concern that somehow the gyrokinetic law, even if correct, is somehow irrelevant. Since one is ultimately concerned with the collection of actual particles that constitute the true plasma, surely one must examine the momentum conservation law for the true particles; perhaps its content is distinct from the gyrokinetic law. However, we will argue that this is not the case. There is just one conservation law, although there are multiple paths to its derivation and its form can be written in different ways. The connection between the various forms is the transformation law T between the particle and the gyrocenter variables ($\overline{\boldsymbol{z}} = \mathrm{T}\boldsymbol{z}$) or between the particle and gyrocenter distributions ($f = \mathrm{T}^{*}\overline{F}$).

4.1 Momentum equation for particles

First, we follow Parra & Catto by studying the momentum equation for the true particles. The derivation of that equation is well known but is worth reviewing. Let us begin with the Vlasov equation[43]

$$\frac{\partial f_s}{\partial t} + \boldsymbol{v}\cdot\boldsymbol{\nabla} f_s + \left(\frac{q}{m}\right)_s (\boldsymbol{E} + c^{-1}\boldsymbol{v}\times\boldsymbol{B})\cdot\frac{\partial f_s}{\partial \boldsymbol{v}} = 0. \tag{39}$$

We use the normalization convention

$$\int d\boldsymbol{v}\, \overline{n}_s f_s(\boldsymbol{x},\boldsymbol{v},t) = \breve{n}_s(\boldsymbol{x},t), \tag{40}$$

where $\overline{n}_s$ is the mean (spatially constant) density of species s; the check accent denotes a particle (as opposed to gyrocenter) moment. From this, one finds that

[43] The inclusion of a collision operator does not change the final momentum conservation law because that operator conserves momentum when summed over species.

the mean particle flow $\breve{\boldsymbol{u}}$ is obtained from

$$\breve{n}_s(\boldsymbol{x},t)\breve{\boldsymbol{u}}_s(\boldsymbol{x},t) = \int d\boldsymbol{v}\, \overline{n}_s \boldsymbol{v} f_s(\boldsymbol{x},\boldsymbol{v},t). \tag{41}$$

(In this convention, $d\boldsymbol{v}\, f$ is dimensionless, so all other dimensions are explicit.) Upon multiplying Eq. (39) by $m_s\overline{n}_s\boldsymbol{v}$ and integrating over $\boldsymbol{v}$, one is led to the familiar moment equation

$$\partial_t[(\breve{n}m\breve{\boldsymbol{u}})_s] + \boldsymbol{\nabla}\cdot\breve{\Pi}_s = (\breve{n}q)_s(\boldsymbol{E} + c^{-1}\breve{\boldsymbol{u}}_s\times\boldsymbol{B}), \tag{42}$$

where

$$\breve{\Pi}_s \doteq [[(\overline{n}m)_s \boldsymbol{v}\,\boldsymbol{v}]] \tag{43}$$

and the double-bracket notation denotes the velocity average with f_s. [In many discussions, such as that of Braginskii (1965), the peculiar velocity $\delta\boldsymbol{v} \doteq \boldsymbol{v} - \breve{\boldsymbol{u}}$ is introduced and one writes $\breve{\Pi} = \breve{n}m\breve{\boldsymbol{u}}\breve{\boldsymbol{u}} + \breve{\boldsymbol{\pi}}$, where $\breve{\boldsymbol{\pi}}$ is conventionally called the stress tensor; however, it is unnecessary to introduce $\breve{\boldsymbol{\pi}}$ here.] Although the kinetic part on the left-hand side of Eq. (42) is in locally conservative form, the electromagnetic part on the right is not obviously of that form. We will use the quasineutrality approximation $\breve{\rho} \doteq \sum_s(\breve{n}q)_s = 0$, Ampère's law $\boldsymbol{\nabla}\times\boldsymbol{B} = (4\pi/c)\breve{\boldsymbol{j}}$, and the identity (for arbitrary $\boldsymbol{V}$)

$$(\boldsymbol{\nabla}\times\boldsymbol{V})\times\boldsymbol{V} = (\boldsymbol{V}\cdot\boldsymbol{\nabla})\boldsymbol{V} - \frac{1}{2}\boldsymbol{\nabla}(V^2). \tag{44}$$

Then summing Eq. (42) over species annihilates the $\boldsymbol{E}$ term, introduces $\boldsymbol{j}\times\boldsymbol{B}$, and with the identity (44) leads to the momentum conservation law

$$\frac{\partial}{\partial t}\left(\sum_s(\breve{n}m)_s\breve{\boldsymbol{u}}_s\right) = \boldsymbol{\nabla}\cdot\breve{\mathsf{T}}, \tag{45}$$

where

$$\breve{\mathsf{T}} \doteq \mathsf{T}_M - \breve{\Pi} \tag{46}$$

and

$$\mathsf{T}_M \doteq \frac{1}{4\pi}\left(\boldsymbol{B}\,\boldsymbol{B} - \frac{1}{2}B^2\mathsf{I}\right) \tag{47}$$

is the magnetic part of the Maxwell stress tensor. (If one does not invoke quasineutrality and does include the displacement current, the Poynting momentum flux and the electric part of the Maxwell tensor also appear.)

Equation (45) is a local conservation law (it is not integrated over space), and it is valid irrespective of any spatial symmetries. It is not optimally useful, however, because the kinetic momentum fluxes are coupled to the Maxwell stresses, and it does not exploit any simplifications that result from toroidal symmetry. Furthermore, it does not describe the evolution of angular momentum (plasma rotation). To obtain an equation for angular momentum density, we use the flux coordinates $\{\psi,\theta,\varphi\}$ (see Appendix A) and dot Eq. (45) with

the basis vector (see Sec. A.1) $\boldsymbol{e}_\varphi = R\widehat{\boldsymbol{\varphi}}$ in order to obtain an evolution equation for the covariant toroidal component u_φ of $\boldsymbol{u}$:

$$\frac{\partial}{\partial t}\left(\sum_s (\check{n}m)_s u_{\varphi s}\right) = \boldsymbol{e}_\varphi \cdot (\boldsymbol{\nabla}\cdot\check{\mathsf{T}}). \tag{48}$$

Because $\boldsymbol{e}_\varphi$ is not a constant in toroidal geometry, it cannot be trivially moved inside of the divergence. We adopt the convention that the divergence is a contraction on the second tensor index (in fact, the tensors of interest here are symmetric). Then, in *Cartesian components only*, one has

$$\boldsymbol{e}\cdot(\boldsymbol{\nabla}\cdot\mathsf{T}) = e^i(\partial_k T_i^k) = \partial_k(e^i T_i^k) - T_i^k \partial_k e^i, \tag{49}$$

The last term expresses the fact that the divergence of a tensor in a general coordinate system involves the covariant derivative, which is expressed in terms of the connection coefficients or Christoffel symbols Γ^i_{jk}. In coordinate-free notation, for a mixed second-rank tensor [see Eq. (A27)],

$$\boldsymbol{e}\cdot(\boldsymbol{\nabla}\cdot\mathsf{T}) = \boldsymbol{\nabla}\cdot(\boldsymbol{e}\cdot\mathsf{T}) - \mathsf{T} : (\boldsymbol{\nabla}\boldsymbol{e}). \tag{50}$$

Now $\boldsymbol{\nabla}\boldsymbol{e}_\varphi \neq \mathsf{0}$. However, we show in footnote 49 that $\boldsymbol{\nabla}\boldsymbol{e}_\varphi$ is antisymmetric, so its contraction with the symmetric tensor T vanishes. To further simplify Eq. (48), using Eq. (50), one can apply the flux-surface average (see Sec. B.3), which we denote by $\langle\ldots\rangle$:

$$\frac{\partial}{\partial t}\left(\sum_s (\check{n}m)_s \langle\check{u}_{\varphi s}\rangle\right) = \langle\boldsymbol{\nabla}\cdot(\boldsymbol{e}_\varphi\cdot\check{\mathsf{T}})\rangle. \tag{51}$$

The quantity inside the divergence is a vector,[44] so one may use formula (B11) for the flux-surface average of a divergence: for arbitrary vector $\boldsymbol{A}$,

$$\langle\boldsymbol{\nabla}\cdot\boldsymbol{A}\rangle = \frac{1}{V'}\frac{\partial}{\partial\psi}(V'\langle A^\psi\rangle), \tag{52}$$

where V' is defined by Eq. (B9). Thus

$$\frac{\partial}{\partial t}\left(\sum_s (\check{n}m)_s \langle\check{u}_{\varphi s}\rangle\right) = \frac{1}{V'}\frac{\partial}{\partial\psi}(V'\langle\check{T}_\varphi^\psi\rangle). \tag{53}$$

Furthermore, $(\mathsf{T}_M)_\varphi^\psi = 0$ since $B^\psi = 0$. Therefore, one finds

$$\frac{\partial}{\partial t}\left(\sum_s (\check{n}m)_s \langle\check{u}_{\varphi s}\rangle\right) = -\frac{1}{V'}\frac{\partial}{\partial\psi}(V'\langle\check{\Pi}_\varphi^\psi\rangle), \tag{54}$$

a well-known result that is usually obtained by different means. It is usually summarized by saying, somewhat incompletely, that "the [flux-surface-averaged] toroidal rotation depends on the off-diagonal component of the stress tensor."

[44] In an arbitrary coordinate system, one has $\mathsf{T} = T_i^k \boldsymbol{e}^i \otimes \boldsymbol{e}_k$, so $\boldsymbol{e}_\varphi\cdot\mathsf{T} = T_\varphi^k \boldsymbol{e}_k$.

Although the real-space conservation law (54) is correct, it is not necessarily in optimal form. It is defined in terms of the $\check{\Pi}$ defined by Eq. (43), which includes both mean flows and fluctuating stresses. Also, it evolves the momentum of all particle flows, including the diamagnetic flow $\boldsymbol{u}_*$. Diamagnetic current is, of course, a real physical effect. On the other hand, it is well known that manipulations involving the particle moment equations inevitably involve one with the so-called *gyroviscous cancellation* (Hinton & Horton, 1971), which removes $\boldsymbol{u}_*$ from the advective nonlinearities. Brizard (1992) showed that the gyroviscous cancellation can be avoided altogether in the gyrokinetic formulation. A drawback to Eq. (54) is that it is not formulated in terms of the gyrocenter quantities that one knows to be physically important. Nevertheless, it is a correct equation. Although the stress tensor is expressed in terms of a second velocity moment of the particle distribution function f, one can instead calculate it from the gyrocenter PDF $\overline{F}$ by using the pullback transformation $f = \mathrm{T}^*\overline{F}$. That is the approach followed by Parra and coworkers. That is, although they use the gyrokinetic equation for intermediate calculations, they bring the results back to the laboratory space of the real particles. That is a sensible and direct way to proceed. They explicate the algebraic details in various of their papers; we shall not reproduce them here.

4.2 Angular momentum conservation from the gyrokinetic equation

The alternate way of proceeding is to examine momentum conservation from the point of view of gyrocenters rather than particles. Indeed, Scott & Smirnov (2010) and Brizard & Tronko (2011) have employed Noether methods to obtain a conservation law for toroidal angular momentum density written in terms of the gyrocenter quantities. To introduce the form of that law, we will first consider its derivation for an important and physically transparent special case. That leads one to the general form of the result with a minimum of algebra. The general form is discussed further in Appendixes I and J, and its implications are discussed in Sec. 5.

To motivate the final form of the gyrokinetic conservation law for toroidal angular momentum, we give in the next few sections a simplified derivation using the assumptions of cold ions (no FLR effects), the lowest-order approximation to the ion polarization drift velocity, and constant $\boldsymbol{B}$. Remarkably, the approximate calculation leads to a form of the conservation law that is identical to that which follows from the complete theory [recorded in Eq. (92) and derived in full detail in Appendix I]; only the definitions of some of the variables and fluxes need to be slightly generalized.

The calculation we shall give here is not the shortest. [For the most complete derivation, which proceeds from the conservation of canonical angular momentum, see Brizard & Tronko (2011).] However, it has the virtue of being explicit and elementary. We shall merely consider (with the benefit of hindsight) plausible quantities that ought to participate in the evolution of toroidal angular momentum, and show that their sum leads to a proper local conservation law.

Importantly, our direct calculation does not proceed from a variational principle. That allows us to explain what happens when the automatic consistency built into the variational approach is violated.

4.2.1 Useful identities

In the manipulations, the identity[45]

$$(\boldsymbol{E}\times\boldsymbol{B})_\varphi = E^\psi \tag{56}$$

(true for arbitrary $\boldsymbol{E}$) will be important. When $\boldsymbol{E}$ is in fact the electric field, this relates the toroidal (covariant) component of the $\boldsymbol{E}\times\boldsymbol{B}$ velocity to the radial (contravariant) component of the field (which is related to the electric polarization). Another useful identity that follows by inverting Eq. (56) is[46]

$$(B^{-1}\widehat{\boldsymbol{b}}\times\boldsymbol{E})^\psi = (\boldsymbol{E}_\perp)_\varphi. \tag{57}$$

4.2.2 $\boldsymbol{E}\times\boldsymbol{B}$ flows and polarization current

An alternate way of expressing Eq. (56) is to consider the time rate of change of the associated angular momentum density (of the ion gyrocenter $\boldsymbol{E}\times\boldsymbol{B}$ flow):

$$\frac{\partial}{\partial t}\left((Nm)_i\frac{c}{B^2}(\boldsymbol{E}\times\boldsymbol{B})_\varphi\right) = \frac{(Nm)_i c}{B^2}\frac{\partial E^\psi}{\partial t} \approx \frac{1}{c}[(Nq)_i u_p^\psi] \equiv \frac{1}{c}J_p^\psi, \tag{58}$$

where J_p is the (lowest-order approximation to the) polarization current derived from the ion polarization drift velocity

$$\boldsymbol{u}^{\rm pol} = \frac{1}{\omega_{ci}}\frac{\partial}{\partial t}\left(\frac{c\boldsymbol{E}_\perp}{B}\right). \tag{59}$$

Now polarization current $J^{\rm pol}$ and gyrocenter current J^G are related by the current continuity equation, which under quasineutrality is $\boldsymbol{\nabla}\cdot\boldsymbol{J} = \boldsymbol{\nabla}\cdot(\boldsymbol{J}^G + \boldsymbol{J}^{\rm pol}) = 0$ or

$$\boldsymbol{\nabla}\cdot\boldsymbol{J}^{\rm pol} = -\boldsymbol{\nabla}\cdot\boldsymbol{J}^G = \partial_t\rho^G. \tag{60}$$

[45] With the representation for an axisymmetric field $\boldsymbol{B} = I\boldsymbol{\nabla}\varphi + \boldsymbol{\nabla}\varphi\times\boldsymbol{\nabla}\psi$ discussed in Sec. A.6, one has

$$(\boldsymbol{E}\times\boldsymbol{B})_\varphi = -\boldsymbol{e}_\varphi\times\boldsymbol{B}\cdot\boldsymbol{E} \tag{55a}$$
$$= -\boldsymbol{e}_\varphi\times(I\boldsymbol{\nabla}\varphi + \boldsymbol{\nabla}\varphi\times\boldsymbol{\nabla}\psi)\cdot\boldsymbol{E} \tag{55b}$$
$$= -[\boldsymbol{\nabla}\varphi(\boldsymbol{e}_\varphi\cdot\boldsymbol{\nabla}\psi) - \boldsymbol{\nabla}\psi(\boldsymbol{e}_\varphi\cdot\boldsymbol{\nabla}\varphi)]\cdot\boldsymbol{E} \tag{55c}$$
$$= \boldsymbol{\nabla}\psi\cdot\boldsymbol{E} \equiv E^\psi. \tag{55d}$$

[46] The $\perp$ is important. If $\boldsymbol{E}$ is replaced by the gradient of a scalar s, the identity reads $(B^{-1}\widehat{\boldsymbol{b}}\times\boldsymbol{\nabla}s)^\psi = (\boldsymbol{\nabla}_\perp s)_\varphi$. That is not the same as $(\boldsymbol{\nabla}s)_\varphi = \boldsymbol{e}_\varphi\cdot\boldsymbol{\nabla}s = \partial_\varphi s$ because the full gradient also has a parallel component that in general has a projection on the toroidal direction.

Under quasineutrality, the gyrocenter charge $\rho^G \equiv \rho$ (no check accent) is related to gyrocenter polarization by

$$\rho^G = \boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{P}} \tag{61}$$

(an exact result provided that $\boldsymbol{\mathcal{P}}$ is defined appropriately, including nonlinear corrections), so Eq. (60) can be written as

$$\boldsymbol{\nabla} \cdot \boldsymbol{J}^{\rm pol} = \boldsymbol{\nabla} \cdot (\partial_t \boldsymbol{\mathcal{P}}). \tag{62}$$

Upon flux-surface averaging this equation, one finds

$$\frac{1}{V'} \frac{\partial}{\partial \psi} [V' (\partial_t \langle \mathcal{P}^\psi \rangle - \langle J_p^\psi \rangle)] = 0. \tag{63}$$

Integration leads to

$$\langle J_p^\psi \rangle = \partial_t \langle \mathcal{P}^\psi \rangle, \tag{64}$$

where the integration constant can be argued to vanish by virtue of regularity at the origin. By integration of Eq. (58) in time, we thus learn that the toroidal angular momentum density of the averaged perpendicular gyrocenter flow *is* to lowest order the averaged radial polarization:

$$\langle (Nm)_i U_{E\varphi} \rangle \approx c^{-1} \langle \mathcal{P}^\psi \rangle. \tag{65}$$

(In the detailed derivation, the left-hand side is summed over species, but the ions dominate.) From this argument, one anticipates that the final gyrokinetic momentum conservation law will involve the radial polarization as an important constituent.

4.2.3 Gyrocenter parallel flow

Gyrocenter parallel flow also has a projection onto the toroidal direction (unless $\boldsymbol{B}$ is fully poloidal). Note that $\boldsymbol{U}_\parallel \cdot \boldsymbol{e}_\varphi = U_\parallel \widehat{\boldsymbol{b}} \cdot \boldsymbol{e}_\varphi = U_\parallel b_\varphi$, so consider the quantity $\partial_t \langle \sum_s (Nm)_s U_\parallel b_\varphi \rangle$. An equation for the parallel flow may be obtained from the gyrokinetic equation, which can be conveniently written in conservative form:

$$\frac{\partial (\mathcal{J} F)}{\partial t} + \boldsymbol{\nabla} \cdot [(U \widehat{\boldsymbol{b}} + \overline{\boldsymbol{V}}_E^n) \mathcal{J} F] + \frac{\partial}{\partial U} \left(\frac{q}{m} \overline{E}_\parallel^n \mathcal{J} F \right) = 0. \tag{66}$$

Here $\mathcal{J}$ is the Jacobian of the gyrocenter transformation, the overline means a effective, gyroaveraged quantity, and the n superscript indicates that in principle one must include nonlinear corrections. Upon multiplying by $\overline{N} U$, where $\overline{N}$ is the average density of gyrocenters,[47] and integrating over the momentum coordinates U and μ, one obtains

$$\frac{\partial (N U_\parallel)}{\partial t} + \boldsymbol{\nabla} \cdot ([[\overline{N} U^2]] \widehat{\boldsymbol{b}}) + \boldsymbol{\nabla} \cdot [[\overline{N} U \overline{\boldsymbol{V}}_E^n]] = \frac{q}{m} [[\overline{N}\, \overline{E}_\parallel^n]], \tag{67}$$

[47] One may assume that $\overline{N} = \overline{n}$. Although in some formulas we use $\overline{N}$ to emphasize that we are dealing with gyrocenters, in others the upper-case notation is obstrusive and we just use $\overline{n}$.

where for gyrocenter quantities $[[A]] \doteq \int dU\, d\mu\, \mathcal{J}\, AF \equiv \int d\boldsymbol{p}\, AF$. The term in $\widehat{\boldsymbol{b}}$ vanishes under a flux-surface average. The latter two terms cannot be evaluated trivially because of the momentum dependence of the effective potential and the nonlinear correction terms. Upon multiplying Eq. (67) by m_s, summing over species, and flux-surface averaging, one obtains

$$\frac{\partial}{\partial t}\left(\langle\sum_s (Nm)_s U_{\parallel s}\rangle\right) + \frac{1}{V'}\frac{\partial}{\partial\psi}(V'\Pi_\parallel^\psi) = \sum_s \langle[[(\overline{N}q)_s \overline{E}_\parallel^n]]\rangle, \tag{68}$$

where

$$\Pi_\parallel^\psi \doteq \langle\sum_s [[(\overline{N}m)_s U \overline{V}_E^{n,\psi}]]\rangle. \tag{69}$$

To obtain an equation for the toroidal component of the parallel gyrocenter flow, one must multiply Eq. (68) by b_φ. In the general case with nonconstant $\boldsymbol{B}$, b_φ can be passed through the ∂_ψ and the flux-surface average in the second term only at the price of a correction term involving $\partial_\psi b_\varphi$. In the full derivation (Appendix I), that correction is canceled by another one relating to the evolution of the polarization due to magnetic drifts. Here we shall ignore such terms in the interest of exposing the basic structure of the theory. Then

$$\frac{\partial}{\partial t}\left(\langle\sum_s (Nm)_s U_{\parallel s} b_\varphi\rangle\right) + \frac{1}{V'}\frac{\partial}{\partial\psi}(V'\Pi_{\parallel\varphi}^\psi) \approx \sum_s \langle[[(\overline{N}q)_s \overline{E}_\parallel^n]] b_\varphi\rangle \equiv \mathcal{R}, \tag{70}$$

where

$$\Pi_{\parallel\varphi}^\psi \doteq \langle\sum_s [[(\overline{N}m)_s U b_\varphi \overline{V}_E^{n,\psi}]]\rangle. \tag{71}$$

The term in $\Pi_{\parallel\varphi}^\psi$ is obviously in conservative form and will persist in the final conservation law; it is a parallel–perpendicular Reynolds stress.

Now we shall manipulate the right-hand side $\mathcal{R}$ of Eq. (70), which describes the toroidal force on the gyrocenter due to parallel acceleration. Ignore FLR effects and write $E_\parallel^n = E_\parallel + \Delta E_\parallel$, where nonlinear corrections are represented by $\Delta E_\parallel$. Since $E_\parallel^n$ is independent of species, the $E_\parallel$ contribution gives $\langle\rho^G E_\parallel b_\varphi\rangle = \langle(\boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{P}})E_\parallel b_\varphi\rangle$. Now

$$E_\parallel b_\varphi = \boldsymbol{E}\cdot\widehat{\boldsymbol{b}}\,\widehat{\boldsymbol{b}}\cdot\boldsymbol{e}_\varphi = \boldsymbol{E}\cdot[\mathsf{I} - (\mathsf{I} - \widehat{\boldsymbol{b}}\,\widehat{\boldsymbol{b}})]\cdot\boldsymbol{e}_\varphi = E_\varphi - (\boldsymbol{E}_\perp)_\varphi. \tag{72}$$

Furthermore, we have from Eq. (57) that $(\boldsymbol{E}_\perp)_\varphi = -c^{-1}u_E^\psi$. Note that the terms in $\mathcal{R}$ stemming from ρ^G are of second order ($\mathcal{P}\times E_\parallel$) in the electric field, so for consistency one must keep second-order nonlinear corrections. Thus, one has

$$\sum_s (\overline{N}q)_s \Delta E_\parallel = -\sum_s N_s \widehat{\boldsymbol{b}}\cdot\boldsymbol{\nabla}\Delta H_s, \tag{73}$$

where ΔH is the nonlinear correction to the Hamiltonian. To second order and in the absence of FLR effects, the form of this term can be guessed correctly by noting that the Lagrangian density contains the terms $-H + \frac{1}{8\pi}E^2$. One expects

that the kinetic energy of fluid motion should add to the electrostatic energy; thus one can argue heuristically (a much more detailed discussion is given in Appendix F) that

$$\Delta H_i \approx -\frac{1}{2} m_i u_E^2 \tag{74}$$

and $\Delta H_e \approx 0$, so one has

$$\mathcal{R} = \langle (\boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{P}})(E_\varphi + c^{-1} u_E^\psi) \rangle - \sum_s \langle N_s \widehat{\boldsymbol{b}} \cdot \boldsymbol{\nabla} \Delta H_s b_\varphi \rangle. \tag{75}$$

4.2.4 Simplified momentum equation

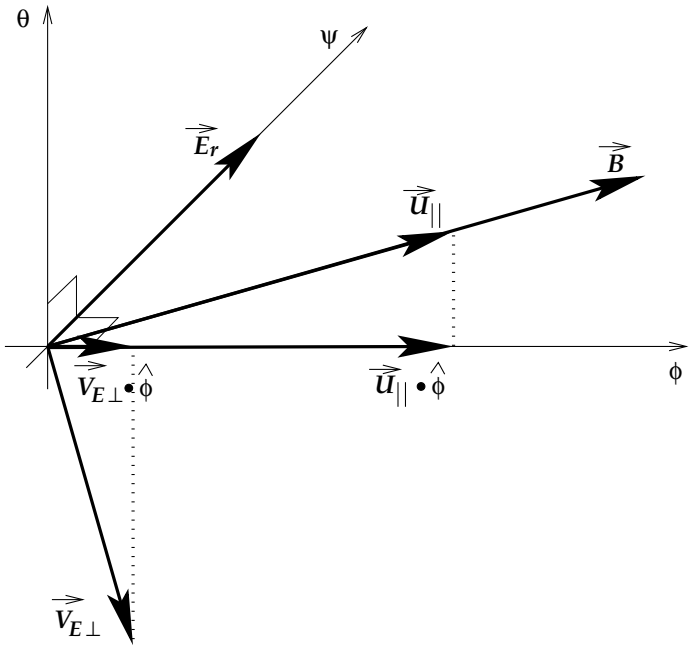

Figure 2: Important vectors related to toroidal flow. The total flow in the toroidal direction is the toroidal projection of the sum of the parallel gyrocenter flow and the perpendicular $\boldsymbol{E} \times \boldsymbol{B}$ flow arising from the radial polarization field. (The reader may question why the toroidal projection of the magnetic drifts does not appear. This is a subtle point; see foonote 54 on p. 56.)

This result for $\mathcal{R}$ is not in conservative form. However, we have argued that the time rate of change of the radial polarization should also enter the final momentum equation. To find that equation, we will sum the parallel and perpendicular forces. Both of the parallel and perpendicular flows have projections onto the toroidal direction (Fig. 2). First, recall Eq. (64). Explicitly,

the radial gyrocenter current is (no FLR effects)

$$J^\psi_G = \sum_s (Nq)_s \left(u^\psi_E + \frac{c}{q_s B}(\widehat{\boldsymbol{b}} \times \boldsymbol{\nabla}\Delta H_s)^\psi \right). \tag{76}$$

(We have assumed here that ΔH has no dependence on velocity.) The first term produces the gyrocenter charge, which can be replaced by $\boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{P}}$. For the second term, use the identity (57). Thus

$$c^{-1}\langle J^\psi_G\rangle = c^{-1}\langle(\boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{P}})u^\psi_E\rangle + \sum_s \langle N_s(\boldsymbol{\nabla}_\perp \Delta H_s)_\varphi\rangle. \tag{77}$$

Upon adding Eq. (64) to Eq. (70) and using the identity (57), one finds

$$\begin{aligned}\frac{\partial}{\partial t}\left(c^{-1}\langle \mathcal{P}^\psi\rangle + \left\langle \sum_s (Nm)_s U_{\parallel s} b_\varphi \right\rangle\right) + \frac{1}{V'}\frac{\partial}{\partial\psi}(V'\Pi^\psi_{\parallel\varphi}) \\ = \langle(\boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{P}})E_\varphi\rangle - \sum_s [\langle N_s[b_\varphi \widehat{\boldsymbol{b}}\cdot\boldsymbol{\nabla}\Delta H + (\boldsymbol{\nabla}_\perp \Delta H)_\varphi]\rangle.\end{aligned} \tag{78}$$

Note the cancellation of the terms involving u^ψ_E.

To simplify the terms involving ΔH, note that $b_\varphi\widehat{\boldsymbol{b}} = \boldsymbol{e}_\varphi\cdot\widehat{\boldsymbol{b}}\,\widehat{\boldsymbol{b}}$ and that $(\boldsymbol{\nabla}_\perp\Delta H)_\varphi = \boldsymbol{e}_\varphi\cdot[(\mathsf{I}-\widehat{\boldsymbol{b}}\,\widehat{\boldsymbol{b}})\cdot\boldsymbol{\nabla}\Delta H]$. Therefore, the terms in $\widehat{\boldsymbol{b}}$ cancel and one has[48]

$$b_\varphi\widehat{\boldsymbol{b}}\cdot\boldsymbol{\nabla}\Delta H + (\boldsymbol{\nabla}_\perp\Delta H)_\varphi = \boldsymbol{e}_\varphi\cdot\boldsymbol{\nabla}\Delta H = \partial_\varphi \Delta H. \tag{80}$$

Finally, note that (in the present example of the cold-ion limit; see Appendix F) $-\Delta H = \frac{1}{2}mu^2_E \propto |\boldsymbol{\nabla}_\perp\phi|^2$, and consider

$$\frac{1}{2}\partial_\varphi|\boldsymbol{\nabla}_\perp\phi|^2 = \boldsymbol{\nabla}_\perp\phi\cdot\partial_\varphi\boldsymbol{\nabla}_\perp\phi \propto -\boldsymbol{\mathcal{P}}\cdot\partial_\varphi\boldsymbol{\nabla}_\perp\phi, \tag{81}$$

where one can easily verify that the dimensional factors work out correctly. If one could pass the ∂_φ through the $\boldsymbol{\nabla}_\perp$, the resulting term (which would involve $-\boldsymbol{e}_\varphi\cdot\boldsymbol{\nabla}\phi = E_\varphi$) would add to the first term on the right-hand side of Eq. (78) and create the perfect divergence $\boldsymbol{\nabla}\cdot(\boldsymbol{\mathcal{P}}\,E_\varphi)$, thereby completing the derivation of the local conservation law. Therefore, consider the commutator $[\partial_\varphi, \boldsymbol{\nabla}_\perp]$. Unfortunately, this commutator *does not vanish*[49] because $\partial_\varphi = \boldsymbol{e}_\varphi\cdot\boldsymbol{\nabla}$ and

[48] Note that $E_\varphi = -\partial_\varphi\phi$, so the right-hand side of Eq. (78) can be written as $-\sum_s\langle N_s\partial_\varphi H_s\rangle$, where $H \doteq q\phi + \Delta H$. In the exact theory, this term generalizes trivially to

$$-\sum_s \overline{N}_s\langle\int d\boldsymbol{v}\, F_s\partial_\varphi H_s\rangle. \tag{79}$$

[49] In Appendix A.1 of Scott & Smirnov (2010), it appears to be stated that the vanishing of this and similar commutators is a requirement for the success of the theory. In the usual interpretation of the gradient operator as $\boldsymbol{\nabla} = \boldsymbol{e}^i\partial_i$, that statement is incorrect. A simple illustration is obtained by considering an (R,φ) polar coordinate system. One has

$$\partial_\varphi\boldsymbol{\nabla} = \partial_\varphi(\boldsymbol{\nabla}R\,\partial_R + \boldsymbol{\nabla}\varphi\,\partial_\varphi) = \boldsymbol{\nabla}R\,\partial_R\partial_\varphi + \boldsymbol{\nabla}\varphi\,\partial_\varphi\partial_\varphi + \widehat{\boldsymbol{\varphi}}\,\partial_R - \widehat{\boldsymbol{R}}\,R^{-1}\partial_\varphi; \tag{82}$$

both $\boldsymbol{e}_\varphi$ and $\widehat{\boldsymbol{b}}$ depend on space (even in an axisymmetric torus). However, one can show[50] that in an axisymmetric torus one does have

$$\boldsymbol{\nabla}_\perp s \cdot \partial_\varphi \boldsymbol{\nabla}_\perp s = \boldsymbol{\nabla}_\perp s \cdot \boldsymbol{\nabla}_\perp (\partial_\varphi s) \tag{89}$$

for an arbitrary scalar field s. Thus, the net right-hand side reduces to $\langle \boldsymbol{\nabla} \cdot (\boldsymbol{\mathcal{P}} E_\varphi) \rangle$, and one obtains the axisymmetric conservation law for toroidal angular momentum

$$\frac{\partial}{\partial t}\left(\mathcal{P}_{\varphi\parallel} + c^{-1}\langle \mathcal{P}^\psi \rangle\right) + \frac{1}{V'}\frac{\partial}{\partial \psi}(V' \Pi^\psi_{\parallel\varphi}) + \frac{1}{V'}\frac{\partial}{\partial \psi}\left(V' \left\langle \mathcal{P}^\psi \frac{\partial \phi}{\partial \varphi} \right\rangle\right) = 0, \tag{90}$$

where

$$\mathcal{P}_{\varphi\parallel} \doteq \langle \sum_s (\overline{N} m)_s U_{\parallel s} b_\varphi \rangle. \tag{91}$$

In the present approximate derivation, which omits FLR corrections, the polarization is proportional to the field (a first-order quantity). Its time derivative, however, is small, and an important lesson is that in order to obtain the consistent time evolution equation one had to consider second-order gyrocenter drifts

---

thus

$$[\partial_\varphi, \boldsymbol{\nabla}] = \widehat{\boldsymbol{\varphi}}\, \partial_R - \widehat{\boldsymbol{R}} R^{-1} \partial_\varphi \neq \boldsymbol{0}. \tag{83}$$

As a consistency check, consider the scalar field $s = \varphi$. One has $\frac{1}{2}\partial_\varphi |\boldsymbol{\nabla}\varphi|^2 = \frac{1}{2}\partial_\varphi (R^{-2}) = 0$. This result also follows by introducing the commutator:

$$\frac{1}{2}\partial_\varphi |\boldsymbol{\nabla}\varphi|^2 = \boldsymbol{\nabla}\varphi \cdot \partial_\varphi \boldsymbol{\nabla}\varphi = \boldsymbol{\nabla}\varphi \cdot \boldsymbol{\nabla}(1) + \boldsymbol{\nabla}\varphi \cdot [\partial_\varphi, \boldsymbol{\nabla}]\varphi = 0 + R^{-1}\widehat{\boldsymbol{\varphi}} \cdot (-\widehat{\boldsymbol{R}} R^{-1}) = 0. \tag{84}$$

Although the commutator itself does not vanish, it is orthogonal to the required gradient.

As another example, consider $\boldsymbol{\nabla} \boldsymbol{e}_\varphi$, which arises in the derivation of the real-space conservation law and, more generally, in the theory of covariant differentiation. One has $\boldsymbol{e}_\varphi = \partial_\varphi \boldsymbol{x}$, so

$$\boldsymbol{\nabla} \boldsymbol{e}_\varphi = \boldsymbol{\nabla} \partial_\varphi \boldsymbol{x} = \partial_\varphi \boldsymbol{\nabla} \boldsymbol{x} + [\boldsymbol{\nabla}, \partial_\varphi]\boldsymbol{x} = [\boldsymbol{\nabla}, \partial_\varphi]\boldsymbol{x} = (\widehat{\boldsymbol{R}} R^{-1} \partial\varphi - \widehat{\boldsymbol{\varphi}} \partial_R)\boldsymbol{x} = \widehat{\boldsymbol{R}}\,\widehat{\boldsymbol{\varphi}} - \widehat{\boldsymbol{\varphi}}\widehat{\boldsymbol{R}}. \tag{85}$$

Obviously this quantity does not vanish, but importantly it is antisymmetric.

[50] To prove this, write $|\boldsymbol{\nabla}_\perp s|^2 = |\boldsymbol{\nabla} s|^2 - (\nabla_\parallel s)^2$ and separately consider the cases for the full and parallel gradients. One has

$$\frac{1}{2}\partial_\varphi |\boldsymbol{\nabla} s|^2 = \frac{1}{2}\frac{\partial}{\partial \varphi}\left(\partial_i s\, g^{ij} \partial_j s\right) = \partial_i s\, \partial_\varphi \left(g^{ij} \partial_j s\right) \tag{86a}$$

$$= \partial_i s\, g^{ij} \partial_\varphi \partial_j s = \partial_i s\, g^{ij} \partial_j \partial_\varphi s \tag{86b}$$

$$= \boldsymbol{\nabla} s \cdot \boldsymbol{\nabla}(\partial_\varphi s). \tag{86c}$$

We used $\partial_\varphi g^{ij} = 0$, valid for axisymmetry. Similarly for the parallel gradients,

$$\frac{1}{2}\partial_\varphi (\nabla_\parallel s)^2 = \frac{1}{2}\frac{\partial}{\partial \varphi}[(b^i \partial_i s)^2] \tag{87a}$$

$$= b^i \partial_i s\, \partial_\varphi (b^j \partial_j s) = b^i \partial_i s\, b^j \partial_j (\partial_\varphi s) \tag{87b}$$

$$= \nabla_\parallel s \nabla_\parallel (\partial_\varphi s). \tag{87c}$$

Here we used $\partial_\varphi b^i = 0$ for axisymmetry. Explicitly, from Eq. (B2), one has $b^i = B^i/B$ with

$$B^\psi = 0, \quad B^\theta = J^{-1}, \quad B^\varphi = I/R^2; \tag{88}$$

all of I, J, R, and B are φ-independent for axisymmetry.

in the equations of motion. More generally, the fact that nth-order polarization is related to terms of order $n+1$ in the drifts is a consequence of the fact that the gyrokinetic–Poisson system is derivable from a Lagrangian variational principle, as we will discuss later in more detail.

This approximate derivation captures the essence of the manipulations. It is remarkable that the generalization obtained using Noether methods by Scott & Smirnov (2010) and later by Brizard & Tronko (2011) retains this same form. Explicitly, those authors find

$$\frac{\partial(\mathcal{P}_{\varphi\parallel}+c^{-1}\mathcal{P}^{\psi})}{\partial t}+\frac{1}{V'}\frac{\partial}{\partial\psi}V'(\Gamma^{\psi}_{\varphi\parallel}+\Gamma^{\psi}_{\varphi\perp})=0, \tag{92}$$

where

$$\Gamma^{\psi}_{\varphi\parallel}\doteq\sum_s(\overline{N}m)_s\left\langle\int_{\boldsymbol{P}}FV^{\psi}U_{\parallel}b_{\varphi}\right\rangle, \tag{93a}$$

$$\Gamma^{\psi}_{\varphi\perp}\equiv\Gamma\doteq\sum_s\overline{N}_s\left\langle\int_{\boldsymbol{P}}F\left(\frac{\partial H}{\partial\boldsymbol{E}}\right)^{\psi}\frac{\partial\phi}{\partial\varphi}\right\rangle+\cdots. \tag{93b}$$

The centered dots in Eq. (93b) signify additional terms involving derivatives of H with respect to higher-order gradients of the potential. This formula differs from Eq. (90) only in that the complete polarization $\mathcal{P}^{\psi}=(\partial H/\partial\boldsymbol{E})^{\psi}$ is kept, quantities are defined in terms of the full Hamiltonian, and additional terms appear in the definition of $\Gamma^{\psi}_{\varphi\perp}$.

Although we have gotten to the proper form of the answer, the manipulations in this section and in Appendix I may leave one cold. Although one suspects that conservation of canonical toroidal angular momentum should play a role, we did not begin with that constraint. In the more formal derivations of Scott & Smirnov and Brizard & Tronko canonical momentum conservation is brought to the fore. We will discuss that further in Sec. 5.3.

4.2.5 Approximate derivation of the angular momentum conservation law using particle variables

Further insight into the content of the conservation law can be obtained by rederiving it from the particle equations of motion. Although this is substantially more tedious, it demonstrates that there are not two distinct conservation laws (i.e., one for particles and one for gyrocenters); there is just one that is appropriate for slow motions in a magnetized plasma.

We begin with the real-space conservation law (54), repeated here for convenience:

$$\frac{\partial}{\partial t}\left(\sum_s(\check{n}m)_s\langle\check{u}_{\varphi s}\rangle\right)=-\frac{1}{V'}\frac{\partial}{\partial\psi}(V'\langle\check{\Pi}^{\psi}_{\varphi}\rangle). \tag{94}$$

There are three ways in which Eq. (94) differs from the gyrokinetic conservation law:

1. It involves the particle density n rather than the gyrocenter density N.
2. It evolves the particle fluid flow $\check{\boldsymbol{u}} = U\widehat{\boldsymbol{b}} + \boldsymbol{u}_{\boldsymbol{E}} + \boldsymbol{u}_* + \cdots$; gyrocenters do not move with the diamagnetic velocity u_*.
3. The stress tensor $\check{\Pi} \doteq \sum_s (\overline{n}m)_s \boldsymbol{v}\boldsymbol{v} f$ describes the particle-based stresses, which differ from those appearing in the gyrokinetic conservation law.

What needs to be done is to transform Eq. (94) into the gyrocenter representation. The calculations bear some resemblance to those performed by Scott (2007). They are tedious and are omitted from the present report. The conclusion, however, is clear: momentum conservation can be expressed in either a particle or a gyrocenter representation. Those are connected by the pullback transformation T^*.

4.2.6 Energetic consistency and momentum conservation

One can now discuss what happens when the field-theoretic consistency between the gyrokinetic equation and the gyrokinetic Poisson equation are violated. First, we summarize the four-step derivation of the momentum conservation equation given by Brizard & Tronko (2011) (and which closely tracks the procedure of Scott & Smirnov).

1. **Write the equation for parallel–toroidal canonical momentum.** One derives (the details are omitted)

$$\frac{\partial \mathcal{P}_{\varphi\parallel}}{\partial t} + \boldsymbol{\nabla}\cdot\Pi_\varphi = \frac{1}{c}\psi\frac{\partial \rho^G}{\partial t} - \sum_s \overline{N}\int d\boldsymbol{p}\, F\frac{\partial H}{\partial \phi}, \tag{95}$$

   where

$$\Pi_\varphi \doteq \sum \int F\frac{d_{\mathrm{gy}}\boldsymbol{X}}{dt}p_{\mathrm{gy}\varphi} \tag{96}$$

   is the canonical gyrocenter Reynolds stress tensor. Here $p_{\mathrm{gy}\varphi}$ is the gyrocenter canonical momentum [Eq. (108)].

2. **Flux-surface average.** The flux-surface average of Eq. (95) is

$$\frac{\partial \langle \mathcal{P}_{\varphi\parallel}\rangle}{\partial t} = -\frac{1}{V'}\frac{\partial}{\partial \psi}\left(V'\langle \Pi_\varphi^\psi\rangle\right) + \frac{1}{c}\psi\frac{\partial\langle\rho^G\rangle}{\partial t} - \sum \overline{N}\left\langle \int F\frac{\partial H}{\partial\varphi}\right\rangle, \tag{97}$$

   where V' is given by Eq. (B9) and where one finds

$$\langle \Pi_\varphi^\psi\rangle = \langle \Pi_{\parallel\varphi}^\psi\rangle - \frac{1}{c}\psi\langle J_{\mathrm{gy}}^\psi\rangle. \tag{98}$$

3. **Use the gyrocenter quasineutrality relation to relate current and polarization.** We have already done that; see Eq. (64).

4. **Derive the parallel-toroidal momentum transport equation.** One finds

$$\frac{\partial}{\partial t}\left(\langle \mathcal{P}_{\varphi\parallel}\rangle + c^{-1}\langle \mathcal{P}^{\psi}\rangle\right) + \frac{1}{V'}\frac{\partial}{\partial \psi}\left(V'\langle \Pi^{\psi}_{\parallel\varphi}\rangle\right) = -\sum \overline{N}\left\langle \int F \frac{\partial H}{\partial \varphi}\right\rangle. \quad (99)$$

In the methodology outlined here, this result follows simply,[51] in general geometry. If, on the other hand, one directly calculates the time derivatives of $\langle \mathcal{P}_{\varphi\parallel}\rangle$ and $\langle \mathcal{P}^{\psi}\rangle$, then sums them, the algebra is tedious and a variety of cancellations occur. Those are detailed in Appendix I; it is instructive to see them explicitly.

To complete the derivation, the crux of the argument, as emphasized by Scott & Smirnov, is to show that the last term of Eq. (99) can be written as a divergence. One finds (see Sec. I.2 for more discussion) that

$$\sum \overline{N} \int F \frac{\partial H}{\partial \varphi} = (\rho^G - \boldsymbol{\nabla}\cdot \boldsymbol{\mathcal{P}}_{\mathrm{GKE}})\frac{\partial \phi}{\partial \varphi} + \boldsymbol{\nabla}\cdot(\dots). \quad (102)$$

Here $\rho^G \doteq \sum(\overline{N}q)\int F$ is the gyrocenter charge, i.e., whatever charge density follows by integrating the gyrocenter distribution function over velocity. By quasineutrality, it is in balance with whatever polarization effect $\boldsymbol{\mathcal{P}}_{\mathrm{Poisson}}$ is kept in the gyrokinetic Poisson equation. [For example, the pullback transformation could be truncated to second order, as was done by Dubin et al. (1983).] The polarization $\boldsymbol{\mathcal{P}}_{\mathrm{GKE}}$, which is calculated from the drift effects in the gyrokinetic equation, is given as a series, which in practice is truncated to some order. Thus, one can write the key factor on the right-hand side of Eq. (102) as

$$\rho^G - \boldsymbol{\nabla}\cdot \boldsymbol{\mathcal{P}}_{\mathrm{GKE}} = \boldsymbol{\nabla}\cdot \boldsymbol{\mathcal{P}}_{\mathrm{Poisson}} - \boldsymbol{\nabla}\cdot \boldsymbol{\mathcal{P}}_{\mathrm{GKE}}. \quad (103)$$

Now when the gyrokinetic–Poisson system is derived variationally from a single gyrokinetic Hamiltonian, the approximations to the two polarizations in Eq. (103) are identical, the potentially nonconservative term vanishes, and one recovers a properly local conservation law. However, if that consistency is violated, spurious nonconservative terms will arise. A simple

[51] Begin with Eq. (97) and focus on the two terms explicitly proportional to ψ — the $\partial_t \rho^G$ term in Eq. (97) and the $\langle J^{\psi}_{\mathrm{gy}}\rangle$ term in Eq. (98). The time derivative of the quasineutrality equation (61) yields after flux-surface averaging

$$\frac{\partial \langle \rho^G\rangle}{\partial t} = \frac{1}{V'}\frac{\partial}{\partial \psi}V'\frac{\partial \langle \mathcal{P}^{\psi}\rangle}{\partial t}. \quad (100)$$

The $\langle J^{\psi}_{\mathrm{gy}}\rangle$ term was replaced by $\partial_t\langle \mathcal{P}^{\psi}\rangle$ by using Eq. (64) (the gyrocenter current is the negative of the polarization current). The explicit ψ terms on the right-hand side of Eq. (97) [including the last term of Eq. (98)] then sum to

$$-\frac{1}{V'}\frac{\partial}{\partial \psi}V'\left(\psi\frac{1}{c}\frac{\partial \langle \mathcal{P}^{\psi}\rangle}{\partial t}\right) + \frac{1}{c}\psi\frac{1}{V'}\frac{\partial}{\partial \psi}V'\frac{\partial \langle \mathcal{P}^{\psi}\rangle}{\partial t} = -\frac{1}{c}\frac{\partial \langle \mathcal{P}^{\psi}\rangle}{\partial t}. \quad (101)$$

The ψ terms have canceled, a consequence of quasineutrality.

but important example was given by Parra & Catto (2010a), who considered [in the context of the slab calculation done by Dubin et al. (1983)] the implications of retaining second-order terms in Poisson's equation and no more than second-order terms in the gyrokinetic equation. They showed explicitly how that inconsistent approximation led to a spurious nonconservative term (related to the derivative of a third-order Hamiltonian) and argued that the spurious term would disappear if one included third-order drifts. That is the energetic consistency discussed more generally by Scott & Smirnov.

The work of Parra & Catto (2010a) made clear that the third-order Hamiltonian was required in a slab calculation. They worried (see the excerpt in Sec. L.8) that in the presence of magnetic inhomogeneity one might have to work to fourth order. That does not happen, however; cancellations occur such that working to third order is adequate. For the details, see Sec. I.1.

5 Discussion of the ordering issues, and Answers to the Charge questions

With Secs. 3 and 4, we have provided the background required for an appreciation of the controversy regarding momentum conservation and ordering. We now turn to a discussion of their consequences.

5.1 The original ordering estimates

Let us begin by considering Parra's original ordering estimates, which are summarized in Table 1.

Method		Order of f_i	Chapter
Gyrokinetic quasineutrality equation		$\epsilon_i^4 f_{Mi}$	3
Radial transport of toroidal angular momentum	Evaluated directly from f_i	$\epsilon_i^3 f_{Mi}$	2
	Moment equation	$\epsilon_i^2 f_{Mi}$	5
	Moment equation and $B/B^{\rm pol} \gg 1$	$(B/B^{\rm pol})\epsilon_i^2 f_{Mi}$	5

Table 1: Comparison of different methods to obtain the long-wavelength axisymmetric radial electric field. From the PhD dissertation of Parra (2009, Table 1.1). (In Parra's original notation, the ordering parameter was δ; we use ϵ in this report.)

Parra couches the discussion in terms of the calculation of the long-wavelength axisymmetric part of the radial electric field; he states the precision of the ion

(gyrocenter) distribution function (which he calles f_i) that is required in various methodologies, which we shall first define, then discuss.

1. **Gyrokinetic quasineutrality equation:** This pertains to full-F simulations using the gyrokinetic equation and GK Maxwell equations. The simulation is supposed to be run to transport times scales. In this method, no subsidiary fluid equation is used.

2. **Radial transport of toroidal angular momentum:** In the following methods, it is supposed that momentum transport is obtained by integrating the fluid equation for toroidal angular momentum; the GK system is used to evaluate the requisite stress tensor.

   (a) **Evaluated directed from f_i:** Here the definition of the stress tensor Π is used directly. As we explained in Sec. 2.2, that stress tensor is $O(\epsilon^3)$ in the low-flow ordering. Since Π is a velocity moment of F, one must calculate F to $O(\epsilon^3)$ in the direct approach.

   (b) **Moment equation:** Instead of calculating Π directly, Parra showed that one can use a moment equation to calculate the required off-diagonal components of Π. In that equation, the off-diagonal components arise from the Lorentz force term, which is proportional to $\omega_{ci} \sim \epsilon^{-1}$. The solution then has the form $\epsilon \times$ (driving term), so the driving term need be calculated to just second order in order to obtain a third-order result. Thus, the moment method *saves an order*.

   (c) **Moment equation and $B/B^{\rm pol} \gg 1$:** Further simplifications ensue when $B^{\rm pol}/B$ is taken to be small. (Although that is usually satisfactory, it does not hold for spherical tokamaks such as NSTX.) We will not discuss that approximation in this report.

5.1.1 Discussion of the moment method

With method 2(b), one sees that the distribution function need be calculated only through second order in order to obtain the third-order fluxes. But working at least to second order is essential; first-order drifts (e.g., the $\boldsymbol{E}\times\boldsymbol{B}$ drift) and first-order polarization in Poisson's equation are inadequate. Note that although the $\boldsymbol{E}\times\boldsymbol{B}$ drift is nominally of first order, since $\boldsymbol{V}_E \propto \phi$ it can pick up a second-order piece from a second-order potential.

As Parra & Catto have argued, it would appear that use of the fluid equation for toroidal angular momentum, in conjunction with a second-order description of the gyrokinetics, is the most efficient way to proceed to a theory of the radial electric field and momentum transport on the long transport time scale. Ongoing research is proceeding in that direction (Barnes et al., 2013).

5.1.2 Full-F: Primitive ordering considerations

In this section we present a version of the ordering argument of Parra & Catto that provides a relatively simple demonstration that trying to use the gyrokinetic quasineutrality equation to calculate the mean radial electric field will require keeping the Hamiltonian through third order, $H_3 \sim \epsilon^3 T$. This supports the claim of Parra & Catto that previous approaches are not sufficiently accurate to determine rotation profiles on the transport time scale in the low-flow gyro-Bohm regime (since previous approaches use at most only the second-order Hamiltonian or partial third-order terms).

In Secs. 5.3 and 5.4 we will give similar ordering arguments in the notation of the gyrokinetic momentum conservation law of Scott & Smirnov; we will find again that a third-order Hamiltonian is required. For a more detailed discussion, see Appendix I.

The original concern, reflected in Table 1, was that the F in this approach would even be required to fourth order, so that a fourth-order Hamiltonian would be needed. It was the extreme smallness of ϵ^4 that led to some of the initial skepticism. Here we reproduce that conclusion, but then discuss heuristically why it turns out to be overly pessimistic because of a cancellation that occurs. In fact, the quasineutrality constraint implies that at most only third-order effects are necessary, in agreement with the estimated size of the momentum flux. But even this level of accuracy is beyond what present codes can do or seems practical to do (particularly given that there is a more efficient alternative approach that Parra & Catto point out).

In the long-wavelength limit, $k_\perp \rho_i \ll 1$, the gyrokinetic quasineutrality equation that is usually used to determine the electrostatic potential ϕ reduces to

$$-(nq)_i \rho_{\rm s}^2 \nabla_\perp^2 \left(\frac{e\phi}{T_e}\right) = \sum_s (\overline{n} q)_s \int d\boldsymbol{v}\, F_s. \tag{104}$$

Here the left-hand side is the (negative of) the ion polarization charge density (the precise position of the n_i with respect to the divergence operator is irrelevant for the following considerations), and the right-hand side is the charge density of gyrocenters. Here q_s, n_s, and F_s are the charge, density, and gyrocenter distribution function of species s, $\rho_{\rm s} \doteq c_{\rm s}/\omega_{ci}$ is the ion sound gyroradius, $\overline{n}$ is the mean density, $c_{\rm s} \doteq (ZT_e/m_i)^{1/2}$ is the ion sound speed, and T_e is the electron temperature. (FLR corrections have been neglected. Those can be included, but they do not affect the following primitive scaling argument.) Following the procedure used by Parra & Catto in various of their publications, take the time derivative of both sides and substitute on the right-hand side using the gyrokinetic equation in the conservative form $\partial_t F_s = -\boldsymbol{\nabla} \cdot (\boldsymbol{V}_{d,s} F_s) + \cdots$, where $\boldsymbol{V}_{d,s}$ is the gyrocenter drift velocity, to get

$$-\frac{\partial}{\partial t}\left[(nq)_i \rho_{\rm s}^2 \nabla_\perp^2 \left(\frac{e\phi}{T_e}\right)\right] = -\boldsymbol{\nabla} \cdot \left(\sum_s (\overline{n} q)_s \int d\boldsymbol{v}\, F_s \boldsymbol{V}_{d,s}\right). \tag{105}$$

(In the general theory, one must worry about a nontrivial Jacobian, but the present arguments capture the proper flavor.) Consider the long-wavelength

(equilibrium-scale) mean component of the electrostatic potential found from this equation. Assume (for the reasons given in Sec. 2.2) that the equilibrium-scale mean toroidal momentum evolves on a gyro-Bohm transport time scale, $\partial_t \sim D_{\rm gB}\nabla_\perp^2 \sim (c_{\rm s}/L)\rho_{\rm s}^2/L^2 \sim \epsilon^2(c_{\rm s}/L)$, where L is the equilibrium scale length (i.e., typically of the order of the minor radius here), and $\epsilon \doteq \rho_{\rm s}/L$ is the fundamental gyrokinetic expansion parameter. (As stated in Sec. 2.2, the assumption of a gyro-Bohm transport time scale does not mean that the momentum fluxes are necessarily diffusive, only that they are of this magnitude; they may contain a pinch term, residual stresses, off-diagonal transport terms, *etc.*) In the low-flow ordering regime, the toroidal rotation scales as $u_\varphi \sim \epsilon c_{\rm s}$ by assumption, and the corresponding long-wavelength equilibrium-scale potential ϕ_0 satisfies $e\phi_0/T_e \sim 1$ (here we are taking all geometric factors like ι (the inverse of the safety factor q), r/R, or B_θ/B_φ to be of order unity and focusing on the ρ/L scaling). Upon taking $\nabla_\perp \sim 1/L$, the order of magnitude of the left-hand side of this equation is seen to be

$$\text{LHS} \sim \frac{c_{\rm s}}{L} e n_i \epsilon^4 \tag{106}$$

and the right-hand side (using $\int d\boldsymbol{v}\, F_s \overline{n}_s \boldsymbol{V}_{d,s} \sim n_s V_{d,s}$) is of the order

$$\text{RHS} \sim \frac{1}{L} e n_i V_{d,s}. \tag{107}$$

Balancing these tells us that one apparently needs to calculate very small drifts of order $V_{d,s} \sim \epsilon^4 c_{\rm s}$ in order to accurately calculate the evolution of the momentum on the very slow gyro-Bohm transport time scale. This reproduces Parra's original estimate. [See also the discussion in Parra & Catto (2010a).]

However, if one examines the argument more deeply, one finds that one only needs the drifts through third order, not fourth. To see this, consider the following *plausibility* argument. Because it is only the mean potential that evolves on the slow transport time scale, what one actually needs is the average of the radial flux, $\langle \int d\boldsymbol{v}\, F_s V_{d,s}^\psi \rangle$. (This average includes both a flux-surface average and a time average over the turbulent fluctuations.) Now the drifts are derivable as the Poisson bracket of the gyrocenter position with the Hamiltonian. In the same way as the $\boldsymbol{E}\times\boldsymbol{B}$ drift is $\propto \widehat{\boldsymbol{b}}\times\boldsymbol{\nabla}\phi$, a general Hamiltonian gives rise to a contribution $\boldsymbol{V}_d \propto \widehat{\boldsymbol{b}}\times\boldsymbol{\nabla}H$. For simplicity, consider only this term. (For inhomogeneous magnetic fields, there are other terms related to the magnetic curvature, but they do not change the final conclusion; see Appendix I.) The gradient operator involves derivatives with respect to poloidal angle θ and toroidal angle φ. If one considers the fourth-order drifts, i.e., those derivable from H_4, then it is adequate to approximate $F \approx F_0(\psi)$, the lowest-order distribution. Because F_0 is constant on a flux surface, it can be removed from the integral. Then under the flux-surface average the ∂_φ vanishes for axisymmetric geometry. The ∂_θ does not integrate away exactly in general geometry; however, integration by parts shows that the ∂_θ acts on background magnetic quantities, which have only macroscopic variation on scale L. Therefore the ∂_θ introduces another power of ϵ and is negligible.

The argument so far does not exclude the possibility that fourth-order drifts might arise from slowly varying gradients of an H_3. However, some terms were neglected in this discussion. In general geometry, the toroidal angular momentum has contributions from both the toroidal component of the perpendicular drift from the radial electric field as well as the toroidal component of the parallel velocity. This leads to cancellations between the perpendicular and parallel components of the momentum flux such that fourth-order drifts do not matter. Nevertheless, the argument above illustrates the main ordering problem, which in the Scott–Smirnov general-geometry calculation of Sec. 5.3 is also found to be in the perpendicular component of the momentum flux. Furthermore, the above argument demonstrating the need for H_3 is rigorous in the limit of an infinitely elongated tokamak, where toroidal flow is due entirely to the perpendicular component. In general geometry, the rigorous version of the argument is given in Appendix I; analysis of the final equation is given in Appendix J. Those appendices lead to the conclusion that at most only effects arising from the third-order Hamiltonian need to be kept in order to satisfy the balance implied by the quasineutrality constraint. This agrees with the conclusion one reaches from the Scott–Smirnov conservation law, as discussed in Secs. 5.3 and 5.4.

As we also noted in Sec. 2.5, a third-order Hamiltonian gives rise to two kinds of effects: directly, to third-order drifts; more subtly, to *second-order* polarization effects. (Polarization is related to a derivative of the Hamiltonian with respect to potential.) For polarization, it is not necessary to actually construct H_3; one can instead calculate the pullback transformation (Appendix D) directly to second order (which is just one order higher than the first-order effect that all authors retain). Formally, the reason that third-order drifts also play a role is that if one retains only second-order drifts and second-order polarization, there is a small violation of energy conservation. For more details, see the discussion in Sec. I.2.

Clearly small terms can only be important if lower-order fluxes such as $\langle \int d\boldsymbol{v}\, F_1 V_{D,1} \rangle$ either vanish or are much smaller after flux-surface/ensemble averaging than they first appear to be. As demonstrated in Appendix K, in some simple limits it is straightforward to show that this is in fact true (because of quasineutrality constraints and other effects), as it must be in order for gyro-Bohm scaling of momentum transport to hold.

5.2 The Chapman–Enskog-like approach of Parra & Calvo

Recently Calvo & Parra (2012) have spelled out many of the details of the low-flow problem in a definitive manuscript describing the order-by-order solution of the gyrokinetic equation at long wavelengths. Although there is no substitute for study of that comprehensive paper, we can give the flavor of the argument here. The basic idea is reminiscent of the Chapman–Enskog procedure for solution of the kinetic equation for classical transport. In that method, solvability conditions at higher order constrain the time evolution of lower-order quantities. [This is a familiar technique in perturbation theory, and indeed was the method used by Frieman & Chen (1982) to derive the nonlinear gyrokinetic equation.]

In the present context, the theory proceeds order-by-order as follows:

- $O(\epsilon^0)$: The lowest-order distribution F_0 is determined to be a Maxwellian with slowly varying density and temperature. (Because of the low-flow ordering, flow is absent from F_0.)
- $O(\epsilon^1)$: F_1 is determined in terms of F_0.
- $O(\epsilon^2)$: The solvability condition for F_2 constrains the slow-time evolution of n_0 and T_0. Then F_2 is determined. No constraint on the radial electric field E_r is obtained at this order.
- $O(\epsilon^3)$: The solvability condition for F_3 constrains the slow-time evolution of n_1 and T_1.[52] (Again, no constraint on E_r is obtained.) Then F_3 is determined.[53]
- $O(\epsilon^4)$: The solvability condition for F_4 constrains the slow-time evolution of n_2 and T_2. Only at this order does a constraint on E_r appear, by taking the difference of the equations for $\partial_t n_i$ and $\partial_t n_e$. In other words, one deduces the form of the momentum equation at $O(\epsilon^4)$. One does not need to explicitly calculate f_4.[53]

This calculation is systematic. It shows that the plasma is intrinsically ambipolar through second order and that details of F_3 must be determined. It provides a very complete alternate way of demonstrating that third-order calculations are required in a standard full-F treatment of gyrokinetic momentum conservation.

5.3 The gyrokinetic momentum conservation law of Scott & Smirnov

The methodology of Parra and coworkers is complete in itself; the formalism can be developed without explicit reference to Hamiltonians, Lagrangians, field theory, Noether theorems, *etc.* However, there is ample evidence from many fields of physics that such concepts can be greatly unifying, and they furthermore cast a sufficiently different perspective (both physical and mathematical) on the gyrokinetic formalism that it is nontrivial to establish the consistency between the approaches. Thus we turn to a discussion of the Lagrangian field-theoretic conservation law of Scott & Smirnov. One expects that ordering issues will arise there as well, since the particle and the gyrocenter dynamics are connected by the particle-to-gyrocenter transformation.

We have already quoted the form of their result in Eq. (92), and motivated it with the approximate calculation in Sec. 4.2. But hidden in the derivation of the

[52] The true density is $n = n_0 + n_1 + \cdots$. Usually Chapman–Enskog theory is developed with multiple-scale perturbation theory. In that case there is sufficient freedom that one can choose to carry all of the hydrodynamic quantities in the lowest-order distribution. That does not happen in the particular method chosen by Parra & Calvo, who do not use multiple-scale expansion. Of course, one must come to the same physical conclusions regardless of the mathematical technique.

[53] The discussion here of the $O(\epsilon^3)$ and $O(\epsilon^4)$ calculations is not given in the paper by Calvo & Parra (2012); it has been distilled from private discussions (2012) with those authors.

final equation is a crucial cancellation (see footnote 51); without it, one would have to work even harder, to $O(\epsilon^4)$. To understand the origin of the cancellation, we recall the formula for single-particle canonical momentum $\boldsymbol{p} = m\boldsymbol{v} + q\boldsymbol{A}/c$. For particle motion in axisymmetric fields, the covariant toroidal component of $\boldsymbol{p}$, $p_\varphi \doteq \boldsymbol{p}\cdot\boldsymbol{e}_\varphi$ is conserved according to standard Lagrangian mechanics. Because $\boldsymbol{e}_\varphi = R\widehat{\boldsymbol{\varphi}}$, p_φ is the toroidal component of canonical *angular* momentum. For gyrocenters, the relevant canonical momentum is[54] $\boldsymbol{P} = mU_\parallel\widehat{\boldsymbol{b}}+q\boldsymbol{A}/c$. Then the toroidal component is obtained by dotting with $\boldsymbol{e}_\varphi$ and defining $b_\varphi \doteq \widehat{\boldsymbol{b}}\cdot\boldsymbol{e}_\varphi$:

$$P_\varphi = mU_\parallel b_\varphi + \frac{q}{c}A_\varphi = \underbrace{mU_\parallel b_\varphi}_{O(1)} - \underbrace{\frac{q}{c}\psi}_{O(\epsilon^{-1})} \ , \tag{108}$$

where ψ is the poloidal flux. One readily establishes that the field term is one order larger than the gyrocenter term. (That is, of course, the entire basis for the gyrokinetic expansion.) But the gyrocenter conservation law describes the evolution of the physical momentum of the gyrocenter, not the canonical momentum. If raw factors of ψ were to remain in the final equation, as in, say, $\Pi \sim \psi Q$ for some quantity Q, a third-order contribution to Π could arise from a fourth-order Q: $\epsilon^3 = \epsilon^{-1}\times\epsilon^4$. Scott & Smirnov demonstrated that this does not happpen by virtue of the quasineutrality condition. Thus, their final equation contains no raw factors of ψ; this is one of their most important conclusions.[55]

In the equation of Scott & Smirnov, all of the fluxes are now determined in terms of the gyrocenter Hamiltonian H, which has so far not been specified. Scott & Smirnov concluded that a consistent treatment of angular momentum conservation could be obtained by using a Hamiltonian that was accurate through second order, i.e., $H \approx H_0+\epsilon H_1+\epsilon^2 H_2$. Due to the way the gyrokinetic Poisson equation is constructed in the variational formalism of gyrokinetic field theory (involving the functional derivative of H with respect to potential, which is equivalent to implementing the pullback transformation), that means that the associated Poisson equation would retain polarization effects through only first order. That is the order of accuracy that is retained in most current codes, so Scott & Smirnov concluded that those codes (perhaps with small modifications) could be satisfactory from the point of view of the momentum issues.

[54] The use of $\boldsymbol{P}$ follows the work of Scott & Smirnov (2010) and Brizard & Tronko (2011), but it disguises an important subtlety. The expression used here does not explicitly contain the angular momentum associated with the magnetic drifts of the gyrocenter; although small, that is a physical effect. Brizard (unpublished, 2013) has addressed this issue by reconsidering the choice that has been conventionally made for the vector generating function (Sec. D.3.2) for the guiding-center displacement. Beginning with the work of Littlejohn (1982), the perpendicular part of that function has been set to zero for convenience. This implies a particular *definition* of the gyrocenter position (and the other gyrocenter variables). However, other choices are possible, and Brizard has demonstrated that a particular nonzero choice leads to expressions in which the magnetic drifts appear explicitly in the momentum conservation law for gyrocenters. The implications of this observation remain to be fully explicated at the time of writing.

[55] With the benefit of hindsight, it is obvious that this cancellation must happen because an equation for the time evolution of physical momentum cannot contain a term involving a raw vector potential. But that it indeed does happen is an important consistency check on the manipulations.

[For strict energetic consistency, use of H_2 implies either that second-order drifts should be included in the kinetic equation (most codes do not do that) or that an energy-conserving partial linearization should be done along lines that Scott & Smirnov have discussed (not all codes implement that).]

It is important to understand the meaning of 'consistent.' In the language of Scott & Smirnov, a consistent treatment means that the formalism obeys energy and momentum conservation laws (derived from the same Lagrangian). However, just because a conservation law exists does not mean that the fluxes in that law describe all of the relevant plasma physics or are calculated with sufficient accuracy. With this observation in mind, we argue that the logic of Scott & Smirnov regarding the adequacy of H_2 is incomplete. A particular difficulty relates to determining the size of the terms related to the perpendicular Reynolds stress. We give an abbreviated discussion here, deferring more detailed analysis to Appendix J.

Schematically,[56] the form of the gyrokinetic momentum conservation law is [see Eqs. (80) and (83) of Scott & Smirnov]

$$\frac{\partial}{\partial t}\left[\begin{pmatrix}\text{toroidal projection of}\\ \text{gyrocenter parallel}\\ \text{angular momentum}\end{pmatrix}+\begin{pmatrix}\text{toroidal projection of}\\ \text{the angular momentum of}\\ \text{the gyrocenter (effective)}\\ \boldsymbol{E}\times\boldsymbol{B}\text{ drift}\end{pmatrix}\right] = -\frac{\partial}{\partial\psi}(\Gamma_\parallel+\Gamma_\perp), \tag{109}$$

where, as shown by Scott & Smirnov,

$$\Gamma_\parallel \sim \left\langle \int d\boldsymbol{v}\, F V^\psi v_\varphi \right\rangle, \tag{110a}$$

$$\Gamma_\perp \sim \left\langle \int d\boldsymbol{v}\, F \left(\frac{\partial H}{\partial \boldsymbol{E}}\right)^\psi \frac{\partial \phi}{\partial \varphi} \right\rangle + \begin{pmatrix}\text{similar}\\ \text{terms}\end{pmatrix}, \tag{110b}$$

and $v_\varphi \doteq v_\parallel b_\varphi$ is the toroidal projection of the parallel velocity of the gyrocenter. As stated in the last section, V^ψ is directly related to the Hamiltonian. To intuitively understand the perpendicular term, which is the most subtle one in the theory, note that

$$\partial_\varphi \phi \sim E_\varphi \sim u_E^\psi, \tag{111}$$

where $\boldsymbol{u}_{\boldsymbol{E}}$ is the $\boldsymbol{E}\times\boldsymbol{B}$ velocity. (In this discussion, signs, dimensional factors, and some metric or geometric factors are dropped. The rigorous manipulations with covariant and contravariant indices are spelled out in Appendix J.) Now in the simplest case (cold-ion limit, constant B), H_2 is quadratic in $\boldsymbol{E}_\perp$. [In Appendix F, this ponderomotive effect is explained in terms of the conservation

[56] More generally, the second term on the left-hand side of Eq. (109) involves the radial polarization. In detail, that includes pressure-related contributions that need not vanish when the electric field vanishes.

of the true magnetic moment $\overline{\mu}$, which is defined in the frame moving with the $\boldsymbol{E}\times\boldsymbol{B}$ velocity (thus ensuring Galilean invariance).] One then finds that

$$\left(\frac{\partial H_2}{\partial \boldsymbol{E}}\right)^{\psi} \sim E^{\psi} \sim u_{E\varphi} = O(\epsilon). \tag{112}$$

[The differentiation of H_n with respect to $\boldsymbol{E} = -\boldsymbol{\nabla}\phi$, and other similar manipulations involving higher-order gradients of ϕ, is equivalent to the calculation of the $(n-1)$st term of the pullback transformation T^*.] Upon inserting Eqs. (112) and (111) into Eq. (110b), one sees that the contribution from H_2 to $\Gamma_\perp$ is the conventional perpendicular Reynolds stress:

$$\Gamma_\perp^{[2]} \sim (nm)_i \langle u_{E\varphi} u_E^{\psi} \rangle = O(\epsilon^2). \tag{113}$$

The bracketed superscripts denote the order of H from which the term derives, and the ordering estimate follows just from the raw sizes of the associated flows. Similarly, there is an nth-order generalized Reynolds stress $\Gamma_\perp^{[n]}$ that stems from H_n (or T_{n-1}).

The basic ordering argument is now the following. (We amplify here the discussion in Sec. 2.2.2.) Although $\Gamma_\perp^{[2]}$ is nominally one order larger than $\Gamma_\perp^{[3]}$, in the low-flow ordering (and assuming an up–down symmetric tokamak) the statistical correlation between $u_{E\varphi}$ and u_E^{ψ} is very weak, making the net Reynolds stress arising from H_2 one order smaller. The Reynolds stress arising from H_3 is thus comparable and must be considered, unless for some reason (which has not yet been found) that stress also experiences a near cancellation.[57] It does not appear that it is small due to a microscopic symmetry. It may be numerically small in subsidiary orderings involving such things as small $B^{\mathrm{pol}}/B_{\mathrm{tor}}$, or due to particular properties of the underlying microturbulence.

Similar statements about the role of correlations apply to $\Gamma_\parallel$, which also has a second-order piece $\Gamma_{\parallel,2}$. However, it is not hard to argue that $\Gamma_{\parallel,2}$ stems from H_1 [F_0 does not contribute for the same reasons as in Sec. 5.1.2, so the effect involves the product of δF ($\doteq F - F_0$) and H_1], and similarly H_2 produces a third-order flux; thus H_3 is negligible for $\Gamma_\parallel$.

5.4 The Ordering Problem Restated

To try to be as clear as possible, we will restate and summarize the ordering problem here in a self-contained way, directly using the Scott & Smirnov momentum conservation law to demonstrate it, and keeping track of all dimensional factors. Their conservation law is given here in Eqs. (92) and (93) [those are equivalent to Eqs. (80) and (83) of Scott & Smirnov (2010), just with slightly different notation]. We focus on the mean long-wavelength component of the flux-surface-averaged toroidal rotation in the gyro-Bohm regime, which evolves at the transport rate $\partial/\partial t \sim D_{\mathrm{gB}}\nabla_\perp^2 \sim (c_s/L)\rho^2/L^2 \sim \epsilon^2 c_s/L$. [By 'mean' long-wavelength component, we mean the ensemble or time-averaged

[57] B. Scott (private communication, 2013) believes that such a cancellation exists.

equilibrium-scale component. There are, of course, components of the toroidal rotation with shorter radial wavelengths, related to zonal flows and geodesic acoustic modes driven by turbulence, that fluctuate on shorter time scales, just as there are zonal components of the flux-surface-averaged density and temperature.] In the low-flow regime where the parallel flow $U_\parallel \sim \epsilon c_s$, and using $b_\varphi = R\hat{\varphi}\cdot\mathbf{B}/B \sim L$ (all equilibrium scales are represented as order L), the first term in Eq. (92) is of the order

$$\frac{\partial \mathcal{P}_{\varphi\parallel}}{\partial t} \sim \epsilon^2 \frac{c_s}{L} m_i n_i \epsilon c_s L \sim \epsilon^3 n_i T_e. \tag{114}$$

The troublesome term in Eq. (92) is the last one, which is related to the radial flux of the toroidal component of the $\boldsymbol{E}\times\boldsymbol{B}$ momentum. Upon using $\partial_\varphi \phi \sim k_\perp L\phi$ and $(\partial H/\partial \mathbf{E})^\psi = (\partial H/\partial \mathbf{E})\cdot\nabla\psi$, with $\nabla\psi \sim \psi/L$, one finds that the last term in Eq. (92) is of the order

$$\frac{\partial}{\partial\psi}\sum_s \overline{N}_s \left\langle \int_P F \left(\frac{\partial H}{\partial \mathbf{E}}\right)^\psi \frac{\partial\phi}{\partial\varphi}\right\rangle \sim \frac{1}{\psi} n_i \frac{H}{k_\perp\phi}\frac{\psi}{L} k_\perp L\phi \sim n_i H. \tag{115}$$

Thus balancing this with the first term means that one needs to keep the third-order Hamiltonian, $H_3 \sim \epsilon^3 T$, in order to keep all effects that can contribute to the momentum flux in the low-flow gyro-Bohm ordering. This is the same result found from the previous quasineutrality order-of-magnitude argument, Sec. 5.1.2, and it brings one back to the original concern of Parra & Catto.

One might have thought that the momentum flux from H_2 would be even bigger than the momentum flux from H_3, so that one would not need to worry about H_3. But as we have discussed elsewhere (see for example Sec. 2.2.2 and Appendix K), the momentum flux from H_2 almost averages to zero due to symmetry properties, so the correlation between the radial and binormal components of the $\boldsymbol{E}\times\boldsymbol{B}$ velocity is almost zero, with only a weak correlation that introduces another factor of ϵ, so the net contribution to the last term in Eq. (92) from H_2 turns out to also be of order $\epsilon^3 n_i T_e$, as it should be if a low-flow gyro-Bohm ordering holds.

Similarly, it is conceivable that the nominally third-order fluxes stemming from H_3 are also in fact of higher order. However, it does not appear that one can argue those away by using symmetry considerations; some other mechanism must be invoked. For the special case of slab geometry, an unpublished calculation of Krommes (2011; see Sec. J.2.2) suggests that in the cold-ion limit (i.e., no FLR effects) the nominally third-order Reynolds stresses are in fact of fourth order and hence negligible; that result has been verified by Parra & Calvo (private communication, 2011). However, for the realistic case of $k_\perp\rho_i \sim \mathcal{O}(1)$, it has not been possible to argue that those terms are negligible, so in general it is necessary to include them (or to use the alternative formulation of Parra & Catto that does not need them) unless proven otherwise.

5.5 Answers to the Charge

With the previous discussion, one can address the questions posed to the Study Group. The answers were summarized in Sec. 1.4; we repeat them here for convenience, in some cases with some embellishment.

In addition to the questions identified in the Charge, a more basic question must be answered first:

[0] Are the physics contents of the particle and the gyrocenter conservation laws equivalent or different?

The physics contents are equivalent. The mathematical representations differ. The connection is given by the pullback transformation $f = \mathrm{T}^*\overline{F}$.

With question [0] answered, a major source of uncertainty is removed and one can focus on the original questions:

[1] Are the gyrokinetic equations used in existing codes (including `GTS`, `GYRO`, and `GTC`) adequate to simulate the evolution of turbulence over transport time scales?

No existing gyrokinetic code contains all of the terms that are required for completely consistent simulations of momentum transport in the low-flow gyro-Bohm regime on transport time scales. However, efforts in that direction are being made in the context of coupling turbulence-time-scale gyrokinetic simulations with long-time transport solvers.

It should be noted that most present codes are adequate (perhaps with some extension to fully incorporate second-order Hamiltonian effects into the fluxes) to study turbulent momentum fluxes on the shorter turbulence time scale in the high-flow regime, such as if there is strong beam injection. Present codes are also adequate to study the particle and heat fluxes on the shorter turbulence time scale. The equilibrium-scale toroidal rotation in the low-flow, gyro-Bohm regime appears to be too weak to affect the turbulence level, but even these small levels of flow may be important for the MHD stability of resistive wall modes, and it is of general interest to understand the higher-order terms that may contribute to residual stress in this regime.

An important area of current research is developing codes to handle turbulence in the edge/pedestal region, where eddy sizes might not be much smaller than local gradient scale lengths and effects like ion orbit loss to the wall may be important so that a low-flow gyro-Bohm ordering might not hold locally. In such cases it might be sufficient to use a second-order Hamiltonian.

[2] Are second- and/or third-order corrections in the normalized Larmor radius required for (a) the gyrokinetic equation; (b) the gyrokinetic Maxwell equations?

In principle, under a particular set of assumptions relevant to a specific physics regime (see Sec. 2.1), third-order corrections are required for the

gyrokinetic equation with a global full-F approach, while second-order corrections are necessary in the gyrokinetic Maxwell equations. If a fluid momentum equation is adjoined to the gyrokinetic equation, it is possible to work to just second order in the kinetic equation (by using a moment method described by Parra & Catto).

[3] If the answer to [2] is positive, is the vorticity equation proposed by Parra & Catto a possible fix to extend the validity of gyrokinetic codes to long time scales?

Yes. This must be implemented carefully. Scott (private communication, 2012) has emphasized that whereas the momentum conservation law based on the gyrokinetic equation is firm and clean since it is derived variationally, the conservation properties of a coupled system of gyrokinetic and fluid system may be subtle and certainly need to be demonstrated.

[4] If second-order gyrokinetics is sufficient, what needs to be done to fully implement this into codes with sufficient numerical accuracy, either directly or through separate gyrokinetic transport equations for the long time scale?

Second-order gyrokinetics should be sufficient if it is coupled directly to a momentum transport equation, or if one is in a parameter regime that relaxes the low-flow gyro-Bohm ordering. Second-order effects could be added in principle to existing codes (along with adding the momentum equation). The polarization term in the gyrokinetic Poisson equation should be calculated in a particular way to be consistent. Calculating the effects of small second-order drifts accurately when they are combined with larger first-order drifts in the same equation may require higher numerical accuracy than is typical at present. Analysis of the accuracy requirements for the several types of gyrokinetic codes (which use various PIC and continuum discretization and time-integration methods that possess various conservation properties) should be carried out. The accuracy requirements might be more manageable near the plasma edge where the local $\epsilon \doteq \rho/L$ (the ratio of the gyroradius ρ to the local gradient scale length L) is not as small and there is less separation between the turbulence and profile time-evolution scales.

6 Discussion

6.1 Outstanding questions

- **If fluxes related to H_3 are important, what do they mean?** This question is still not fully resolved. Parra hypothesizes that they are related to residual stress and intrinsic rotation. This is, of course, another area of intense current interest. Extant calculations of residual stress are difficult and controversial in their own right. A thorough analysis of those

calculations and an understanding of their relation, if any, to the confusing third-order fluxes is beyond the immediate scope of the present Study Group. It is, however, a very interesting area for future research.

At least part of the significance of the H_3-related contributions to momentum flux relates to the fact that the gyrocenter velocity is not exactly the laboratory flow velocity because of nonlinear and FLR corrections. To make contact with laboratory diagnostics, one must ultimately decompose the natural gyrocenter momentum fluxes into, say, diffusive, pinch, and residual stress terms expressed in terms of laboratory quantities. The subtleties involved with such decomposition in the presence of all nonlinearities are not yet understood and are not likely to be for some time.

- **What are the implications for the fidelity of current codes?** In general, one can designate a truncation of the gyrokinetic–Maxwell system as (n, m), where n refers to the highest order of the drifts that are retained in the gyrokinetic equation and m refers to the highest order of the (polarization) terms retained in the gyrokinetic Poisson equation. The basic truncation that is implied by the variational principle is $(n, n-1)$ because the gyrokinetic Poisson equation is obtained by the functional derivative of the Hamiltonian with respect to potential, which lowers the order by one. Thus if H_3 is important, nontrivial second-order terms must be included in the polarization contribution to the gyrokinetic Poisson equation. To ensure the most straightforward form of energetic consistency, third-order drifts stemming from H_3 would have to be included in the gyrokinetic equation, i.e., one would have a $(3, 2)$ truncation. Implementing the third-order drifts appears to be daunting. One may speculate whether simplifications are possible. For example, Scott & Smirnov have shown how to back-construct an effective Lagrangian that provides an energetically consistent $(1, 1)$ closure[58] (used by many of the extant codes). Related techniques may be useful at higher order.

- **What issues remain in developing the hybrid approach?**

  Parra & Catto (2008) advocate an alternate approach in which the real-space momentum conservation law is used directly to determine toroidal flows. A second moment may be taken that relates the required stress tensor to a third moment that can then be evaluated from the solution of the gyrokinetic equation. This is a well-known trick in neoclassical theory that reduces the required order of accuracy by one power of the expansion parameter. A related trick is used in the standard derivation of drift-kinetic MHD (Kulsrud, 1964, 1983). By explicitly using a momentum conservation law (closed by calculating the pressure tensor from moments of the solution of the drift kinetic equation), one only needs to include the lowest-order $\boldsymbol{E} \times \boldsymbol{B}$ drift (of order c_s in the high-flow MHD ordering)

[58] Dubin et al. (1983) were the first to identy the energy constant associated with the $(1, 1)$ closure.

in the drift-kinetic equation. The higher-order $\boldsymbol{\nabla} B$ and curvature drifts are neglected in that drift-kinetic equation, even though they would be necessary if one were going to try to calculate currents (and resulting $\boldsymbol{j} \times \boldsymbol{B}$ forces) directly from particle drifts, since $j \sim (c/B)\nabla p \sim (\rho_s/L)enc_s$.

GS2 and GYRO (and GENE, though we focus on the first two here) already do something similar to the real-space approach. Those codes use a local δF formulation instead of a global full-F one. (GYRO and GENE can also operate with finite-radial-width simulation domains, which capture some global effects like turbulence spreading.) The long-time evolution of the equilibrium-scale particle-, momentum-, and energy-density profiles are then determined by transport equations that are of one order higher than that of the physics incorporated in the δF evolution. A systematic derivation of those equations was presented in an important paper by Sugama & Horton (1998), and a comprehensive discussion has recently been given by Abel et al. (2013). As the GYRO and GS2 teams have emphasized, their present equations are valid for an intermediate-flow ordering, $1 \gg \mathcal{M} \gg \epsilon$, where terms of order $\mathcal{M} \doteq u_\varphi/c_\mathrm{s}$ are retained but terms of order $\mathcal{M}^2$ are neglected, and it is acknowledged that not all of the terms required for the low-flow ordering, $\mathcal{M} \sim \epsilon$, are yet included. Parra is working with M. Barnes to perturbatively calculate the second-order terms required for the low-flow ordering [for recent results, see Barnes et al. (2013)]. (Here second order refers to the order to which the pullback transformation T^* must be calculated in order to relate the gyrocenter and particle phase spaces. As discussed in the previous paragraph, that is equivalent to saying that H_3 is important.)

Similarly to these δF codes, most full-F codes could use their present formulation to study momentum fluxes in the high- or intermediate-flow regimes on the turbulence time scales (short compared to the transport time scale so profiles do not evolve significantly). This can be used to study diffusive and pinch contributions to the momentum flux. The issue of an 'intrinsic' momentum flux (separate from pinch or diffusive terms) is more subtle, as according to recent papers by Sugama et al. (2011) and Parra et al. (2011b) it can be shown that the intrinsic momentum flux should vanish in high- or intermediate-flow regimes in an up–down symmetric tokamak in the gyro-Bohm regime. The implication is that any intrinsic momentum flux that arises from profile variations in a global code should get weak when extrapolating to smaller ρ_* in the gyro-Bohm regime.

- **How do collisions modify the considerations?** In collisionless theory, equilibrium distributions must be constructed from the constants of motion (for example, the magnetic moment μ). The modern gyrokinetic formalism is ideally suited to construct and work with those constants of motion. Unfortunately, a workable extension of the Lagrangian methods to include collisions in the Lagrangian itself is not yet at hand. However, there are various ways that collisions have been incorporated in gyrokinetic

codes, indicating that it is possible to treat collisions (thus neoclassical effects) and turbulence on equal footing. [See, for example, the recent work of Calvo & Parra (2012).]

- **What is the impact of numerical errors?** Is there a best way to formulate gyrokinetic algorithms in order to minimize numerical errors, particularly in view of the presence of small but physically important terms?

- **Are other ordering regimes of importance?** What is the impact of flows in the low-flow gyro-Bohm ordering? Are there subsidiary orderings (like large aspect ratio or expansions in weak poloidal field) that can usefully simplify the equations in various parameter regimes?

  In the standard low-flow gyro-Bohm regime, flows are too weak by themselves to affect energy or particle transport, since the resulting shearing rate $\nabla \mathbf{u_E} \sim \epsilon c_s/L$ is small compared to the drift-wave growth rates $\sim c_s/L$. This point is also made in the conclusions of Sugama et al. (2011). [Perhaps further thought should be given to modes at longer wavelengths and/or very close to marginal stability where the growth rates are small compared to c_s/L, but those would be in a different ordering regime.] However, those flows might serve as triggers or boundary conditions for other mechanisms, so perhaps further study is needed, particularly in intermediate regimes such as near a transition between gyro-Bohm and Bohm scaling. Furthermore, even such small flows might be important for MHD stability because of the importance of mode locking and resistive wall effects. Also, it is of general interest to understand the effects of H_3 because that is the only way to obtain a residual stress in the low-flow gyro-Bohm regime that could drive intrinsic rotation (other than edge effects).

  It is possible that transport barriers or the edge region might be in a Bohm regime because the eddy sizes are not much smaller than the local steep gradient scales and because of effects like ion orbit losses to the wall. In this case it might be sufficient for edge gyrokinetic codes to only keep second-order Hamiltonian effects. Most existing gyrokinetic codes use only a first-order Hamiltonian or approximate second-order effects, but they could be extended. One of the outcomes of the study of momentum transport stimulated by the Parra & Catto work is a better appreciation of the high accuracy requirements. Before that, the need for second-order terms even in the Bohm regime was not appreciated. Future work that studies residual stress or general momentum transport should be careful to study the scaling of the observed rotation and momentum fluxes in order to understand whether all of the terms of the relevant order are being kept and how the rotation will scale to larger machines.

  Existing core gyrokinetic codes (perhaps with extensions to handle second-order Hamiltonian effects) can also be used to study momentum transport in a high- or intermediate-flow regime, such as can occur with beam injection. While beam-driven rotation gets weaker in larger reactor-scale machines, there are some recent indications (Staebler & St. John, 2006;

Budny, 2009) that beam-driven rotation might still have a significantly favorable effect in ITER. It would be interesting to investigate this possibility further.

- **What are the consequences of asymmetry?** The main discussion has been conducted in the context of axisymmetric tokamaks. In this case, the turbulent transport is intrinsically ambipolar through second order, so it is natural that third-order effects must be considered. In the presence of asymmetry, this conclusion does not hold, as emphasized by Sugama et al. (2011). He concluded that further research is necessary to properly understand the interaction between small amounts of asymmetry and the higher-order effects.

6.2 Some history

To be fair, we should relate some of the history and acknowledge a contribution of one of us to the confusion on this topic. In the early summer of 2010, B. Scott gave a talk at PPPL presenting an early version of a momentum conservation law in which the vector potential $\boldsymbol{A}$ for the background field $\boldsymbol{B}$ appeared explicitly, without a gradient operating on it. Because $\boldsymbol{A}$ is large, $O(\epsilon^{-1})$ relative to the kinetic term in the canonical momentum, the presence of such a term would have increased the accuracy requirements by one power of ϵ, as we explained in Sec. 5.3. Although one of us (GWH) demonstrated that there were ordering problems with that equation, he did not at that time give a complete description of the ordering arguments that Parra & Catto have made. Scott's later momentum-conservation equation solved the problem of a 'bare' $\boldsymbol{A}$ by identifying an important cancellation resulting from the quasineutrality constraint. His equation is important in that it provides a proof of the existence of momentum conservation in Hamiltonian gyrokinetics in the absence of ordering assumptions. However, as we have seen, it does not avoid the fundamental ordering problems that Parra & Catto have been pointing out.

6.3 Summary

In summary, Parra & Catto have pointed the community to an extremely subtle and difficult issue, namely the proper treatment of the transport of toroidal angular momentum in an axisymmetric torus. As Parra, Catto, Sugama, and Calvo have shown in multiple publications, an axisymmetric tokamak is ambipolar through $O(\epsilon^2)$. That implies that in a full-F simulation of a low-flow gyro-Bohm regime that relies only on the gyrokinetic Poisson equation for the potential, one must retain terms of third order in the gyrokinetic equation and of second order in the gyrokinetic Poisson equation. A hybrid scheme in which the fluid momentum equation is adjoined to the gyrokinetic simulation appears to be fruitful; in that case, it is possible to work to just second order in the kinetic equation.

The flows in the standard low-flow gyro-Bohm regime are too weak by themselves to affect energy or particle transport, though they might be important for MHD stability. There are other parameter regimes that are also of interest, such as near transport barriers or the edge where local gyro-Bohm assumptions might break down and flows can have a stronger impact on turbulence; there it would be sufficient in a direct approach to retain second-order terms in the gyrokinetic equation and first-order terms in the gyrokinetic Poisson equation.

Progress with this topic requires a deep understanding of the interpretation of the gyrokinetic formalism. Some difficult issues have been resolved; others remain. The original concerns involving ordering and truncations have led to other important questions regarding the physics of residual stress and intrinsic rotation. The subject is sufficiently complex and multifaceted that a comprehensive understanding will be achieved only over a period of years, not months. The field continues to evolve; an overview by Peeters et al. (2011) summarized the status as of a few years ago, and we noted in Sec. 1 recent and ongoing unpublished research by various authors.

Finally, it is worth emphasizing that none of the conclusions of this report imply a wholesale indictment of existing gyrokinetic codes. Those have been and will continue to be very productive for the elucidation of diverse basic physics phenomena, quantitative validation tests with experimental data, and simulations of future machine designs. Some upgrades may be necessary in order to deal with particular issues relating to momentum transport, but that is the nature of an ongoing research program. Should a question regarding continuing funding of gyrokinetic simulations arise, the answer is that they deserve more support — certainly not less.

Acknowledgements

We are grateful to R. Betti for suggesting[59] that we investigate this topic. We have benefited from physics discussions with A. Boozer, A. Brizard, J. Burby, I. Calvo, P. Catto, C.-S. Chang, Y. Idomura, W. W. Lee, J. Parker, F. Parra, H. Qin, B. Scott, J. Squire, and H. Sugama. Useful suggestions about, and corrections to various drafts of the manuscript were received from many of those people and were greatly appreciated. This work was supported by the U. S. Department of Energy Contract DE-AC02-09CH11466.

A Generalized coordinates

Most of the following formulas are taken directly from D'haeseleer et al. (1991). Additional, more advanced material from differential geometry can be found in Fecko (2006).

[59] "Since God had commanded it, it was necessary that I do it. Since God commanded it, even if I had a hundred fathers and mothers, even if I had been a King's daughter, I would have gone nevertheless." — Joan of Arc

We use generalized coordinates z^i and assume that we are given a transformation to the Cartesian coordinate system $\boldsymbol{x}(\boldsymbol{z})$. We begin by using physicists' notation, e.g., boldfaced vectors that have space-dependent coefficients. However, in Sec. A.6 we introduce the more compact notation of differential geometry in which one speaks of vector and covector *fields*, which are represented by first-order differential operators.

A.1 Basis vectors

Vectors $\boldsymbol{V}$ are represented as

$$\boldsymbol{V} = V^i \boldsymbol{e}_i \quad \text{(raised index} \equiv \text{contravariant coordinates)}, \tag{A1}$$

where the summation convention for repeated indices is used and the *basis vectors* are

$$\boldsymbol{e}_i \doteq \frac{\partial \boldsymbol{x}}{\partial z^i} \equiv \partial_i \boldsymbol{x}. \tag{A2}$$

Note that these are not necessarily normalized; their length is given by the scale factors (useful mostly for orthogonal coordinate systems) $h_i \doteq |\boldsymbol{e}_i|$. One may visualize the components of a vector as the elements of a column vector.

Covectors $\boldsymbol{\alpha}$ are represented as

$$\boldsymbol{\alpha} = \alpha_i \boldsymbol{e}^i \quad \text{(lowered index} \equiv \text{covariant coordinates)}, \tag{A3}$$

where the *cobasis vectors* (sometimes called *reciprocal* or *dual basis vectors*) are

$$\boldsymbol{e}^i \doteq \boldsymbol{\nabla} z^i. \tag{A4}$$

One may visualize the components of a covector as the elements of a row vector.

The basis and cobasis vectors are orthonormal: $\boldsymbol{e}^i \cdot \boldsymbol{e}_j = \delta^i_j$. In three dimensions, one has

$$\boldsymbol{e}_i = \frac{\boldsymbol{e}^j \times \boldsymbol{e}^k}{\boldsymbol{e}^i \cdot (\boldsymbol{e}^j \times \boldsymbol{e}^k)} \tag{A5}$$

and analogously with the subscripts and superscripts reversed. Here (i, j, k) is a cylic permutation.

Strictly speaking, it is not legitimate to form the scalar product of two vectors; one should use one covector and one vector. Thus

$$\boldsymbol{\alpha} \cdot \boldsymbol{V} = \alpha_i V^i. \tag{A6}$$

A.2 Metric coefficients

Vectors and covectors can be defined on general manifolds that possess no concept of distance. When distance is defined, one has a Riemannian manifold and a metric tensor g. The fully covariant metric coefficients are defined by

$$g_{ij} \doteq \boldsymbol{e}_i \cdot \boldsymbol{e}_j. \tag{A7}$$

The square of the length element is

$$(dl)^2 = g_{ij} dz^i \, dz^j. \tag{A8}$$

Also,

$$\boldsymbol{e}_i = g_{ij} \boldsymbol{e}^j, \tag{A9}$$

Furthermore, one can introduce the fully contravariant metric coefficients as the inverse of g_{ij}:

$$g^{ik} g_{kj} = \delta^i_j. \tag{A10}$$

Those can be used to raise and lower indices according to

$$V^i = g^{ij} V_j, \quad V_i = g_{ij} V^j. \tag{A11}$$

A.3 The Jacobian

The Jacobian J is defined by

$$d\boldsymbol{x} = dz^1 \, dz^2 \, dz^3 \, J, \tag{A12}$$

or

$$J \doteq \frac{\partial(\boldsymbol{x})}{\partial(\boldsymbol{z})}. \tag{A13}$$

One has

$$J = \boldsymbol{e}_1 \cdot (\boldsymbol{e}_2 \times \boldsymbol{e}_3), \tag{A14a}$$

$$J^{-1} = \boldsymbol{e}^1 \cdot (\boldsymbol{e}^2 \times \boldsymbol{e}^3). \tag{A14b}$$

Thus, Eq. (A5) becomes

$$\boldsymbol{e}_i = J(\boldsymbol{e}^j \times \boldsymbol{e}^k) \quad \text{(indices and } J \text{ upstairs)}, \tag{A15a}$$

$$\boldsymbol{e}^i = \frac{1}{J}(\boldsymbol{e}_j \times \boldsymbol{e}_k) \quad \text{(indices and } J \text{ downstairs)}. \tag{A15b}$$

With $g \doteq \det[g_{ij}]$, one also has

$$J = \sqrt{g}. \tag{A16}$$

In an orthogonal coordinate system, $J = h_1 h_2 h_3$.

A.4 Dot and cross products

The scalar product is

$$\boldsymbol{A} \cdot \boldsymbol{B} = A^i B_i. \tag{A17}$$

The cross product can be represented as

$$(\boldsymbol{A} \times \boldsymbol{B})_i = J \epsilon_{ijk} A^j B^k, \tag{A18a}$$

$$(\boldsymbol{A} \times \boldsymbol{B})^i = \frac{1}{J} \epsilon^{ijk} A_j B_k; \tag{A18b}$$

the last expression can be written as the determinant

$$\boldsymbol{A}\times\boldsymbol{B}=\frac{1}{J}\begin{vmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3\\ A_1 & A_2 & A_3\\ B_1 & B_2 & B_3\end{vmatrix}. \tag{A19}$$

As a simple generalization, the dot product of a covector and a vector cross product can be represented as the determinant

$$\boldsymbol{\alpha}\cdot(\boldsymbol{A}\times\boldsymbol{B})=\alpha_i(\boldsymbol{A}\times\boldsymbol{B})^i=\frac{1}{J}\begin{vmatrix}\alpha_1 & \alpha_2 & \alpha_3\\ A_1 & A_2 & A_3\\ B_1 & B_2 & B_3\end{vmatrix}. \tag{A20}$$

A.5 Gradient, curl, and divergence

The gradient is naturally a covector (operator):

$$\boldsymbol{\nabla}=\boldsymbol{e}^i\partial_i; \tag{A21}$$

for example, $(\boldsymbol{\nabla}f)_i=\partial_i f$.

The contravariant components of the curl are [cf. Eq. (A18b)]

$$(\boldsymbol{\nabla}\times\boldsymbol{A})^i=J^{-1}\epsilon^{ijk}\partial_j A_k \tag{A22a}$$

$$\equiv J^{-1}(A_{k,j}-A_{j,k}), \tag{A22b}$$

which can be represented as the determinant

$$\boldsymbol{\nabla}\times\boldsymbol{A}=\frac{1}{J}\begin{vmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3\\ \partial_1 & \partial_2 & \partial_3\\ A_1 & A_2 & A_3\end{vmatrix}. \tag{A23}$$

The divergence of a vector is

$$\boldsymbol{\nabla}\cdot\boldsymbol{A}=J^{-1}\partial_i(JA^i). \tag{A24}$$

The formula for the divergence of a vector is a special case of the covariant gradient $\boldsymbol{\nabla}$ of an arbitrary tensor field. That is defined in terms of the coefficients of linear connection defined by

$$\boldsymbol{\nabla}_a\boldsymbol{e}_b=\Gamma^c_{ba}\boldsymbol{e}_c. \tag{A25}$$

When the $\boldsymbol{e}_a$'s are a coordinate basis $\boldsymbol{e}_i$, as was assumed above [see Eq. (A2)], then the Γ^k_{ji}'s are called the Christoffel symbols of the second kind, and the covariant derivative in the direction $m\equiv\boldsymbol{e}_m$ is realized by ∂_m. For a general tensor $\mathsf{T}\to T^{i\ldots j}_{k\ldots l}$, one has

$$(\boldsymbol{\nabla}\mathsf{T})^{i\ldots j}_{k\ldots lm}=(\nabla_m\mathsf{T})^{i\ldots j}_{k\ldots l}\equiv T^{i\ldots j}_{k\ldots l;m}, \tag{A26}$$

where

$$T^{i\ldots j\ldots k}_{l\ldots m\ldots n;r}\doteq T^{i\ldots j\ldots k}_{l\ldots m\ldots n,r}+\cdots+\Gamma^j_{sr}T^{i\ldots s\ldots k}_{l\ldots m\ldots n}+\cdots-\cdots-\Gamma^s_{mr}T^{i\ldots j\ldots k}_{l\ldots s\ldots n}. \tag{A27}$$

The divergence of a tensor is then defined as the contraction

$$(\boldsymbol{\nabla}\cdot\mathsf{T})^{i\ldots j\ldots}_{l\ldots m\ldots} = T^{i\ldots j\ldots k}_{l\ldots m\ldots;k}. \tag{A28}$$

As a specific example, consider the contravariant component of the divergence of a second-rank mixed tensor T^i_l. The covariant gradient is

$$(\boldsymbol{\nabla}\mathsf{T})^i_{lr} = T^i_{l,r} + \Gamma^i_{sr}T^s_l - \Gamma^s_{lr}T^i_s \tag{A29}$$

and the divergence is

$$(\boldsymbol{\nabla}\cdot\mathsf{T})_l = T^k_{l,k} + \Gamma^k_{sk}T^s_l - \Gamma^s_{lk}T^k_s. \tag{A30}$$

The first two terms can be recognized as the divergence of the vector $\boldsymbol{\tau}_l \doteq T^k_l$ for fixed l, so the formula reads

$$(\boldsymbol{\nabla}\cdot\mathsf{T})_l = \boldsymbol{\nabla}\cdot\boldsymbol{\tau}_l - (\partial_k\boldsymbol{e}_l)^s T^k_s. \tag{A31}$$

This follows also from the prosaic manipulation

$$\boldsymbol{e}_l\cdot(\boldsymbol{\nabla}\cdot\mathsf{T}) = \boldsymbol{\nabla}\cdot(\boldsymbol{e}_l\cdot\mathsf{T}) - (\boldsymbol{\nabla}\boldsymbol{e}_l):\mathsf{T}, \tag{A32}$$

which is spelled out in more detail in Eq. (49).

A.6 Vector fields

A *vector field* is a generalization to arbitrary differentiable manifolds of the familiar tangent vector. It provides a way of smoothly assigning one vector to each point of the manifold. In differential geometry, instead of writing $\boldsymbol{V} = V^i\boldsymbol{e}_i$ with $\boldsymbol{e}_i \doteq \partial_i\boldsymbol{x}$, one introduces the vector field V as the operator

$$V = V^i(z)\partial_i; \tag{A33}$$

thus a vector field is a first-order differential operator with nonconstant coefficients. The special vector fields ∂_i are called the *coordinate basis* (for vector fields). In this notation, the directional derivative of a function in the direction of $\boldsymbol{V}$ is represented concisely as $Vf \equiv V^i\partial_i f$.

A *covector field* or *differential 1-form* (see Appendix C) is written as

$$\alpha = \alpha_i(z)dz^i. \tag{A34}$$

The dz^i's are a coordinate basis for covectors, dual to the ∂_i's.

The gradient operation leads to a natural covector field. The gradient (field) df of a function f is a covector field whose components are $\partial_i f$, i.e.,

$$df = (\partial_i f)dz^i. \tag{A35}$$

Further discussion of the gradient is given in Sec. C.1.1.

B Magnetic fields

B.1 Representation of magnetic fields

For a field with nested flux surfaces (guaranteed by toroidal axisymmetry), we use the flux coordinates $\boldsymbol{z} = (\psi, \theta, \varphi)$. It is frequently convenient to use the mixed representation

$$\boldsymbol{B} = I(\psi)\boldsymbol{\nabla}\varphi + \boldsymbol{\nabla}\varphi \times \boldsymbol{\nabla}\psi, \tag{B1}$$

or

$$\boldsymbol{B} = I\boldsymbol{e}^{\varphi} + J^{-1}\boldsymbol{e}_{\theta}. \tag{B2}$$

It follows from Eq. (B2) that

$$B^{\theta} = J^{-1}, \quad B_{\varphi} = I. \tag{B3}$$

B.2 Magnetic field lines

The equation of a magnetic field line is

$$\boldsymbol{B} \times d\boldsymbol{x} = \boldsymbol{0} \quad \text{or} \quad \boldsymbol{B} = c\, d\boldsymbol{x}, \tag{B4}$$

where c is a constant. This becomes

$$\frac{B}{dl} = \frac{B^1}{dz^1} = \frac{B^2}{dz^2} = \frac{B^3}{dz^3}. \tag{B5}$$

Here, by definition of contravariant component, $B^i = \boldsymbol{B} \cdot \boldsymbol{\nabla} z^i = \boldsymbol{B} \cdot \boldsymbol{e}^i$. In an axisymmetric torus, one has $B^{\psi} = 0$ (since $\boldsymbol{e}^{\psi} \cdot \boldsymbol{e}^{\varphi} = 0$).

B.3 Flux-surface average

In this section flux surfaces will be denoted by the generalized coordinate ρ (for tokamak applications, frequently the poloidal flux $\psi \equiv \frac{1}{2\pi}\psi_{\text{pol}}$; for stellarator work, frequently the toroidal flux ψ_{tor}). From D'haeseleer et al. (1991, p. 85),

> The flux surface average of a function $\Phi(\boldsymbol{x})$ is defined by the *volume average* over an infinitesimally small shell with volume ΔV, where ΔV lies between two neighboring flux surfaces with volumes V and $V + \Delta V$. It is denoted by $\langle \Phi(\boldsymbol{x}) \rangle$, and is equal to

$$\langle \Phi(\boldsymbol{x}) \rangle = \lim_{\Delta V \to 0} \frac{1}{\Delta V} \int \Phi\, d^3R. \tag{B6}$$

One has

$$\langle \Phi \rangle = \frac{1}{V'(\rho)} \int_S \Phi \frac{dS}{|\boldsymbol{\nabla}\rho|}, \tag{B7}$$

where $V' \doteq dV/d\rho$. Because in a toroidal coordinate system $\boldsymbol{z} = (\rho, \theta, \zeta)$, one has $dS = J\,|\boldsymbol{\nabla}\rho|\, d\theta\, d\zeta$ and

$$\langle \Phi \rangle = \frac{1}{V'} \int_0^{2\pi} d\theta\, d\zeta\, J\Phi. \tag{B8}$$

This implies (using $\Phi = 1$) that

$$V' = \int_0^{2\pi} d\theta \, d\zeta \, J \tag{B9}$$

and that

$$\langle \Phi \rangle = \frac{\int_0^{2\pi} d\theta \, d\zeta \, J\Phi}{\int_0^{2\pi} d\theta \, d\zeta \, J}. \tag{B10}$$

It immediately follows that (for toroidal systems)

$$\langle \boldsymbol{\nabla} \cdot \boldsymbol{A} \rangle = \frac{1}{V'} \frac{\partial}{\partial \rho} (V' \langle A^\rho \rangle) \quad \text{(toroidal)}. \tag{B11}$$

If $\boldsymbol{B}$ has good flux surfaces (guaranteed for an axisymmetric system), then

$$\langle \boldsymbol{B} \cdot \boldsymbol{\nabla} \Phi \rangle = \langle \boldsymbol{\nabla} \cdot (\boldsymbol{B} \Phi \rangle = 0 \quad \text{(good surfaces)}, \tag{B12}$$

since in that situation $B^\rho = 0$.

C Differential forms

One of the tools of modern mathematics that has found increasing applications to physics, including plasma physics, is the method of *differential forms*. That method is not 'new' to physics in any sense; see, for example, the lengthy paper of Misner & Wheeler (1957), now more than a half-century old. (Cartan's research on differential forms dates back to at least 1889.) However, it is becoming increasingly more appreciated and practically useful. The beautifully pedagogical textbook *Gravitation* by Misner et al. (1973) uses differential forms in a fundamental way to provide both mathematical conciseness and a deep physical picture of the geometry of space-time. In plasma physics, 'modern' ($\gtrsim$ 1988[60]) derivations of the nonlinear gyrokinetic equation frequently employ differential forms, as do modern perturbation techniques based on noncanonical coordinates (the latter being employed in the former). Here we give a brief sketch of the essential ideas, providing only enough detail to support the later discussion of noncanonical perturbation theory in Appendix E. For further information, see the readable (yet nontrivial) book by Flanders (1963); the textbook of Misner et al. (1973); Appendix A of the important paper on noncanonical perturbation theory by Cary & Littlejohn (1983); and any textbook on differential geometry, such as the one by Fecko (2006).

C.1 Definition of differential form

Loosely speaking,[61] a *differential form* is the quantity that occurs under the integral sign of a line integral, surface integral, volume integral, *etc.* Thus, in

[60] This year is the publication date of the paper by Hahm (1988).

[61] In this section, we sacrifice some mathematical precision in favor of intuition; the discussion is intended for newcomers to the field. For example, it is really measures that appear

the integral $I_1 \doteq \int f(x)dx$, the quantity $f(x)dx$ is a differential 1-form. In $I_2 \doteq \int\int f(x,y)dx\,dy$, the quantity $f(x,y)dx\,dy$ is a *differential 2-form*. One of the reasons that a systematic formalism for differential p-forms is useful is that it provides elegant and efficient proofs of powerful theorems of integration, such as Stokes' Theorem and Gauss's Law. Electromagnetism is also neatly represented in terms of differential forms.

One of the tricky issues with integration in higher dimensions, say over a 2D surface, relates to *orientation*. That is, in vector calculus line or surface elements are usually considered to be vectors, e.g., $d\boldsymbol{l}$ or $d\boldsymbol{S}$. The systematic theory of differential forms handles that issue by introducing *basis 1-forms* $\boldsymbol{\omega}^\mu \equiv \boldsymbol{d}z^\mu$ (note the boldfaced font for $\boldsymbol{d}$), which endow the basic differentials with some orientation; and the *exterior product* or *wedge product* $\wedge$, which enables one to create higher-order forms from lower-order ones in a useful way. For example, $\boldsymbol{d}x \wedge \boldsymbol{d}y$ is an area element oriented in the z direction.

While the use of a bold font for the basis 1-forms may prevent some confusion with the differentials of ordinary calculus, a consistent notation would make essentially everything in sight bold. It is common to drop the bold, and we shall do so from now on.

C.1.1 1-forms

According to Misner et al. (1973), "a 1-form is a linear, real-valued function of vectors; i.e., a linear machine that takes in a vector and puts out a number." In their notation, that output is denoted by $\langle \alpha, v\rangle$, where α is the 1-form (with components α_μ) and v is the vector (with components v^μ) on which the 1-form acts. Specifically, $\langle \alpha, v\rangle = \alpha_\mu v^\mu$. Another common notation is $\langle \alpha, v\rangle \equiv \alpha(v)$.

A general *differential* 1-form α is just a linear superposition of basis 1-forms: $\alpha = \alpha_\mu(z)dz^\mu$. Although the *components* α_μ of the 1-form α can often[62] be interpreted as the covariant components of a vector, modern notation prefers to call α_μ a *covector*.

An important example of a differential 1-form is the energy-momentum form $p = -E\,dt + p_x dx + p_y dy + p_z dz$. A closely related object, the Poincaré–Cartan differential 1-form, is encountered in noncanonical Hamiltonian mechanics; see Appendix E.

Ordinary functions are 0-forms. The gradient or differential df of a 0-form f is the 1-form

$$df = \partial_\mu f\, dz^\mu. \tag{C1}$$

In ordinary calculus, one would write this as $df = d\boldsymbol{x} \cdot \boldsymbol{\nabla} f$; i.e., df is the directional derivative of f in the $d\boldsymbol{x}$ direction. Specifically, df is the change in f over a short distance $d\boldsymbol{x}$. The point is that the 1-form df contains information about the (linear-order) behavior of f in all possible directions relative to some

under integral signs, and the transformation properties of measures and differential forms differ by the presence or absence of an absolute value. Also, the second paragraph of the section does not adequately distinguish between global orientation and the local properties of basis 1-forms.

[62] (when a metric tensor exists)

particular point. Suppose one wants to know how f changes over an infinitesimal distance Δs in the $\widehat{\boldsymbol{y}}$ direction. To answer that, one constructs the vector (field) $v = \Delta s\, \partial_y$ and runs it through the 1-form df, obtaining

$$\langle df, v\rangle = \langle df, \Delta s\, \partial_y\rangle = \Delta s\, \widehat{\boldsymbol{y}} \cdot \boldsymbol{\nabla} f. \tag{C2}$$

C.1.2 p-forms

A *p-form* is (or has as components) a completely antisymmetric tensor of rank[63] $\binom{0}{p}$. For example,

$$\alpha = \frac{1}{p!}\alpha_{\mu_1\mu_2\ldots\mu_p} dz^{\mu_1} \wedge dz^{\mu_2} \wedge \cdots \wedge dz^{\mu_p}. \tag{C3}$$

Here $\wedge$ denotes the *exterior product*, which is defined as follows. Given an n-dimensional linear space and p-forms $\alpha^{(p)}$ and $\beta^{(q)}$ in that space, then $\alpha^{(p)} \wedge \beta^{(q)}$ is a $(p+q)$-form. The $\wedge$ behaves like an ordinary multiplication sign (it is distributive and associative) except for the commutation law

$$\alpha^{(p)} \wedge \beta^{(q)} = (-1)^{pq} \beta^{(q)} \wedge \alpha^{(p)}. \tag{C4}$$

For two 1-forms $\alpha = \alpha_\mu dz^\mu$ and $\beta = \beta_\nu dz^\nu$, some consequences of Eq. (C4) are

$$\alpha \wedge \beta = -\beta \wedge \alpha, \tag{C5a}$$

$$\alpha \wedge \alpha = 0, \tag{C5b}$$

$$\alpha \wedge \beta = \alpha_\mu \beta_\nu dz^\mu \wedge dz^\nu = \frac{1}{2}(\alpha_\mu\beta_\nu - \beta_\mu\alpha_\nu) dz^\mu \wedge dz^\nu. \tag{C5c}$$

C.2 Interior product

A p-form can be thought of as a multi-linear machine that accepts p vectors and returns a number. The *interior product* of a vector field X with a p-form α is written as $i_X\alpha$; it is a $(p-1)$-form defined by

$$(i_X\alpha)(X_1, \ldots, X_{p-1}) = \alpha(X, X_1, \ldots, X_{p-1}), \tag{C6}$$

i.e., the vector is just inserted into the first slot of the p-form. The components of $i_X\alpha$ are $X^k \alpha_{k\, i_1\, i_2\, \ldots i_{p-1}}$.

C.3 Exterior derivative

The *exterior derivative* d turns a p-form into a $(p+1)$-form by using the definition of df $(p = 0)$ plus the rules

$$d(\alpha + \beta) = d\alpha + d\beta, \tag{C7a}$$

$$d(\alpha^{(p)} \wedge \beta^{(q)}) = d\alpha^{(p)} \wedge \beta^{(q)} + (-1)^p \alpha^{(p)} \wedge d\beta^{(q)}, \tag{C7b}$$

$$d(d\alpha) = 0. \tag{C7c}$$

[63] The upper index refers to the number of contravariant (vector) indices; the lower index refers to the number of covariant (covector) indices.

For example, suppose one has the 1-form $\gamma = \gamma_\mu dz^\mu$. Then $d\gamma$ is a 2-form:

$$d\gamma = d\gamma_\mu \wedge dz^\mu = \frac{\partial \gamma_\mu}{\partial z^\nu} dz^\nu \wedge dz^\mu = \frac{\partial \gamma_\nu}{\partial z^\mu} dz^\mu \wedge dz^\nu = \frac{1}{2}\omega_{\mu\nu} dz^\mu \wedge dz^\nu, \quad \text{(C8)}$$

where the components of the 2-form are

$$\omega_{\mu\nu} \doteq \frac{\partial \gamma_\nu}{\partial z^\mu} - \frac{\partial \gamma_\mu}{\partial z^\nu} \equiv \partial_\mu \gamma_\nu - \partial_\nu \gamma_\mu. \quad \text{(C9)}$$

A differential 2-form plays a central role in the study of noncanonical Hamiltonian mechanics; see Appendix E.

A form α is said to be *closed* if $d\alpha = 0$.

D Pullback transformations

The fundamental connection between the particle and the gyrocenter phase spaces is the pullback transformation. A proper exposition of the concepts underlying the notion of a pullback would require writing the first few chapters of a book on differential geometry, but that has already been done; one enjoyable example aimed at physicists is by Fecko (2006). Thus the following discussion will be brief, incomplete, and informal. For applications to gyrokinetics, an important paper is by Qin & Tang (2004). The basic ideas relating to variable and function transformations outlined below are covered in an old but still useful review by Cary (1981) (who did not use the modern language of differential geometry).

Let us refer to the particle phase space as the manifold $\mathcal{M}$ (coordinates z) and the gyrocenter phase space as the manifold $\overline{\mathcal{M}}$ (coordinates $\overline{z}$). Assume that one is given the transformation $\overline{z} = \varphi(z)$, or in operator notation $\overline{z} = \mathrm{T}z$. φ is a nonlinear map from $\mathcal{M}$ to $\overline{\mathcal{M}}$ (Fig. 3).

D.1 Pullback transformation on functions

On $\mathcal{M}$, consider a scalar function $f(z)$. The value of that scalar is independent of the choice of variables. Thus, one has $f(z) = \overline{f}(\overline{z})$ for some function $\overline{f}$ of the gyrocenter variables. This implies

$$f(z) = \overline{f}(\varphi(z)) \doteq (\varphi^*(\overline{f}))(z), \quad \text{(D1)}$$

which defines the pullback map φ^* on functions. The corresponding operator representation is cleaner:

$$f(z) = \overline{f}(\mathrm{T}z) \doteq (\mathrm{T}^*\overline{f})(z), \quad \text{(D2)}$$

or, since both sides are evaluated at the same arbitrary point,

$$f = \mathrm{T}^*\overline{f}. \quad \text{(D3)}$$

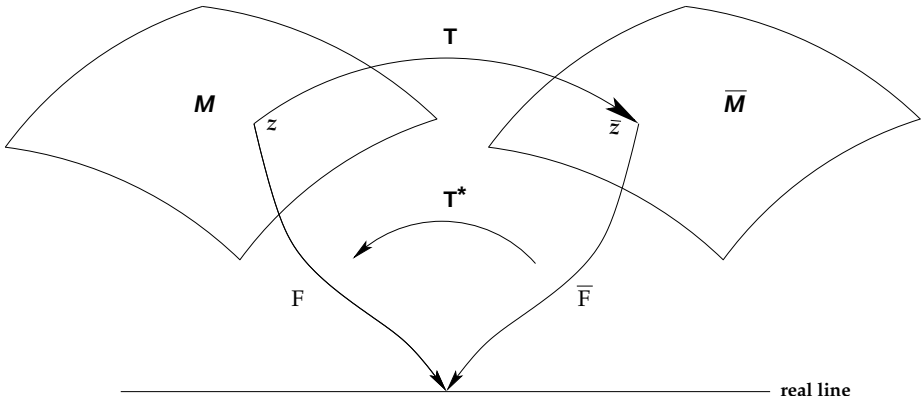

Figure 3: Illustration of the pullback T^* of the gyrokinetic distribution function $\overline{F}$ from gyrocenter phase space $\overline{\mathcal{M}}$ to the particle phase space $\mathcal{M}$. The transformation T takes points z in $\mathcal{M}$ to points $\overline{z}$ in $\overline{\mathcal{M}}$: $\overline{z} = \mathrm{T}z$. The scalar function $F(z)$ on $\mathcal{M}$ has the same value as the function $\overline{F}(\overline{z})$ evaluated on the transformed point. The functions are related by the pullback transformation: $F = \mathrm{T}^*\overline{F}$.

T^* (which acts here on functions) is called the pullback transformation induced by the transformation T (which acts on points in $\mathcal{M}$). More generally, T^* acts on p-forms. (A function is a 0-form.) Typically Lie transforms (Sec. D.3.2) are used to construct T perturbatively. To go in the other direction, $\overline{f} = \overline{\mathrm{T}}^* f$, one can use the inverse function: $\overline{\mathrm{T}}^* = (\mathrm{T}^*)^{-1} = (\mathrm{T}^{-1})^*$.

D.2 Pushforward of vectors

Note that although T transforms z to $\overline{z}$, it is not called a 'pushforward' operator; rather, it is said to induce the *pullback* transformation T^*. (One must take care, as some early plasma literature is not consistent on this point.) The *pushforward* operator T_* and the corresponding map φ_* transform *vectors* on the tangent space of $\mathcal{M}$ to their corresponding representation in the tangent space of $\overline{\mathcal{M}}$.[64] Vectors are associated with infinitesimal displacements, so consider

$$\overline{z} + \Delta\overline{z} = \varphi(z + \Delta z) \approx \varphi(z) + \Delta z^j \partial_j \varphi(z), \tag{D4}$$

[64]Note that with the definitions adopted here, $\mathrm{T}_* \neq (\mathrm{T}^*)^{-1}$; indeed, the operators T_* and T^* act on different objects (vectors and forms, respectively). However, some people will loosely call $(\mathrm{T}^*)^{-1}$ a 'pushforward' (Brizard & Hahm, 2007), and generally at the practical level one gets into no trouble by doing so. Fecko (2006, p. 58, Ex. 3.1.6) discusses the subtleties and the possibility of *defining* the pushforward to be $(\mathrm{T}^*)^{-1}$.

One way of seeing the relationship of $(\mathrm{T}_*)^{-1}$ to the pushforward of a vector is to recognize that the components of a vector, $V^i(z)$, are functions (0-forms), which therefore transform (going forward) with $(\mathrm{T}^*)^{-1}$. For example, in a 1D Cartesian coordinate system consider the simple guiding-center transformation $\overline{x}\widehat{\boldsymbol{x}} = x\widehat{\boldsymbol{x}} - \rho\widehat{\boldsymbol{x}}$, where ρ is a constant. The numerical value of $\overline{x}$ can be realized in two ways: $\overline{x} = f(x) = \overline{f}(\overline{x})$, where $f(x) = x - \rho$ and $\overline{f}(\overline{x}) = \overline{x}$. Upon referring to Fig. 3, one sees that $\overline{f} = (T^*)^{-1} f$, or $\overline{x}\widehat{\boldsymbol{x}} = (T^*)^{-1}(x - \rho)\widehat{\boldsymbol{x}}$.

or

$$\Delta\overline{z} = \Delta z^j \partial_j \varphi(z) \doteq \varphi_*(\Delta z) \equiv \mathrm{T}_* \Delta z. \tag{D5}$$

One can replace $\Delta \boldsymbol{z}$ by an arbitrary vector $\boldsymbol{V}$. It can be readily seen that the coordinate representation of φ_* is the Jacobian matrix $J^i_j = \partial \overline{z}^i / \partial z^j$:

$$\overline{V}^i = J^i_j V^j. \tag{D6}$$

In the notation of vector *fields*, in which vectors $\boldsymbol{V} = V^i(z)\boldsymbol{e}_i$ are written as $V = V^i(z)\partial_i$, the corresponding statement is

$$\varphi_*(\partial_a) \equiv \mathrm{T}_* \partial_\alpha = J^b_a \overline{\partial}_b. \tag{D7}$$

D.3 Lie derivatives

The notion of *flow* is important to gyrokinetics in two ways: in the Hamiltonian time evolution of the particle phase-space coordinates; and, more abstractly, in the development of a perturbative transformation in ϵ. For both of these applications, the notion of the *Lie derivative* is fundamental. Let us define it first in the more familiar context of time evolution. The equations of motion $dz^i/dt = V^i(z)$ define a diffeomorphism $\Phi_t : \mathcal{M} \to \mathcal{M}$ that gives the time-evolved variables in terms of their initial values. The V^i's define a vector field $V = V^i \partial_i$ that gives the direction of the flow. Tensor fields A are also evolved by the transformation. The *Lie derivative* of a tensor field A in the direction V is defined by

$$L_V A \doteq \frac{d\Phi_t^* A}{dt}\Big|_{t=0}. \tag{D8}$$

Fecko (2006, p. 71) describes this as follows:

> This derivative "palpates" the changes of tensor fields induced by a tiny Lie transport along V: first, the value of the field A at the "slightly drained away" point $\Phi_\epsilon(x)$ is transported back into x and then it is compared with the initial value of A in x. The comparison := their difference, ... divided by the increment of the parameter ϵ, resulting in a quantity measuring just the "change of the tensor field per unit value of the parameter t" (or the "rate of change of the field" A along V).

Note that it is crucial that this operation is defined in terms of the pullback; that allows one to make a comparison at a single point in time.

It is an important theorem that

$$[L_{V_1}, L_{V_2}] = L_{[V_1, V_2]}, \tag{D9}$$

where brackets denote commutator. In principle, the left-hand side would be a second-order differential operator; however, according to the theorem the

second-order terms cancel and the result is of first order.[65]

D.3.1 Lie derivative of functions and of 1-forms

It is easy to establish that the Lie derivative of a function f (a 0-form) is

$$L_V f = Vf \equiv V^i \partial_i f \equiv \langle df, V \rangle; \tag{D11}$$

i.e., it is just the directional derivative.

The action of the Lie derivative on a 1-form α will be important for the noncanonical perturbation theory described in the next section. One can find the result by working directly from the definition:

$$L_V \alpha = \lim_{\Delta t \to 0} \frac{\Phi_{\Delta t}^* \alpha - \alpha}{\Delta t}. \tag{D12}$$

One can determine the result of the pullback of the 1-form by representing α in terms of coordinates; coordinates are functions, and the pullback of a function is easy (see Appendix D.1). Thus

$$\begin{aligned} \Phi_{\Delta t}^* \alpha &= \Phi_{\Delta t}^* (\alpha_i dz^i) && \text{(D13a)} \\ &\approx \alpha_i (z + V\Delta t) d[z + V(z)\Delta t]^i && \text{(D13b)} \\ &\approx \alpha_i(z) dz^i + \Delta t (V^j \partial_j \alpha_i) dz^i + \alpha_i \Delta t (\partial_j V^i) dz^j. && \text{(D13c)} \end{aligned}$$

After the interchange of i and j in the last term, one finds the desired result,

$$L_V \alpha = (V^j \partial_j \alpha_i + \alpha_j \partial_i V^j) dz^i. \tag{D14}$$

By antisymmetrizing the $\partial_j \alpha_i$ in the first term, one obtains

$$L_V \alpha = [V^j (\partial_j \alpha_i - \partial_i \alpha_j) + \underbrace{V^j \partial_i \alpha_j + (\partial_i V^j) \alpha_j}_{\partial_i (V^j \alpha_j)}] dz^i. \tag{D15}$$

In terms of the interior product defined in Sec. C.2, one can write Eq. (D15) as

$$L_V \alpha = i_V d\alpha + d(i_V \alpha). \tag{D16}$$

This is a special case of *Cartan's formula* (sometimes called *Cartan's magic formula*). In fact, it holds for all p-forms; it relates the Lie derivative L, interior product i, and exterior derivative d.

[65] This can be checked readily when the Lie derivatives act on functions, since the Lie derivative is then merely the directional derivative [see Eq. (D11)]:

$$\begin{aligned} [L_V, L_W] f &= V^j \partial_j (W^i \partial_i f) - W^j \partial_j (V^i \partial_i f) && \text{(D10a)} \\ &= V^j (\partial_j W^i) \partial_i f - W^j (\partial_j V^i) \partial_i f && \text{(D10b)} \\ &= (V^j \partial_j W^i \partial_i f - W^j \partial_j V^i \partial_i f && \text{(D10c)} \\ &= L_{[V,W]} f. && \text{(D10d)} \end{aligned}$$

It is clear from the definition that tL_V is the first-order part of Φ_t^*: $\Phi_t^* = 1 + tL_V + \cdots$. But one can say more. It follows from $d\Phi_t^*/dt = d\Phi_{t+s}^*/ds_{|s=0}$ that

$$\frac{d\Phi_t^*}{dt} = \Phi_t^* L_V, \tag{D17}$$

from which one can establish that[66]

$$\Phi_t^* = \exp(t\, L_V). \tag{D18}$$

D.3.2 Lie transforms

So far we have discussed flows that evolve in time. Consider instead fixed time, but a flow that develops in a perturbation parameter ϵ. Given a vector generating field W (independent of ϵ), one can therefore define a flow $\Phi_\epsilon^* = \exp(\epsilon L_W) = \exp(L_{\epsilon W})$. This effects a particular coordinate transformation $\overline{z} = \mathrm{T}_{\epsilon W} z$, called a *Lie transform*, and induces the corresponding pullback $\mathrm{T}_{\epsilon W}^*$. Lie transforms with more degrees of freedom can be obtained by iterating elementary transforms. In the method of Dragt & Finn (1976), one writes

$$\mathrm{T}_\epsilon = \ldots e^{\epsilon^3 \mathcal{L}_{W_3}} e^{\epsilon^2 \mathcal{L}_{W_2}} e^{\epsilon \mathcal{L}_{W_1}} \equiv \ldots e^{\mathcal{L}_3} e^{\mathcal{L}_2} e^{\mathcal{L}_1} \tag{D19a}$$

$$= 1 + \mathcal{L}_1 + \mathcal{L}_2 + \frac{1}{2}\mathcal{L}_1^2 + \cdots . \tag{D19b}$$

Such near-identity transformations will be used in Sec. E.2 to implement noncanonical perturbation theory. Given T_ϵ, the corresponding pullback transformation T_ϵ^* is obtained from T_ϵ by replacing $\mathcal{L}_n$ by L_n.

E Symplectic structure and noncanonical coordinates

The fact that Hamiltonian dynamics is symplectic frees one from the necessity of employing canonical coordinates. Great progress has been made in gyrokinetics by using noncanonical variables, a technique first advocated by Littlejohn (1982).

From Arnold (1978):

> A symplectic structure on a manifold is a closed nondegenerate differential 2-form. The phase space of a mechanical system has a natural symplectic structure.
>
> On a symplectic manifold, as on a riemannian manifold, there is a natural isomorphism between vector fields and 1-forms. A vector

[66] A common example is the Taylor expansion of $\psi(x+t) = \psi(x) + t\psi'(x) + \frac{1}{2}t^2\psi''(x) + \cdots = \exp(t\,\partial_x)\psi$. With $V = \partial_x$ (i.e., $V^x = 1$), this can be written as $\psi(x+t) = \Phi_t^*\psi$ with the pullback given by Eq. (D18).

field on a symplectic manifold corresponding to the differential of a function is called a hamiltonian vector field. A vector field on a manifold determines a phase flow, i.e., a one-parameter group of diffeomorphisms. The phase flow of a hamiltonian vector field on a symplectic manifold preserves the symplectic structure of phase space.

The vector fields on a manifold form a Lie algebra. The hamiltonian vector fields on a symplectic manifold also form a Lie algebra. The operation in this algebra is called the Poisson bracket.

From Gotay & Isenberg (1992):

What is the origin of the unusual name "symplectic"? It is derived from the Greek $\sigma\nu\mu\pi\lambda\epsilon\kappa\tau\iota\kappa\acute{\upsilon}s$, which is the antecedent of the Latin "complex." Its mathematical usage is due to Hermann Weyl who, in an effort to avoid a certain semantic confusion, renamed the then obscure "line complex group" the "symplectic group." But whatever its etymology, the adjective "symplectic" means "plaited together" or "woven." This is wonderfully apt, for it is this intertwining—already evident in the above expression for the form $[\Omega = \sum_{i=1}^{n} dq_i \wedge dp_i]$—that most characterizes, and is in fact the essence of, both symplectic geometry and Hamiltonian mechanics. And it is the intricate plaiting together of mathematics and physics which gives symplectic geometry its power and its promise.

E.1 Noncanonical Hamiltonian mechanics

In Appendix C we pointed out that Lagrange's variational principle can be written as $\delta \int \gamma = 0$, where the 'fundamental' or Poincaré–Cartan differential 1-form is $\gamma \doteq \boldsymbol{p} \cdot d\boldsymbol{q} - H\, dt$. The $\boldsymbol{p} \cdot d\boldsymbol{q}$ term is called the *symplectic part*. The exterior derivative of γ defines the fundamental 2-form $\widehat{\omega} = d\gamma$. This is trivially closed: $d\widehat{\omega} = dd\gamma = 0$. However, it is degenerate: $\det[\widehat{\omega}] = 0$ since $\widehat{\omega}$ is antisymmetric and the dimension of the extended phase space is odd.

Although in principle one can transform γ to any set of new variables one pleases (including transformations of the time t), an important subset of transformations consists of those that leave t invariant. Then the phase-space components of $\widehat{\omega}$ define an even-dimensional submatrix that is called the *Lagrange tensor* ω. This defines a closed, nondegenerate, differential 2-form that determines the natural symplectic structure of Hamiltonian mechanics. The Euler–Lagrange equations are

$$\omega_{ij} \frac{dz^j}{dt} = \frac{\partial H}{\partial z^i} + \frac{\partial \gamma_i}{\partial t}. \tag{E1}$$

It is conventional to introduce the *Poisson tensor* J (not to be confused with a Jacobian) as the (fully contravariant) inverse of ω. Thus the equations of motion are explicitly

$$\frac{dz^i}{dt} = J^{ij} \left(\frac{\partial H}{\partial z^j} + \frac{\partial \gamma_j}{\partial t} \right). \tag{E2}$$

In a canonical coordinate system, $\partial\gamma_i/\partial t = 0$ and one recovers the usual form of Hamiltonian's equations

$$\frac{dz}{dt} = \mathsf{J} \cdot \frac{\partial H}{\partial z}, \tag{E3}$$

with the canonical representation of the Poisson tensor being

$$\mathsf{J} = \begin{pmatrix} 0 & \mathsf{I} \\ -\mathsf{I} & 0 \end{pmatrix}. \tag{E4}$$

The form of Eq. (E3) is also recovered for noncanonical representations in which the transformation is explicitly time-independent.

Equation (E3) can be written as

$$\frac{dz^i}{dt} = \{z^i, H\}, \tag{E5}$$

where the Poisson bracket of A and B is

$$\{A, B\} \doteq (\partial_i A) J^{ij} (\partial_j B). \tag{E6}$$

E.2 Noncanonical Lie perturbation theory

The strategy of noncanonical Lie perturbation theory is to construct a variable transformation T and the induced pullback transformation[67] T^* (Appendix D) that transforms γ to a new 1-form $\overline{\gamma}$ chosen such that all of its components are independent of (transformed) gyrophase $\overline{\zeta}$. If $\overline{\zeta}$ is chosen as one of the gyrocenter coordinates, it is then a theorem[68] (Cary & Littlejohn, 1983) that the coefficient $\overline{\mu}$ of $d\overline{\zeta}$ is conserved: $\overline{\gamma} = \cdots + \overline{\mu}\, d\overline{\zeta}$. $\overline{\mu}$ is known as the first adiabatic invariant. The transformation law (for a differential 1-form γ) is

$$\overline{\gamma} = (\mathrm{T}^*)^{-1}\gamma + d\widehat{S}, \tag{E7}$$

[67] For gyrokinetics, Brizard has found it useful to construct the pullback (more precisely, the pushforward) transformation in two steps (Brizard, 1990; Brizard & Hahm, 2007): he writes $(\mathrm{T}^*)^{-1} = (\mathrm{T}^*_{\mathrm{gy}})^{-1}(\mathrm{T}^*_{\mathrm{gc}})^{-1}$. Here gc stands for guiding center and refers to the physics of guiding-center motion in the presence of inhomogeneous magnetic field but in the absence of fluctuation potential; gy stands for gyrokinetic and refers to the further physics induced by the presence of the potential. The guiding-center transformation involves an asymptotic expansion in the inhomogeneity parameter $\epsilon_B = \rho/L_B$, while the gyrokinetic transformation involves expansion in the size ϵ_δ of the fluctuations. In a maximal ordering, $\epsilon_B \sim \epsilon_\delta \sim \epsilon$.

Use of this two-step procedure is technically convenient, although in practice it has been typically carried out only to $O(\epsilon_B, \epsilon_\delta^2)$. While this might be adequate in some cases, Parra & Calvo (2011) were interested in the complete second-order Hamiltonian through $O(\epsilon^2)$ in the maximal ordering, in which case one needs terms through $O(\epsilon_\delta^2, \epsilon_\delta\epsilon_B, \epsilon_B^2)$. They argued for a one-step procedure, especially to be sure that the cross terms of $O(\epsilon_\delta\epsilon_B)$ were not overlooked.

A question is, Is there any inherent advantage to a one-step vs a two-step approach? In principle, the answer is no; given any two successively applied transformations T_1 and T_2, it is always possible to define the one-step transformation $\mathrm{T} = \mathrm{T}_2\mathrm{T}_1$. Trouble can arise if one of the transformations is not worked out to the necessary order, or (closely related) if gyrophase information is neglected prematurely in a truncated expansion; however, that need not happen if one is careful. For the perspective of Parra & Calvo, see the excerpt from their paper in Sec. L.13.

[68] This is a special case of Noether's theorem; see Appendix H.

where $\widehat{S}$ is a gauge scalar (whose presence can be shown to not affect the equations of motion). The idea is to express the coordinate transformation T ($\overline{z} = \mathrm{T}z$) perturbatively in terms of unknown Lie generators (vector fields) $w_i(z)$ (i denoting the order in ϵ), then to choose those (and $\widehat{S}$) to enforce the $\overline{\zeta}$ independence of $\overline{\gamma}$. An important technical simplification, first used in this context by Littlejohn (1982), can be made by invoking Cartan's formula (Sec. D.3)

$$L_X\gamma = i_X d\gamma + d(i_X\gamma). \tag{E8}$$

This is useful because we saw in Sec. D.3.2 that T^* can be written as a series involving Lie derivatives: $(\mathrm{T}^*)^{-1} = 1 - L_{w_1} + \cdots$. At first order one has $L_{w_1}\gamma = i_{w_1}\gamma + d(i_{w_1}\gamma)$. Because the last term on the right-hand side of Eq. (E8) is a pure derivative, it can be combined with $d\widehat{S}$ to give a new quantity dS (which still does not affect the equations of motion). Thus as far as the dynamics are concerned one can interpret $L_{w_1}\gamma$ as $i_{w_1}d\gamma = w_1^k\omega_{ki}dz^i$. This can be applied recursively to products such as $L_{w_2}L_{w_1}$ using the fact that the Lie derivative on a pure differential vanishes. It turns out that there is sufficient freedom in the choice of S and the w_i's to write $\overline{\gamma}$, through arbitrary order, as

$$\overline{\gamma} = \left(\frac{q}{c}\boldsymbol{A}(\overline{\boldsymbol{X}}) + m\overline{U}\,\widehat{\boldsymbol{b}}\right)\cdot d\overline{\boldsymbol{X}} + (\overline{\mu}\,d\overline{\zeta} - \boldsymbol{K}\cdot d\overline{\boldsymbol{X}}) - \overline{H}\,dt \tag{E9}$$

(i.e., two perpendicular components of kinetic momentum have been eliminated). Here $\boldsymbol{K} \doteq (\boldsymbol{\nabla}\boldsymbol{e}_1)\cdot\boldsymbol{e}_2$ is the gyrogauge vector[69] (Littlejohn, 1984). There is further freedom in how information is partitioned between the symplectic part and the Hamiltonian part. In a purely *symplectic representation*, one preserves the form of the symplectic part of the 1-form to have its guiding-center form; one then has a nontrivial representation $\overline{U} = \mathrm{T}U$. In a purely *Hamiltonian representation*, one preserves $\overline{U} = U$ and absorbs higher-order corrections into $\overline{H}$. Mixed representations are also possible. The calculation performed by Parra & Calvo (2011) of a second-order Hamiltonian used a symplectic representation at first order but a Hamiltonian one at second order. The physical content of all such representations is invariant, although expressed in terms of different Poisson brackets and transformations. This is discussed in a recent paper by Brizard & Tronko (2012).

The virtue of preserving the symplectic structure is that the Poisson brackets retain their guiding-center form [a particular specialization of Eq. (E6)]. They were derived by Littlejohn (1983) and are recorded in the review article of

[69] Littlejohn explains why the gyrogauge vector is necessary in order to guarantee that the equations of motion do not depend on the arbitrary choice of the perpendicular unit vectors. He also shows that $\boldsymbol{K}$ can be expressed in terms of differential operations on $\boldsymbol{B}$.

Brizard & Hahm (2007, p. 448)[70]:

$$\{\mathcal{F},\mathcal{G}\} = \frac{\partial\mathcal{F}}{\partial\zeta}\frac{\partial\mathcal{G}}{\partial\mu} - \frac{\partial\mathcal{F}}{\partial\mu}\frac{\partial\mathcal{G}}{\partial\zeta} + \frac{\boldsymbol{B}_*}{B_*}\cdot\left(\boldsymbol{\nabla}\mathcal{F}\frac{\partial\mathcal{G}}{\partial p_\parallel} - \frac{\partial\mathcal{F}}{\partial p_\parallel}\boldsymbol{\nabla}\mathcal{G}\right) - \frac{c\widehat{\boldsymbol{b}}}{eB_*}\cdot\boldsymbol{\nabla}\mathcal{F}\times\boldsymbol{\nabla}\mathcal{G} + \left(\frac{\partial\mathcal{F}}{\partial w}\frac{\partial\mathcal{G}}{\partial t} - \frac{\partial\mathcal{F}}{\partial t}\frac{\partial\mathcal{G}}{\partial w}\right). \tag{E10}$$

F Ponderomotive nonlinearities[71]

The equations of motion for a gyrocenter follow from the guiding-center Poisson bracket of the gyrocenter position and the gyrocenter Hamiltonian $\overline{H}$: $\dot{\overline{\boldsymbol{x}}} = \{\overline{\boldsymbol{x}},\overline{H}\}$. $\overline{H}$ also determines the polarization in Poisson's equation as well as the Reynolds stresses in the gyrokinetic momentum conservation law. In the present section we discuss $\overline{H}$ in its simplest nonlinear manifestation: its second-order form in the absence of both magnetic-field gradients and FLR effects. In that case, one finds[72]

$$\overline{H}_2 = -\frac{1}{2}m_i u_E^2. \tag{F1}$$

In addition to elucidating some subtleties about the distinction between the lowest-order and the true magnetic moment, and between the unbarred and barred cooordinate systems, the calculation that we describe here serves as an explicit example of the noncanonical 1-form perturbation theory described in the previous section.

The interpretation of $\overline{H}_2$ is somewhat subtle and nontrivial. Frequently the program for obtaining $\overline{H}$ is described as a systematic, order-by-order elimination of gyrophase ζ, which inevitably conjures up the idea of an average over a rapidly rotating gyroradius vector $\boldsymbol{\rho}$. On the other hand, one usually associates the cold-ion limit $T_i \to 0$ with a zero-gyroradius limit $\boldsymbol{\rho} \to \boldsymbol{0}$. In that limit, one could inquire whether there is anything left to rotate.

One of the difficulties with such discussion is an imprecision about exactly to which coordinates one is referring. Let us define the 'lowest-order' gyrocenter $\overline{\boldsymbol{X}}$ by $\overline{\boldsymbol{X}} \approx \boldsymbol{x} - \boldsymbol{\rho}$, where $\boldsymbol{\rho} \doteq \widehat{\boldsymbol{b}}\times\boldsymbol{v}_\perp/\omega_c$; gyrophase is defined in terms of $\boldsymbol{v}_\perp$ (or, equivalently, $\boldsymbol{\rho}$). This definition is natural, and the approximation becomes exact for purely circular motion. However, in the presence of an electrostatic potential ϕ, $\overline{\boldsymbol{X}}$ of course moves with (at least) the $\boldsymbol{E}\times\boldsymbol{B}$ velocity; the particle motion is not precisely circular. Relative to the instantaneous center of gyration, defined in some systematic way, one can introduce a 'true' gyrophase $\overline{\zeta} \neq \zeta$ and

[70] The terms involving $\partial/\partial\mu$ have been modified to conform with our definition $\mu \doteq \frac{1}{2}mv_\perp^2/\omega_c$.

[71] Some of the material of this section has been distilled into the Brief Communication by Krommes (2013).

[72] This result is well known (Lee, 1983; Dubin et al., 1983; Mishchenko & Brizard, 2011); see, for example, discussion by Scott & Smirnov (2010). Some of the underlying physics, as discussed in this section, is perhaps not so well known.

a true magnetic moment $\overline{\mu} \neq \mu$. One must be careful about the meaning of the 'zero-gyroradius limit'; at fixed ω_c, does one mean $\rho(\mu, \zeta) \to 0$ or $\overline{\rho}(\overline{\mu}, \overline{\zeta}) \to 0$?

The definitions of the new, barred variables are not unique in the absence of a further constraint (which amounts to giving a precise meaning to 'instantaneous center of gyration'). That constraint is the adiabatic conservation of the true magnetic moment $\overline{\mu}$. To guess the form of $\overline{\mu}$, one can invoke the idea of Galilean invariance. The lowest-order quantity $\mu \doteq \frac{1}{2} m v_\perp^2 / \omega_c$ is not Galilean-invariant; it changes its value under a shift in velocity. Galilean invariance is restored if $\boldsymbol{v}_\perp$ is referred to a reference velocity, which is naturally chosen to be $\boldsymbol{u_E}$. Thus, one guesses that

$$\overline{\mu} \approx \frac{1}{2} m |\boldsymbol{v}_\perp - \boldsymbol{u_E}(\boldsymbol{x})|^2 / \omega_c. \tag{F2}$$

Indeed, it is easy to demonstrate from the equation of motion

$$\frac{d\boldsymbol{v}}{dt} = \frac{q}{m}(\boldsymbol{E} + c^{-1}\boldsymbol{v} \times \boldsymbol{B}) \tag{F3}$$

that $\overline{\mu}$ is exactly conserved when $\boldsymbol{u_E}$ is constant. Putting that another way, if one makes the transformation $\boldsymbol{v}_\perp = \boldsymbol{u_E} + \delta\boldsymbol{v}_\perp$, one finds for constant $\boldsymbol{u_E}$ that the equation of motion reduces to

$$\frac{d\delta\boldsymbol{v}_\perp}{dt} - \omega_c \delta\boldsymbol{v}_\perp \times \widehat{\boldsymbol{b}} = \boldsymbol{0}, \tag{F4}$$

the solution of which is purely circular motion. That is, the motion is purely circular in a frame moving with $\boldsymbol{u_E}$.

To discuss the cold-ion limit, one usually invokes the idea of an equilibrium Maxwellian distribution. An equilibrium solution of the gyrokinetic equation should be a function of the constants of motion. Indeed, a perpendicular ion Maxwellian shifted by $\boldsymbol{u_E}$ is proportional to

$$\exp\left(-\frac{1}{2}|\boldsymbol{v}_\perp - \boldsymbol{u_E}|^2 / v_{ti}^2\right) = e^{-\overline{\mu}\omega_c / T_i}. \tag{F5}$$

Thus the limit $T_i \to 0$ constrains $\overline{\mu}$ to vanish [the PDF is proportional to $\delta(\overline{\mu})$]. With gyroradius defined by $\rho \doteq v_\perp/\omega_c = (2\mu/m\omega_c)^{1/2}$ and similarly $\overline{\rho} \doteq (2\overline{\mu}/m\omega_c)^{1/2}$, one sees that $\overline{\rho} \to 0$ in the cold-ion limit. ρ, however, does not vanish, since $\overline{\mu} = 0$ constrains $\boldsymbol{v}_\perp$ to be equal to $\boldsymbol{u_E}$, or $\rho = u_E/\omega_c$. This length is not associated with circular motion.

Thus it is crucial to understand the distinction between the barred and unbarred coordinates, and the definitions of each. Although at first order it is not hard to use geometric reasoning to obtain the barred coordinates, the most systematic procedure is to employ the symplectic methodology described in the last section. The constant-$\boldsymbol{B}$, zero-FLR limit provides the simplest example of that procedure.

We use Lie transforms (Sec. D.3.2) to represent the transformation T between the unbarred and barred coordinates: $\overline{z} = \mathrm{T}z$, where (Dragt & Finn,

1976)

$$\mathrm{T} = \ldots e^{\mathcal{L}_3} e^{\mathcal{L}_2} e^{\mathcal{L}_1} = 1 + \mathcal{L}_1 + \left(\mathcal{L}_2 + \frac{1}{2}\mathcal{L}_1^2\right) + \left(\mathcal{L}_3 + \mathcal{L}_2\mathcal{L}_1 + \frac{1}{6}\mathcal{L}_1^3\right) + \cdots . \quad \text{(F6)}$$

Here $\mathcal{L}_n \equiv \epsilon^n w_n$, where w_n is the nth-order vector field: $w_n = w_n^\nu \partial_\nu$. Therefore

$$\overline{z}^\nu = z^\nu + w_1^\nu + w_2^\nu + \frac{1}{2}\mathcal{L}_1 w_1^\nu + \cdots . \quad \text{(F7)}$$

We do *not* make a preparatory transformation to a gyrocenter position variable; we begin with $\boldsymbol{z} = \{\boldsymbol{x}, U, \mu, \zeta\}$, where $U \doteq \boldsymbol{v} \cdot \widehat{\boldsymbol{b}}$ and $\zeta = \tan^{-1}(-v_y/v_x)$, and let the transformation determine the proper gyrocenter position.

F.1 Heuristic discussion of the first-order generating functions

For constant $\boldsymbol{B}$, it is easy to guess the first-order generating functions w_1^ν. In that case, parallel and perpendicular dynamics cleanly decouple, and at first order a long-wavelength spatial dependence of $\boldsymbol{E}$ should just enter parametrically. Thus one anticipates $w_1^U = 0$ and $\boldsymbol{w}_1^{\boldsymbol{x}} \cdot \widehat{\boldsymbol{b}} = 0$. Also, on the basis of the previous observation that for constant $\boldsymbol{u}_{\boldsymbol{E}}$ the motion is purely circular in a frame moving with $\boldsymbol{u}_{\boldsymbol{E}}$, one expects that

$$\boldsymbol{w}_{1\perp}^{\boldsymbol{x}} = -\boldsymbol{\rho} \quad \text{(F8)}$$

so that $\overline{\boldsymbol{x}} = \boldsymbol{x} - \boldsymbol{\rho} + \cdots$. From Eq. (F2), the first-order term in the expected $\overline{\mu}$ is proportional to

$$-\omega_c^{-1} \boldsymbol{v}_\perp \cdot \boldsymbol{u}_{\boldsymbol{E}} \propto -\omega_c^{-1} v_\perp \widehat{\boldsymbol{c}} \cdot \boldsymbol{E} \times \widehat{\boldsymbol{b}} = -\rho \widehat{\boldsymbol{a}} \cdot \boldsymbol{E} = \boldsymbol{\rho} \cdot \boldsymbol{\nabla}\phi; \quad \text{(F9)}$$

thus we expect to find

$$w_1^\mu \propto \boldsymbol{\rho} \cdot \boldsymbol{\nabla}\phi. \quad \text{(F10)}$$

Finally, to obtain w_1^ζ we observe that a particle *with fixed gyrocenter* in circular motion will feel a periodically varying force that is maximum when $\boldsymbol{v}_\perp \propto \widehat{\boldsymbol{c}}$ is aligned with $\boldsymbol{E}$. That force, of course, is responsible for translation of the gyrocenter at a speed u_E and also for periodic variations of $\dot{\zeta}$ with respect to the constant $\dot{\overline{\zeta}} = \omega_c$. Thus the effect should be proportional to $\widehat{\boldsymbol{c}} \cdot \boldsymbol{\nabla}\widehat{\phi}$, where $\widehat{\phi} \doteq q\phi/m$ is a convenient variable used in the later calculations, i.e., $w_1^\zeta = \beta(v_\perp)\widehat{\boldsymbol{c}} \cdot \boldsymbol{\nabla}\widehat{\phi}$, where $\beta(v_\perp)$ is a dimensional constant to be determined. Because $\overline{\zeta}$ is dimensionless, it is easy to see that the dimensions of β are $[\beta] = [t][v]$. The natural time scale is ω_c^{-1}, and if one argues that the size of the effect should depend on the gyroradius $\rho \propto v_\perp$, one concludes that $\beta \propto (\omega_c v_\perp)^{-1}$. In fact, the proportionality constant turns out to be unity; we will find

$$w_1^\zeta = (\omega_c v_\perp)^{-1} \widehat{\boldsymbol{c}} \cdot \boldsymbol{\nabla}\widehat{\phi}. \quad \text{(F11)}$$

F.2 Systematic derivation of the generating functions

We now rederive these and also (most of) the second-order results by proceeding systematically from the particle 1-form

$$\gamma = \underbrace{\frac{q}{mc}\boldsymbol{A}\cdot d\boldsymbol{x}}_{\gamma^{(-1)}} + \underbrace{(U\widehat{\boldsymbol{b}} + v_\perp \widehat{\boldsymbol{c}})\cdot d\boldsymbol{x} - \left(\frac{1}{2}U^2 + \widehat{\mu}\widehat{\omega}_c + \widehat{\phi}\right) dt}_{\gamma^{(0)}}, \tag{F12}$$

where mass m has been divided out for convenience, $\widehat{\phi} \doteq q\phi(\boldsymbol{x})/m$, $\widehat{\mu} \doteq \frac{1}{2}v_\perp^2$, and $\widehat{\omega}_c = 1$. The reason for introducing $\widehat{\mu}$ is so that we can deal with an order-unity quantity. (In the presence of a non-constant $\boldsymbol{B}$, we would define $\widehat{\mu} \doteq \mu\,\omega_{c0}$, where ω_{c0} is a constant reference gyrofrequency. Then $\widehat{\omega}_c = \omega_c/\omega_{c0}$.) Subsequently we will just write μ instead of $\widehat{\mu}$ for convenience. We order the vector potential $\boldsymbol{A}$ large, $O(\epsilon^{-1})$. We consider a time-independent electrostatic potential ϕ, which we consider to be $O(1)$ but to have only long-wavelength variation; then the derived $\boldsymbol{E}\times\boldsymbol{B}$ velocity will be $O(\epsilon)$. The fundamental equation is

$$\overline{\gamma} = (\mathrm{T}^*)^{-1}\gamma + dS, \tag{F13}$$

where[73]

$$(\mathrm{T}^*)^{-1} = 1 - L_1 + \left(-L_2 + \frac{1}{2}L_1^2\right) + \left(-L_3 + L_1L_2 - \frac{1}{6}L_1^3\right) + \cdots \tag{F14}$$

and S is a gauge scalar that will also be expanded in ϵ. Order by order, one has

$$\overline{\gamma}^{(-1)} = \gamma^{(-1)} + dS^{(-1)}, \tag{F15a}$$

$$\overline{\gamma}^{(0)} = \gamma^{(0)} - L_1\gamma^{(-1)} + dS^{(0)}, \tag{F15b}$$

$$\overline{\gamma}^{(1)} = \gamma^{(1)} - L_1\gamma^{(0)} + \left(-L_2 + \frac{1}{2}L_1^2\right)\gamma^{(-1)} + dS^{(1)}, \tag{F15c}$$

$$\overline{\gamma}^{(2)} = \gamma^{(2)} - L_1\gamma^{(1)} + \left(-L_2 + \frac{1}{2}L_1^2\right)\gamma^{(0)} + \left(-L_3 + L_1L_2 - \frac{1}{6}L_1^3\right)\gamma^{(-1)} + dS^{(2)}, \tag{F15d}$$

$$\vdots$$

For this problem, $\gamma^{(n\geq 1)} = 0$.

We also need the action of an arbitrary L on an arbitrary 1-form. In Sec. D.3.1 we showed that effectively (from the point of view of the equations of motion)

$$L_w\gamma = w^\nu(\partial_\nu\gamma_\lambda - \partial_\lambda\gamma_\nu)dz^\lambda. \tag{F16}$$

[73] The recipe for constructing $(\mathrm{T}^*)^{-1}$ from T^* is simple: Change the sign of each L and swap the order of all noncommuting operators.

We assume that time is not transformed: $w^t = 0$. Also, all of the coefficients of the 1-forms are time-independent, so $\partial_t = 0$. Then

$$\begin{aligned} L_w\gamma = &\{\boldsymbol{w}^{\boldsymbol{x}} \cdot [\boldsymbol{\nabla}\gamma_{\boldsymbol{x}} - (\boldsymbol{\nabla}\gamma_{\boldsymbol{x}})^T] + w^U(\partial_U\gamma_{\boldsymbol{x}} - \boldsymbol{\nabla}\gamma_U) \\ &\quad + w^\mu(\partial_\mu\gamma_{\boldsymbol{x}} - \boldsymbol{\nabla}\gamma_\mu) + w^\zeta(\partial_\zeta\gamma_{\boldsymbol{x}} - \boldsymbol{\nabla}\gamma_\zeta)\} \cdot d\boldsymbol{x} \\ &+ [\boldsymbol{w}^{\boldsymbol{x}} \cdot (\boldsymbol{\nabla}\gamma_U - \partial_U\gamma_{\boldsymbol{x}}) + w^\mu(\partial_\mu\gamma_U - \partial_U\gamma_\mu) + w^\zeta(\partial_\zeta\gamma_U - \partial_U\gamma_\zeta)]dU \\ &+ [\boldsymbol{w}^{\boldsymbol{x}} \cdot (\boldsymbol{\nabla}\gamma_\mu - \partial_\mu\gamma_{\boldsymbol{x}}) + w^U(\partial_U\gamma_\mu - \partial_\mu\gamma_U) + w^\zeta(\partial_\zeta\gamma_\mu - \partial_\mu\gamma_\zeta)]d\mu \\ &+ [\boldsymbol{w}^{\boldsymbol{x}} \cdot (\boldsymbol{\nabla}\gamma_\zeta - \partial_\zeta\gamma_{\boldsymbol{x}}) + w^U(\partial_U\gamma_\zeta - \partial_\zeta\gamma_U) + w^\mu(\partial_\mu\gamma_\zeta - \partial_\zeta\gamma_\mu)]d\zeta \\ &+ (\boldsymbol{w}^{\boldsymbol{x}} \cdot \boldsymbol{\nabla} + w^U\partial_U + w^\mu\partial\mu + w^\zeta\partial_\zeta)\gamma_t\, dt. \end{aligned} \tag{F17}$$

A common operation will be the action of L_w on $\gamma^{(-1)}$, which has only an $\boldsymbol{x}$ component that depends only on $\boldsymbol{x}$. Thus

$$L_w\gamma^{(-1)} = \boldsymbol{w}^{\boldsymbol{x}} \cdot [\boldsymbol{\nabla}\gamma^{(-1)}_{\boldsymbol{x}} - (\boldsymbol{\nabla}\gamma^{(-1)}_{\boldsymbol{x}})^T] \cdot d\boldsymbol{x}. \tag{F18}$$

Upon noting that

$$\partial_i A_j - \partial_j A_i = \epsilon_{ijk}(\boldsymbol{\nabla} \times \boldsymbol{A})_k, \tag{F19}$$

one can readily verify that

$$\boldsymbol{w} \cdot [(\boldsymbol{\nabla}\boldsymbol{A}) - (\boldsymbol{\nabla}\boldsymbol{A})^T] = -\boldsymbol{w} \times \boldsymbol{\nabla} \times \boldsymbol{A}, \tag{F20}$$

so one finds

$$L_w\gamma^{(-1)} = -\omega_c \boldsymbol{w}^{\boldsymbol{x}} \times \widehat{\boldsymbol{b}} \cdot d\boldsymbol{x}. \tag{F21}$$

F.3 $O(\epsilon^{-1})$

At dominant order, one can choose $S^{(-1)} = 0$, so

$$\overline{\gamma}^{(-1)} = \gamma^{(-1)} = \frac{q}{mc}\boldsymbol{A} \cdot d\boldsymbol{x}. \tag{F22}$$

F.4 $O(\epsilon^0)$

From Eqs. (F12) and (F15b), one has

$$\begin{aligned} \overline{\gamma}^{(0)} = &(U\widehat{\boldsymbol{b}} + v_\perp\widehat{\boldsymbol{c}} + \omega_c\boldsymbol{w}_1^{\boldsymbol{x}} \times \widehat{\boldsymbol{b}} + \boldsymbol{\nabla}S^{(0)}) \cdot d\boldsymbol{x} \\ &+ \partial_U S^{(0)}\, dU + \partial_\mu S^{(0)}\, d\mu + \partial_\zeta S^{(0)}\, d\zeta - \underbrace{\left(\frac{1}{2}U^2 + \mu\widehat{\omega}_c + \widehat{\phi}\right)}_{H_0} dt. \end{aligned} \tag{F23}$$

One can choose $S^{(0)} = 0$ without encountering any contradictions. We want to eliminate the ζ dependence from $\overline{\gamma}^{(0)}_{\boldsymbol{x}}$. That dependence arises from the terms in $\widehat{\boldsymbol{c}}$ and $\boldsymbol{w}_1^{\boldsymbol{x}}$. Then crossing the $\boldsymbol{x}$ component with $\widehat{\boldsymbol{b}}$ determines the perpendicular part of $\boldsymbol{w}_1^{\boldsymbol{x}}$ to be

$$\boldsymbol{w}_{1\perp}^{\boldsymbol{x}} = -\boldsymbol{\rho}, \tag{F24}$$

where $\boldsymbol{\rho} \doteq \rho\widehat{\boldsymbol{a}}$ is the lowest-order gyroradius. This leaves

$$\overline{\gamma}^{(0)} = U\widehat{\boldsymbol{b}} \cdot d\boldsymbol{x} - H_0\, dt. \tag{F25}$$

F.5 $O(\epsilon^1)$

At higher orders, simplifications arise by writing as many of the formulas as possible in terms of already-determined barred quantities. For example, applying L_1 to Eq. (F15b) gives[74]

$$L_1\overline{\gamma}^{(0)} = L_1\gamma_0 - L_1^2\gamma^{(-1)} \tag{F27}$$

or

$$\frac{1}{2}L_1^2\gamma^{(-1)} = \frac{1}{2}L_1(\gamma^{(0)} - \overline{\gamma}^{(0)}). \tag{F28}$$

Upon inserting this result into Eq. (F15c) and recalling that $\gamma^{(-1)} = \overline{\gamma}^{(-1)}$, one obtains

$$\overline{\gamma}^{(1)} = -L_2\overline{\gamma}^{(-1)} - \frac{1}{2}L_1(\gamma^{(0)} + \overline{\gamma}^{(0)}) + dS^{(1)}. \tag{F29}$$

Now with $\gamma^{(0)} = (U\widehat{\boldsymbol{b}} + v_\perp\widehat{\boldsymbol{c}})\cdot d\boldsymbol{x} - [\frac{1}{2}U^2 + \mu\widehat{\omega}_c + \widehat{\phi}(\boldsymbol{x})]dt$, one has

$$\begin{aligned} L_1\gamma^{(0)} &= (w_1^U\widehat{\boldsymbol{b}} + w_1^\mu\partial_\mu v_\perp\widehat{\boldsymbol{c}} + w_1^\zeta\partial_\zeta v_\perp\widehat{\boldsymbol{c}})\cdot d\boldsymbol{x} \\ &\quad - \boldsymbol{w}_1^{\boldsymbol{x}}\cdot\partial_U(U\widehat{\boldsymbol{b}})dU - \boldsymbol{w}_1^{\boldsymbol{x}}\cdot\partial_\mu(v_\perp\widehat{\boldsymbol{c}})d\mu - \boldsymbol{w}_1^{\boldsymbol{x}}\cdot\partial_\zeta(v_\perp\widehat{\boldsymbol{c}})d\zeta \\ &\quad - (\boldsymbol{w}_1^{\boldsymbol{x}}\cdot\boldsymbol{\nabla} + w_1^U\partial_U + w_1^\mu\partial_\mu)\left(\frac{1}{2}U^2 + \mu\widehat{\omega}_c + \widehat{\phi}\right)dt \qquad \text{(F30a)} \\ &= (w_1^U\widehat{\boldsymbol{b}} + w_1^\mu v_\perp^{-1}\widehat{\boldsymbol{c}} - w_1^\zeta v_\perp\widehat{\boldsymbol{a}})\cdot d\boldsymbol{x} \\ &\quad - \boldsymbol{w}_1^{\boldsymbol{x}}\cdot\widehat{\boldsymbol{b}}\,dU - \underbrace{\boldsymbol{w}_1^{\boldsymbol{x}}\cdot\widehat{\boldsymbol{c}}}_{0}v_\perp^{-1}d\mu + \underbrace{\boldsymbol{w}_1^{\boldsymbol{x}}\cdot\widehat{\boldsymbol{a}}v_\perp}_{2\mu/\omega_{c0}}d\zeta \\ &\quad - (\boldsymbol{w}_1^{\boldsymbol{x}}\cdot\boldsymbol{\nabla}\widehat{\phi} + w_1^U U + w_1^\mu\widehat{\omega}_c)dt. \qquad \text{(F30b)} \end{aligned}$$

The terms tracing their origin from $v_\perp\widehat{\boldsymbol{c}}\cdot d\boldsymbol{x}$ are absent from $\overline{\gamma}^{(0)}$, so

$$L_1\overline{\gamma}^{(0)} = w_1^U\widehat{\boldsymbol{b}}\cdot d\boldsymbol{x} - \boldsymbol{w}_1^{\boldsymbol{x}}\cdot\widehat{\boldsymbol{b}}\,dU - (\boldsymbol{w}_1^{\boldsymbol{x}}\cdot\boldsymbol{\nabla}\widehat{\phi} + w_1^U U + w_1^\mu\widehat{\omega}_c)dt. \tag{F31}$$

Thus, upon recalling that $\boldsymbol{w}_{1\perp}^{\boldsymbol{x}} = -\boldsymbol{\rho}$,

$$\begin{aligned} \overline{\gamma}^{(1)} &= \left(\omega_c\boldsymbol{w}_2^{\boldsymbol{x}}\times\widehat{\boldsymbol{b}} - w_1^U\widehat{\boldsymbol{b}} - \frac{1}{2}w_1^\mu v_\perp^{-1}\widehat{\boldsymbol{c}} + \frac{1}{2}w_1^\zeta v_\perp\widehat{\boldsymbol{a}} + \boldsymbol{\nabla}S^{(1)}\right)\cdot d\boldsymbol{x} \\ &\quad + (-w_{1\parallel}^{\boldsymbol{x}} + \partial_U S^{(1)})dU + \partial_\mu S^{(1)}d\mu + (\mu\omega_{c0}^{-1} + \partial_\zeta S^{(1)})d\zeta \\ &\quad + [(-\boldsymbol{\rho} + w_{1\parallel}^{\boldsymbol{x}}\,\widehat{\boldsymbol{b}})\cdot\boldsymbol{\nabla}\widehat{\phi} + w_1^U U + w_1^\mu\widehat{\omega}_c]dt. \end{aligned} \tag{F32}$$

From the ζ component, one finds that $\partial_\zeta S^{(1)} = 0$. It is consistent to choose $S^{(1)} = 0$. From the $\boldsymbol{x}$ component, the parallel projection gives $w_1^U = 0$. One can

[74] In writing Eq. (F27), we have omitted the term $L_1 dS^{(0)}$. From Cartan's magic formula (D16), one has

$$L_w dS = i_w ddS + d(i_w dS) = d(i_w dS). \tag{F26}$$

Thus the contribution from $L_w dS^{(0)}$ is a pure differential. Since that can be absorbed into $S^{(1)}$, we do not write it explicitly. We proceed similarly at higher orders.

annihilate the U component by choosing $w_{1\|}^{\boldsymbol{x}} = 0$. By eliminating ζ dependence from the t component, $\langle w_1^\mu \rangle$ is not determined, but the fluctuating part is found to be

$$\delta w_1^\mu = \widehat{\omega}_c^{-1} \boldsymbol{\rho} \cdot \boldsymbol{\nabla} \widehat{\phi}. \tag{F33}$$

By crossing the $\boldsymbol{x}$ component with $\widehat{\boldsymbol{b}}$, the perpendicular part of $\boldsymbol{w}_2^{\boldsymbol{x}}$ is determined to be

$$\boldsymbol{w}_{2\perp}^{\boldsymbol{x}} = \omega_c^{-1} \frac{1}{2} \left(w_1^\mu v_\perp^{-1} \widehat{\boldsymbol{a}} + w_1^\zeta v_\perp \widehat{\boldsymbol{c}} \right). \tag{F34}$$

One is left with

$$\overline{\gamma}^{(1)} = \mu \omega_{c0}^{-1} \, d\zeta + \langle w_1^\mu \rangle \widehat{\omega}_c \, dt. \tag{F35}$$

F.6 $O(\epsilon^2)$

We proceed to express the operations involved in $\overline{\gamma}^{(2)}$ in terms of quantities already known (as much as possible). First, one can eliminate $\gamma^{(0)}$ from Eq. (F15c) by using Eq. (F15b):

$$\overline{\gamma}^{(1)} = -L_1(\overline{\gamma}^{(0)} + L_1 \gamma^{(-1)}) + \left(-L_2 + \frac{1}{2} L_1^2 \right) \gamma^{(-1)} + dS^{(1)} \tag{F36a}$$

$$= -L_1 \overline{\gamma}^{(0)} - \left(L_2 + \frac{1}{2} L_1^2 \right) \gamma^{(-1)} + dS^{(1)}. \tag{F36b}$$

Eliminating $\gamma^{(0)}$ from Eq. (F15d) gives

$$\overline{\gamma}^{(2)} = \left(-L_2 + \frac{1}{2} L_1^2 \right) (\overline{\gamma}^{(0)} + L_1 \overline{\gamma}^{(-1)}) + \left(-L_3 + L_1 L_2 - \frac{1}{6} L_1^3 \right) \overline{\gamma}^{(-1)} + dS^{(2)} \tag{F37a}$$

$$= \left(-L_2 + \frac{1}{2} L_1^2 \right) \overline{\gamma}^{(0)} - L_2 L_1 \overline{\gamma}^{(-1)} + \frac{1}{2} L_1^3 \overline{\gamma}^{(-1)} + \left(-L_3 + L_1 L_2 - \frac{1}{6} L_1^3 \right) \overline{\gamma}^{(-1)} + dS^{(2)} \tag{F37b}$$

$$= -L_3 \overline{\gamma}^{(-1)} + [L_1, L_2] \overline{\gamma}^{(-1)} + \frac{1}{3} L_1^3 \overline{\gamma}^{(-1)} + \left(-L_2 + \frac{1}{2} L_1^2 \right) \overline{\gamma}^{(0)}. \tag{F37c}$$

One can find the L_1^3 term by applying L_1 to Eq. (F36b); that gives (recalling that $\gamma^{(1)} = 0$)

$$L_1 \overline{\gamma}^{(1)} = -L_1^2 \overline{\gamma}^{(0)} - \left(L_1 L_2 + \frac{1}{2} L_1^3 \right) \gamma^{(-1)}, \tag{F38}$$

or

$$\frac{1}{3} L_1^3 \gamma^{(-1)} = -\frac{2}{3} (L_1 L_2 \overline{\gamma}^{(-1)} + L_1^2 \overline{\gamma}^{(0)} + L_1 \overline{\gamma}^{(1)}). \tag{F39}$$

Then

$$\overline{\gamma}^{(2)} = (-L_3 + [L_1, L_2])\overline{\gamma}^{(-1)} - \frac{2}{3}(L_1 L_2 \overline{\gamma}^{(-1)} + L_1^2 \overline{\gamma}^{(0)} + L_1 \overline{\gamma}^{(1)}) + \left(-L_2 + \frac{1}{2}L_1^2\right)\overline{\gamma}^{(0)} + dS^{(2)} \tag{F40a}$$

$$= (-L_3 + [L_1, L_2])\overline{\gamma}^{(-1)} - \frac{2}{3}(L_1 L_2 \overline{\gamma}^{(-1)} + L_1 \overline{\gamma}^{(1)}) - \left(L_2 + \frac{1}{6}L_1^2\right)\overline{\gamma}^{(0)} + dS^{(2)}. \tag{F40b}$$

Now

$$L_2 \overline{\gamma}^{(-1)} = -\omega_c \boldsymbol{w}_2^{\boldsymbol{x}} \times \widehat{\boldsymbol{b}} \cdot d\boldsymbol{x} \tag{F41a}$$

$$= \frac{1}{2}(-w_1^\mu v_\perp^{-1} \widehat{\boldsymbol{c}} + w_1^\zeta v_\perp \widehat{\boldsymbol{a}}) \cdot d\boldsymbol{x}, \tag{F41b}$$

where we used Eq. (F34). Then [specializing Eq. (F17) to the case of only an $\boldsymbol{x}$ component and recalling Eq. (F24)]

$$\begin{aligned} L_1 L_2 \overline{\gamma}^{(-1)} = \Bigg[& - \boldsymbol{w}_1^{\boldsymbol{x}} \times \boldsymbol{\nabla} \times (-\omega_c \boldsymbol{w}_2^{\boldsymbol{x}} \times \widehat{\boldsymbol{b}}) \\ & + w_1^\mu \partial_\mu \left(\frac{1}{2}(-w_1^\mu v_\perp^{-1} \widehat{\boldsymbol{c}} + w_1^\zeta v_\perp \widehat{\boldsymbol{a}}) \right) \\ & + w_1^\zeta \partial_\zeta \left(\frac{1}{2}(-w_1^\mu v_\perp^{-1} \widehat{\boldsymbol{c}} + w_1^\zeta v_\perp \widehat{\boldsymbol{a}}) \right) \Bigg] \cdot d\boldsymbol{x} \\ & - \boldsymbol{w}_1^{\boldsymbol{x}} \cdot \partial_U \left(\frac{1}{2}(-w_1^\mu v_\perp^{-1} \widehat{\boldsymbol{c}} + w_1^\zeta v_\perp \widehat{\boldsymbol{a}}) \right) dU \\ & - \boldsymbol{w}_1^{\boldsymbol{x}} \cdot \partial_\mu \left(\frac{1}{2}(-w_1^\mu v_\perp^{-1} \widehat{\boldsymbol{c}} + w_1^\zeta v_\perp \widehat{\boldsymbol{a}}) \right) d\mu \\ & - \boldsymbol{w}_1^{\boldsymbol{x}} \cdot \partial_\zeta \left(\frac{1}{2}(-w_1^\mu v_\perp^{-1} \widehat{\boldsymbol{c}} + w_1^\zeta v_\perp \widehat{\boldsymbol{a}}) \right) d\zeta \end{aligned} \tag{F42}$$

$$= [\dots] \cdot d\boldsymbol{x} + \frac{1}{2}\rho v_\perp \partial_U w_1^\zeta dU + \frac{1}{2}\rho \partial_\mu (w_1^\zeta v_\perp) d\mu + \frac{1}{2}(\omega_c^{-1} w_1^\mu + \rho v_\perp \partial_\zeta w_1^\zeta) d\zeta. \tag{F43}$$

Upon recalling that $\overline{\gamma}^{(1)} = \mu \omega_{c0}^{-1}\, d\zeta$, one finds that

$$L_1 \overline{\gamma}^{(1)} = \omega_{c0}^{-1}(-w_1^\zeta\, d\mu + w_1^\mu\, d\zeta). \tag{F44}$$

Upon recalling that $\overline{\gamma}^{(0)} = U\widehat{\boldsymbol{b}} \cdot d\boldsymbol{x} - H_0 dt$, one has

$$L_2 \overline{\gamma}^{(0)} = w_2^U \widehat{\boldsymbol{b}} \cdot d\boldsymbol{x} - \boldsymbol{w}_2^{\boldsymbol{x}} \cdot \widehat{\boldsymbol{b}}\, dU - (\boldsymbol{w}_2^{\boldsymbol{x}} \cdot \boldsymbol{\nabla}\widehat{\phi} + w_2^U U + w_2^\mu \widehat{\omega}_c) dt. \tag{F45}$$

Finally, we calculate $L_1^2 \overline{\gamma}^{(0)}$. We had [Eqs. (F31) and (F33)]

$$L_1 \overline{\gamma}^{(0)} = w_1^U \widehat{\boldsymbol{b}} \cdot d\boldsymbol{x} - \boldsymbol{w}_1^{\boldsymbol{x}} \cdot \widehat{\boldsymbol{b}}\, dU - (\boldsymbol{w}_1^{\boldsymbol{x}} \cdot \boldsymbol{\nabla}\widehat{\phi} + w_1^U U + w_1^\mu \widehat{\omega}_c) dt = -\langle w_1^\mu \rangle \widehat{\omega}_c\, dt. \tag{F46}$$

Then from the last line of Eq. (F17),

$$L_1^2\overline{\gamma}^{(0)} = (-\boldsymbol{\rho}\cdot\boldsymbol{\nabla} + w_1^\mu\partial_\mu)(-\langle w_1^\mu\rangle\omega_c)dt. \tag{F47}$$

Now we add up all the terms contributing to Eq. (F40b). Note the fundamental result from the theory of Lie derivatives that [Eq. (D9)]

$$[L_1, L_2] = L_{[w_1,w_2]}, \tag{F48}$$

so the commutator term just contributes to the $\boldsymbol{x}$ component. The ζ component is

$$\overline{\gamma}_\zeta^{(2)} = -\frac{2}{3}\left(\frac{1}{2}(\omega_{c0}^{-1}w_1^\mu + 2\mu\omega_{c0}^{-1}\partial_\zeta w_1^\zeta) + w_1^\mu\omega_{c0}^{-1}\right) + \partial_\zeta S^{(2)} \tag{F49a}$$

$$= -\omega_{c0}^{-1}\left(w_1^\mu + \frac{2}{3}\mu\,\partial_\zeta w_1^\zeta\right) + \partial_\zeta S^{(2)}. \tag{F49b}$$

We preserve the symplectic structure by demanding that $\overline{\gamma}_\zeta^{(2)} = 0$, which requires $\langle w_1^\mu\rangle = 0$; then [Eq. (F46)] $L_1\overline{\gamma}^{(0)} = 0$. Equation (F49b) then determines the fluctuating part of $S^{(2)}$ to be

$$\delta S^{(2)} = \omega_{c0}^{-1}\int d\zeta\left(\delta w_1^\mu + \frac{2}{3}\mu\,\partial_\zeta w_1^\zeta\right) \tag{F50a}$$

$$= -\omega_c^{-1}\rho\widehat{\boldsymbol{c}}\cdot\boldsymbol{\nabla}\widehat{\phi} + \frac{2}{3}\omega_{c0}^{-1}\mu w_1^\zeta. \tag{F50b}$$

To determine w_1^ζ, we turn to the μ component. That is

$$\overline{\gamma}_\mu^{(2)} = -\frac{2}{3}\left(\frac{1}{2}\rho\partial_\mu(w_1^\zeta v_\perp) - w_1^\zeta\right) + \partial_\mu S^{(2)} \tag{F51a}$$

$$= -\frac{2}{3}\left[\partial_\mu\left(\frac{1}{2}\rho v_\perp w_1^\zeta\right) - \frac{3}{2}w_1^\zeta\right] + \partial_\mu S^{(2)} \tag{F51b}$$

$$= w_1^\zeta - \omega_c^{-2}v_\perp^{-1}\widehat{\boldsymbol{c}}\cdot\boldsymbol{\nabla}\widehat{\phi}, \tag{F51c}$$

where we used Eq. (F50b). One finds $\langle w_1^\zeta\rangle = 0$ and

$$\delta w_1^\zeta = (\omega_c v_\perp)^{-1}\widehat{\boldsymbol{c}}\cdot\boldsymbol{\nabla}\widehat{\phi}, \tag{F52}$$

consistent with our previous intuitive and geometrical arguments.

Finally, we consider $\overline{\gamma}_t$, which stems solely from $-L_2\overline{\gamma}^{(0)}$ [Eq. (F45)]:

$$\overline{\gamma}_t^{(2)} = (\boldsymbol{w}_2^{\boldsymbol{x}}\cdot\boldsymbol{\nabla}\widehat{\phi} + w_2^U U + w_2^\mu\widehat{\omega}_c)dt, \tag{F53}$$

where Eqs. (F33) and (F52) were used. One sees from Eq. (F34) that

$$\boldsymbol{w}_{2\perp}^{\boldsymbol{x}} = \frac{1}{2}\omega_c^{-2}\boldsymbol{\nabla}_\perp\widehat{\phi}. \tag{F54}$$

Consideration of the $\widehat{\boldsymbol{b}}$ projection of $\overline{\gamma}_{\boldsymbol{x}}$ leads one to conclude that $w_2^U = 0$. One must go to third order to find that $\langle w_2^\mu \rangle = 0$. Thus w_2^μ is purely fluctuating, but there are no fluctuating terms in $\overline{\gamma}_t$; therefore $w_2^\mu = 0$ and

$$\overline{\gamma}_t^{(2)} = \tfrac{1}{2}\omega_c^{-2}|\boldsymbol{\nabla}_\perp \widehat{\phi}|^2 dt \equiv -m^{-1}\overline{H}^{(2)} dt. \tag{F55}$$

(Recall that we divided out an m at the beginning.) One readily finds that

$$\overline{H}^{(2)} = -\frac{1}{2} m u_E^2. \tag{F56}$$

We will given an interpretation of this result in the next section.

F.7 Interpretation of $\overline{H}^{(2)}$

Let us check the consistency of these results by examining the resulting variable transformation. From Eq. (F7),

$$\overline{\boldsymbol{x}} = \boldsymbol{x} + \boldsymbol{w}_1^{\boldsymbol{x}} + \boldsymbol{w}_2^{\boldsymbol{x}} + \frac{1}{2}\mathcal{L}_1 \boldsymbol{w}_1^{\boldsymbol{x}} + \cdots . \tag{F57}$$

We have

$$\begin{aligned}
\mathcal{L}_1 \boldsymbol{w}_1^{\boldsymbol{x}} &= (\boldsymbol{w}_1^{\boldsymbol{x}} \cdot \underbrace{\boldsymbol{\nabla}}_{0} + w_1^\mu \partial_\mu + w_1^\zeta \partial_\zeta)(-\boldsymbol{\rho}) && \text{(F58a)}\\
&= -\{\boldsymbol{\rho} \cdot \boldsymbol{\nabla}\widehat{\phi}[(\omega_c v_\perp)^{-1}\widehat{\boldsymbol{a}}] + (\omega_c v_\perp)^{-1}\widehat{\boldsymbol{c}} \cdot \boldsymbol{\nabla}\widehat{\phi}(\rho\widehat{\boldsymbol{c}})\} && \text{(F58b)}\\
&= -\omega_c^{-2}[(\widehat{\boldsymbol{a}}\,\widehat{\boldsymbol{a}} + \widehat{\boldsymbol{c}}\,\widehat{\boldsymbol{c}}) \cdot \boldsymbol{\nabla}\widehat{\phi}] && \text{(F58c)}\\
&= -\omega_c^{-2}\boldsymbol{\nabla}_\perp \widehat{\phi}. && \text{(F58d)}
\end{aligned}$$

Upon recalling Eq. (F54), we see that the second-order terms of Eq. (F57) cancel and one has

$$\overline{\boldsymbol{x}} = \boldsymbol{x} - \boldsymbol{\rho} + O(\epsilon^3). \tag{F59}$$

Similarly, one has

$$\overline{\mu} = \mu + w_1^\mu + w_2^\mu + \frac{1}{2}\mathcal{L}_1 w_1^\mu + \cdots , \tag{F60}$$

with

$$\begin{aligned}
\mathcal{L}_1 w_1^\mu &= (w_1^\mu \partial_\mu + w_1^\zeta \partial_\zeta)(\omega_c^{-1}\boldsymbol{\rho} \cdot \boldsymbol{\nabla}\widehat{\phi}) && \text{(F61a)}\\
&= \omega_c^{-2}|\boldsymbol{\nabla}_\perp \widehat{\phi}|^2. && \text{(F61b)}
\end{aligned}$$

(We already calculated $\mathcal{L}_1\boldsymbol{\rho}$ above.) Since $w_2^\mu = 0$, one readily sees that Eq. (F60) reproduces Eq. (F2):

$$\overline{\mu} = \frac{1}{2} m|\boldsymbol{v}_\perp - \boldsymbol{u}_{\boldsymbol{E}}|^2/\omega_c + O(\epsilon^3). \tag{F62}$$

Because we chose the variable transformation to preserve the symplectic structure, the Poisson brackets retain their guiding-center forms (E10). In particular,

$$\frac{d\overline{\boldsymbol{x}}}{dt} = \{\overline{\boldsymbol{x}}, \overline{H}\} = \frac{c}{B}\widehat{\boldsymbol{b}} \times \boldsymbol{\nabla}\overline{H}(\overline{\zeta}) + \cdots . \tag{F63}$$

The contribution from $\overline{H}^{(0)} = q\phi$ gives the usual $\boldsymbol{E} \times \boldsymbol{B}$ drift $\boldsymbol{u}_{\boldsymbol{E}}$. From $\overline{H}^{(2)}$, one obtains a ponderomotive correction

$$\Delta\overline{\boldsymbol{u}} = \frac{c}{B}\widehat{\boldsymbol{b}} \times \boldsymbol{\nabla}\left(-\frac{1}{2}mu_E^2\right). \tag{F64}$$

stemming from the ponderomotive potential $q^{-1}[-\frac{1}{2}mu_E^2(\overline{\boldsymbol{x}})]$. To interpret that potential, we turn to the equation of motion

$$m\frac{d\boldsymbol{v}}{dt} = q\boldsymbol{E}(\boldsymbol{x}) + m\omega_c\boldsymbol{v} \times \widehat{\boldsymbol{b}}, \tag{F65}$$

which is written in terms of particle variables. Now

$$\boldsymbol{E}(\boldsymbol{x}) = \boldsymbol{E}(\overline{\boldsymbol{x}} + \boldsymbol{\rho}) = \boldsymbol{E}(\overline{\boldsymbol{x}}) + \boldsymbol{\rho}(z) \cdot \overline{\boldsymbol{\nabla}}\boldsymbol{E} + \cdots . \tag{F66}$$

Here $\boldsymbol{\rho}(z) = \rho(\mu)\widehat{\boldsymbol{a}}(\zeta)$, again written in terms of particle variables. We encounter here the same paradox alluded to earlier, which is that if one were to equate the cold-ion limit with the limit $\rho(z) \to 0$, the gyroradius correction (which will ultimately contain the ponderomotive effect) would vanish. We know, however, that is incorrect because $\boldsymbol{v}_\perp$ contains a $\boldsymbol{u}_{\boldsymbol{E}}$ part. What one must do is write $\boldsymbol{\rho}(z)$ in terms of the barred variables, then take the limit. We have

$$\begin{aligned}
\boldsymbol{\rho}(z) &= \rho(\mu)\widehat{\boldsymbol{a}}(\zeta) && \text{(F67a)}\\
&= \rho(\overline{\mu} - \delta\mu)\widehat{\boldsymbol{a}}(\overline{\zeta} - \delta\zeta) && \text{(F67b)}\\
&= \rho(\overline{\mu})\widehat{\boldsymbol{a}}(\overline{\zeta}) - \delta\mu\,\partial_{\overline{\mu}}\rho(\overline{\mu})\widehat{\boldsymbol{a}}(\overline{\zeta}) - \delta\zeta\,\rho(\overline{\mu})\partial_{\overline{\zeta}}\widehat{\boldsymbol{a}}(\overline{\zeta}) + \cdots && \text{(F67c)}\\
&= \overline{\boldsymbol{\rho}} - w_1^\mu(\omega_c\overline{v}_\perp)^{-1}\widehat{\boldsymbol{a}}(\overline{\zeta}) - w_1^\zeta\rho(\overline{\mu})\widehat{\boldsymbol{c}}(\overline{\zeta}) + O(\epsilon^3) && \text{(F67d)}\\
&= \overline{\boldsymbol{\rho}} - \omega_c^{-2}\boldsymbol{\nabla}_\perp\widehat{\phi}(\overline{\boldsymbol{x}}) + O(\epsilon^3). && \text{(F67e)}
\end{aligned}$$

(The operations are the same as those involved in the calculation of $\mathcal{L}_1\boldsymbol{w}_1^{\boldsymbol{x}}$.) In the cold-ion limit, $\overline{\boldsymbol{\rho}} \to \boldsymbol{0}$ and the $\boldsymbol{\rho}$ correction in Eq. (F66) is at second order proportional to $-\boldsymbol{\nabla}_\perp\phi \cdot \boldsymbol{\nabla}(-\boldsymbol{\nabla}\phi) = \frac{1}{2}\boldsymbol{\nabla}|\boldsymbol{\nabla}_\perp\phi|^2$. Upon comparing with Eq. (F64), one see that this corresponds exactly to the negative of the ponderomotive potential $q^{-1}\overline{H}^{(2)}$.

This calculation shows in detail how ponderomotive nonlinearities are intimately related to conservation of the magnetic moment $\overline{\mu}$. Those nonlinearities are most easily obtained by the systematic calculation of gyrokinetic $\overline{\mu}$-conserving variables, but the effect is a physical one and the ponderomotive force shows up in a particle-based calculation as well. This observation has implications for the derivation of the momentum-conservation law. Although that can be obtained most easily from the gyrokinetic–Poisson system, it also follows from the particle-based moment equations provided that one deals properly with the ponderomotive nonlinearities.

G Eulerian variational principle for gyrokinetics

The seminal derivation of a gyrokinetic variational principle was by Sugama (2000), who used a Lagrangian formulation. Here we summarize instead the Eulerian formulation of Brizard (2000), which is based on constrained variations.

G.1 Eulerian and Lagrangian variations

In the discussion to follow, it is important to distinguish between Eulerian and Lagrangian variations. Those are formulated in terms of general transformations of both the coordinates x^μ and the fields ψ:

$$x^\mu \to \xi^\mu \doteq x^\mu + \delta x^\mu, \tag{G1a}$$

$$\psi(x) \to \psi'(\xi) \doteq \psi(x) + \delta\psi_L(x). \tag{G1b}$$

The subscript L denotes *Lagrangian*. It is important to note that the change $\delta\psi$ contains an intrinsic Eulerian part $\delta\psi$ arising from a change in shape of ψ at fixed x (the same variation that is used in the derivation of the Euler–Lagrange equations), as well as a change arising from the coordinate transformation; thus the Eulerian variation is defined by

$$\delta\psi(x) \doteq \psi'(x) - \psi(x). \tag{G2}$$

If ψ is a scalar field, one can expand the $\psi'(\xi)$ in Eq. (G1b) as

$$\psi'(\xi) = \psi'(x) + \delta x^\mu \partial_\mu \psi + \cdots, \tag{G3}$$

so

$$\delta\psi = \delta\psi_L - \delta x^\mu \partial_\mu \psi. \tag{G4}$$

More generally, the directional derivative should be replaced by the *Lie derivative* in the direction of the flow defined by δx^μ. To first order in ϵ, write

$$\delta x^\mu = \epsilon X^\mu, \tag{G5a}$$

$$\delta\psi_L = \epsilon \Psi_L. \tag{G5b}$$

Then the Eulerian variation is through first order

$$\delta\psi = \epsilon\Psi_L - \epsilon L_X \psi. \tag{G6}$$

A *constrained Eulerian variation* is defined to be one for which $\delta\psi_L = 0$.

G.2 Brizard's variational principle

Brizard writes the gyrokinetic action in the form

$$\mathcal{A}_{\rm gy} = \int d^4x\, L_{\rm gy}(x), \tag{G7}$$

where the *Lagrangian density* is

$$L_{\rm gy}(x) \doteq L_M(x) - \sum_s \overline{n}_s \int d^4p\, \mathcal{F}_{\rm gy}(z)\mathcal{H}_{\rm gy}(\phi;z) \tag{G8}$$

and (for electrostatics)

$$L_M \doteq \frac{1}{8\pi}(|\boldsymbol{\nabla}\phi|^2 - B^2). \tag{G9}$$

B is here the background magnetic field and does not participate in variations. Also, $\mathcal{H}$ is the gyrocenter Hamiltonian in extended phase space,

$$\mathcal{H}_{\rm gy}(\phi;z) \doteq H_{\rm gy}(\phi;z) - w, \tag{G10}$$

and $\mathcal{F}$ is the extended gyrocenter (Vlasov) distribution. For physical solutions,

$$\mathcal{F}(z) \doteq c\delta(w - H_{\rm gy})F(z) = c\delta(\mathcal{H})F. \tag{G11}$$

However, for use in the variational principle its form must be allowed to be arbitrary.

The argument dependence in these formulas may be confusing. The gyrokinetic Hamiltonian is derived by the sequence of transformations $\{\boldsymbol{x},\boldsymbol{p}\} \equiv z \to Z \to \overline{Z}$, where Z denotes the lowest-order guiding-center coordinates. Strictly speaking, then, the gyrokinetic Lagrangian should depend on $\overline{Z}$, and the action should involve $\int d\overline{X}$ and $\int d\overline{P}$. However, all of the variables are under integrals, so they are dummies and can be renamed. Brizard has chosen to call them x and p rather than $\overline{X}$ and $\bar{P}$ (z rather than $\overline{Z}$). This has advantages later where a proliferation of overlines would clutter up the notation.

Eulerian variation of the functional (G7) gives straightforwardly $\delta\mathcal{A}_{\rm gy} = \int d^4x\, \delta L_{\rm gy}$, where

$$\begin{aligned}\delta L_{\rm gy}(x) \doteq \frac{1}{4\pi}\boldsymbol{E}\cdot\delta\boldsymbol{E} &- \sum_s \overline{n}_s \int d^4p\, \delta\mathcal{F}_{\rm gy}(z)\mathcal{H}_{\rm gy}(z) \\ &- \sum_s \overline{n}_s \int d^4p \int d^4X\, \mathcal{F}_{\rm gy}(z)\frac{\delta H_{\rm gy}(z)}{\delta\phi(X)}\delta\phi(X).\end{aligned} \tag{G12}$$

(In the last term, H rather than $\mathcal{H}$ appears here because w is independent of ϕ. Also, X is a dummy variable of space-time integration.) Note that $\delta\mathcal{F}$ is not considered to be a totally independent variation. Rather, it is taken to be a *constrained Eulerian variation* $\mathcal{F}(Z) \to \mathcal{F}'(Z + \delta Z)$, where δZ is a virtual displacement in the extended phase space. From formula (G4) with ψ replaced by the scalar distribution $\mathcal{F}$ and $x \to Z$, one finds

$$\delta\mathcal{F} = -\delta Z^a \frac{\partial\mathcal{F}}{\partial Z^a}. \tag{G13}$$

Brizard (2001) has shown that path independence in a two-dimensional orbit–parameter space leads to the formula

$$\delta Z^a = \{Z^a, \mathcal{S}\}, \tag{G14}$$

where S is a scalar generating function for an infinitesimal transformation. Thus

$$\delta\mathcal{F} = -\{Z^a, \mathcal{S}\}\frac{\partial \mathcal{F}}{\partial Z^a} = -\{\mathcal{F}, \mathcal{S}\}. \tag{G15}$$

$\mathcal{S}$ plays the role of an effective Hamiltonian for virtual displacements; note the similarity between Eqs. (G14) and (G15) and the equations for Hamiltonian time development

$$\dot{z}^i = \{z^i, H\}, \tag{G16a}$$

$$\partial_t F = -\{F, H\}. \tag{G16b}$$

Now consider the $\delta\mathcal{F}$ term of Eq. (G12) in more detail. Using canonical coordinates for simplicity, one has

$$-\sum_s \overline{n}_s \int d^4p\, \delta\mathcal{F}\, \mathcal{H} = -\sum_s \overline{n}_s \int d^4p\, \{\mathcal{S}, \mathcal{F}\}\mathcal{H} \tag{G17a}$$

$$= -\sum_s \overline{n}_s \int d^4p \left(\frac{\partial \mathcal{S}}{\partial x^\mu}\frac{\partial \mathcal{F}}{\partial p_\mu} - \frac{\partial \mathcal{S}}{\partial p_\mu}\frac{\partial \mathcal{F}}{\partial x^\mu}\right)\mathcal{H} \tag{G17b}$$

$$\begin{aligned} = &-\sum_s \overline{n}_s \int d^4p \frac{\partial}{\partial x^\mu}\left(\mathcal{S}\frac{\partial \mathcal{F}}{\partial p_\mu}\mathcal{H}\right) + \sum_s \overline{n}_s \int d^4p\, \mathcal{S}\frac{\partial}{\partial x^\mu}\left(\frac{\partial \mathcal{F}}{\partial p_\mu}\mathcal{H}\right) \\ &+ \sum_s \overline{n}_s \int d^4p \frac{\partial}{\partial p_\mu}\left(\mathcal{S}\frac{\partial \mathcal{F}}{\partial x^\mu}\mathcal{H}\right) - \sum_s \overline{n}_s \int d^4p\, \mathcal{S}\frac{\partial}{\partial p_\mu}\left(\frac{\partial \mathcal{F}}{\partial x^\mu}\mathcal{H}\right) \end{aligned} \tag{G17c}$$

$$= \frac{\partial}{\partial x^\mu}\sum_s \overline{n}_s \int d^4p\, \mathcal{F}\left(\frac{\partial \mathcal{S}}{\partial p_\mu}\mathcal{H} + \mathcal{S}\frac{\partial \mathcal{H}}{\partial p_\mu}\right) - \sum_s \overline{n}_s \int d^4p\, \mathcal{S}\{\mathcal{F}, \mathcal{H}\} \tag{G17d}$$

$$= \frac{\partial}{\partial x^\mu}\sum_s \overline{n}_s \int d^4p\, \mathcal{F}\mathcal{S}v^\mu - \sum_s \overline{n}_s \int d^4p\, \mathcal{S}\{\mathcal{F}, \mathcal{H}\}, \tag{G17e}$$

where $v^\mu \doteq \partial\mathcal{H}/\partial p_\mu$. In obtaining the last line, one used the fact that $\mathcal{F}\mathcal{H} \equiv 0$ [see Eq. (G11)].

Next, consider the evaluation of the functional derivative with respect to ϕ that is required for Eq. (G12). Here it is important to recall that the derivation of the gyrokinetic 1-form proceeds (or, more precisely, can proceed; see footnote 67 on p. 81) in two stages. First, one implements a nonperturbative transformation T_{gc} to lowest-order guiding-center coordinates. That is,

$$\mathrm{T}_{\mathrm{gc}}\colon \{\boldsymbol{x}, \boldsymbol{p}\} \equiv z \to \{\boldsymbol{X}, U, \mu, \zeta\} \equiv Z, \tag{G18}$$

which is written as $Z = \mathrm{T}_{\mathrm{gc}} z$. Thus

$$\gamma = \Gamma_{\mathrm{gc}} - \mathcal{H}_{\mathrm{gc}}\, d\tau, \tag{G19}$$

where

$$\mathcal{H}_{\mathrm{gc}} \doteq \mathcal{H}_0 + \delta\mathcal{H} - W, \tag{G20a}$$

$$\mathcal{H}_0 \doteq (2m)^{-1}p_\parallel^2 + \mu\omega_c, \tag{G20b}$$

$$\delta\mathcal{H} \doteq q\phi(\boldsymbol{X} + \boldsymbol{\rho}_0). \tag{G20c}$$

($p_\parallel \doteq m v_\parallel$.) By the standard result that the functional forms of scalar fields transform inversely to the coordinates, one has

$$\phi(\boldsymbol{X} + \boldsymbol{\rho}_0) \equiv \phi_{\rm gc}(Z) = (\mathrm{T}^*)^{-1}_{\rm gc}\phi(Z). \tag{G21}$$

Next, one implements a perturbative Lie transformation $\mathrm{T}_{\rm gy}\colon Z \to \overline{Z}$ to remove any residual ζ dependence. Some of that dependence may still reside in Γ. However, there is enough freedom to preserve the symplectic structure (and thus the form of the guiding-center Poisson brackets). Therefore, although it may be tedious to construct the Lie transformation at high order, the *form* of the gyrocenter Hamiltonian is simple. If one defines $(\mathrm{T}^*)^{-1} \doteq (\mathrm{T}^*_{\rm gy})^{-1}(\mathrm{T}^*_{\rm gc})^{-1}$, then

$$\delta\overline{\mathcal{H}}_{\rm gy}(\overline{Z}) = q\langle(\mathrm{T}^*)^{-1}\phi(\overline{Z})\rangle. \tag{G22}$$

Since this is a statement about a functional form, all of the bars may be dropped, so we will work with

$$\delta\mathcal{H}_{\rm gy}(z) = q\langle(\mathrm{T}^*_z)^{-1}\phi(x)\rangle \tag{G23}$$

(the subscript on T^* indicates on which variables it operates). Then the last term in Eq. (G12) (including the integration over x that defines the action) is

$$-\int d^4x \sum_s \overline{n}_s \int d^4p \int d^4X\, \mathcal{F}_{\rm gy}(z) q\langle(\mathrm{T}^*)^{-1}_z \delta^4(x-X)\rangle\delta\phi(X) \tag{G24a}$$

$$= -\int d^4X\, \delta\phi(X) \sum_s (\overline{n}q)_s \int d^4x \int d^4p\, \mathcal{F}_{\rm gy}(z)\langle(\mathrm{T}^*_z)^{-1}\delta^4(x-X)\rangle \tag{G24b}$$

$$= -\int d^4x\, \delta\phi(x) \sum_s (\overline{n}q)_s \int d^3\boldsymbol{X} \int d^3\boldsymbol{p}\, F_{\rm gy}(t,\boldsymbol{X},\boldsymbol{p})\langle(\mathrm{T}^*_Z)^{-1}\delta^3(\boldsymbol{X}-\boldsymbol{x})\rangle. \tag{G24c}$$

In obtaining the last line, the names of the variables were interchanged, integration over the temporal delta function was done (T does not change time), and the $\delta(W - H_{\rm gy})$ in $\mathcal{F}$ was integrated away (T does not depend on W).

Upon integrating the $\delta\boldsymbol{E}$ term in Eq. (G12) by parts in order to isolate the $\delta\phi$, one has found in summary that $\delta\mathcal{A}_{\rm gy} = \int d^4x\, \delta L_{\rm gy}$, where

$$\begin{aligned}\delta L_{\rm gy} = &-\sum_s \overline{n}_s \int d^4p\, \mathcal{S}\{\mathcal{F}_{\rm gy}, \mathcal{H}_{\rm gy}\} \\ &+ \left(\frac{1}{4\pi}\boldsymbol{\nabla}\cdot\boldsymbol{E} - \sum_s \overline{n}q \int d\boldsymbol{X}\, d\boldsymbol{p}\, F(t,\boldsymbol{X},\boldsymbol{p})\langle(\mathrm{T}^*_Z)^{-1}\delta^3(\boldsymbol{X}-\boldsymbol{x})\rangle\right)\delta\phi \\ &+ \frac{\partial}{\partial x^\mu}\sum_s \overline{n}_s \int d^3\boldsymbol{p}\, v^\mu F\mathcal{S} - \frac{1}{4\pi}\boldsymbol{\nabla}\cdot(\boldsymbol{E}\,\delta\phi).\end{aligned} \tag{G25}$$

Assuming that variations vanish on the boundary, the terms in the last line do not contribute to the Euler–Lagrange equations (they will, however, be crucial

in the Noether method to follow). Since $\delta\phi$ and $\mathcal{S}$ are independent, one finds from the coefficient of $\mathcal{S}$ the result

$$\{\mathcal{F}_{\rm gy}, \mathcal{H}_{\rm gy}\} = 0, \tag{G26}$$

which is the representation of the gyrokinetic equation in extended phase space. From the coefficient of $\delta\phi$, one finds a general representation of the gyrokinetic Poisson equation:

$$\boldsymbol{\nabla}\cdot\boldsymbol{E} = 4\pi\rho, \tag{G27}$$

where

$$\rho = \sum_s (\overline{n}q)_s \int d\boldsymbol{X}\, d\boldsymbol{p}\, F(t,\boldsymbol{X},\boldsymbol{p})\langle(\mathrm{T}_Z^*)^{-1}\delta(\boldsymbol{X}-\boldsymbol{x})\rangle. \tag{G28}$$

We have recovered the representation (29a). See the subsequent discussion of that equation for the formally exact definition of polarization.

H Noether's theorem

Noether's theorem states that for every symmetry of the Lagrangian there is an associated conservation law (for which an explicit expression is provided). Particular examples include conservation of energy (invariance under translation in time) and conservation of toroidal angular momentum (invariance under rotation in the toroidal angle).

H.1 General proof of Noether's theorem

The version of the proof given here is identical to the one given in Wikipedia:Noether's theorem except for changes in notation. Begin with a Lagrangian density $L[\psi(x), \partial_\mu\psi(x); x]$, where $x \equiv x^\mu$ ($\mu = 0,\ldots,3$), such that the action is

$$\mathcal{A} = \int_\Omega d^4x\, L(\psi, \partial_\mu\psi; x). \tag{H1}$$

Here Ω is the appropriate 4D domain in space-time. First, let us recover the conventional Euler–Lagrange equation. The variation of $\mathcal{A}$ under a change of functional form of ψ ($\psi \to \psi + \delta\psi$) is

$$\delta\mathcal{A} = \int_\Omega d^4x \left(\frac{\partial L}{\partial\psi}\delta\psi + \frac{\partial L}{\partial(\partial_\mu\psi)}\delta(\partial_\mu\psi)\right) \tag{H2a}$$

$$= \int_\Omega d^4x \left[\left(\frac{\partial L}{\partial\psi} - \partial_\mu\frac{\partial L}{\partial(\partial_\mu\psi)}\right)\delta\psi + \partial_\mu\left(\frac{\partial L}{\partial(\partial_\mu\psi)}\delta\psi\right)\right]. \tag{H2b}$$

Here one noted that because the variation is taken at fixed x (i.e., it is *Eulerian*), one can commute the δ and the ∂ operations. One assumes that the variations vanish on the boundary; then because $\delta\psi$ is arbitrary one obtains the Euler–Lagrange equation

$$-\partial_\mu\frac{\partial L}{\partial(\partial_\mu\psi)} + \frac{\partial L}{\partial\psi} = 0. \tag{H3}$$

Noether's theorem relates to symmetries of the Lagrangian density. Therefore, consider the infinitesimal transformation (G1) of both the coordinates and the fields. (This variation is arbitrary; it need not vanish on the boundary.) We postulate that the action is unchanged under this transformation, i.e., $\delta\mathcal{A} = 0$, where

$$\delta\mathcal{A} = \int_{\Omega'} d^4\xi\, L[\psi'(\xi), \partial_\mu\psi'(\xi); \xi] - \int_{\Omega} d^4x\, L[\psi(x), \partial_\mu\psi(x); x]. \tag{H4}$$

We noted that the domain Ω has changed to a new domain Ω' under the coordinate transformation. Because ξ is a dummy variable of integration, one can write

$$\delta\mathcal{A} = \int_{\Omega'} d^4x\, L[\psi'(x), \partial_\mu\psi'(x); x] - \int_{\Omega} d^4x\, L[\psi(x), \partial_\mu\psi(x); x]. \tag{H5}$$

Write $\Omega' = \Omega + \delta\Omega$ and integrate separately over each part of Ω'. Because the transformation is infinitesimal, the integral over $\delta\Omega$ can be evaluated with the unperturbed ψ:

$$\begin{aligned}\delta\mathcal{A} = &\int_{\Omega} d^4x\, \{L[\psi'(x), \partial_\mu\psi'(x), x] - L[\psi(x), \partial_\mu\psi(x), x]\} \\ &+ \int_{\delta\Omega} d^4x\, L[\psi(x), \partial_\mu\psi(x), x].\end{aligned} \tag{H6}$$

The volume $\delta\Omega$ comprises the boundary $\partial\Omega$ of the original domain and the displacement δx^μ. One has

$$\int_{\delta\Omega} d^4x\, L[\psi(x), \partial_\mu\psi(x); x] = \int_{\partial\Omega} d^3x\, \delta x^\mu L(\psi, \partial_\mu\psi; x); \tag{H7}$$

it then follows by Stokes' theorem that

$$\int_{\delta\Omega} d^4x\, L[\psi(x), \partial_\mu\psi(x); x] = \int_{\Omega} d^4x\, \partial_\mu[\delta x^\mu L(\psi, \partial_\nu\psi; x)]. \tag{H8}$$

Thus

$$\delta\mathcal{A} = \int_{\Omega} d^4x\, \{L(\psi', \partial_\mu\psi'; x) - L(\psi, \partial_\mu\psi, x) + \partial_\mu[\delta x^\mu L(\psi, \partial_\nu\psi; x)\} \tag{H9a}$$

$$= \int_{\Omega} d^4x \left(\frac{\partial L}{\partial\psi}\delta\psi + \frac{\partial L}{\partial(\partial_\mu\psi)}\delta(\partial_\mu\psi) + \partial_\mu[\delta x^\mu L(\psi, \partial_\nu\psi; x)]\right). \tag{H9b}$$

Because the δ variations are Eulerian, the δ and ∂ operations commute. One may replace the $\partial L/\partial\psi$ by using the Euler–Lagrange equation. Then the first two terms combine to become a perfect derivative, and one finds

$$\delta\mathcal{A} = \int_{\Omega} d^4x\, \partial_\mu \left(\frac{\partial L}{\partial(\partial_\mu\psi)}\delta\psi + L(\psi, \partial_\nu\psi; x)\delta x^\mu\right). \tag{H10}$$

Since $\delta\mathcal{A} = 0$ and Ω is arbitrary, the integrand must vanish and one finds a conserved 4-current:

$$0 = \partial_\mu \left(\frac{\partial L}{\partial(\partial_\mu \psi)} \delta\psi + L(\psi, \partial_\nu \psi; x) \delta x^\mu \right). \tag{H11}$$

We now reformulate Eq. (H11) in terms of Lagrangian variations. Using the result (G6) one finds that Eq. (H11) becomes, after a change of sign,

$$0 = \partial_\mu j^\mu, \tag{H12}$$

where the conserved Noether current is

$$j^\mu \doteq \frac{\partial L}{\partial(\partial_\mu \psi)} L_X \psi - L X^\mu - \frac{\partial L}{\partial(\partial_\mu \psi)} \Psi_L. \tag{H13}$$

In the simplest applications to translational or rotational symmetry, the Lagrangian field variation vanishes, i.e.,

$$\psi(x) = \psi'(\xi), \tag{H14}$$

or $\Psi_L = 0$.

H.2 Constrained variations and Noether's theorem

The essence of Noether's method is, given a known symmetry transformation, to find an explicit expression for the Eulerian variation of L at fixed x and equate that to the form that follows from Stokes' theorem. That is, Eq. (H9a) for $\delta\mathcal{A} = 0$ is

$$L(\psi', \partial_\mu \psi'; x) - L(\psi, \partial_\mu \psi, x) \equiv \delta L = -\partial_\mu [\delta x^\mu L(\psi, \partial_\nu \psi; x)]. \tag{H15}$$

Here the variations do not vanish on the boundary but are those associated with the particular symmetry transformation. We have already worked out the left-hand side of Eq. (H15) in the calculations leading to Eq. (G25). When the Euler–Lagrange equations are asserted, Eq. (G25) reduces to its last line. Also, $\mathcal{F}\mathcal{H}$ vanishes, so $L = L_M$. Therefore, Eq. (H15) reduces to

$$\frac{\partial}{\partial x^\mu} \sum_s \overline{n}_s \int d^3\boldsymbol{p}\, v^\mu F \mathcal{S} - \frac{1}{4\pi} \boldsymbol{\nabla} \cdot (\boldsymbol{E}\, \delta\phi) = -\partial_\mu (\delta x^\mu L_M). \tag{H16}$$

What remains is to choose $\mathcal{S}$ to generate the desired δx^μ according to Eq. (G14). As Brizard has observed, the generator of an arbitrary spatial transformation is

$$\mathcal{S} = \boldsymbol{p}_{\rm gy}^{\rm can} \cdot \delta\boldsymbol{x}, \tag{H17}$$

where $\boldsymbol{p}_{\rm gy}$ is the canonical momentum of the gyrocenter. From Eq. (G6) one has

$$\delta\phi = -\delta x^\mu \partial_\mu \phi = -\delta\boldsymbol{x} \cdot \boldsymbol{\nabla}\phi, \tag{H18}$$

since the Lagrangian variations vanish in Brizard's method. Finally, the specific $\delta\boldsymbol{x}$ associated with an infinitesimal rotation by an amount $\delta\varphi$ in the toroidal direction is

$$\delta\boldsymbol{x} = \frac{\partial \boldsymbol{x}}{\partial \varphi}\delta\varphi = (\widehat{\boldsymbol{z}} \times \boldsymbol{x})\delta\varphi. \tag{H19}$$

The Noether equation (H16) is obviously already in local conservative form, i.e., it can be written as

$$\frac{\partial P_\varphi^{\rm can}}{\partial t} + \boldsymbol{\nabla}\cdot\Pi_\varphi^{\rm can} = 0, \tag{H20}$$

where

$$P_\varphi^{\rm can} \doteq \sum_s \overline{n}_s \int d\boldsymbol{p}\, \boldsymbol{p}_{{\rm gy},\varphi}^{\rm can} F \tag{H21}$$

and one has a related definition for the canonical momentum flux. (Note that for arbitrary vector $\boldsymbol{V}$, one has $\boldsymbol{V}\cdot\partial_\varphi\boldsymbol{x} = V_\varphi$, the covariant component.) It is important that this equation evolves the *canonical* momentum. Equation (H20) must be processed further in order to find an equation for the plasma momentum. That is explained by Brizard & Tronko (2011).

I Direct derivation of the gyrokinetic momentum conservation law, including all magnetic inhomogeneity effects

The methods of Scott & Smirnov (2010) and Brizard & Tronko (2011) lead one very efficiently to a local momentum conservation law, in general geometry, that evolves the sum of the toroidal projections of the parallel momentum and perpendicular polarization of the gyrocenter; certain cancellations happen almost magically. It is instructive to derive the result in an alternate way, by considering the parallel and perpendicular evolutions separately. Neither obeys a local conservation law by itself, but important cancellations occur when they are summed, as we demonstrate explicitly. That leads one (Sec. I.1) to an equation in which just one term, involving the derivative of the gyrokinetic Hamiltonian with respect to toroidal angle, is not obviously conservative. Further manipulations reviewed in Sec. I.2 lead one to the final conservative result provided that both the gyrokinetic equation and the gyrokinetic Poisson equation are derived consistently from the same Lagrangian.

I.1 Summing the parallel and perpendicular evolutions

We will write the gyrocenter distribution function as $F(\boldsymbol{X}, \mu, U, t)$, where μ is conserved. The gyrocenter equations of motion in general geometry and for

arbitrary H are [in the symplectic representation; see Eqs. (E10)]

$$\frac{d\boldsymbol{X}}{dt} = B_*^{-1}\left(\boldsymbol{B}_* U + \frac{c}{q}\widehat{\boldsymbol{b}}\times\boldsymbol{\nabla}H\right), \tag{I1a}$$

$$m\frac{dU}{dt} = -\frac{1}{B_*}\boldsymbol{B}_*\cdot\boldsymbol{\nabla}H. \tag{I1b}$$

Here

$$\boldsymbol{B}_* \doteq \boldsymbol{B} + \underbrace{\frac{mc}{q}U\boldsymbol{\nabla}\times\widehat{\boldsymbol{b}}}_{\delta\boldsymbol{B}_*}, \tag{I2a}$$

$$B_* \doteq \widehat{\boldsymbol{b}}\cdot\boldsymbol{B}_* = B + \frac{mc}{q}U\widehat{\boldsymbol{b}}\cdot\boldsymbol{\nabla}\times\widehat{\boldsymbol{b}}. \tag{I2b}$$

Velocity integrals introduce B_* as the Jacobian of the transformation from particle to gyrocenter coordinates:

$$\int d\boldsymbol{v}\, f \to \int d\mu\, dU\, B_*\, F \equiv \int B_*\, F. \tag{I3}$$

As we know from the work of Scott & Smirnov (2010) and the approximate calculation in Sec. 4.2, two terms contribute to the toroidal momentum conservation: $c^{-1}\langle\mathcal{P}^\psi\rangle$, and $\mathcal{P}_{\varphi\parallel} \doteq \sum_s (m\overline{n})_s \int B_* F U b_\varphi$. We will split the time derivatives of each of those into various pieces, then show how most of those cancel when the two terms are added together. First, one has

$$\frac{1}{c}\frac{\partial\langle\mathcal{P}^\psi\rangle}{\partial t} = -\frac{1}{c}\sum_s (\overline{n}q)_s\left\langle\int B_* F V^\psi\right\rangle \tag{I4a}$$

$$= -\frac{1}{c}\sum_s (\overline{n}q)_s\left\langle\int F\left(\boldsymbol{B}_*^\psi U + \frac{c}{q}(\widehat{\boldsymbol{b}}\times\boldsymbol{\nabla}H)^\psi\right)\right\rangle \tag{I4b}$$

$$\equiv \dot{\mathcal{P}}_\parallel + \delta\dot{\mathcal{P}}_\parallel + \dot{\mathcal{P}}_\perp. \tag{I4c}$$

Here $\dot{\mathcal{P}}_\parallel \sim \boldsymbol{B}^\psi U$, $\delta\dot{\mathcal{P}}_\parallel \sim \delta\boldsymbol{B}_*^\psi U$, and $\dot{\mathcal{P}}_\perp \sim (\widehat{\boldsymbol{b}}\times\boldsymbol{\nabla}H)^\psi$. The $\dot{\mathcal{P}}_\parallel$ term (parallel streaming along the field line) vanishes because the ψ component is required and $\boldsymbol{B}\cdot\boldsymbol{\nabla}\psi = 0$. For $\delta\dot{\mathcal{P}}_\parallel$, we use Eqs. (I2a) and (A22b) to find

$$(\boldsymbol{\nabla}\times\widehat{\boldsymbol{b}})^\psi = \frac{1}{J}\left(\frac{\partial b_\varphi}{\partial\theta} - \frac{\partial b_\theta}{\partial\phi}\right) = \frac{1}{J}\frac{\partial b_\varphi}{\partial\theta}, \tag{I5}$$

where axisymmetry was used. Thus

$$\delta\dot{\mathcal{P}}_\parallel = -\sum_s (\overline{n}m)_s\left\langle\int F U^2 J^{-1}\partial_\theta b_\varphi\right\rangle. \tag{I6}$$

Also, from Eq. (A18b), one has $(\widehat{\boldsymbol{b}}\times\boldsymbol{\nabla}H)^\psi = J^{-1}(b_\theta\partial_\varphi H - b_\varphi\partial_\theta H)$, so

$$\dot{\mathcal{P}}_\perp = -\sum_s \overline{n}_s\left\langle\int F J^{-1}(b_\theta\partial_\varphi H - b_\varphi\partial_\theta H)\right\rangle. \tag{I7}$$

Now let us evaluate $\langle \dot{\mathcal{P}}_{\varphi\parallel} \rangle$. We use the conservative form of the GKE,

$$\frac{\partial (B_* F)}{\partial t} + \boldsymbol{\nabla} \cdot (B_* F \boldsymbol{V}) + \frac{\partial}{\partial U}(B_* F \dot{U}) = 0, \tag{I8}$$

and find

$$\frac{\partial \langle \mathcal{P}_{\varphi\parallel} \rangle}{\partial t} = -\sum_s (\overline{n} m)_s \left\langle \int U b_\varphi \left(\boldsymbol{\nabla} \cdot (B_* F \boldsymbol{V}) + \frac{\partial}{\partial U}(B_* F \dot{U}) \right) \right\rangle. \tag{I9}$$

For the first term, use $\boldsymbol{\nabla} \cdot (s\boldsymbol{A}) = (\boldsymbol{\nabla} s) \cdot \boldsymbol{A} + s \boldsymbol{\nabla} \cdot \boldsymbol{A}$ to pull the b_φ inside the divergence (at the price of an extra term). For the second term, integrate by parts in U. Then

$$\langle \dot{\mathcal{P}}_{\varphi\parallel} \rangle = -\sum_s (\overline{n} m)_s \left\langle \boldsymbol{\nabla} \cdot \int B_* F \boldsymbol{V} U b_\varphi \right\rangle + \underbrace{\sum_s (\overline{n} m)_s \left\langle \int B_* F U \boldsymbol{V} \cdot \boldsymbol{\nabla} b_\varphi \right\rangle}_{\dot{\chi}_\parallel + \delta \dot{\chi}_\parallel + \dot{\chi}_\perp}$$
$$+ \underbrace{\sum_s (\overline{n} m)_s \left\langle \int B_* F \dot{U} b_\varphi \right\rangle}_{\dot{\mathcal{Q}}_\parallel + \delta \dot{\mathcal{Q}}_\parallel}. \tag{I10}$$

(The pieces $\dot{\chi}_\parallel$, $\delta\dot{\chi}_\parallel$, and $\dot{\chi}_\perp$ arise by writing $\boldsymbol{V} = U\widehat{\boldsymbol{b}} + \boldsymbol{V}_\perp$; they are defined explicitly below.) The first term can be written as $-V'^{-1}\partial_\psi (V' \Gamma^\psi_{\varphi\parallel})$, which is one of the terms that survives in the conservation law. Each of the second and third terms have various pieces, as indicated.

We have

$$\dot{\chi}_\parallel \doteq \sum_s (\overline{n} m)_s \left\langle \int F U^2 \boldsymbol{B} \cdot \boldsymbol{\nabla} b_\varphi \right\rangle. \tag{I11}$$

Now

$$\boldsymbol{B} \cdot \boldsymbol{\nabla} b_\varphi = B^i \partial_i b_\varphi = B^\theta \partial_\theta b_\varphi = J^{-1} \partial_\theta b_\varphi. \tag{I12}$$

Thus

$$\dot{\chi}_\parallel = \sum_s (\overline{n} m)_s \left\langle \int F U^2 J^{-1} \partial_\theta b_\varphi \right\rangle = -\delta \dot{\mathcal{P}}_\parallel. \tag{I13}$$

This demonstrates the first cancellation: $\dot{\chi}_\parallel + \delta \dot{\mathcal{P}}_\parallel = 0$.

Now consider

$$\delta \dot{\chi}_\parallel \doteq \sum_s (\overline{n} m)_s \left\langle \int F U^2 \delta \boldsymbol{B}_* \cdot \boldsymbol{\nabla} b_\varphi \right\rangle. \tag{I14}$$

We have $\delta \boldsymbol{B}_* \propto U \boldsymbol{\nabla} \times \widehat{\boldsymbol{b}}$. The dot product with the curl can be conveniently represented [Eq. (A23)] as the determinant (specialized for axisymmetry)

$$(\boldsymbol{\nabla} \times \widehat{\boldsymbol{b}}) \cdot \boldsymbol{\nabla} b_\varphi = J^{-1} \begin{vmatrix} \partial_\psi b_\varphi & \partial_\theta b_\varphi & 0 \\ \partial_\psi & \partial_\theta & 0 \\ b_\psi & b_\theta & b_\varphi \end{vmatrix} = 0. \tag{I15}$$

It is fortunate that this term vanishes by itself because it is the only term that involves a factor of U^3.

Next,

$$\dot{\mathcal{Q}}_\parallel \doteq -\sum_s \overline{n}_s \left\langle \int F(\boldsymbol{B}\cdot\boldsymbol{\nabla}H) b_\varphi \right\rangle . \tag{I16}$$

We have $\boldsymbol{B}\cdot\boldsymbol{\nabla}H = B^\theta \partial_\theta H + B^\varphi \partial_\varphi H$ (and $B^\theta = J^{-1}$), so

$$\dot{\mathcal{Q}}_\parallel = -\sum_s \overline{n}_s \left\langle \int F(J^{-1}\partial_\theta H + B^\varphi \partial_\varphi H) b_\varphi \right\rangle . \tag{I17}$$

The first term of this cancels against the second term of $\dot{\mathcal{P}}_\perp$ [Eq. (I7)]. Thus

$$\dot{\mathcal{P}}_\perp + \dot{\mathcal{Q}}_\parallel = -\sum_s \overline{n}_s \left\langle \int F(J^{-1} b_\theta + B^\varphi b_\varphi)\partial_\varphi H \right\rangle . \tag{I18}$$

Now one has $b_\theta = B^\theta/B = JB_{\rm pol}^2/B$ [Eq. (B3)] and $B^\varphi b_\varphi = B^\varphi B_\varphi/B = B_{\rm tor}^2/B$. Thus $J^{-1}b_\theta + B^\varphi b_\varphi = (B_{\rm pol}^2 + B_{\rm tor}^2)/B = B$ and

$$\dot{\mathcal{P}}_\perp + \dot{\mathcal{Q}}_\parallel = -\sum_s \overline{n}_s \left\langle \int BF\partial_\varphi H \right\rangle . \tag{I19}$$

This is close to the term in $\partial_\varphi H$ that survives in the momentum conservation law, except that we are expecting the full Jacobian B_* instead of B. Fortunately there are more terms.

We have

$$\dot{\chi}_\perp \doteq \sum_s \overline{n}_s \left(\frac{mc}{q}\right)_s \left\langle \int FU\widehat{\boldsymbol{b}}\times\boldsymbol{\nabla}H\cdot\boldsymbol{\nabla}b_\varphi \right\rangle . \tag{I20}$$

Now

$$\widehat{\boldsymbol{b}}\times\boldsymbol{\nabla}H\cdot\boldsymbol{\nabla}b_\varphi = J^{-1}\begin{vmatrix} \partial_\psi b_\varphi & \partial_\theta b_\varphi & 0 \\ b_\psi & b_\theta & b_\varphi \\ \partial_\psi H & \partial_\theta H & \partial_\varphi H \end{vmatrix} \tag{I21a}$$

$$= J^{-1}[b_\varphi(\partial_\theta b_\varphi)(\partial_\psi H) - b_\varphi(\partial_\psi b_\varphi)(\partial_\theta H) + (b_\theta\partial_\psi b_\varphi - b_\psi\partial_\theta b_\varphi)\partial_\varphi H]. \tag{I21b}$$

Also,

$$\delta\dot{\mathcal{Q}}_\parallel \doteq -\sum_s \overline{n}_s \left(\frac{mc}{q}\right)_s \left\langle \int FU(\boldsymbol{\nabla}\times\widehat{\boldsymbol{b}}\cdot\boldsymbol{\nabla}H) b_\varphi \right\rangle . \tag{I22}$$

Now

$$\boldsymbol{\nabla}\times\widehat{\boldsymbol{b}}\cdot\boldsymbol{\nabla}H = J^{-1}\begin{vmatrix} \partial_\psi H & \partial_\theta H & \partial_\varphi H \\ \partial_\psi & \partial_\theta & 0 \\ b_\psi & b_\theta & b_\varphi \end{vmatrix} \tag{I23a}$$

$$= J^{-1}[(\partial_\theta b_\varphi)(\partial_\psi H) - (\partial_\psi b_\varphi)(\partial_\theta H) + (\partial_\psi b_\theta - \partial_\theta b_\psi)\partial_\varphi H]. \tag{I23b}$$

Upon adding $\dot{\chi}_\perp$ and $\delta\dot{\mathcal{Q}}_\parallel$ and using Eq. (75), one finds that the terms in $\partial_\psi H$ and $\partial_\theta H$ cancel. For the remaining terms, recognize that

$$\widehat{\boldsymbol{b}} \cdot \boldsymbol{\nabla} \times \widehat{\boldsymbol{b}} = J^{-1} \begin{vmatrix} b_\psi & b_\theta & b_\varphi \\ \partial_\psi & \partial_\theta & 0 \\ b_\psi & b_\theta & b_\varphi \end{vmatrix}. \tag{I24}$$

The sum of the remaining terms is proportional to the negative of this; upon using Eq. (I2b), one finds

$$\dot{\chi}_\perp + \delta\dot{\mathcal{Q}}_\parallel = -\sum_s \overline{n}_s \left\langle \int F \delta B_* \partial_\varphi H \right\rangle. \tag{I25}$$

This is just what one needs to get the full Jacobian:

$$\dot{\mathcal{P}}_\perp + \dot{\mathcal{Q}}_\parallel + \dot{\chi}_\perp + \delta\dot{\mathcal{Q}}_\parallel = -\sum_s \overline{n}_s \left\langle \int B_* F \partial_\varphi H \right\rangle. \tag{I26}$$

Thus, in summary, one has

$$\frac{1}{c}\frac{\partial \langle \mathcal{P}^\psi \rangle}{\partial t} = \dot{\mathcal{P}}_\parallel + \delta\dot{\mathcal{P}}_\parallel + \dot{\mathcal{P}}_\perp, \tag{I27a}$$

$$\frac{\partial \langle \mathcal{P}_{\varphi\parallel} \rangle}{\partial t} = -\frac{1}{V'}\frac{\partial}{\partial \psi} V' \Gamma^\psi_{\varphi\parallel} + (\dot{\chi}_\parallel + \delta\dot{\chi}_\parallel + \dot{\chi}_\perp) + (\dot{\mathcal{Q}}_\parallel + \delta\dot{\mathcal{Q}}_\parallel), \tag{I27b}$$

and

$$\begin{aligned} \frac{\partial (c^{-1}\langle \mathcal{P}^\psi \rangle + \langle \mathcal{P}_{\varphi\parallel} \rangle)}{\partial t} = &-\frac{1}{V'}\frac{\partial}{\partial \psi} V' \Gamma^\psi_{\varphi\parallel} + \underbrace{\dot{\mathcal{P}}_\parallel}_{0} + \underbrace{(\delta\dot{\mathcal{P}}_\parallel + \dot{\chi}_\parallel)}_{0} + \underbrace{\delta\dot{\chi}_\parallel}_{0} \\ &+ \underbrace{[(\dot{\mathcal{P}}_\perp + \dot{\mathcal{Q}}_\parallel) + (\dot{\chi}_\perp + \delta\dot{\mathcal{Q}}_\parallel)]}_{-\sum_s \overline{n}_s \left\langle \int B_* F \, \partial_\varphi H \right\rangle}. \end{aligned} \tag{I28}$$

I.2 Writing the pure Reynolds stress in conservative form

At this point we have recovered the most important intermediate equation in the calculations of Scott & Smirnov. It remains to be shown that the last term on the right-hand side of Eq. (I28) can be written as a divergence. Scott & Smirnov demonstrated that by relating the ordinary derivative $\partial_\varphi H$ to the functional derivative of H with respect to potential, then recognizing that the Euler–Lagrange equation for potential eliminated all nonconservative terms. Brizard & Tronko (2011) used a different (but equivalent) technique. They observed that the gyrocenter Hamiltonian can be expressed in terms of the pushforward transformation: $H^G = H_0 + q(\mathrm{T}^*)^{-1}\phi$. Now

$$(\mathrm{T}^*)^{-1}\phi(\boldsymbol{X}) = \phi(\boldsymbol{X} + \boldsymbol{\rho}_\epsilon), \tag{I29}$$

where $\boldsymbol{\rho}_\epsilon$ is given as a series whose lowest-order term coincides with the usual gyroradius vector. An expansion in small $\boldsymbol{\rho}_\epsilon$ then gives[75] (Brizard & Tronko, 2011, Eq. (117))

$$\sum \overline{n} \int F \left(\frac{\partial H^G}{\partial \varphi} \right) = \rho^G \partial_\varphi \phi + \left(\sum \overline{n} q \int F \langle \boldsymbol{\rho}_\epsilon \rangle \right) \cdot \boldsymbol{\nabla} \partial_\varphi \phi + \frac{1}{2} \left(\sum \overline{n} q \int F \langle \boldsymbol{\rho}_\epsilon \, \boldsymbol{\rho}_\epsilon \rangle \right) : \boldsymbol{\nabla} \boldsymbol{\nabla} \partial_\varphi \phi + \cdots . \tag{I30}$$

Now the exact polarization is given by Eq. (32), also an expansion in $\boldsymbol{\rho}_\epsilon$. For the terms of $O(\boldsymbol{\rho}_\epsilon)$ or smaller, one can move all gradients to the left at the price of correction terms and establish that the sole nonconservative term involving $\boldsymbol{\rho}_\epsilon$ is *exactly* the divergence of the polarization. Ultimately, one finds

$$\sum \overline{n} \int F \left(\frac{\partial H^G}{\partial \varphi} \right) = (\rho^G - \boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{P}}) \partial_\varphi \phi + \boldsymbol{\nabla} \cdot \left[\boldsymbol{\mathcal{P}} \partial_\varphi \phi + \frac{1}{2} \left(\sum \overline{n} q \int F \langle \boldsymbol{\rho}_\epsilon \, \boldsymbol{\rho}_\epsilon \rangle \right) \cdot \boldsymbol{\nabla} \partial_\varphi \phi + \cdots \right] . \tag{I31}$$

If the truncation of the Hamiltonian in the kinetic equation is at $O(\epsilon^n)$ and terms through $O(\epsilon^m)$ are retained in Poisson's equation, then the nonconservative term in Eq. (I31) is $(\boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{P}}^{[m]} - \boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{P}}^{[n-1]}) \partial_\varphi \phi$. When both the kinetic equation and Poisson's equation are derived variationally from the same Lagrangian with a Hamiltonian truncated at $O(\epsilon^n)$, then $m = n - 1$ and the (approximate) quasineutrality condition annihilates the nonconservative term; one arrives at a properly local conservation law. Otherwise, there is a mismatch and a nonconservative term remains. For example, if one retains second-order drifts in the kinetic equation ($n = 2$) and also second-order terms in Poisson's equation ($m = 2$), a conservative residual $\boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{P}}^{[2]} \partial_\varphi \phi = O(\epsilon^3)$ remains, as demonstrated explicitly in the slab calculation of Parra & Catto (2010a).

J Detailed analysis of the gyrokinetic conservation law for toroidal angular momentum

From Appendix I as well as the Noether methods employed by Scott & Smirnov (2010) and later by Brizard & Tronko (2011), one has the exact result

$$\frac{\partial (\langle \mathcal{P}_{\varphi \parallel} \rangle + c^{-1} \langle \mathcal{P}^\psi \rangle)}{\partial t} + \frac{1}{V'} \frac{\partial}{\partial \psi} V' (\Gamma^\psi_{\varphi \parallel} + \Gamma^\psi_{\varphi \perp}) = 0, \tag{J1}$$

[75] In obtaining the final result, the ∂_φ has been commuted past the pushforward. This can be justified (A. Brizard, private communication, 2013).

where the parallel–perpendicular Reynolds stress is

$$\Gamma^{\psi}_{\varphi\parallel} \doteq \sum_s (\overline{n}m)_s \left\langle \int_{\boldsymbol{P}} F V^{\psi} v_{\parallel} b_{\varphi} \right\rangle \tag{J2}$$

and the pure Reynolds stress is

$$\Gamma^{\psi}_{\varphi\perp} \equiv \Gamma \doteq \sum_s \overline{n}_s \left\langle \int_{\boldsymbol{P}} F \left(\frac{\partial H}{\partial \boldsymbol{E}} \right)^{\psi} \frac{\partial \phi}{\partial \varphi} \right\rangle + \cdots . \tag{J3}$$

From the discussion in Sec. 2.2, we know that the particle momentum fluxes must be $O(\epsilon^3)$ for gyro-Bohm scaling in the low-flow ordering. It is clear that that ordering applies as well to the gyrocenter fluxes. Now the integrands of both $\Gamma^{\psi}_{\varphi\parallel}$ and $\Gamma^{\psi}_{\varphi\perp}$ derive from the gyrocenter Hamiltonian H and the gyrocenter PDF F, so the question is how accurately those quantities must be calculated.

J.1 The parallel–perpendicular Reynolds stresses

We have

$$\Gamma^{\psi}_{\varphi\parallel} \doteq \sum_s (\overline{n}m)_s \left\langle \int_{\boldsymbol{P}} B_* F V^{\psi} U b_{\varphi} \right\rangle , \tag{J4}$$

where we have explicitly inserted the Jacobian B_*. Here the gyrocenter drift is

$$\boldsymbol{V} = B_*^{-1} \left(\boldsymbol{B}_* U + \frac{c}{q} \widehat{\boldsymbol{b}} \times \boldsymbol{\nabla} H \right) , \tag{J5}$$

where

$$\boldsymbol{B}_* \doteq \boldsymbol{B} + \frac{mc}{q} U \boldsymbol{\nabla} \times \widehat{\boldsymbol{b}} \equiv \boldsymbol{B} + \delta \boldsymbol{B}_*, \tag{J6a}$$

$$B_* \doteq \widehat{\boldsymbol{b}} \cdot \boldsymbol{B}_* = B + \frac{mc}{q} U \widehat{\boldsymbol{b}} \cdot \boldsymbol{\nabla} \times \widehat{\boldsymbol{b}}. \tag{J6b}$$

Since $B^{\psi} = 0$, one has

$$\Gamma^{\psi}_{\varphi\parallel} = \sum_s (\overline{n}m)_s \left\langle \int_{\boldsymbol{P}} F U b_{\varphi} [\delta \boldsymbol{B}_* U + (c/q) \widehat{\boldsymbol{b}} \times \boldsymbol{\nabla} H]^{\psi} \right\rangle . \tag{J7}$$

Now $(\boldsymbol{\nabla} \times \widehat{\boldsymbol{b}})^{\psi}$ is the coefficient of $\boldsymbol{e}_{\psi}$ in Eq. (A22b) (with $\partial_{\varphi} = 0$ for axisymmetric geometry):

$$(\boldsymbol{\nabla} \times \widehat{\boldsymbol{b}})^{\psi} = J^{-1} \partial_{\theta} b_{\varphi}. \tag{J8}$$

Also, $(\widehat{\boldsymbol{b}} \times \boldsymbol{\nabla} H)^{\psi}$ is the coefficient of $\boldsymbol{e}_{\psi}$ in Eq. (A19):

$$(\widehat{\boldsymbol{b}} \times \boldsymbol{\nabla} H)^{\psi} = J^{-1} (b_{\theta} \partial_{\varphi} H - b_{\varphi} \partial_{\theta} H). \tag{J9}$$

Thus

$$\Gamma^{\psi}_{\varphi\parallel} = \sum_s (\overline{n}m)_s \left(\frac{c}{q} \right) \left\langle J^{-1} \int_{\boldsymbol{P}} F U b_{\varphi} [m U^2 \partial_{\theta} b_{\varphi} + (b_{\theta} \partial_{\varphi} H - b_{\varphi} \partial_{\theta} H)] \right\rangle . \tag{J10}$$

We write F as a zeroth-order part $F_0(\psi)$ plus a perturbation δF, where δF is taken to be $O(\epsilon)$. For the $F_0(\psi)$ contribution, the $\partial_\varphi H$ term vanishes under the flux-surface average (the J^{-1} nicely cancels the J in the definition of the average, and the b_i's are independent of φ). The first term, $\propto b_\varphi \partial_\theta b_\varphi = \frac{1}{2}\partial_\theta b_\varphi^2$, also vanishes for F_0. For F_0, one is thus left with a term $\propto -F_0 U b_\varphi^2 \partial_\theta H \to F_0 U \partial_\theta(b_\varphi^2) H$ after integration by parts under the flux-surface average. The θ derivative is $O(\epsilon)$ since it acts on background geometry with macroscopic variation. Therefore, one needs to use no smaller than H_2 in order to obtain a result of $O(\epsilon^3)$. Note that the contributions from H_0 and H_1 are nominally larger. One has $H_0 = \frac{1}{2}mU^2 + \mu B$; since B depends on θ, there is a contribution from H_0. The integral $\int FU$ either vanishes or is taken to be of first order. Thus the contribution from H_0 is nominally of second order. The contribution from H_1 is second order as well.

Since $\delta F = O(\epsilon)$, it is easy to see that all contributions from δF to $\Gamma^\psi_{\varphi\|}$ are nominally of second order. Furthermore, one needs to use no smaller than H_2 in order to recover $O(\epsilon^3)$. Note that the second-order terms need to become smaller by one order in order to recover gyro-Bohm scaling. This can be argued on the basis of symmetry arguments.

J.2 The pure Reynolds stresses

Now we discuss

$$\Gamma^\psi_{\varphi\perp} \doteq \sum_s \overline{n}_s \left\langle \int_{\boldsymbol{P}} F \left(\frac{\partial H}{\partial \boldsymbol{E}}\right)^\psi \frac{\partial \phi}{\partial \varphi} \right\rangle + \cdots . \tag{J11}$$

This behaves somewhat differently than $\Gamma^\psi_{\varphi\|}$ because evaluating it with $F_0(\psi)$ does not change the nominal order of the integrand.

J.2.1 The nominal size of the pure Reynolds stresses

Focus on the explicit (dipolar) terms displayed in Eq. (J11). There is no contribution of H_1 to the derivative with respect to $\boldsymbol{E}$ because $H_1 = q\phi + O(\nabla_\perp^2 \phi)$ and derivatives with respect to $\nabla_\perp^2 \phi$ are taken care by another unwritten term in Eq. (J11). Assume that the explicit $\boldsymbol{E}$ dependence of H_2 is written as

$$H_2 = \frac{1}{2}\mathsf{N}^{(0)}_{\boldsymbol{EE}} : \boldsymbol{EE} + \boldsymbol{N}^{(1)}_{\boldsymbol{BE}} \cdot \boldsymbol{E} + \cdots . \tag{J12}$$

Here $N^{(1)}_{\boldsymbol{BE}}$ involves macroscopic gradients of the magnetic field and is $O(\epsilon)$. The unwritten terms in Eq. (J12) would be of second order but do not involve $\boldsymbol{E}$ explicitly; for example, they could be quadratic in gradients of the background magnetic field, involve FLR corrections like $(\nabla_\perp^2\phi)^2$, *etc.* One then has

$$\frac{\partial H_2}{\partial \boldsymbol{E}} = \mathsf{N}^{(0)}_{\boldsymbol{EE}} \cdot \boldsymbol{E} + \boldsymbol{N}^{(1)}_{\boldsymbol{BE}} = O(\epsilon). \tag{J13}$$

Similar results hold at higher order. Eq. (J13) demonstrates the obvious fact that differentiation with respect to $\boldsymbol{E}$ *reduces the order by one.*

We shall denote the contribution to $\Gamma^{\psi}_{\varphi\perp} \equiv \Pi$ from H_n as $\Pi^{[n]}$ (the bracketed superscript implies that the actual order in ϵ is yet to be determined because flux-surface or statistical averaging can change the order). Thus one has

$$\Pi^{[2]} \sim \int_{\boldsymbol{P}} \langle F E^{\psi} \partial_{\varphi} \phi \rangle, \tag{J14}$$

where here E^{ψ} is a proxy for the ψ component of Eq. (J13). One may apply a cumulant expansion[76] to the average:

$$\begin{aligned}\langle F E^{\psi} \partial_{\varphi} \phi \rangle = \langle\langle F \rangle\rangle \langle\langle E^{\psi} \rangle\rangle \underbrace{\langle\langle \partial_{\varphi} \phi \rangle\rangle}_{0} \\ + \langle\langle F \rangle\rangle \langle\langle E^{\psi} \partial_{\varphi} \phi \rangle\rangle + \langle\langle E^{\psi} \rangle\rangle \langle\langle F \partial_{\varphi} \phi \rangle\rangle + \underbrace{\langle\langle \partial_{\varphi} \phi \rangle\rangle}_{0} \langle\langle F E^{\psi} \rangle\rangle \\ + \langle\langle F E^{\psi} \partial_{\varphi} \phi \rangle\rangle.\end{aligned} \tag{J15}$$

The terms in $\langle\langle \partial_{\varphi} \phi \rangle\rangle = \langle \partial_{\varphi} \phi \rangle$ vanish under the flux-surface average (or by statistical symmetry of the turbulence). One has $\langle\langle F \rangle\rangle = \langle F \rangle = F_0(\psi)$. Thus

$$\langle F E^{\psi} \partial_{\varphi} \phi \rangle = \underbrace{F_0(\psi)}_{O(1)} \underbrace{\langle \delta E^{\psi} \partial_{\varphi} \delta\phi \rangle}_{O(\epsilon^2)} + \underbrace{\langle E^{\psi} \rangle}_{O(\epsilon)} \underbrace{\langle \delta F \, \partial_{\varphi} \delta\phi \rangle}_{O(\epsilon^2)} + \underbrace{\langle \delta F \, \delta E^{\psi} \partial_{\varphi} \delta\phi \rangle}_{O(\epsilon^3)}. \tag{J16}$$

Here we note that the $\delta\phi$'s must be at the short scales, where $e\delta\phi/T_e = O(\epsilon)$], and that $\delta F = O(\epsilon)$. The last two terms are already $O(\epsilon^3)$, and these stem from H_2. But the first term is $O(\epsilon^2)$. There are now two possibilities:

1. The Reynolds stress $\mathcal{R}^{[2]} \doteq \langle \delta E^{\psi} \partial_{\varphi} \delta\phi \rangle$ is in fact of second order even in the low-flow ordering. In that case the assumption of gyro-Bohm scaling is false. All physics follows from no more than H_2; H_3 is not needed.

2. $\mathcal{R}^{[2]}$ *dominantly vanishes* due to some underlying symmetry. Within a standard power-law expansion to measure the size of the symmetry breaking, this implies that $\mathcal{R} = O(\epsilon^3)$, which is consistent with the low-flow gyro-Bohm scaling.

We assume that possibility 2 is realized; we expand on this assumption in Appendix K. It is then clear that a third-order term $\mathcal{R}^{[3]} = O(\langle \delta E^2 \partial_{\varphi} \delta\phi \rangle)$ will arise from H_3 in the same way as $\mathcal{R}^{(2)}$ arose from H_2. There is no reason to expect

[76] Formally, the Fourier transform of a PDF is a moment generating function. The logarithm of that transform is a cumulant generating function. Properties of the logarithm lead to the rule relating nth-order moments and cumulants (Kubo, 1962), which is that the nth-order moment should be partitioned into all possible products of cumulants of nth order or lower such that the sum of the cumulant orders in any one term equals n. If two random variables are statistically independent, any cumulant involving them vanishes.

that that term will also drop an order by symmetry. Therefore, it appears that H_3 may contribute to the gyrocenter momentum flux.

This general argument is based on the explicit term in formula (J11). It is not complete because it does not proceed from an explicit form of H_3. In principle, it could happen that while an arbitrary H_3 behaves as suggested above, all physically relevant H_3's have special properties. However, an explicit example suggests otherwise. Consider formula (35) from Idomura (2012),

$$H_3 \propto \frac{1}{4}(\boldsymbol{\nabla}\phi\cdot\boldsymbol{\nabla})|\boldsymbol{\nabla}\phi|^2 \tag{J17}$$

($\boldsymbol{\nabla} \equiv \boldsymbol{\nabla}_\perp$). (This is a special case valid only in the cold-ion limit.) Because this involves second derivatives on ϕ, one must also consider the "other terms" in Eq. (J11). Scott & Smirnov (2010) derive those terms by assuming the dependence $H(\phi, \boldsymbol{\nabla}\phi, \nabla^2\phi)$. We consider the more general dependence $H(\phi, \partial_i\phi, \partial_i\partial_j\phi)$ and find that the divergence term in the representation of the flux associated with $\partial_\varphi H$ [see, for example, Eq. (I31)] is replaced by

$$\partial_i\left(\frac{\partial H}{\partial(\partial_i\phi)}\frac{\partial\phi}{\partial\varphi} + \frac{\partial P^{ij}}{\partial\phi}\partial_j\phi + (\partial_j P^{ij})\frac{\partial\phi}{\partial\varphi}\right), \tag{J18}$$

where (temporarily ignoring flux-surface averages)

$$P^{ij} \doteq -\frac{\partial H}{\partial(\partial_i\partial_j\phi)}. \tag{J19}$$

One has

$$H_3 = \frac{1}{4}\partial^k\phi\,\partial_k(\partial_l\phi\,\partial^l\phi) = \frac{1}{2}\partial_k\phi\,\partial_l\phi\,\partial^k\partial^l\phi. \tag{J20}$$

Then

$$\frac{\partial H_3}{\partial(\partial_i\phi)} = \partial_l\phi\,\partial^i\partial^l\phi \tag{J21}$$

and

$$P^{ij} \doteq -\frac{\partial H_3}{\partial(\partial_i\partial_j\phi)} = -\partial^i\phi\,\partial^j\phi. \tag{J22}$$

Also,

$$\partial_j P^{ij} = -(\partial_j\partial^i\phi\,\partial^j\phi + \partial^i\phi\,\partial_j\partial^j\phi). \tag{J23}$$

The first term cancels against Eq. (J21). Thus one is left with the second term of Eq. (J18) and the last term of Eq. (J23). For definiteness, consider the last term. That contributes a flux proportional to

$$\left\langle F\partial^\psi\phi\,\nabla^2\phi\,\frac{\partial\phi}{\partial\varphi}\right\rangle. \tag{J24}$$

That has the same form as the explicit term in Eq. (J11) except for an extra factor of $\nabla^2\phi$. That new factor changes sign under the symmetry transformation T. Thus, whereas under T the explicit term in Eq. (J11) changes sign

(and thus vanishes in the absence of symmetry breaking), formula (J24) is unchanged under T. It is thus also of third order. This is an example of the general argument that says that H_3 should be necessary in general.

In the work of Scott & Smirnov (2010), the authors concluded that a consistent, energy- and momentum-conserving theory could be obtained by retaining no more than H_2. That is correct, but it does not conflict with our conclusions above. One could use essentially *any* H_2 and still have a theory that conserved energy and momentum; as emphasized by Scott & Smirnov, that is the virtue of gyrokinetic field theory, in which both the kinetic equation and the Maxwell equations derive from the same action functional. But consistency does not imply that all physical processes are included or that the physics of the model Hamiltonian has anything to do with reality. If it is true that H_3 must be included, the implication is that physics processes described by H_3 compete on equal footing with other terms stemming from H_0, H_1, and H_2 to describe the macroscopic transport of momentum in the low-flow ordering.

J.2.2 Are the contributions from H_3 smaller than they appear?

No argument has been advanced[77] to show that a symmetry of the underlying turbulence will make $\mathcal{R}^{(3)} = O(\epsilon^4)$. However, it is possible that other specific properties of H_3 may make its contribution to Γ smaller than it nominally appears to be.

Example: A simple fluid model:

To illustrate some of the points about symmetry and structure of the third-order Hamiltonian, in this section we discuss a simple fluid model that captures some of the features of the full problem. We do not suggest that predictions of this model apply literally to the much more complicated tokamak problem; it merely serves as an example to motivate further analysis.

Consider an x–y–z coordinate system with x being a direction of inhomogeneity, z being a statistically homogeneous direction of mean flow, and y being a periodic coordinate. We will just discuss a general flow; however, if one wishes to think of that as an $\boldsymbol{E} \times \boldsymbol{B}$ flow, then in this model $\boldsymbol{B}$ would be in the y direction (i.e., purely poloidal).

Assume the random flow obeys

$$\partial_t \boldsymbol{u} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} + \widehat{D} \boldsymbol{u} = \delta \boldsymbol{f}(t) \tag{J25}$$

with $\boldsymbol{u}$ and $\delta \boldsymbol{f}$ being purely in the x and z directions and $\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0$ (i.e., we are considering a 2D, nonzonal flow). $\delta \boldsymbol{f}$ is a zero-mean stirring force (not necessarily Gaussian) that replaces the stirring associated with linear instability, and $\widehat{D}$ is a positive-definite linear dissipation operator. Let $\langle \dots \rangle$ denote an ensemble average over the turbulence. By the assumptions, one has statistical homogeneity in y and z, so ensemble averages can depend at most on x. We

[77] B. Scott (private communication, 2013) believes that an argument exists. However, the details have not been communicated to us.

have ignored pressure forces in this equation. We do not believe that is relevant for the points we wish to make, which are rather general in nature.

If boundary effects are ignored, a microscopic symmetry of Eq. (J25) is

$$\mathrm{T}\colon t \to t, \tag{J26a}$$
$$x \to -x \tag{J26b}$$
$$z \to z, \tag{J26c}$$
$$u_x \to -u_x, \tag{J26d}$$
$$u_z \to u_z \tag{J26e}$$

provided that $\delta f_x \to -\delta f_x$ and $\delta f_z \to \delta f_z$ under the transformation. Periodicity in the y direction replaces the up–down symmetry in a tokamak.

The average of Eq. (J25) is

$$\partial_t \langle \boldsymbol{u} \rangle + \partial_x (\langle u_x \rangle \langle \boldsymbol{u} \rangle) + \partial_x (\langle \delta u_x \, \delta \boldsymbol{u} \rangle) + \widehat{D} \langle \boldsymbol{u} \rangle = \langle \delta \boldsymbol{f} \rangle = \boldsymbol{0}. \tag{J27}$$

A mean flow can be excited through either a boundary condition or various kinds of forcing. First, let us apply a boundary condition as it is done in planar Couette flow, by moving the top surface (at $x = \frac{1}{2}a$) relative to the bottom one at $x = -\frac{1}{2}a$ in the z direction at speed U:

$$u_z\left(\frac{1}{2}a\right) = \frac{1}{2}U, \quad u_z\left(-\frac{1}{2}a\right) = -\frac{1}{2}U. \tag{J28}$$

To develop an analogy to the low-flow gyrokinetics problem, one may take $U = O(\epsilon)$. One is interested in $\langle u_z \rangle$. From $\boldsymbol{\nabla} \cdot \langle \boldsymbol{u} \rangle = 0$ and translational symmetry in z, it follows that $\partial_x \langle u_x \rangle = 0$. Since u_x must vanish at the walls, one concludes that $\langle u_x \rangle = 0$. Then

$$\partial_t \langle u_z \rangle + \partial_x \Gamma + \widehat{D} \langle u_z \rangle = 0, \tag{J29}$$

where the Reynolds stress

$$\Gamma \doteq \langle \delta u_x \, \delta u_z \rangle \tag{J30}$$

has appeared. This quantity changes sign under T. However, one cannot conclude that Γ vanishes for this problem. The wall boundary condition introduces an asymmetry that induces a nonvanishing stress. One way of saying this is that the macroscopic boundary condition does not obey the microscopic symmetry in x. That is, if one were to change $a \to -a$ in Eq. (J28), one would have to change $U \to -U$ in order that the constraint on u_z remained invariant. However, that conflicts with the requirement that z velocities remain unchanged under the transformation.

(We will return later to discuss internal rather than boundary forcing.)

More quantitatively, one can attempt to evaluate Γ by considering the fluctuations, which exactly obey

$$\partial_t \delta \boldsymbol{u} + \delta \boldsymbol{u} \cdot \boldsymbol{\nabla} \langle \boldsymbol{u} \rangle + \langle \boldsymbol{u} \rangle \cdot \boldsymbol{\nabla} \delta \boldsymbol{u} + \boldsymbol{\nabla} \cdot (\delta \boldsymbol{u} \, \delta \boldsymbol{u} - \langle \delta \boldsymbol{u} \, \delta \boldsymbol{u} \rangle) + \widehat{D} \delta \boldsymbol{u} = \delta \boldsymbol{f} \tag{J31}$$

or, because $\langle \boldsymbol{u} \rangle = \langle u_z \rangle(x) \widehat{\boldsymbol{z}}$,

$$\partial_t \delta \boldsymbol{u} + \delta u_x \partial_x \langle u_z \rangle \widehat{\boldsymbol{z}} + \langle u_z \rangle \partial_z \delta \boldsymbol{u} + \boldsymbol{\nabla} \cdot (\delta \boldsymbol{u}\, \delta \boldsymbol{u} - \langle \delta \boldsymbol{u}\, \delta \boldsymbol{u} \rangle) + \widehat{D} \delta \boldsymbol{u} = \delta \boldsymbol{f}. \quad \text{(J32)}$$

Let us consider δu_z:

$$\partial_t \delta u_z + \delta u_x \partial_x \langle u_z \rangle + \langle u_z \rangle \partial_z \delta u_z + \boldsymbol{\nabla} \cdot (\delta \boldsymbol{u}\, \delta u_z - \langle \delta \boldsymbol{u}\, \delta u_z \rangle) + \widehat{D} \delta u_z = \delta f_z. \quad \text{(J33)}$$

In the spirit of renormalized turbulence theory, the terms not involving $\langle u_z \rangle$ (the 'eddy–eddy interactions') are represented *via* a response function $\widehat{g}$:

$$\widehat{g}^{\,-1} \delta u_z = -(\delta u_x \partial_x \langle u_z \rangle + \langle u_z \rangle \partial_z \delta u_z) + \delta f_z, \quad \text{(J34)}$$

or, upon ignoring an initial-condition term,

$$\delta u_z = -\widehat{g}\, (\delta u_x \partial_x \langle u_z \rangle + \langle u_z \rangle \partial_z \delta u_z) + \widehat{g}\, \delta f_z. \quad \text{(J35)}$$

(Note that $\widehat{g}$ integrates over a lagged time.) This gives

$$\Gamma = -\langle \delta u_x \widehat{g}\, \delta u_x \rangle \partial_x \langle u_z \rangle - \langle \delta u_x \widehat{g}\, \partial_z \delta u_z \rangle \langle u_z \rangle + \langle \delta u_x \widehat{g}\, \delta f_z \rangle. \quad \text{(J36)}$$

For the first term, substitute a turbulence autocorrelation time τ_{ac} for $\widehat{g}$. That gives a positive-definite viscosity coefficient $\mu \doteq \langle \delta u_x^2 \rangle \tau_{\mathrm{ac}}$. If one invokes the microscopic symmetry on the last two terms, one can argue that they vanish. The result $\Gamma = -\mu\, \partial_x \langle u_z \rangle$ shows how the macroscopic boundary condition breaks the microscopic symmetry (and the influence of that boundary condition is felt throughout the domain; it does not just extend over a thin boundary layer).

What about the size of higher-order cumulants? There is no direct analog of a third-order Hamiltonian in this problem, but one can ask about the kind of cumulant to which a Hamiltonian theory would lead. One representive term (which would arise from the cold-ion limit) would be

$$\Gamma^{[3]} \sim \left\langle \delta u_x \delta u_z \left(\frac{\partial \delta u_z}{\partial x} - \frac{\partial \delta u_x}{\partial z} \right) \right\rangle, \quad \text{(J37)}$$

the parenthesized term being proportional to the y component of vorticity. This term is invariant under the microscopic symmetry, so its size is unclear at this point. However, upon integrating by parts and using $\boldsymbol{\nabla} \cdot \delta \boldsymbol{u} = 0$, one has

$$\begin{aligned}
\Gamma^{[3]} &= \frac{1}{2} \left\langle \delta u_x \frac{\partial \delta u_z^2}{\partial x} - \delta u_z \frac{\partial \delta u_x^2}{\partial z} \right\rangle & \text{(J38a)} \\
&= \frac{1}{2} \left(\partial_x \langle \delta u_x \delta u_z^2 \rangle - \left\langle \frac{\partial \delta u_x}{\partial x} \delta u_z^2 \right\rangle - \partial_z \langle \delta u_z \delta u_x^2 \rangle + \left\langle \frac{\partial \delta u_z}{\partial z} \delta u_x^2 \right\rangle \right) & \text{(J38b)} \\
&= \frac{1}{2} \left(\partial_x \langle \delta u_x \delta u_z^2 \rangle + \left\langle \frac{\partial \delta u_z}{\partial z} \delta u_z^2 \right\rangle - \partial_z \langle \delta u_z \delta u_x^2 \rangle - \left\langle \frac{\partial \delta u_x}{\partial x} \delta u_x^2 \right\rangle \right) & \text{(J38c)} \\
&= \frac{1}{2} \left(\partial_x \langle \delta u_x \delta u_z^2 \rangle + \frac{1}{3} \partial_z \langle \delta u_z^3 \rangle - \partial_z \langle \delta u_z \delta u_x^2 \rangle - \frac{1}{3} \partial_x \langle \delta u_x^3 \rangle \right). & \text{(J38d)}
\end{aligned}$$

The terms in ∂_z vanish by the translational invariance in the z direction. The terms in ∂_x are of fourth order, not third, because the ∂_x acts on an averaged quantity, and all such quantities have at most only weak, macroscopic spatial variation. Thus $\Gamma^{[3]}$, which looked like it might be of the same order as $\Gamma^{[2]} \equiv \Gamma$, is actually negligible.

Now ask what happens when one removes the wall forcing. Even in the absence of an initial $\langle u_z \rangle$, there could be an intrinsic contribution to Γ from the stirring: $\Gamma \propto \langle \delta f_x \, \delta f_z \rangle$. Under the simplest assumption, that would vanish due to the microscopic symmetry. But within the context of this simple model, one can impose a small correlation, of order ϵ. (In reality, that would again come from some sort of macroscopic symmetry breaking.) Then one would obtain an intrinsic stress of order $\epsilon^2 \times \epsilon = \epsilon^3$. That stress comes from $\Gamma^{[2]}$, not $\Gamma^{[3]}$. (One could also discuss pinch terms in the same way; cf. the α effect in MHD turbulence.)

H_3 contributions:

Similar manipulations can be applied to the cold-ion limit of terms arising in H_3. As an example, consider again the formula (J24). Written in a local Cartesian coordinate system, that becomes $\langle \partial_x\phi \, \partial_z\phi \, (\partial_x^2 + \partial_z^2)\phi \rangle$. This can be manipulated into the same form as Eq. (J38d) with the definitions $\delta u_x = -\partial_y \phi$ and $\delta u_z = \partial_x \phi$. This term is therefore expected to be very small, in agreement with the numerical results of Idomura (2012).

Unfortunately, such manipulations do not appear to work for FLR terms, whose expansion involves $(\nabla_\perp^2 \phi)^n$, including $n > 1$. No argument has yet been given why those terms should be smaller than they nominally appear to be.

K Example of Near-Cancellation of Lower-Order Terms

As we have discussed in several places, the reason that higher-order terms can be important is that the momentum flux from lower-order terms experiences near-cancellations that make the net flux of higher order than they at first appear to be. Here we demonstrate in a concrete way the possibility of that occurring. We consider a slab limit for simplicity, where $\widehat{\boldsymbol{x}}$ corresponds to the radial coordinate and $\widehat{\boldsymbol{y}}$ is the binormal direction (the direction perpendicular to both the radial direction and the magnetic field direction $\widehat{\boldsymbol{z}}$). (One can consider this as a rigorous limit for an infinitely elongated, large-aspect-ratio tokamak, so that the magnetic field is vertical and the binormal direction is toroidal.) Upon denoting the flux-surface and ensemble/time average by $\langle \ldots \rangle = L_y^{-1} \int_0^{L_y} dy \ldots$ (with an implied ensemble or time average) and using the lowest-order $\boldsymbol{E} \times \boldsymbol{B}$ drift for the radial drift velocity, one finds that the mean flux-surface average of the right-hand side of Eq. (105) is

$$\text{RHS} = \frac{\partial}{\partial x} \sum_s (\overline{n} q)_s \frac{c}{B} \int d\mathbf{v} \left\langle F_s \frac{\partial \phi}{\partial y} \right\rangle . \tag{K1}$$

Since $\boldsymbol{E}\times\boldsymbol{B}$ drifts are of order ϵc_s, the nominal order of this term appears to be $\sim \epsilon n q c_s/L$. However the equilibrium component of F_{s0} gives no net contribution since $\langle F_{s0}\,\partial\phi/\partial y\rangle = F_{s0}\langle\partial\phi/\partial y\rangle = 0$, so only the smaller fluctuating component involving $\delta F_s \sim \epsilon F_{s0}$ contributes to this flux. Thus this term now appears to be of order $\epsilon^2 n q c_s/L$. This is much larger than the mean flux-surface average of the LHS of Eq. (105), which was shown in Sec. 2.2.2 to be of order $\epsilon^4 n q c_s/L$ if a low-flow gyro-Bohm ordering holds. There are two key steps to show that the actual order of the RHS term is in fact smaller by another factor of ϵ^2.

The first step starts by noting that one can use the quasineutrality equation (104) to eliminate the guiding-center charge density in terms of the polarization term to get

$$\text{RHS} = -\frac{\partial}{\partial x}\left(n_{i0} e\rho_s^2 \frac{ec}{T_e B}\left\langle (\nabla_\perp^2\phi)\frac{\partial\phi}{\partial y}\right\rangle\right).$$

where we have pulled the leading factor of the ion density n_i out of the flux-surface average for simplicity, replacing it with the equilibrium density n_{i0} (this is also done in some forms of the gyrokinetic Poisson equation implemented in certain codes). Upon expanding the Laplacian $\nabla_\perp^2 = \partial_x^2 + \partial_y^2$ and using the identities $\langle(\partial_y^2\phi)\partial_y\phi\rangle = \frac{1}{2}\langle\partial_y(\partial_y\phi)^2\rangle = 0$ and $\langle(\partial_x^2\phi)\partial_y\phi\rangle = \partial_x\langle(\partial_x\phi)\partial_y\phi\rangle - \langle(\partial_x\phi)\partial_y(\partial_x\phi)\rangle = \partial_x\langle(\partial_x\phi)\partial_y\phi\rangle$, one gets

$$\text{RHS} = -\frac{\partial}{\partial x}\left(n_{i0} e\rho_s^2 \frac{ec}{T_e B}\frac{\partial}{\partial x}\left\langle \frac{\partial\phi}{\partial x}\frac{\partial\phi}{\partial y}\right\rangle\right).$$

Because a gradient has now been pulled outside the flux-surface/ensemble average, which varies only on the slower equilibrium scale $\partial/\partial x \sim 1/L$ instead of on the faster turbulence scale $\partial\phi/\partial x \sim \phi/(\epsilon L)$, this term is now of apparent order $\epsilon^3 n q c_s/L$, one order smaller than it previously seemed. Note that the quantity inside the angle brackets is of the form of a Reynolds stress, $\propto \langle\delta v_y \delta v_x\rangle$. At this point we make contact with the simple fluid model discussed in Sec. J.2.2.

The second key step is to note that δv_x and δv_y are very weakly correlated, so that while the ensemble average of $\langle\delta v_y\delta v_x\rangle$ appears to be nominally of order $\epsilon^2 c_s^2$, in fact $\delta v_y \delta v_x$ almost averages to zero, resulting in the ensemble average $\langle\delta v_y\delta v_x\rangle$ actually being of order $\epsilon^3 c_s^2$. This near-cancellation is illustrated physically by the eddies shown in Fig. 4.

One way to break the symmetry of these eddies is if there is an existing background sheared flow. One can think of this as a two-step process in which eddies pointing in the radial direction are created in a decorrelation time $\tau_c \sim L/c_s$ by the background turbulence drive, but then are subjected for a decorrelation time to a background shear flow with a shearing rate $\partial v_{y0}/\partial x \sim \epsilon c_s/L$. This results in eddies with a tilt angle of $\delta\theta \sim \tau_c\,\partial v_{y0}/\partial x \sim \epsilon$. Thus the mean momentum flux is of order $\langle\delta v_x\delta v_y\rangle \sim \langle\delta\theta|\delta v_x|^2\rangle \sim \epsilon^3 c_s^2$, one factor of ϵ smaller than its nominal order.

Another way to see the weak correlation between δv_y and δv_x is to consider a fluid estimate for the evolution of the binormal flow in the presence of a background sheared flow, $\partial\delta v_y/\partial t = -\delta v_x \partial v_{y0}/\partial x - \delta v_y/\tau_c$, where the first term on the RHS is a linear driving term and the second term is a decorrelation

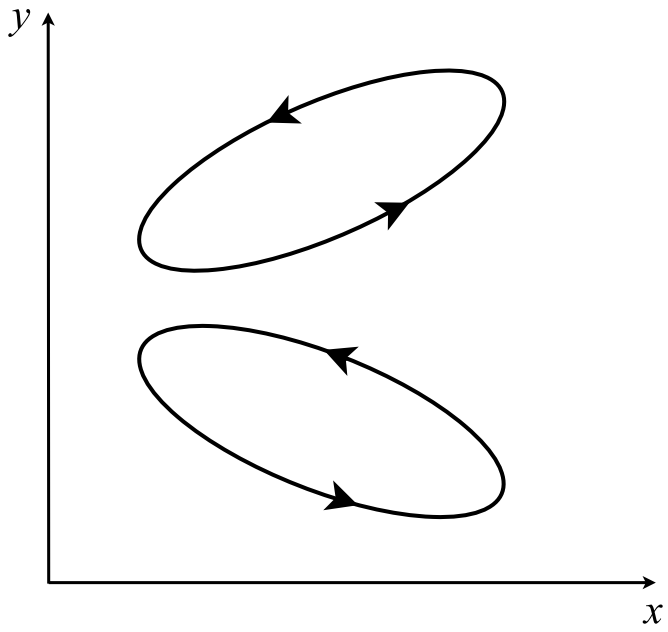

Figure 4: Illustration of the cancellation of momentum flux for symmetrically distributed eddies. For the top eddy that is tilted counterclockwise, δv_x and δv_y are both positive in the lower half of the eddy, while δv_x and δv_y are both negative in the upper half of the eddy, so the momentum flux $\delta v_x \delta v_y$ from this eddy is positive. For the bottom eddy that is tilted clockwise, the reverse happens, so its momentum flux is negative. In the absence of effects (like up–down asymmetry) that would break the symmetry, there is no preferred direction, so eddies tilted either way are equally likely and the mean momentum flux $\langle \delta v_x \delta v_y \rangle$ averages to 0.

model of the nonlinear term. This gives $\delta v_y \sim \tau_c \delta v_x \partial v_{y0}/\partial x \sim \tau_c \delta v_x \epsilon c_s/L$ in the low-flow ordering, again giving $\langle v_y v_x \rangle \sim \epsilon^3 c_s^2$. This is the same argument that was given somewhat more formally in Sec. J.2.2.

Idomura (2012) refers to a diamagnetic shear mechanism described by (Camenen et al., 2011) that might give a stronger tilt to the eddies, with a tilt angle $\delta\theta \sim \epsilon^{1/3}$. This is based on high-order '$1/n$' corrections (of order $\epsilon \sim \rho/L$) from the radial envelope equation for the ballooning representation. While this is an interesting hypothesis, it is based on only linear/quasilinear calculations and has not yet been demonstrated nonlinearly. There are other examples where it is known that higher-order $1/n$ corrections are not relevant to the nonlinear dynamics (basically because nonlinear effects are much stronger than the weak linear corrections that affect a formal radial eigenmode structure). For example, they sometimes predict a linear radial mode width that scales as $\sqrt{\rho L}$, while nonlinearly the eddies break up and have a radial width

that scales with ρ when one is sufficiently deep into the small-ρ/L gyro-Bohm regime. [Also, if the effect of weak diamagnetic shear is somehow amplified to have a stronger fractional order, one might wonder if other weak terms (such as from H_3) also have a stronger effect than one first believes.]

L Excerpts from papers related to momentum conservation

There is no better way to understand the historical development of the momentum-conservation topic than to read the authors' original words. Here we include a sampling, in some cases appending some remarks to place the excerpt in the context of the present report.

L.1 Excerpt from Parra (2009) — "Extension of gyrokinetics to transport time scales"

> "In recent years, several groups have begun to build codes that evolve the full distribution function, without splitting it into a slowly varying Maxwellian and a fast, fluctuating piece. These simulations, known as full f models, are employing the traditional gyrokinetic formulation. In this thesis, I will argue that this approach is inadequate since it is unable to solve for the self-consistent radial electric field that is crucial for the turbulence."

> **Remark:** See Table 1 and our discussion in Sec. 5.1 for Parra's summary of the various approaches.
>
> To be clear, it is not the case that a full-f model is theoretically *unable* to solve for the radial electric field. Rather, it is that working it out to the requisite order is onerous and probably impractical.

L.2 Excerpts from Catto et al. (2008) — "Electrostatic turbulence in tokamaks on transport time scales"

This is the first mention of the hybrid fluid–kinetic description.

1. "The quasineutrality equation used in [the] more familiar gyrokinetic descriptions is unable to ensure intrinsic ambipolarity in the axisymmetric, long radial wavelength limit. ... In our hybrid description the global electric field behavior is not an issue because intrinsic ambipolarity is automatically recovered in the appropriate axisymmetric limit since finite orbit polarization effects are correctly retained by the ion inertial term in the total momentum conservation equation and by the charge conservation or vorticity equation."

Remark: The phrase "more familiar gyrokinetic descriptions" presumably refers to those that use a kinetic equation correct only to first order in ϵ, i.e., those that just use an H_1.

2. "Flux surface averaging $\boldsymbol{\nabla} \cdot \boldsymbol{J} = 0$ and integrating once in ψ gives the global ambipolarity constraint $\langle \boldsymbol{J} \cdot \boldsymbol{\nabla}\psi \rangle_\psi = 0 \ldots$."

3. "Note that we must not improperly determine the axisymmetric radial electric field by adjusting it until $\langle n\boldsymbol{V} \cdot \boldsymbol{\nabla}\psi \rangle_\psi \approx 0$ or $\langle \boldsymbol{J} \cdot \boldsymbol{\nabla}\psi \rangle_\psi = 0$, as is sometimes mistakenly done in tokamaks."

Remark: This remark is explained in more detail in excerpt 2 of Sec. L.4.

L.3 Excerpts from Parra & Catto (2008) — "Limitations of gyrokinetics on transport time scales"

1. "The electric field is of special importance since the poloidal zonal flow ... induced by its radial structure can act to control the saturated amplitude of turbulence."

Remark: This is one important (well-known) reason for worrying about the (radial) electric field. In later publications, the focus turns more to the connection between the radial field and the profile of toroidal rotation.

2. "[The quasineutrality condition

$$\frac{Z^2 e}{M} \int d^3v \, \widetilde{\varphi} \left(\frac{\partial f_i}{\partial E_0} + \frac{1}{B} \frac{\partial f_i}{\partial \mu_0} \right) \approx -Z\widehat{N}_i(\boldsymbol{r}, t) - n_e(\boldsymbol{r}, t), \tag{L1}$$

where $\widetilde{\varphi} \doteq \varphi(\boldsymbol{x} + \boldsymbol{\rho}) - \langle \varphi(\boldsymbol{x} + \boldsymbol{\rho}) \rangle_\zeta$,] may be used to calculate ϕ for wavelengths of the order of the gyroradius, including zonal flow, as is normally done in δf turbulence codes such as `GS2`... or `GYRO`.... However, the equation is not useful for long wavelengths. In the limit $k_\perp L \sim 1$, the average of $\widetilde{\phi}$ holding $\boldsymbol{r}$, $v_\parallel$ and $v_\perp$ fixed becomes the same order as terms already neglected ... As a result, the terms on the left side of [Eq. (L1)] vanish to the order δ the equation has been derived, leaving

$$Z\widehat{N}_i(\boldsymbol{r}, t) = n_e(\boldsymbol{r}, t) \tag{L2}$$

as the quasineutrality equation. This equation does not depend on ϕ explicitly, and even though it could depend implicitly through the ion and electron distribution functions, that dependence should be negligible to $O(\delta^2 n_e)$ due to intrinsic ambipolarity [36, 37]. Therefore, we cannot solve for the electrostatic potential at long wavelengths."

Remark: This is their basic argument. The "terms neglected" include higher-order terms in the pullback transformation, i.e., more precise representations of the polarization.

3. See their Eq. (55), which contains the spurious term commented on by Lee & Kolesnikov (2009). For PC's comment on Lee & Kolesnikov, see Parra & Catto (2009a).

   Remark: PC explain how they obtained this term, which is proportional to T_i^{-1} in Parra & Catto (2009a). It arises from an unusual choice of energy variable and leads to a representation of the physics in which the $T_i \to 0$ limit may not be taken. That restriction should not be present in the final GK–Poisson system [and was not present in the formulation used by Dubin et al. (1983)] because one should be able to straightforwardly extract cold-ion physics (which is nonsingular).

4. "It might seem that keeping more terms in the gyrokinetic equation to obtain a higher order solution for the distribution fnction would be enough to find the potential, but finding such a gyrokinetic equation for general geometry is difficult and its solution by numerical means requires high numerical precision since terms smaller than $O(\delta^2 f_M)$ must be recovered without appreciable error to calculate the full axisymmetric potential to lowest order."

L.4 Excerpts from Catto et al. (2009) — "Limitations, insights and improvements to gyrokinetics"

1. "We have recently shown that rather than evaluating quasineutrality to higher order it is more sensible to replace it by a toroidal angular momentum or vorticity conservation equation that does not require as accurate an ion distribution function . . . ."

   Remark: There are clear advantages to such hybrid approaches. One concern is the status of overall conservation laws for the hybrid system. (Clearly the form of the momentum conservation equation is automatically satisfied.)

2. "Satisfying the neoclassical relation for the parallel ion flow

$$V_{i\|} = \frac{IT_i}{M\Omega_i}\left(\frac{kB^2}{\langle B^2\rangle_\psi}\frac{\partial}{\partial\psi}\ln T_i - \frac{\partial}{\partial\psi}\ln p_i - \frac{e}{T_i}\frac{\partial\Phi}{\partial\psi}\right) \tag{L3}$$

   merely verifies that the ion gyrokinetic equation is being solved consistently to the order employed in codes, where the numerical coefficient k depends on the regime of collisionality. Equation (L3) simply provides a relation between $\partial\Phi/\partial\psi$ and $V_{i\|}$, and confirms that parallel ion momentum is being satisfied through leading order in the pressure anisotropy. It

does not determine the radial electric field (unless it is mistakenly and arbitrarily assumed that the parallel ion flow can be set to zero or specified in some other ad hoc manner) and satisfying it is not a test that the correct rotation profile is obtained. Equation (L3) places such a strong constraint on the relation between the parallel ion flow and the global radial electric field that it has recently been shown to be valid even at long (nonzonal flow) wavelengths in turbulent plasmas [Parra & Catto (2009c)]. In brief, the ambipolarity constraint $\langle n\boldsymbol{V}_i\cdot\boldsymbol{\nabla}\psi - n\boldsymbol{V}_e\cdot\boldsymbol{\nabla}\psi\rangle_\psi = 0$ should be automatically satisfied to very high order for any long wavelength radial electric field and should not determine the global axisymmetric radial electric field.

"Gyrokinetic simulations using quasineutrality with polarization effects retained through second order in the ion poloidal gyroradius ($\rho_{\rm pol}$) expansion and a guiding centre density valid only through first order in $\rho_{\rm pol}$ determine an incorrect global axisymmetric radial electric field that is different from the one obtained by conservation of toroidal angular momentum. Effectively, such a gyrokinetic quasineutrality treatment adds charge sources and sinks that result in an incorrect global axisymmetric radial electric field. Indeed, in the pioneering banana regime evaluation of Rosenbluth et al [15] quasineutrality is not employed to determine the radial electric field since the second order in $(\rho_{\rm pol}/L)^2$ corrections to quasineutrality exactly cancel. Instead, toroidal angular momentum conservation is used to determine the radial electric field at higher order in the combined gyroradius and collisionality expansions. More generally, a carefully constructed full f turbulent gyrokinetic code should result in no net radial transport of toroidal momentum through order $(\rho_{\rm pol}/L)^2$ in the absence of sources and sinks if quasineutrality is self-consistently treated to the same order [Parra & Catto (2009c)]. In the remainder of this section we give an improved and streamlined proof that $\partial\Phi/\partial\psi$ cannot be determined in a turbulence-free tokamak without a moment approach if the ion distribution function is only known to second order in the gyroradius expansion.

"A moment procedure for the electron particle flux using $C_{1e}\{f_{1e}\} = C_{1ee}\{f_{1e}\} + C_{ei}\{f_{1e}\}$ with C_{1ee} the electron–electron operator and $C_{ei}\{f_{1e}\} = L_{ei}\{f_{1e} - (m/T_e)V_{\parallel i}v_\parallel f_{0e}\}$ the unlike electron–ion Lorentz operator gives the electron radial particle flux as $\langle n\boldsymbol{V}_e\cdot\boldsymbol{\nabla}\psi\rangle_\psi = (mcI/E)\langle B^{-1}\int d^3v\, v_\parallel C_{1e}\{f_{1e} - (m/T_e)V_{\parallel i}v_\parallel f_{0e}\}\rangle_\psi$, with f_{1e} the leading order correction to the electron Maxwellian f_{0e} and T_e and m the electron temperature and mass. The electron drift kinetic equation can be written as $v_\parallel\boldsymbol{n}\cdot\boldsymbol{\nabla}g_e = C_{1e}\{g_e + (Iv_\parallel/\Omega_e)(\partial f_{0e}/\partial\psi) - (m/T_e)V_{\parallel i}v_\parallel f_{oe}\}$ with $g_e = f_{1e} - (Iv_\parallel/\Omega_e)(\partial f_{0e}/\partial\psi)$. The $\partial\Phi/\partial\psi$ drive terms in the collision operator cancel, making g_e independent of the radial electric field so that $\langle n\boldsymbol{V}_e\cdot\boldsymbol{\nabla}\psi\rangle_\psi = (mcI/e)\langle B^{-1}\int d^3v\, v_\parallel C_{1e}\{g_e + (Iv_\parallel/\Omega_e)(\partial f_{0e}/\partial\psi) - (m/T_e)V_{\parallel i}v_\parallel f_{eo}\}\rangle_\psi$ cannot depend on the radial electric field to order $(\rho_i/L)^2$ since $C_{ii}/C_{ee}\sim\nu_{ii}/\nu_{ee}\sim(m/M)^{1/2}\sim\rho_i/L$ is normally assumed, with ν_{ii} and ν_{ee} the ion–ion and electron–electron collision frequencies.

"Alternately, a moment description can be used to further demonstrate

that intrinsic ambipolarity must be satisfied to order ρ_i^2/L^2 and demonstrate that it is the flux surface average of conservation of toroidal angular momentum that must give the radial electric field (pressure anisotropy does not enter this constraint). To order ρ_i^2/L^2 the cross field viscosity is diamagnetic (and so collisionless to lowest order) and the radial flux of toroidal angular momentum may be written in terms of the ion gyroviscosity $\boldsymbol{\pi}_{ig}$ (within small up–down asymmetric contributions) as [16, 17] $\langle R^2\boldsymbol{\nabla}\zeta\cdot\boldsymbol{\pi}_{ig}\cdot\boldsymbol{\nabla}\psi\rangle_\psi = \langle(MI/B)\int d^3v\, v_\parallel f_{1i}\boldsymbol{v}_{di}\cdot\boldsymbol{\nabla}\psi\rangle_\psi$, with R the major radius. Inserting $f_{1i} = g_i - (Iv_\parallel/\Omega_i)(\partial f_{0i}/\partial\psi)$, using $\langle\int d^3v\, f_{0i}(v_\parallel/B)^2 v_\parallel \boldsymbol{n}\cdot\boldsymbol{\nabla}(v_\parallel/B)\rangle_\psi = 0$, and recalling g_i depends only on $\partial T_i/\partial\psi$ gives a $\partial\Phi/\partial\psi$ independent result for $\langle R^2\boldsymbol{\nabla}\zeta\cdot\boldsymbol{\pi}_{ig}\cdot\boldsymbol{\nabla}\psi\rangle_\psi$. Hence, we have proven that the correct neoclassical radial electric field cannot be determined directly from toroidal angular momentum conservation knowing f_i to second order. As a result, a direct determination of $\partial\Phi/\partial\psi$ requires knowing the ion distribution function through third order in the gyroradius expansion, or a moment approach as outlined in section 4 must be used to save an order.

"By considering a steady-state theta pinch using a model collision operator, we have explicitly shown that gyrokinetic quasineutrality cannot determine the axisymmetric, long radial wavelength electrostatic potential to order ρ_i^2/L^2 [1]. Here we have proven the same situation occurs in axisymmetric tokamaks. In modern gyrokinetic treatments intrinsic ambipolarity is violated when the ion distribution function is retained to only order ρ_i/L in the guiding centre density, while being kept to order ρ_i^2/L^2 in the finite orbit polarization term in gyrokinetic quasineutrality. However, when f_i is kept to order ρ_i^2/L^2 in both places, the radial electric field does not enter and therefore cannot be determined, and no inconsistency arises. To determine this axisymmetric radial electric field higher order effects must be retained. The same conclusion holds in a turbulent tokamak but the proof is substantially more involved [Parra & Catto (2009c)]."

> **Remark:** This is probably the most detailed and coherent explanation of the 'intrinsic ambipolarity' issue. Focus on the last paragraph, where they state
>
> > "In modern gyrokinetic treatments intrinsic ambipolarity is violated when the ion distribution function is retained to only order ρ_i/L in the guiding centre density, while being kept to order ρ_i^2/L^2 in the finite orbit polarization term in gyrokinetic quasineutrality."
>
> They intend this remark to apply to neoclassical theory, but it can also be considered for the turbulent case. Then note:
>
> - The cancellations that give rise to the $\langle F\,\partial_\varphi H\rangle$ term happen for any H.
> - Reducing the $\langle F\,\partial_\varphi H\rangle$ term is therefore crucial. Note that to get even lowest-order polarization one must keep H_2. So

then the question becomes: If one uses at most H_2 in the kinetic equation, does one retain enough information to get the third–order terms in the momentum conservation law? Parra & Catto say no for a full-f formulation, and we agree.

L.5 Excerpts from Parra & Catto (2009a) — "Comment on "On higher order corrections to gyrokinetic Vlasov-Poisson equations in the long wavelength limit" [Phys. Plasmas 16, 044506 (2009)]"

1. "...that higher order terms are formally smaller does not invalidate our conclusions because we find that the formally larger terms cancel at long wavelengths."

2. "...the higher order terms are crucial for the final result because the lower order polarization density exactly cancels with other contributions without determining the long wavelength radial electric field. Our second order calculation for a nonturbulent θ-pinch in [Parra & Catto (2008)] shows that the lowest order polarization density and the solution of the second order gyrokinetic equation cannot determine the axisymmetric long wavelength electrostatic potential. ... The long wavelength, axisymmetric piece of the electrostatic potential must remain undetermined unless the ion distribution function is determined to higher order than second in an expansion of the small ratio $\delta_i = \rho_i/L \ll 1$ of the ion gyroradius ρ_i over the characteristic length L. Indeed, we have proven that there is no implementable way to calculate this contribution to the potential directly from quasineutrality; a fact that has been understood from drift kinetic theory since at least the early 1970s (Rosenbluth et al., 1971; Hazeltine, 1974).

   Remark: The word 'implementable' means 'practically implementable.'

3. "The radial electric field can be obtained from two different equations, namely, the gyrokinetic Poisson equation and the transport of momentum. The transport of momentum determines the velocity profile and the velocity profile is uniquely related to the long wavelength electric field. These two approaches of obtaining the electric field will give conflicting results at long wavelengths unless [the equation for $\partial_t \rho_{\rm tot}$] is satisfied. That is, the time derivative of the flux surface averaged charge density $\langle e(Zn_i - n_e)\rangle_\psi$ must be obtained to the order of the right side, requiring it to be very small and of order $\partial\langle e(Zn_i - n_e)\rangle_\psi/\partial t \sim \delta_i^4 e n_e v_{ti}/L$."

   Remark: For discussion of the relevance of fourth-order terms, see Sec. 5.1.2.

L.6 Parra & Catto (2009b) — "Gyrokinetic equivalence"

This paper compares the Hamiltonian approach of Dubin *et al.* with the recursive methodology of Parra & Catto (2008); as expected, they find order-by-order agreement. There are no explicit remarks about momentum conservation in this paper.

L.7 Excerpts from Parra & Catto (2009c) — "Vorticity and intrinsic ambipolarity in turbulent tokamaks"

1. "On the one hand, the toroidal rotation, determined by the radial electric field, is believed to play an essential role in the creation and equilibrium of regions of reduced turbulence like the pedestal and internal transport barriers .... On the other hand, the poloidal zonal flow ... induced by the radial structure of the electric field can act to control the saturated amplitude of turbulence."

   Remark: These are good reasons for worrying about the toroidal rotation and the radial electric field. However, it is not necessarily the case that the ordering discussed in this report is relevant to all instances of pedestals and transport barriers.

2. "The distribution function required to directly obtain the viscosity is higher order than second; the order at which intrinsic ambipolarity is maintained."

   Remark: The order-by-order solution for the gyrocenter distribution function is given in detail by Calvo & Parra (2012).

3. "In particular, we are able to show that turbulence dominated tokamaks are also intrinsically ambipolar, meaning that the distribution function is required to at least third order in a gyroradius over scale length expansion to determine the electric field from quasineutrality."

   Remark: See the previous comment.

L.8 Excerpt from Parra & Catto (2010a) — "Non-physical momentum sources in slab geometry gyrokinetics"

"We have shown that the Hamiltonian gyrokinetic formulation of Dubin et al. (1983) results in a non-physical velocity profile in the low flow ordering unless a proper momentum description is employed. If quasineutrality or vorticity are used, it is necessary to keep some third order corrections to [the potential Ψ] to recover the correct transport of momentum. Employing the lowest order version of the same procedure, as is done in full f gyrokinetic codes ..., it is easy to derive that for $\Psi \approx \overline{\phi}$

the non-physical source of momentum becomes large enough to drive the velocity to the high flow ordering.

"Note that in a slab, it is necessary to calculate the gyrokinetic drifts up to $O(\delta_i^3 v_i)$ to recover the correct momentum equation in the low flow ordering.... It might be surprising that the drifts are only needed up to order $\delta_i^3 v_i$ in the drift ordering where in section 1 we argued that $\delta_i^4 v_i$ terms were required. This simplification is a result of the special geometry of the slab. In a collisionless slab, the flux surface averaged current density due to the $O(\delta_i^4 v_i)$ drift is to $O(\delta_i^4 e n_e v_i)$

$$Ze \left\langle \int d\boldsymbol{v}\, f_{i0}^{(0)} \frac{c}{B} \frac{\partial \Psi_0^{(4)}}{\partial y} \right\rangle_x = 0, \tag{L4}$$

since the lowest order piece of the distribution function $f_{i0}^{(0)}$ is independent of y. In a tokamak, on the other hand, there are magnetic geometry effects that may prevent such a concellation from happening.

"In conclusion, solving the quasineutrality equation for all the pieces of the electric field, including the long wavelength pieces, in a tokamak requires a gyrokinetic formulation that keeps the corrections to the drifts up to order $\delta_i^4 v_{ti}$ in the low flow ordering, ... Lagrangian formulations keep drifts to order $\delta_i^2 v_{ti}$ at most. We have shown for a slab that next order corrections are required. The slab case shows how the electric field obtained from quasineutrality introduces an artificial momentum source that will accelerate the plasma in the y direction. The higher order corrections to the drifts studied in this paper appear in general geometries, but in addition there are magnetic geometry effects that make the equations almost intractable to order $\delta_i^2 v_{ti}$, and hopelessly complicated to order $\delta_i^3 v_{ti}$ and $\delta_i^4 v_{ti}$. Therefore, trying to calculate all the contributions to the electric field employing a gyrokinetic quasineutrality equation is impractical. Instead, the momentum transport equation should be explicitly solved to determine the long wavelength velocity profile."

> **Remark:** The basic issue here is the offset between nth-order effects in the kinetic equation and $(n-1)$th-order effects in the Poisson equation. That follows automatically from the variational approach to the gyrokinetic system.
>
> From the second paragraph of the excerpt, one sees that the status of the effects due to magnetic inhomogeneities was unclear at that time (they used the phrase "*may* prevent"). The general form of the gyrocenter conservation law shows that the third-order calculation is sufficient. Exactly how the cancellations happen is detailed in Appendix I.
>
> In defense of the (30-year-old) work of Dubin *et al.*, they were not interested in momentum conservation, and their equations were intended to be integrated only over turbulence time scales.

On the other hand, it is clear that they did not fully understand the ordering issues, which were only clarified with the advent of the Lagrangian field theory.

L.9 Excerpts from Parra & Catto (2010b) — "Transport of momentum in full f gyrokinetics"

This is a recent invited APS talk by Parra, so it should contain a reasonably up-to-date summary of his thinking.

1. "In quasineutrality, any small error in the calculation of a charge density leads to large, nonphysical deviations in the radial electric field. These deviations lead to a toroidal rotation profile that does not satisfy the correct conservation equation for the toroidal angular momentum."

2. "Any error in [the equation for the time rate of change of the charge] appears as a nonphysical force added to [the conservation law of toroidal angular momentum]."

3. "Another way of stating the problem is that the quasineutrality equation gives the wrong long wavelength radial electric field. This incorrect radial electric field then results in an incorrect toroidal rotation satisfying [the balance between toroidal velocity, radial electric field, radial pressure gradient, and poloidal velocity], but not satisfying the transport of toroidal angular momentum. . . "

4. "Thus, even though having some form of conservation of momentum may be arguably better than not having any, . . . , the variational approach has not solved the problem of the long wavelength electric field. It is necessary to obtain the gyrokinetic equation to painfully high orders when using quasineutrality, and even variational approaches cannot escape this fact. Indeed, using a variational formulation may only hide the problem in an incorrect transport of momentum."

   Remark: Scott & Smirnov showed that a consistent statement of momentum conservation can be obtained by using any Hamiltonian in the action principle (see excerpt 1 of Sec. L.11). That H could be truncated to any power in ϵ. It could also contain terms that have nothing to do with the true gyrokinetic physics. A 'consistent' conservation law is not sufficient.

L.10 Excerpts from Parra & Catto (2010c) — "Turbulent transport of toroidal angular momentum in low flow gyrokinetics"

1. "Full f simulations . . . , on the other hand, employ a gyrokinetic quasineutrality equation that has not been derived to high enough order in an ion

gyroradius over scale length expansion to solve for the correct long wavelength radial electric field .... The problem is that the radial electric field profile and the toroidal rotation in the plasma are uniquely related to each other. Due to axisymmetry, the toroidal rotation is determined exclusively by the radial transport of toroidal angular momentum, contained in the small off-diagonal components of the viscosity and Reynolds stress. Obtaining these small terms makes the calculation extremely challenging and not feasible when using the gyrokinetic quasineutrality equation."

2. "... the $O(\delta_i^2 f_{M,i})$ correction to the ion distribution function is not self-consistently calculated in general magnetic geometries."

   Remark: As of that time, the full H_2, including all effects of magnetic inhomogeneity, had not even been derived, let alone implemented in a code. The calculation was first done by Parra & Calvo (2011).

L.11 Excerpt from Scott & Smirnov (2010) — "Energetic consistency and momentum conservation in the gyrokinetic description of tokamak plasmas"

1. "The main results are valid for any ordering scheme which might be used since their demonstration does not depend on ordering, but uses the general functional form of the Hamiltonian on the field variables... ."

2. "...for small scale fluctuation ordering, the magnetic flux term (A_φ/c), which is formally of order ϵ^{-1} ..., does not introduce terms in H at higher order than already are necessary to evaluate the momentum transport equation. This result ... is sufficient to allay recent concerns about the integrity of the treatment of momentum conservation and transport by gyrokinetics which have been voiced by others (Parra & Catto, 2008)."

   Remark: As we have discussed, a third-order Hamiltonian is required for a complete calculation of the momentum fluxes, so the concerns of Parra & Catto remain. If the cancellation referred to by Scott & Smirnov had not occurred, one would have had to work to fourth order.

L.12 Excerpt from Brizard & Tronko (2011) — "Exact momentum conservation laws for the gyrokinetic Vlasov-Poisson equations"

"The major difference between previous works ... and our work is that the gyrokinetic polarization ... considered here includes contributions from the guiding-center and gyrocenter phase-space transformations, while previous works have neglected the guiding-center polarization contribution (which is of the same order as the gyrocenter contribution)."

L.13 Excerpt from Parra & Calvo (2011) — "Phase-space Lagrangian derivation of electrostatic gyrokinetics in general geometry"

"The complete calculation [of the gyrokinetic Hamiltonian] to order ϵ^2 has not been done for a general static magnetic field ... so far. In the most common Lagrangian formulation..., the calculation is done in two steps: first, the turbulent electromagnetic fields are ignored and only the background magnetic field is considered, giving the drift kinetic equation; in the second step, the turbulent electromagnetic fields are added and the corresponding corrections are calculated. ... In the first step, the equations are expanded in the small parameter $\epsilon \sim \rho/L$, whereas in the second step, they are expanded in $\epsilon_\phi \sim Ze\phi/Mv_t^2$. The expansion in ϵ is only performed to first order because the next order results are very tedious to calculate. The expansion in ϵ_ϕ is continued to second order because the pieces quadratic in ϕ are needed to have an energy-like invariant. In the expansion in ϵ_ϕ, the fact that there has been a previous expansion in ϵ is ignored, and as a result the terms of order $\epsilon\,\epsilon_\phi$ are never calculated. The missing terms of order ϵ^2 and $\epsilon\,\epsilon_\phi$ are comparable to the terms of order ϵ_ϕ^2 according to the gyrokinetic ordering..., making this expansion consistent only when $\epsilon_\phi \gg \epsilon$. In addition, since the cross-terms that contain both the background magnetic field and the turbulent electrostatic potential, of order ϵ^2 and $\epsilon\,\epsilon_\phi$, are always neglected when the two-step method is presented, it is not obvious how to calculate them following that procedure. In this paper, we present the complete phasespace Lagrangian calculation with the standard gyrokinetic ordering..., emphasizing the selfconsistent calculation of the terms of order ϵ^2 and $\epsilon\,\epsilon_\phi$. In the gyrokinetic equations that result from the new Lagrangian, the magnetic geometry effects and the fluctuating potential appear together in the second-order terms, showing that geometry and turbulence cannot be separated and dealt with independently. Our main result is [an] explicit expression for the second-order gyrokinetic Hamiltonian .... It clearly exhibits the interplay between geometry and turbulence inherent to gyrokinetic theory, possessing terms of three types: terms quadratic in the electrostatic potential, terms that include both the electrostatic potential and the magnetic geometry and terms that are purely geometrical."

> **Remark:** Parra & Calvo calculate the second-order Hamiltonian using a particular choice of representation (symplectic at first order; Hamiltonian at second order). For a discussion of equivalent representations, see Brizard & Tronko (2012).

L.14 Excerpt from Calvo & Parra (2012) — "Long-wavelength limit of gyrokinetics in a turbulent tokamak and its intrinsic ambipolarity"

1. "The gyrokinetic equations have typically been solved only for the turbulent components of the distribution function and the electrostatic potential (we restrict our discussion to electrostatic gyrokinetics), but in recent years growing supercomputer capabilities have motivated an increasing interest in the extension of gyrokinetic calculations to longer wavelengths and transport time scales. However, at least for [an axisymmetric] tokamak, this is a subtle issue, as F. I. Parra and P. J. Catto have discussed in a series of papers [Parra & Catto (2008, 2009c,a, 2010b,a,c)]. The main lines of the argument can be stated in a succinct way. The perpendicular component of the long-wavelength piece of the plasma velocity depends on the long-wavelength radial electric field through the $\boldsymbol{E} \times \boldsymbol{B}$ drift. The momentum conservation equation can be used to obtain the three components of the velocity, and from it, derive the radial electric field. The plasma velocity is to lowest order parallel to the flux surfaces because the radial particle drift is small. Then, the poloidal and toroidal components of the momentum conservation equation are sufficient to calculate the velocity to the order of interest, and by decomposing it in parallel and perpendicular components, the radial electric field can be obtained by making the perpendicular component equal to the $\boldsymbol{E} \times \boldsymbol{B}$ drift plus the diamagnetic velocity. The poloidal component of the velocity is strongly damped by collisions because the poloidal direction is not a direction of symmetry. The poloidal velocity is determined by setting the collisional viscosity in the poloidal direction equal to zero, giving a poloidal velocity proportional to the ion temperature gradient unless collisionality is really small and turbulence can compete with the collisional damping . . . . Unfortunately, the toroidal component of the momentum equation that would give the toroidal component of the velocity and completely determine the radial electric field is identically satisfied to order ϵ^2 by any toroidal velocity. . . . Since gyrokinetic equations are customarily derived and solved to order ϵ, the tokamak long-wavelength radial electric field cannot be correctly obtained from the standard set of gyrokinetic equations available in the literature.

   "In the limit in which the velocity is of the order of the diamagnetic velocity, known as low flow limit, the calculation of the radial flux of toroidal angular momentum, which we need to compute the radial electric field, is especially demanding because this flux is smaller than the radial flux of particles and energy in the expansion in ϵ. The low flow limit is relevant in the study of intrinsic rotation . . . . In [Parra & Catto (2010c); Parra et al. (2011a)], a method to calculate the toroidal angular momentum conservation equation in the low flow limit to the order in which it is not identically zero is proposed. With the toroidal angular momentum equation to this

order, it is possible to obtain the toroidal rotation and hence calculate the radial electric field. The formula for the radial flux of toroidal angular momentum in [Parra & Catto (2010c); Parra et al. (2011a)] is given as a sum of several integrals over the first- and second-order pieces of the distribution functions and the electrostatic potential. To avoid calculating these second-order pieces in complete detail, a subsidiary expansion in $B_p/B \ll 1$ was employed, where B_p is the poloidal magnetic field and B is the total magnetic field. With the derivation for the first time of the gyrokinetic equations and change of coordinates in general magnetic geometry up to second order [Parra & Calvo (2011)], it has become possible to calculate the second-order pieces without resorting to a subsidiary expansion. In this article, we present the equations that need to be solved to obtain the long-wavelength second order pieces.

"Carrying the expansion to second order in ϵ at long wavelengths also clarifies the issues with the radial electric field ... pointed out at the beginning of this introduction. Along with the derivation of the equations we give an explicit proof of the indeterminacy of the radial electric field, showing that it cannot be found from the long-wavelength gyrokinetic Fokker–Planck and quasineutrality equations correct to second order. This property, known as intrinsic ambipolarity, was first proven for neoclassical transport ... and it was shown to hold for turbulent tokamaks in [Parra & Catto (2009c)] using the identical cancellation of the toroidal angular momentum conservation equation to the order of interest. The intrinsic ambipolarity of purely turbulent particle fluxes was shown to hold in [Sugama et al. (1996)], even electromagnetically and in general magnetic geometry (that is why the long-wavelength radial electric field in non-quasisymmetric stellarators is determined from neoclassical theory). This is, however, the first direct, explicit, and general proof for turbulent tokamaks. Instead of resorting to the toroidal angular momentum equation, we write the long-wavelength equations order by order and show that they can be solved for any radial electric field, leaving it undetermined. Those readers who are familiar with the Chapman–Enskog results on the derivation of fluid equations from kinetic theory ... will find that the approach that we adopt at some stages of the proof is very similar. The analogy becomes especially clear in Section 5.2. In previous sections the long-wavelength Fokker-Planck and quasineutrality equations have been derived up to second order. In Section 5.2 we inspect the second-order piece of the long-wavelength Fokker-Planck equation and learn that it possesses solvability conditions, i.e. the existence of solutions of this equation imposes constraints on lowest-order quantities. These constraints are transport equations for particle and energy density. The way of obtaining them and of showing that we have actually found all the solvability conditions are the aspects particularly reminiscent of the Chapman-Enskog techniques."

2. "... we have given a complete proof that the long-wavelength tokamak

radial electric field cannot be determined by simply using Fokker-Planck and quasineutrality equations accurate to second order in the gyrokinetic expansion parameter. In other words, we have proven that gyrokinetics does not spoil the well-known neoclassical intrinsic ambipolarity property of the tokamak."

M Notation

M.1 Basic physics symbols

– a –

a — minor radius

$\boldsymbol{A}(\boldsymbol{x}, t)$ — Vector potential. $\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}$.

α — 1-form.

– b –

$\widehat{\boldsymbol{b}}$ — Unit vector in direction of magnetic field: $\hat{b} \doteq \boldsymbol{B}/B$.

b_φ — Toroidal (covariant) component of $\widehat{\boldsymbol{b}}$: $b_\varphi \doteq \widehat{\boldsymbol{b}} \cdot \boldsymbol{e}_\varphi$.

$\boldsymbol{B}(\boldsymbol{x}, t), B(\boldsymbol{x}, t)$ — Magnetic field. $B \doteq |\boldsymbol{B}|$.

$\boldsymbol{B}_*$ — Effective magnetic field in gyrokinetics: $\boldsymbol{B}_* \doteq \boldsymbol{B} + (mc/q)U\boldsymbol{\nabla} \times \widehat{\boldsymbol{b}}$.

– c –

c — Speed of light.

c_s — Sound speed: $c_\mathrm{s} \doteq (ZT_e/M)^{1/2}$.

$C_s[f]$ — The nonlinear collision operator for species s. $\partial_t f + \cdots = -C_s[f]$, where $C_s \doteq \sum_{s'} C_{s,s'}$.

– d –

δA — The fluctuation of A around the mean: $\delta A \doteq A - \langle A \rangle$.

δF — Deviation of the gyrocenter PDF F from a given background F_0: $F = F_0 + \delta F$.

$\delta(x - y)$ — Dirac delta function.

$\delta f/\delta\eta(1)$ — Functional derivative of $f[\eta]$ with respect to η. The fundamental derivative is $\delta\eta(1)/\delta\eta(1') = \delta(1 - 1')$.

D — Diffusion coefficient.

D_B — Bohm diffusion coefficient: $D_B \doteq cT/eB$.

D_gB — Gyro-Bohm diffusion coefficient: $D_\mathrm{gB} \doteq (\rho_\mathrm{s}/L)D_B$.

$\chi_\perp$ — Perpendicular dielectric permittivity: $\chi_\perp \doteq \rho_\mathrm{s}^2/\lambda_{De}^2 = \omega_{pi}^2/\omega_{ci}^2 \gg 1$.

– e –

e — A number approximately equal to 2.7 (Candlestickmaker, 1972).

e — Magnitude of electronic charge. $q_e = -e$.

$\boldsymbol{e}_i$ — Basis vector: $\boldsymbol{e}_i \doteq \partial_i \boldsymbol{x}$. A vector $\boldsymbol{v}$ is represented as $\boldsymbol{v} = v^i \boldsymbol{e}_i$.

$\boldsymbol{e}^i$ — Cobasis vector: $\boldsymbol{e}^i \doteq \boldsymbol{\nabla} z^i$. A covector (1-form) $\boldsymbol{\alpha}$ is represented as $\boldsymbol{\alpha} = \alpha_i \boldsymbol{e}^i$.

ϵ — Gyrokinetic ordering parameter: $\epsilon \doteq \rho_{\mathrm{s}}/L$, where L is a macroscopic perpendicular scale length.

$\boldsymbol{E}(\boldsymbol{x}, t),\ E$ — Electric field vector; $E \doteq |\boldsymbol{E}|$.

– f –

$f(\boldsymbol{x}, \boldsymbol{v}, t)$ — Particle distribution function.

$\overline{F}(\overline{\boldsymbol{X}}, \overline{U}, \overline{\mu}, t)$ — Gyrocenter distribution function.

$\mathcal{F}$ — Gyrocenter distribution function in extended phase space.

$f_M(\boldsymbol{v})$ — (Absolute) Maxwellian distribution: $f_M(\boldsymbol{v}) \doteq (2\pi v_t^2)^{-\frac{3}{2}} \exp(-v^2/2v_t^2)$.

$f_{\mathrm{lM}}(\boldsymbol{x}, \boldsymbol{v}, t)$ — Local Maxwellian. $f_{\mathrm{lM}} \doteq [n(\boldsymbol{x}, t)/\overline{n}][2\pi v_t^2(\boldsymbol{x}, t)]^{-3/2} \exp[-[\boldsymbol{v} - \boldsymbol{u}(\boldsymbol{x}, t)]^2/2v_t^2(\boldsymbol{x}, t)]$, where $v_t^2(\boldsymbol{x}, t) \doteq T(\boldsymbol{x}, t)/m$.

– g –

g — Metric tensor. $dl^2 = g_{ij} dz^i\, dz^j$.

γ — Poincaré–Cartan differential 1-form: $\gamma \doteq p \cdot dq - H\, dt$.

– h –

H — Hamiltonian.

H_n — nth-order term in the expansion of H in ϵ.

$\mathcal{H}$ — Gyrocenter Hamiltonian in extended phase space.

– i –

i_X — Interior product with the vector field X.

I — Identity matrix or operator.

– j –

$\boldsymbol{j}(\boldsymbol{x}, t)$ — Current density.

J_0 — Bessel function of order zero.

J — Jacobian of the transformation to generalized spatial coordinates.

$\mathcal{J}$ — Jacobian of the gyrocenter phase-space transformation.

J — Poisson tensor. The equations of motion are $\dot{z} = \mathsf{J} \cdot \partial_z H \equiv \{z, H\}$ (for time-independent transformations).

– k –

k_D — Debye wave number: $k_D^2 \doteq \sum_s k_{\mathrm{D}s}^2$, where $k_{\mathrm{D}s}^2 \doteq 4\pi (ne^2)_s / T_s$.

– l –

λ_D — The Debye length: $\lambda_D \doteq k_D^{-1}$.

$L_n,\ L_T$ — Density and temperature scale lengths: $L_n^{-1} \doteq -\partial_x \ln n$, $L_T^{-1} \doteq -\partial_x \ln T$.

L_w — Lie derivative in the direction of the vector field w.

$\mathcal{L}_w$ — Identical to the vector field w. The Lie transformation is $\mathrm{T} = \exp(w) \equiv \exp(\mathcal{L}_w)$.

$\mathcal{L}_n$ — $\epsilon^n w_n$, where w_n is the nth-order vector field.

– m –

$m_e,\ m_i$ — Electron and ion masses.

μ, μ_0 — Magnetic moment, and its lowest-order approximation.

$\overline{\mu}$ — First adiabatic invariant: $\overline{\mu} = \mu_0 + \epsilon \mu_1 + \cdots$.

$\mathcal{M}$ — Mach number.

– n –

$\overline{n},\ \overline{n}_e,\ \overline{n}_i$ — Mean number density: $\overline{n} \doteq N/V$.

$n_s(\boldsymbol{x}, t)$ — Fluid density. $n_s(\boldsymbol{x}, t) = \int d\boldsymbol{v}\, \overline{n}_s f_s(\boldsymbol{x}, \boldsymbol{v}, t)$.

$N,\ N_s$ — Total number of particles.

$\overline{N}$ — Mean gyrocenter number density. $\overline{N} = \overline{n}$.

ν — Collision frequency.

– ω –

ω — Fundamental differential 2-form: $\omega = d\gamma$; Lagrange tensor.

$\omega_c,\ \omega_{ce},\ \omega_{ci}$ — The gyrofrequency. $\omega_{\mathrm{c}s} \doteq q_s B / m_s c$.

– o –

$O(\epsilon)$ — Asymptotically of order ϵ.

– p –

p — Canonical momentum.

P — The scalar pressure. In the Braginskii equations, $P = nT$.

$\mathcal{P}(\boldsymbol{x}, t)$ — Stress tensor; see $\mathsf{P}(\boldsymbol{x}, t)$.

ψ — Flux (radial) coordinate.

φ — Toroidal angle.

ϕ — Electrostatic potential.

$\boldsymbol{\mathcal{P}}$ — Polarization vector. $\rho = \rho^G - \boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{P}}$.

– q –

q_s — Charge of species s. $q_e = -e$.

– r –

R — Major radius.

$\rho(\boldsymbol{x}, t)$ — Charge density. $\rho(\boldsymbol{x}, t) = \int d\boldsymbol{v}\, (\overline{n} q)_s f_s(\boldsymbol{x}, \boldsymbol{v}, t)$.

$\rho_e,\ \rho_i$ — Gyroradii: $\rho_\nu \doteq v_{t\nu}/\omega_{c\nu}$ for $\nu = e, i$.

ρ_{s} — Sound radius: $\rho_{\mathrm{s}} \doteq c_{\mathrm{s}}/\omega_{ci}$. (The s subscript is in Roman, not italic.)

ρ — Particle charge density. $\rho = \rho^G - \boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{P}}$. The quasineutrality condition is $\rho = 0$.

ρ^G — Gyrocenter charge density.

$\boldsymbol{\rho}$ — Gyroradius vector.

$\boldsymbol{\rho}_\epsilon$ — Lie-transformed gyroradius vector.

$\boldsymbol{\rho}_{\mathrm{pol}}$ — Polarization charge density.

– s –

s — Species label.

– t –

T_s — Temperature of species s.

T — Lie transformation operator (or sometimes a symmetry transformation).

T^* — Pullback operator (on forms) induced from T by replacing $\mathcal{L}$ by L.

T_* — Pushforward operator (on vectors). In the practical gyrokinetic manipulations, $\mathrm{T}_* = (\mathrm{T}^*)^{-1}$.

– u –

$\boldsymbol{u}_s(\boldsymbol{x}, t)$ — Fluid velocity. $(n\boldsymbol{u})_s(\boldsymbol{x}, t) = \int d\boldsymbol{v}\, \overline{n}_s \boldsymbol{v} f_s(\boldsymbol{x}, \boldsymbol{v}, t)$.

$\boldsymbol{u}_E$ — Fluid $\boldsymbol{E} \times \boldsymbol{B}$ velocity.

U — Parallel gyrocenter velocity.

− v −

$\boldsymbol{v}$ — Particle velocity.

$v_t,\ v_{te},\ v_{ti}$ — Thermal velocity: $v_t \doteq (T/m)^{1/2}$.

V — Volume of system.

V' — Derivative of volume inside a flux surface. The flux-surface average of a divergence is $\langle \boldsymbol{\nabla} \cdot \boldsymbol{A} \rangle = V'^{-1} \partial_\psi (V' \langle A^\psi \rangle)$.

V_{ds} — Diamagnetic speed. $V_{ds} \doteq -cT_s/e_s B L_n$.

$\boldsymbol{V}_{\rm pol}$ — Polarization drift velocity: $\boldsymbol{V}_{\rm pol} \doteq \omega_c^{-1} \partial_t (c\boldsymbol{E}_\perp / B)$.

$\boldsymbol{V}_d$ — Gyrocenter drift.

− w −

w — Vector field: $w = w^i \partial_i$. The basic Lie transformation is $\overline{z} = \mathrm{T} z$ with $\mathrm{T} = \exp(w)$ and $w^i = 1 + O(\epsilon)$.

− x −

$\boldsymbol{x},\ \boldsymbol{y}$ — Vector positions in configuration space.

$x,\ y,\ z$ — Cartesian components of $\boldsymbol{x}$.

$\boldsymbol{X}$ — Gyrocenter position (lowest order).

− y −

− z −

Z — Atomic number; see e_i. Charge neutrality demands that $Z\overline{n}_i = \overline{n}_e$.

ζ — Gyrophase.

M.2 Miscellaneous notation

A^i — Contravariant components (or components of a vector).

A_i — Covariant components (or components of a 1-form).

$\widetilde{A}$ — The tilde implies that A is random or that it has dependence on gyrophase.

A^* — Pullback (in differential geometry).

A_* — Pushforward (in differential geometry).

$\langle A \rangle$ — Ensemble average, or statistical mean, of A: $A = \langle A \rangle + \delta A$. Also, flux-surface average.

$\langle\langle x\, y\, z \rangle\rangle$ — Cumulant of x, y, and z.

$A[f]$ — Indicates functional dependence of A on f.

$[A]$ — Dimensions of A.

$A^{[n]}$ — Bracketed superscripts indicated that the quantity is nominally $O(\epsilon^n)$, without considering additional possible cancellations due to averaging.

$\widehat{\boldsymbol{k}}$ — Carets generally denote unit vectors: $\widehat{\boldsymbol{b}} \doteq \boldsymbol{B}/|\boldsymbol{B}|$.

$\doteq$ — Definition.

$\equiv$ — Equivalent to: $\doteq \equiv \stackrel{\text{def}}{=}$.

$\binom{n}{k}$ — Binomial coefficient: $\binom{n}{k} \doteq n!/[k!\,(n-k)!]$.

References

I. G. Abel, G. G. Plunk, E. Wang, M. Barnes, S. C. Cowley, W. Dorland, & A. A. Schekochihin, Multiscale gyrokinetics for rotating tokamak plasmas: Fluctuations, transport and energy flows, Rep. Prog. Phys., October 2013, in press; http://arxiv.org/abs/1209.4782.

V. I. Arnold, *Mathematical Methods of Classical Mechanics*, Springer, New York, 1978, Translated by K. Vogtmann and A. Weinstein.

M. A. Barnes, F. I. Parra, J. P. Lee, E. A. Belli, M. F. F. Nave, & A. E. White, Intrinsic rotation driven by non-Maxwellian equilibria in tokamak plasmas, Phys. Rev. Lett., **111**:055005 (5 pages), 2013, http://link.aps.org/doi/10.1103/PhysRevLett.111.055005.

S. I. Braginskii, Transport processes in a plasma, in *Reviews of Plasma Physics, Vol. 1*, edited by M. N. Leontovich, pp. 205–311, Consultants Bureau, New York, 1965.

A. J. Brizard & T. S. Hahm, Foundations of nonlinear gyrokinetic theory, Rev. Mod. Phys., **79**:421–68, 2007, http://link.aps.org/doi/10.1103/RevModPhys.79.421.

A. Brizard & Tronko, Exact momentum conservation laws for the gyrokinetic Vlasov-Poisson equations, Phys. Plasmas, **18**:082307 (14 pages), 2011, http://dx.doi.org/10.1063/1.3625554.

A. J. Brizard & N. Tronko, Equivalent representations of higher-order Hamiltonian guiding-center theory, Phys. Plasmas, 2012, in preparation.

A. Brizard, *Nonlinear Gyrokinetic Tokamak Physics*, PhD thesis, Princeton University, 1990.

A. Brizard, Nonlinear gyrofluid description of turbulent magnetized plasmas, Phys. Fluids B, **4**:1213–28, 1992, http://dx.doi.org/10.1063/1.860129.

A. J. Brizard, Variational principle for nonlinear gyrokinetic Vlasov–Maxwell equations, Phys. Plasmas, **7**:4816–22, 2000, http://dx.doi.org/10.1063/1.1322063.

A. J. Brizard, A geometric view of Hamiltonian perturbation theory, Phys. Lett. A, **291**:146–9, 2001, http://www.sciencedirect.com/science/article/pii/S0375960101007228#.

A. J. Brizard, A guiding-center Fokker-Planck collision operator for nonuniform magnetic fields, Phys. Plasmas, **11**:4429–38, 2004, http://dx.doi.org/10.1063/1.1780532.

R. V. Budny, Comparisons of predicted plasma performance in ITER H-mode plasmas with various mixes of external heating, Nucl. Fusion, **49**:085008 (14 pages), 2009, http://dx.doi.org/10.1088/0029-5515/49/8/085008.

J. W. Burby & H. Qin, Gyrosymmetry: Global considerations, Phys. Plasmas, **19**:052106 (8 pages), 2012, http://dx.doi.org/10.1063/1.4719700.

J. W. Burby, A. I. Zhmoginov, & H. Qin, The Hamiltonian mechanics of stochastic acceleration, 2013, submitted to Phys. Rev. Lett.

I. Calvo & F. I. Parra, Long-wavelength limit of gyrokinetics in a turbulent tokamak and its intrinsic ambipolarity, Plasma Phys. Control. Fusion, **54**:115007 (33 pages), 2012, http://dx.doi.org/10.1088/0741-3335/54/11/115007.

Y. Camenen, Y. Idomura, S. Jolliet, & A. G. Peeters, Consequences of profile shearing on toroidal momentum transport, Nucl. Fusion, **51**:073039 (11 pages), 2011, http://stacks.iop.org/NF/51/073039.

S. Candlestickmaker, On the imperturbability of elevator operators. LVII, Q. Jl R. astr. Soc., **13**:63–6, 1972, http://articles.adsabs.harvard.edu/cgi-bin/nph-iarticle_query?1972QJRAS..13...63C&data_type=PDF.

J. R. Cary & R. G. Littlejohn, Noncanonical Hamiltonian mechanics and its application to magnetic field line flow, Ann. Phys. (N.Y.), **151**:1–34, 1983, http://dx.doi.org/10.1016/0003-4916(83)90313-5.

J. R. Cary, Lie transform perturbation theory for Hamiltonian systems, Physics Rep., **79**:129–59, 1981, http://dx.doi.org/10.1016/0370-1573(81)90175-7.

P. J. Catto, A. N. Simakov, F. I. Parra, & G. Kagan, Electrostatic turbulence in tokamaks on transport time scales, Plasma Phys. Control. Fusion, **50**:115006 (21 pages), 2008, http://dx.doi.org/10.1088/0741-3335/50/11/115006.

P. J. Catto, F. I. Parra, G. Kagan, & A. N. Simakov, Limitations, insights and improvements to gyrokinetics, Nucl. Fusion, **49**:095026 (8 pages), 2009, http://dx.doi.org/10.1088/0029-5515/49/9/095026.

P. J. Catto, Linearized gyro-kinetics, Plasma Phys., **20**:719–22, 1978, http://dx.doi.org/10.1088/0032-1028/20/7/011.

S. Chandrasekhar, *Plasma Physics*, U. of Chicago Press, Chicago, 1960.

J. S. deGrassie, R. J. Groebner, K. H. Burrell, & W. M. Solomon, Intrinsic toroidal velocity near the edge of DIII-D H-mode plasmas, Nucl. Fusion, **49**:085020 (14 pages), 2009, http://dx.doi.org/10.1088/0029-5515/49/8/085020.

W. D. D'haeseleer, W. N. G. Hitchon, J. D. Callen, & J. L. Shohet, *Flux Coordinates and Magnetic Field Structure*, Springer, Berlin, 1991.

A. M. Dimits & B. I. Cohen, Collision operators for partially linearized particle simulation codes, Phys. Rev. E, **49**:709–21, 1994, http://link.aps.org/doi/10.1103/PhysRevE.49.709.

A. J. Dragt & J. M. Finn, Lie series and invariant functions for analytic symplectic maps, J. Math. Phys., **17**:2215–27, 1976, http://dx.doi.org/10.1063/1.522868.

D. H. E. Dubin & J. A. Krommes, Stochasticity, superadiabaticity, and the theory of adiabatic invariants and guiding center motion, in *Long Time Prediction in Dynamics*, edited by W. Horton, L. E. Reichl, & V. G. Szebehely, pp. 257–80, Wiley, New York, 1982.

D. H. E. Dubin, J. A. Krommes, C. R. Oberman, & W. W. Lee, Nonlinear gyrokinetic equations, Phys. Fluids, **26**:3524–35, 1983, http://dx.doi.org/10.1063/1.864113.

M. Fecko, *Differential Geometry and Lie Groups for Physicists*, Cambridge U. Press, Cambridge, 2006.

H. Flanders, *Differential Forms with Applications to the Physical Sciences*, Academic Press, New York, 1963.

E. A. Frieman & L. Chen, Nonlinear gyrokinetic equations for low-frequency electromagnetic waves in general plasma equilibria, Phys. Fluids, **25**:502–8, 1982, http://dx.doi.org/10.1063/1.863762.

X. Garbet, Y. Idomura, L. Villard, & T. H. Watanabe, Gyrokinetic simulations of turbulent transport, Nucl. Fusion, **50**:043002 (30 pages), 2010, http://dx.doi.org/10.1088/0029-5515/50/4/043002.

H. Goldstein, *Classical Mechanics*, Addison–Wesley, Reading, Massachusetts, 1951.

M. J. Gotay & J. A. Isenberg, The symplectization of science, Gazette des Mathématiciens, **54**:59–79, 1992, http://www.pims.math.ca/~gotay/Symplectization(E).pdf.

T. S. Hahm, Nonlinear gyrokinetic equations for tokamak microturbulence, Phys. Fluids, **31**:2670–3, 1988, http://dx.doi.org/10.1063/1.866544.

R. D. Hazeltine, Rotation of a toroidally confined, collisional plasma, Phys. Fluids, **17**:961–8, 1974, http://dx.doi.org/10.1063/1.1694838.

F. L. Hinton & C. W. Horton, Jr., Amplitude limitation of a collisional drift wave instability, Phys. Fluids, **14**:116–23, 1971, http://dx.doi.org/10.1063/1.1693260.

G. Hu & J. A. Krommes, Generalized weighting scheme for δf particle-simulation method, Phys. Plasmas, **1**:863–74, 1994, http://dx.doi.org/10.1063/1.870745.

Y. Idomura, Accuracy of momentum transport calculations in full-f gyrokinetic simulations, Comp. Sci. Discovery, **5**:014018 (12 pages), 2012, http://dx.doi.org/10.1088/1749-4699/5/1/014018.

J. D. Jackson, *Classical Electrodynamics*, Wiley, New York, 1962.

A. N. Kaufman, Dissipative Hamiltonian systems: A unifying principle, Phys. Lett. A, **100**:419–22, 1984, http://www.sciencedirect.com/science/article/pii/0375960184906340.

J. A. Krommes, W. W. Lee, & C. Oberman, Equilibrium fluctuation energy of gyrokinetic plasma, Phys. Fluids, **29**:2421–5, 1986, http://dx.doi.org/10.1063/1.865534.

J. A. Krommes, Fundamental statistical theories of plasma turbulence in magnetic fields, Phys. Rep., **360**:1–352, 2002, http://dx.doi.org/10.1016/S0370-1573(01)00066-7.

J. A. Krommes, Nonequilibrium gyrokinetic fluctuation theory and sampling noise in gyrokinetic particle-in-cell simulations, Phys. Plasmas, **14**:090501 (26 pages), 2007, http://dx.doi.org/10.1063/1.2759879.

J. A. Krommes, Comment on "Guiding center plasma models in three dimensions" [Phys. Plasmas 15, 092112 (2008)], Phys. Plasmas, **16**:084701 (2 pages), 2009, http://dx.doi.org/10.1063/1.3206670.

J. A. Krommes, Nonlinear gyrokinetics: A powerful tool for the description of microturbulence in magnetized plasmas, Phys. Scripta, **T142**:014035 (13 pages), 2010, http://dx.doi.org/10.1088/0031-8949/2010/T142/014035.

J. A. Krommes, The gyrokinetic description of microturbulence in magnetized plasmas, Annu. Rev. Fluid Mech., **44**:203–25, 2012, http://dx.doi.org/10.1146/annurev-fluid-120710-101223.

J. A. Krommes, The physics of the second-order gyrokinetic MHD Hamiltonian: μ conservation, Galilean invariance, and ponderomotive potential, Phys. Plasmas, 2013, submitted.

R. Kubo, Generalized cumulant expansion method, J. Phys. Soc. Japan, **17**:1100–20, 1962, http://jpsj.ipap.jp/link?JPSJ/17/1100/.

R. M. Kulsrud, General stability theory in plasma physics, in *Proceedings of the International School of Physics Enrico Fermi, Course XXV, Advanced Plasma Theory (Varenna, Italy, 1962)*, New York, 1964, Academic Press.

R. M. Kulsrud, MHD description of plasma, in *Handbook of Plasma Physics*, edited by M. N. Rosenbluth & R. Z. Sagdeev, volume I: Basic Plasma Physics 1, edited by A. A. Galeev and R. N. Sudan, chapter 1.4, pp. 115–45, North-Holland, 1983.

W. W. Lee & R. A. Kolesnikov, On higher order corrections to gyrokinetic vlasov-poisson equations in the long wavelength limit, Phys. Plasmas, **16**:044506 (4 pages), 2009, http://dx.doi.org/10.1063/1.3117482.

W. W. Lee, Gyrokinetic approach in particle simulation, Phys. Fluids, **26**:556–62, 1983, http://dx.doi.org/10.1063/1.864140.

B. Li & D. R. Ernst, Gyrokinetic Fokker-Planck collision operator, Phys. Rev. Lett., **106**:195002 (4 pages), 2011, http://link.aps.org/doi/10.1103/PhysRevLett.106.195002.

R. G. Littlejohn, A guiding center Hamiltonian: A new approach, J. Math. Phys., **20**:2445–58, 1979, http://dx.doi.org/10.1063/1.524053.

R. G. Littlejohn, Hamiltonian perturbation theory in noncanonical coordinates, J. Math. Phys., **23**:742–47, 1982, http://dx.doi.org/10.1063/1.525429.

R. G. Littlejohn, Variational principles of guiding centre motion, J. Plasma Phys., **29**:111–25, 1983, http://dx.doi.org/10.1017/S002237780000060X.

R. G. Littlejohn, Geometry and guiding center motion, in *Contemporary Mathematics*, edited by J. E. Marsden, volume 28, pp. 151–67, American Mathematical Society, Providence, RI, 1984, http://dx.doi.org/10.1090/conm/028/751980.

N. D. Mermin, *Boojums All The Way Through: Communicating Science in a Prosaic Age*, Cambridge University Press, Cambridge, 1990.

A. Mishchenko & A. J. Brizard, Higher-order energy-conserving gyrokinetic theory, Phys. Plasmas, **18**:022305 (13 pages), 2011, http://dx.doi.org/10.1063/1.3552143.

C. W. Misner & J. A. Wheeler, Classical physics as geometry, Ann. Phys., **2**:525–603, 1957, http://dx.doi.org/10.1016/0003-4916(57)90049-0.

C. W. Misner, K. S. Thorne, & J. A. Wheeler, *Gravitation*, W. H. Freeman and Company, San Francisco, 1973.

P. Morrison, Bracket formulation for irreversible classical fields, Phys. Lett. A, **100**:423–7, 1984, http://www.sciencedirect.com/science/article/pii/0375960184906352.

M. F. F. Nave, T. Johnson, L.-G. Eriksson, K. Crombé, C. Giroud, M.-L. Mayoral, J. Ongena, A. Salmi, T. Tala, & M. Tsalas, Influence of magnetic field ripple on the intrinsic rotation of tokamak plasmas, Phys. Rev. Lett., **105**:105005 (4 pages), 2010, http://link.aps.org/doi/10.1103/PhysRevLett.105.105005.

F. I. Parra & I. Calvo, Phase-space Lagrangian derivation of electrostatic gyrokinetics in general geometry, Plasma Phys. Control. Fusion, **53**:045001 (49 pages), 2011, http://dx.doi.org/10.1088/0741-3335/53/4/045001.

F. I. Parra & P. J. Catto, Limitations of gyrokinetics on transport time scales, Plasma Phys. Control. Fusion, **50**:065014 (23 pages), 2008, http://dx.doi.org/10.1088/0741-3335/50/6/065014.

F. I. Parra & P. J. Catto, Comment on "On higher order corrections to gyrokinetic Vlasov-Poisson equations in the long wavelength limit" [Phys. Plasmas **16**, 044506 (2009)], Phys. Plasmas, **16**:124701 (3 pages), 2009, http://dx.doi.org/10.1063/1.3272151.

F. I. Parra & P. J. Catto, Gyrokinetic equivalence, Plasma Phys. Control. Fusion, **51**:065002 (9 pages), 2009, http://dx.doi.org/10.1088/0741-3335/51/6/065002.

F. I. Parra & P. J. Catto, Vorticity and intrinsic ambipolarity in turbulent tokamaks, Plasma Phys. Control. Fusion, **51**:095008 (38 pages), 2009, http://dx.doi.org/10.1088/0741-3335/51/9/095008.

F. I. Parra & P. J. Catto, Non-physical momentum sources in slab geometry gyrokinetics, Plasma Phys. Control. Fusion, **52**:085011 (12 pages), 2010, http://dx.doi.org/10.1088/0741-3335/52/8/085011.

F. I. Parra & P. J. Catto, Transport of momentum in full f gyrokinetics, Phys. Plasmas, **17**:056106 (11 pages), 2010, http://dx.doi.org/10.1063/1.3327127.

F. I. Parra & P. J. Catto, Turbulent transport of toroidal angular momentum in low flow gyrokinetics, Plasma Phys. Control. Fusion, **52**:045004 (32 pages), 2010, http://dx.doi.org/10.1088/0741-3335/52/4/045004. Errata: *ibid.* 059801.

F. I. Parra, M. Barnes, & P. J. Catto, Sources of intrinsic rotation in the low-flow ordering, Nucl. Fusion, **51**:113001 (10 pages), 2011, http://dx.doi.org/10.1088/0029-5515/50/4/043002.

F. I. Parra, M. Barnes, & A. G. Peeters, Up-down symmetry of the turbulent transport of toroidal angular momentum in tokamaks, Phys. Plasmas, **18**:062501 (14 pages), 2011, http://dx.doi.org/10.1063/1.3586332.

F. I. Parra, M. F. F. Nave, A. A. Schekochihin, C. Giroud, J. S. de Grassie, J. H. F. Severo, P. de Vries, K.-D. Zastrow, & JET-EFDA Contributors, Scaling of spontaneous rotation with temperature and plasma current in tokamaks, Phys. Rev. Lett., **108**:095001 (5 pages), 2012, http://link.aps.org/doi/10.1103/PhysRevLett.108.095001.

F. I. Parra, *Extension of gyrokinetics to transport time scales*, PhD thesis, Massachusetts Institute of Technology, 2009, http://arxiv.org/abs/1309.7385.

A. G. Peeters, C. Angioni, A. Bortolon, Y. Camenen, F. J. Casson, B. Duval, L. Fiederspiel, W. A. Hornsby, Y. Idomura, T. Hein, N. Kluy, P. Mantica, F. I. Parra, A. P. Snodin, G. Szepesi, D. Strintzi, T. Tala, G. Tardini, P. de Vries, & J. Weiland, Overview of toroidal momentum transport, Nucl. Fusion, **51**:094027 (13 pages), 2011, http://dx.doi.org/10.1088/0029-5515/51/9/094027.

H. Qin & W. M. Tang, Pullback transformations in gyrokinetic theory, Phys. Plasmas, **11**:1052–63, 2004, http://dx.doi.org/10.1063/1.1640626.

J. E. Rice, A. Ince-Cushman, J. S. deGrassie, L.-G. Eriksson, Y. Sakamoto, A. Scarabosio, A. Bortolon, K. H. Burrell, B. P. Duval, C. Fenzi-Bonizec, M. J. Greenwald, R. J. Groebner, G. T. Hoang, Y. Koide, E. S. Marmar, A. Pochelon, & Y. Podpaly, Inter-machine comparison of intrinsic toroidal rotation in tokamaks, Nucl. Fusion, **47**:1618–24, 2007, http://dx.doi.org/10.1088/0029-5515/47/11/025.

M. N. Rosenbluth, P. H. Rutherford, J. B. Taylor, E. A. Frieman, & L. M. Kovrizhnikh, Neoclassical effects on plasma equilibria and rotation, in *4th Conference on Plasma Physics and Controlled Thermonuclear Fusion*, pp. 495–508, Vienna, 1971, IAEA.

P. H. Rutherford & E. A. Frieman, Drift instabilities in general magnetic field configurations, Phys. Fluids, **11**:569–85, 1968, http://dx.doi.org/10.1063/1.1691954http://dx.doi.org/10.1063/1.1691954.

B. Scott & J. Smirnov, Energetic consistency and momentum conservation in the gyrokinetic description of tokamak plasmas, Phys. Plasmas, **17**:112302 (20 pages), 2010, http://dx.doi.org/10.1063/1.3507920.

B. D. Scott, Nonlinear polarization and dissipative correspondence between low-frequency fluid and gyrofluid equations, Phys. Plasmas, **14**:102318 (12 pages), 2007, http://dx.doi.org/10.1063/1.2783993.

L. Spitzer, Jr., *Physics of Fully Ionized Gases*, Interscience, New York, 1962.

G. M. Staebler & H. E. St. John, Predicted toroidal rotation enhancement of fusion power production in ITER, Nucl. Fusion, **46**:L6–8, 2006, http://dx.doi.org/10.1088/0029-5515/46/8/L02.

T. Stoltzfus-Dueck, Transport-driven toroidal rotation in the tokamak edge, Phys. Rev. Lett., **108**:065002 (5 pages), 2012, http://link.aps.org/doi/10.1103/PhysRevLett.108.065002.

H. Sugama & W. Horton, Nonlinear electromagnetic gyrokinetic equation for plasmas with large mean flows, Phys. Plasmas, **5**:2560–73, 1998, http://dx.doi.org/10.1063/1.872941.

H. Sugama, M. Okamoto, W. Horton, & M. Wakatani, Transport processes and entropy production in toroidal plasmas with gyrokinetic electromagnetic turbulence, Phys. Plasmas, **3**:2379–94, 1996, http://dx.doi.org/10.1063/1.871922.

H. Sugama, T. H. Watanabe, M. Nunami, & S. Nishimura, Momentum balance and radial electric fields in axisymmetric and nonaxisymmetric toroidal plasmas, Plasma Phys. Control. Fusion, **53**:024004 (17 pages), 2011, http://dx.doi.org/10.1088/0741-3335/53/2/024004.

H. Sugama, Gyrokinetic field theory, Phys. Plasmas, **7**:466–80, 2000, http://dx.doi.org/10.1063/1.873832.

L. E. Sugiyama, Guiding center plasma models in three dimensions, Phys. Plasmas, **15**:092112 (13 pages), 2008, http://dx.doi.org/10.1063/1.2972158.

L. E. Sugiyama, Response to "Comment on 'Guiding center plasma models in three dimensions'" [Phys. Plasmas 15, 092112 (2008)], Phys. Plasmas, **16**:084702 (4 pages), 2009, http://dx.doi.org/10.1063/1.3206671.

J. B. Taylor & R. J. Hastie, Stability of general plasma equilibria—I. Formal theory, Plasma Phys., **10**:479–94, 1968, http://dx.doi.org/10.1088/0032-1028/10/5/301.

J. B. Taylor, Magnetic moment under short-wave electrostatic perturbations, Phys. Fluids, **10**:1357–9, 1967, http://dx.doi.org/10.1063/1.1762289.

G. F. van Heijst & H. J. H. Clercx, Studies on qusi-2D turbulence—the effect of boundaries, Fluid Dyn. Res., **41**:064002 (18 pages), 2009, http://dx.doi.org/10.1088/0169-5983/41/6/064002.

Wikipedia:Electret, http://en.wikipedia.org/wiki/Electret.

Wikipedia:Noether's theorem, http://en.wikipedia.org/wiki/Noether's_theorem.

Index

Boldface indicates a major topic;
underlining indicates a definition;
italic indicates an example.