PHYSICS OF PLASMAS VOLUME 11, NUMBER 3 MARCH 2004

Pullback transformations in gyrokinetic theory

H. Qin and W. M. Tang
Plasma Physics Laboratory, Princeton University, Princeton, New Jersey 08543

(Received 21 May 2003; accepted 19 November 2003)

The pullback transformation of the distribution function is a key component of gyrokinetic theory. In this paper, a systematic treatment of this subject is presented, and results from applications of the uniform framework developed are reviewed. The focus is on providing a clear exposition of the basic formalism which arises from the existence of three distinct coordinate systems in gyrokinetic theory. The familiar gyrocenter coordinate system, where the gyromotion is decoupled from the rest of particle's dynamics, is noncanonical and nonfibered. For the phase space (cotangent bundle T^*M) associated with a configuration space M, a nonfibered coordinate system $(\boldsymbol{X}, \boldsymbol{V})$ is a coordinate system where $\boldsymbol{X}$ is not necessarily the coordinates for the configuration space M, and $\boldsymbol{V}$ is not necessarily the coordinates for the cotangent fiber $T_{\boldsymbol{x}}^*M$ at each $\boldsymbol{x}$. On the other hand, Maxwell's equations, which are needed to complete a kinetic system, are initially only defined in the fibered laboratory phase space coordinate system. The pullback transformations provide a rigorous connection between the distribution functions in gyrocenter coordinates and Maxwell's equations in laboratory phase space coordinates. This involves the generalization of the usual moment integrals originally defined on the cotangent fiber of the phase space to the moment integrals on a general six-dimensional symplectic manifold. The resultant systematic treatment of the moment integrals enabled by the pullback transformation is shown to be an important step in the proper formulation of gyrokinetic theory. Without this vital element, a number of prominent physics features, such as the presence of the compressional Alfvén wave and a proper description of the gyrokinetic equilibrium, cannot be readily recovered.
[DOI: 10.1063/1.1640626]

I. INTRODUCTION

Most of the interesting plasmas in laboratory and space environments involve the presence of magnetic field. The particle's motion in a magnetized equilibrium consists of a fast gyromotion and a slow guiding center motion. It is this fast gyromotion which restricts the allowable time step in particle simulations of the associated dynamics in the laboratory phase space coordinate frame. Over the past 20 years, a primary goal of the gyrokinetic theory developed formalism has been to remove the fast gyromotion from the kinetic system for low frequency and long parallel wavelength phenomena.[1–15] Progress in this area has enabled gyrokinetic particle simulations, which use a much larger time step than the time scale of gyromotion,[7,16–22] to be successfully applied in studies of the transport problems of fusion plasmas. In particular, gyrokinetic theory offers a simplified version of the Vlasov–Maxwell system by utilizing the fact that in strongly magnetized plasmas the particle's gyroradius is much smaller than the scale length of the magnetic field: $\epsilon_B \equiv |\rho/L_B| \ll 1$, where $L_B \equiv |B/\nabla \boldsymbol{B}|$. More fundamentally, gyrokinetic theory requires the construction of a gyrocenter coordinate system in which the particle's gyromotion is decoupled from the rest of the particle dynamics. The Vlasov–Maxwell equation system can then be derived in this special coordinate system.[23–28] Guiding center coordinates are employed in the magnetostatic case, while gyrocenter coordinates are employed when there are electromagnetic perturbations in the system. Modern gyrokinetic theory[12–15,23–28] utilizing noncanonical Hamiltonian and phase space Lie perturbation method[4–6] has been carefully established over a number of years. It not only sets up a rigorous and systematic foundation for the gyrokinetic framework, but also clarifies numerous confusing concepts and introduces much more physics content into the theory. Specifically, gyrokinetic theory has been extended to deal with arbitrary frequency, arbitrary wavelength, electromagnetic perturbations in general geometries.[23–28]

One of the key components of modern gyrokinetic theory is the pullback transformation of the distribution function. Premature versions of the pullback transformation have appeared in various formats in the course of the development of the gyrokinetic theory over the last 40 years. It is now clear that the pullback transformation is responsible for many important physics properties in gyrokinetic theory. Prominent examples include the polarization drift density in the gyrokinetic Poisson equation,[7] the polarization current in the gyrokinetic Ampère's law which accounts for the compressional Alfvén wave,[25] and the pressure balance equation for gyrokinetic equilibrium.[27] In this paper, a systematic treatment of this subject will be presented together with a brief summary of key results from applications of the uniform framework developed.

In Sec. II, the general theory of coordinate transformations and their pullback transformations in phase space are developed. The concept of nonfibered phase space coordinates is introduced and the phase space moment integral is

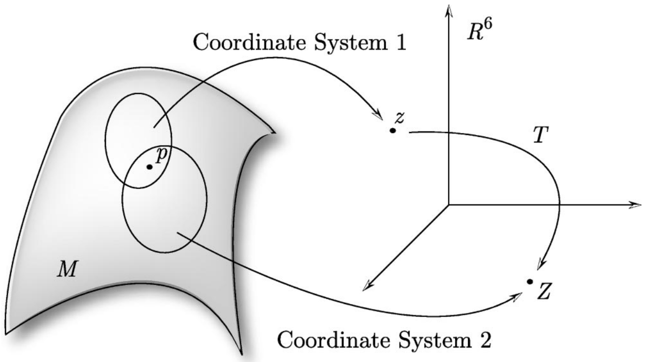

FIG. 1. In Sec. III, we investigate the physics of guiding center pullback transformation. The gyrocenter pullback transformation is studied in Sec. IV.

generalized into a parameterized integral of a moment 6-form in the 6D phase space. It is emphasized that the theoretical formulation in Sec. II is much more general than the gyrokinetic theory. For example, it applies to other physics problems involving phase space coordinate transformation, e.g., Vlasov–Maxwell system for periodically focused charge particle beams. In Sec. III, the pullback associated with the guiding center transformation is described with emphasis placed on the underlying physics. The pullback associated with the nonsymplectic gyrocenter transformation needed to deal with time dependent electromagnetic field is presented in Sec. IV. Finally, in Sec. V, the general implications of the pullback transformation for gyrokinetic theory are summarized.

II. COORDINATE TRANSFORMATIONS AND THEIR PULLBACK TRANSFORMATIONS IN PHASE SPACE

A. Fibered phase space coordinate system

Physics is geometric and independent of coordinates, even though it can be more efficiently described with the help of coordinates. All possible choices of coordinate systems should be equivalent in terms of analyzing the physics of interest. If two coordinate systems are connected through a transformation, then the physics content must be invariant with respect to the transformation. However, the mathematical involvement of different coordinate systems is indeed different when describing the same physics. For a given physics problem, the natural first step is to find the most efficient coordinate system. This can usually be constructed by imposing the desired mathematical structures. More often than not, it is constructed by perturbations around an obvious choice of coordinate system through a near identity coordinate transformation. In this sense, perturbation methods in physics are really about the quest for useful coordinates. In the present analysis, attention will be focused on the coordinate transformations in the 6D phase space for a single nonrelativistic classical particle. A coordinate transformation for the phase space P of dimension 6 can be locally represented by a map between two subsets of the R^6 space, $T:\boldsymbol{z}\mapsto\boldsymbol{Z}=T(\boldsymbol{z})$. As illustrated in Fig. 1, for the same point p in phase space, there could be more than one coordinate system. The correspondence between two different coordinate systems for the same point in phase space is the coordinate transformation. In the present study, it is assumed that a coordinate transformation can be represented by a single map almost everywhere. The subset of coordinates which cannot be covered by the single map has zero measure and does not contribute to the moment integrals.

Kinetic theory deals with a particle distribution function f, which is a function defined on the phase space P, $f:P\rightarrow R$. In addition, kinetic theory in its common form implicitly makes use of the fact that the phase space is the cotangent space (cotangent bundle) of a configuration space (manifold) M, $P=T^*M$. We call a coordinate system $(\boldsymbol{x},\boldsymbol{v})$ a fiber coordinate system if $\boldsymbol{x}$ is the coordinate for M, and $\boldsymbol{v}$ is the coordinate for the cotangent fiber T_x^*M at $\boldsymbol{x}$. Note that the symbol "$\boldsymbol{v}$" is used to represent the cotangent fiber T_x^*M at $\boldsymbol{x}$. It can be viewed as a short form of $\boldsymbol{p}/m$, where $\boldsymbol{p}$ is the momentum and m is the relativistic mass, if $\boldsymbol{p}$ is used to represent the cotangent coordinates. A fiber coordinate system for the cotangent bundle of the laboratory configuration space will be referred to as the laboratory phase space coordinate system. The familiar moment integrals are actually fiber integrals of moment functions $q:P\rightarrow R$ performed on the cotangent fiber $T_{\boldsymbol{x}}^*M$ at each $\boldsymbol{x}$. In the laboratory phase space coordinate system $(\boldsymbol{x},\boldsymbol{v})$, the moment integral of a moment function $q(\boldsymbol{x},\boldsymbol{v})$ has the form

$$q(\boldsymbol{x})=\int_{T_{\boldsymbol{x}}^*M}q(\boldsymbol{x},\boldsymbol{v})f(\boldsymbol{x},\boldsymbol{v})d^3\boldsymbol{v}. \tag{1}$$

The moment integrals yield results which are functions on the laboratory configuration space M. These integrals are themselves independent of the coordinate systems used for the cotangent fiber $T_{\boldsymbol{x}}^*M$ at each $\boldsymbol{x}$. In other words, $q(\boldsymbol{x})$ is invariant under a fiber coordinate transformation, i.e., a coordinate transformation that transforms only the fiber coordinate $\boldsymbol{v}$. Letting $\varphi_{\boldsymbol{x}}:\boldsymbol{v}\rightarrow\boldsymbol{V}$ be a fiber coordinate transformation which in general depends on $\boldsymbol{x}$, the invariance of $q(\boldsymbol{x})$ can be expressed as

$$q(\boldsymbol{x})=\int_{T_{\boldsymbol{x}}^*M}q(\boldsymbol{x},\boldsymbol{v})f(\boldsymbol{x},\boldsymbol{v})d^3\boldsymbol{v}=\int_{T_{\boldsymbol{x}}^*M}Q(\boldsymbol{x},\boldsymbol{V})F(\boldsymbol{x},\boldsymbol{V})d^3\boldsymbol{V}, \tag{2}$$

where $q(\boldsymbol{x},\boldsymbol{v})f(\boldsymbol{x},\boldsymbol{v})d^3\boldsymbol{v}$ and $Q(\boldsymbol{x},\boldsymbol{V})F(\boldsymbol{x},\boldsymbol{V})d^3\boldsymbol{V}$, are the representations of the same 3-forms in $(\boldsymbol{x},\boldsymbol{v})$ and $(\boldsymbol{x},\boldsymbol{V})$, respectively. Here $d^3\boldsymbol{v}$ and $d^3\boldsymbol{V}$ can be viewed as the volume forms of $T_{\boldsymbol{x}}^*M$ at each $\boldsymbol{x}$. For example, in a magnetized plasma, the well-known $(v_\parallel,\mu,\theta)$ velocity coordinates, as well as the Cartesian velocity coordinates, can be used to calculate the same moment integrals,

$$\begin{aligned}q(\boldsymbol{x})&=\int q(\boldsymbol{x},v_1,v_2,v_3)f(\boldsymbol{x},v_1,v_2,v_3)dv_1\wedge dv_2\wedge dv_3\\&=\int\frac{B_\parallel}{m}Q(\boldsymbol{x},v_\parallel,\mu,\phi)F(\boldsymbol{x},v_\parallel,\mu,\phi)dv_\parallel\wedge d\mu\wedge d\theta.\end{aligned} \tag{3}$$

In this example, the volume form $dv_1\wedge dv_2\wedge dv_3$ is canonical whereas the volume form $(B_\parallel/m)dv_\parallel\wedge d\mu\wedge d\theta$ is noncanonical. The most noticeable difference between them is that the former is a constant and the latter depends $\boldsymbol{x}$.

B. Nonfibered phase space coordinate system

In gyrokinetic theory, however, useful coordinate systems are nonfibered. A nonfibered coordinate system ($\boldsymbol{X}$, $\boldsymbol{V}$) is a coordinate system where $\boldsymbol{X}$ is not necessarily the coordinates for the configuration space M, and $\boldsymbol{V}$ is not necessarily the coordinates for the cotangent fiber $T_{\boldsymbol{x}}^{*}M$ at each $\boldsymbol{x}$. A nonfibered coordinate transformation, by definition, transfers a fibered coordinate system into a nonfibered one. In the context of gyrokinetic theory, ($\boldsymbol{X}$, $\boldsymbol{V}$) can be either guiding center coordinates or the gyrocenter coordinates, both of which are nonfibered. The construction of the guiding center coordinates and the gyrocenter coordinates will be described in detail in Secs. III and IV. The discussion in this section applies to any general nonfibered coordinate systems and nonfibered transformations.

No matter which nonfibered coordinate system is used, the moment integrals are still defined on the cotangent fiber $T_{\boldsymbol{x}}^{*}M$ at each $\boldsymbol{x}$, and $q(\boldsymbol{x})$ should be invariant under such a general nonfibered coordinate transformation. For the new coordinate system ($\boldsymbol{X}$, $\boldsymbol{V}$) to be useful, it is necessary to know the construction of $q(\boldsymbol{x})$ in it. To be specific, the scenario studied in this paper is that the distribution function $f(\boldsymbol{x},\boldsymbol{v})$ is known in the transformed nonfibered coordinate system ($\boldsymbol{X}$, $\boldsymbol{V}$) as $F(\boldsymbol{X},\boldsymbol{V})$, whereas $q(\boldsymbol{x},\boldsymbol{v})$ as a physical quantity, such as the position and the velocity, is only meaningfully defined in the laboratory phase space coordinate system ($\boldsymbol{x}$, $\boldsymbol{v}$). Given $q(\boldsymbol{x},\boldsymbol{v})$ and $F(\boldsymbol{X},\boldsymbol{V})$, there are two methods to calculate $q(\boldsymbol{x})$. The first method is to pull back the distribution function $F(\boldsymbol{X},\boldsymbol{V})$ into $f(\boldsymbol{x},\boldsymbol{v})$, and the second method depends on the generalization of the concept of moment integrals.

The first method, where the distribution function $F(\boldsymbol{X},\boldsymbol{V})$ is pulled back[31] into $f(\boldsymbol{x},\boldsymbol{v})$, can be written as

$$q(\boldsymbol{x})=\int_{T_{\boldsymbol{x}}^{*}M} q(\boldsymbol{x},\boldsymbol{v})\varphi^{*}[F(\boldsymbol{X},\boldsymbol{V})]d^{3}\boldsymbol{v}, \tag{4}$$

where

$$\varphi^{*}[F(\boldsymbol{X},\boldsymbol{V})]=F(\boldsymbol{X}(\boldsymbol{x},\boldsymbol{v}),\boldsymbol{V}(\boldsymbol{x},\boldsymbol{v}))=f(\boldsymbol{x},\boldsymbol{v}). \tag{5}$$

In the second method, we consider the generalization of the usual moment integrals originally defined on the cotangent fiber $T_{\boldsymbol{x}}^{*}M$ at each $\boldsymbol{x}$ to the moment integrals on a general 6D symplectic manifold P. This can be accomplished by two different approaches. The first approach is to view $T_{\boldsymbol{x}}^{*}M$ as an orientable 3-subset of P and $q(\boldsymbol{x})$ as an integral of a moment 3-form λ over such an 3-subset,

$$q(\boldsymbol{x})=\int_{T_{\boldsymbol{x}}^{*}M}\lambda. \tag{6}$$

In the laboratory phase space coordinates ($\boldsymbol{x}$, $\boldsymbol{v}$), the moment 3-form λ is defined as

$$\lambda=q(\boldsymbol{x},\boldsymbol{v})f(\boldsymbol{x},\boldsymbol{v})d^{3}\boldsymbol{v}, \tag{7}$$

where $d^{3}\boldsymbol{v}$ is the volume form for $T_{\boldsymbol{x}}^{*}M$ at a fixed $\boldsymbol{x}$. In a general nonfibered coordinate system $(\boldsymbol{X},\boldsymbol{V})=\varphi(\boldsymbol{x},\boldsymbol{v})$, λ is pulled back from its form in ($\boldsymbol{x}$, $\boldsymbol{v}$),

$$\begin{aligned}\Lambda&=\varphi^{-1*}\lambda=\varphi^{-1*}[q(\boldsymbol{x},\boldsymbol{v})f(\boldsymbol{x},\boldsymbol{v})d^{3}\boldsymbol{v}]\\&=[\varphi^{-1*}q(\boldsymbol{x},\boldsymbol{v})][\varphi^{-1*}f(\boldsymbol{x},\boldsymbol{v})][\varphi^{-1*}d^{3}\boldsymbol{v}],\end{aligned} \tag{8}$$

where

$$\varphi^{-1*}q(\boldsymbol{x},\boldsymbol{v})=Q(\boldsymbol{X},\boldsymbol{V})\equiv q(\boldsymbol{x}(\boldsymbol{X},\boldsymbol{V}),\boldsymbol{v}(\boldsymbol{X},\boldsymbol{V})), \tag{9}$$

$$\varphi^{-1*}f(\boldsymbol{x},\boldsymbol{v})=F(\boldsymbol{X},\boldsymbol{V})\equiv f(\boldsymbol{x}(\boldsymbol{X},\boldsymbol{V}),\boldsymbol{v}(\boldsymbol{X},\boldsymbol{V})). \tag{10}$$

Here it has been assumed that the transformation φ is a diffeomorphism (one–one onto and smooth), and φ^{-1*} [or $(\varphi^{-1})^{*}$] is the pullback associated with φ^{-1}, which maps any function on ($\boldsymbol{x}$, $\boldsymbol{v}$) into a function on ($\boldsymbol{X}$, $\boldsymbol{V}$). φ^{-1*} can also be called the pushforward associated with φ if φ is a diffeomorphism. Therefore, in a general nonfibered coordinate system ($\boldsymbol{X}$, $\boldsymbol{V}$), $q(\boldsymbol{x})$ can be expressed as

$$\begin{aligned}q(\boldsymbol{x})&=\int_{T_{\boldsymbol{x}}^{*}M}\lambda=\int_{T_{\boldsymbol{x}}^{*}M}\varphi^{-1*}\lambda\\&=\int_{U=\{(\boldsymbol{X},\boldsymbol{V})|\boldsymbol{x}(\boldsymbol{X},\boldsymbol{V})=\mathrm{const}\}}\Lambda\\&=\int_{U=\{(\boldsymbol{X},\boldsymbol{V})\ |\boldsymbol{x}(\boldsymbol{X},\boldsymbol{V})=\mathrm{const}\}}[\varphi^{-1*}q(\boldsymbol{x},\boldsymbol{v})]F(\boldsymbol{X},\boldsymbol{V})\\&\quad\times[\varphi^{-1*}d^{3}\boldsymbol{v}].\end{aligned} \tag{11}$$

If the coordinates $\boldsymbol{v}=(v_1,v_2,v_3)$ for $T_{\boldsymbol{x}}^{*}M$ are canonical,

$$d^{3}\boldsymbol{v}=dv_1\wedge dv_2\wedge dv_3, \tag{12}$$

then

$$\varphi^{-1*}[d^{3}\boldsymbol{v}]=dv_1(\boldsymbol{X},\boldsymbol{V})\wedge dv_2(\boldsymbol{X},\boldsymbol{V})\wedge dv_3(\boldsymbol{X},\boldsymbol{V}). \tag{13}$$

There are practical difficulties associated with using Eq. (11) to calculate $q(\boldsymbol{x})$. First, the pullback of the volume form $\varphi^{-1*}d^{3}\boldsymbol{v}$ has 20 terms in general, because the dimension of a general 3-form in the 6D phase space is 20. Second, the integration domain expressed in the ($\boldsymbol{X}$, $\boldsymbol{V}$) coordinate system is complicated.

To get around these difficulties, the concept of a moment integral can be generalized by a different approach. Specifically, a moment integral is generalized into a parameterized integral of a moment 6-form $\lambda_{\boldsymbol{r}}$ in the 6D phase space

$$i(\boldsymbol{r})=\int\lambda_{\boldsymbol{r}}, \tag{14}$$

$$\lambda_{\boldsymbol{r}}=i(\boldsymbol{r},\boldsymbol{z})f(\boldsymbol{z})d^{6}\boldsymbol{z}, \tag{15}$$

where $\boldsymbol{z}=(\boldsymbol{x},\boldsymbol{v})$ is phase space coordinate, $\boldsymbol{r}$ is a set of independent parameters, $i(\boldsymbol{r},\boldsymbol{z})$ is a moment function of the phase space and the parameters $\boldsymbol{r}$, and $d^{6}\boldsymbol{z}$ is the Liouville volume from given by the symplectic structure ω in the phase space

$$d^{6}\boldsymbol{z}=\frac{-1}{3!}\,\omega\wedge\omega\wedge\omega. \tag{16}$$

In a canonical coordinate system

$$\omega=\sum_{i=1}^{3}dx^{i}\wedge dv_i, \tag{17}$$

$$d^{6}\boldsymbol{z}=dx^{1}\wedge dx^{2}\wedge dx^{3}\wedge dv_1\wedge dv_2\wedge dv_3. \tag{18}$$

Under a nonfibered coordinate transformation φ, $i(\boldsymbol{r})$ is obtained through the pullback of the $\boldsymbol{r}$-parameterized 6-form $\lambda_{\boldsymbol{r}}$,

$$i(\boldsymbol{r})=\int \varphi^{-1*}\lambda_{\boldsymbol{r}}=\int \varphi^{-1*}[i(\boldsymbol{r},\boldsymbol{z})f(\boldsymbol{z})d^6\boldsymbol{z}]$$
$$=\int \varphi^{-1*}[i(\boldsymbol{r},\boldsymbol{z})f(\boldsymbol{z})]d^6\boldsymbol{Z}, \tag{19}$$

where $d^6\boldsymbol{Z}$ is the Liouville volume form of $(\boldsymbol{X}, \boldsymbol{V})$

$$d^6\boldsymbol{Z}=\varphi^{-1*}d^6\boldsymbol{z}. \tag{20}$$

If $(\boldsymbol{x},\boldsymbol{v})=(x_1,x_2,x_3,v_1,v_2,v_3)$ is a canonical coordinate system, $d^6\boldsymbol{Z}$ can be straightforwardly expressed as

$$d^6\boldsymbol{Z}=\varphi^{-1*}d^6\boldsymbol{z}$$
$$=dx^1(\boldsymbol{X},\boldsymbol{V})\wedge dx^2(\boldsymbol{X},\boldsymbol{V})\wedge dx^3(\boldsymbol{X},\boldsymbol{V})\wedge dv_1(\boldsymbol{X},\boldsymbol{V})\wedge dv_2$$
$$\times(\boldsymbol{X},\boldsymbol{V})\wedge dv_3(\boldsymbol{X},\boldsymbol{V}). \tag{21}$$

But in almost all the cases of practical interest, the Liouville volume form $d^6\boldsymbol{Z}=\varphi^{-1*}d^6\boldsymbol{z}$ of $(\boldsymbol{X}, \boldsymbol{V})$ is more conveniently calculated through the pullback of the symplectic structure

$$d^6\boldsymbol{Z}=\frac{-1}{3!}\Omega\wedge\Omega\wedge\Omega, \tag{22}$$

$$\Omega=\varphi^{-1*}\omega. \tag{23}$$

One major difference between a canonical volume form and a noncanonical one is that the canonical volume form is a constant where the noncanonical volume form is generally a function of the phase space through its dependence on field variables. For the case of the gyrocenter coordinate system which will be discussed in Sec. IV, the volume form nontrivially depends on the perturbed electromagnetic fields. The usual moment integrals are special cases of the generalized moment integrals when

$$i(\boldsymbol{r},\boldsymbol{z})=\delta(\boldsymbol{x}-\boldsymbol{r})q(\boldsymbol{z}). \tag{24}$$

That is

$$q(\boldsymbol{r})=\int_{T_r^*M}q(\boldsymbol{r},\boldsymbol{v})f(\boldsymbol{r},\boldsymbol{v})d\boldsymbol{v}$$
$$=\int \delta(\boldsymbol{x}-\boldsymbol{r})q(\boldsymbol{x},\boldsymbol{v})f(\boldsymbol{x},\boldsymbol{v})d^6\boldsymbol{z}. \tag{25}$$

From Eq. (19), the construction of $q(\boldsymbol{r})$ using the distribution function in the nonfibered coordinate system $F(\boldsymbol{X},\boldsymbol{V})$ is

$$q(\boldsymbol{r})=\int \varphi^{-1*}[\delta(\boldsymbol{x}-r)q(\boldsymbol{x},\boldsymbol{v})f(\boldsymbol{x},\boldsymbol{v})d^6\boldsymbol{z}]$$
$$=\int \varphi^{-1*}[\delta(\boldsymbol{x}-\boldsymbol{r})]\varphi^{-1*}[q(\boldsymbol{x},\boldsymbol{v})]$$
$$\times\varphi^{-1*}[f(\boldsymbol{x},\boldsymbol{v})]d^6\boldsymbol{Z}$$
$$=\int \delta(\boldsymbol{x}(\boldsymbol{X},\boldsymbol{V})-\boldsymbol{r})Q(\boldsymbol{X},\boldsymbol{V})F(\boldsymbol{X},\boldsymbol{V})d^6\boldsymbol{Z}. \tag{26}$$

Equations (4), (11), and (26) are equivalent and can be used interchangeably to simplify the calculation. In practice, the pullbacks involved are often associated with coordinate perturbation transformations, and thus can be further simplified by the perturbation techniques adopted. For example, the term $\delta(\boldsymbol{x}(\boldsymbol{X},\boldsymbol{V})-\boldsymbol{r})$ can be Taylor expanded in terms of the small perturbation parameter such that the integration in Eq. (26) can be carried out order by order. The connection between the single particle dynamics and the moment integral through the symplectic structure as shown in Eqs. (20), (22), and (23) is of significant importance. First of all, the volume form needed to perform phase space integrals is given by the same symplectic form governing the single particle dynamics. No metric is needed for the gyrokinetic Vlasov–Maxwell system, and it is purely symplectic. Second, the Liouville volume form $d^6\boldsymbol{Z}=\varphi^{-1*}d^6\boldsymbol{z}$ of $(\boldsymbol{X}, \boldsymbol{V})$ will assume the same functional form as $d^6\boldsymbol{z}$ of $(\boldsymbol{x}, \boldsymbol{v})$, and the calculation of phase space integral in Eq. (26) is simplified, if φ^{-1*} preserves the symplectic structure of $(\boldsymbol{x}, \boldsymbol{v})$

$$\Omega(\boldsymbol{Z})=\varphi^{-1*}\omega(\boldsymbol{z})=[\omega(\boldsymbol{z})]_{\boldsymbol{z}\rightarrow\boldsymbol{Z}}. \tag{27}$$

If no obvious phase space structure exists, Eq. (4) will be preferred. Such examples will be presented in Secs. III and IV in the context of the gyrokinetic theory.

III. PULLBACK ASSOCIATED WITH THE GUIDING CENTER COORDINATE TRANSFORMATION

This section deals with the pullback associated with the guiding center transformation, which is a necessary theoretical construction for analyzing the gyrokinetic equilibrium.[27] Such an analysis is required for a proper understanding of magnetized plasmas in equilibrium using the guiding center coordinates. The gyrokinetic equilibrium is of fundamental importance for the widely adopted perturbative gyrokinetic particle simulation (δf method),[18–22] where the equilibrium distribution function and the electromagnetic field are assumed to be known. Gyrokinetic equilibria consistent with the well-studied fluid ones are obviously necessary for the perturbative gyrokinetic particle simulations to be reliable. In particular, recent numerical studies of equilibria with zonal flows[22] raise again the question of how to describe the equilibrium flow from the gyrokinetic point of view. The essence of the problem studied here is how to relate the measurable quantities in the laboratory frame to the information in the guiding center coordinates. Given a distribution function $F(\boldsymbol{X},V_\parallel,\mu)$ in the guiding-center coordinates $\boldsymbol{Z}=(\boldsymbol{X},V_\parallel,\mu,\xi)$, how are the fluid density, flow, and current calculated? Do the macroscopic field variables calculated from the gyrokinetic formalism satisfy the fluid equations obtained by taking the moments of the Vlasov equation in the laboratory phase space coordinates? The pullback formulas derived in Sec. II give answers to these questions, when applied to the guiding center transformation.

First, the guiding center transformation $G:\boldsymbol{z}\mapsto\boldsymbol{Z}$, which transforms the laboratory phase space coordinates $\boldsymbol{z}=(\boldsymbol{x},\boldsymbol{v})$ into the guiding center coordinates $\boldsymbol{Z}=(\boldsymbol{X},V_\parallel,\mu,\xi)$ is needed. It turns out that almost all the important terms in the equilibrium equations can be generated by applying the pullback formulas to the simple leading order guiding center transformation

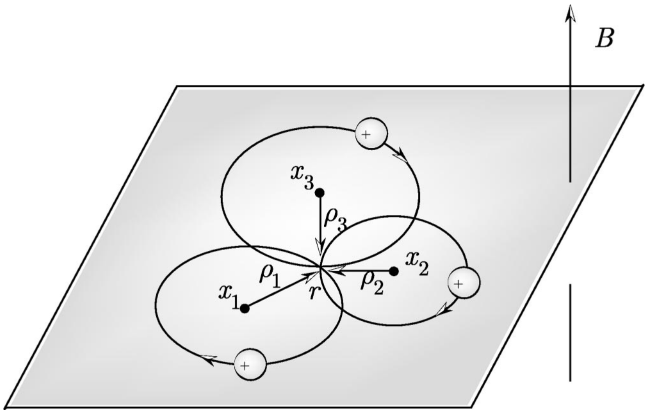

FIG. 2. The physics of the pullback formula.

$$
\begin{aligned}
&\boldsymbol{X}=\boldsymbol{x}-\boldsymbol{\rho}_0+O(\epsilon),\\
&V_{\parallel}=v_{\parallel}+O(\epsilon),\\
&\mu=\mu_0+O(\epsilon),\\
&\xi=\theta+O(\epsilon),
\end{aligned}
\tag{28}
$$

where $(\boldsymbol{x},v_{\parallel},v_{\perp},\theta)$ is the usual laboratory phase space coordinates. $\boldsymbol{\rho}_0$ and μ_0, defined in particle coordinates, are the usual gyroradius and magnetic moment. θ is chosen such that $\hat{\mathbf{v}}_{\perp}=-e/|e|(\boldsymbol{e}_x\sin\theta+\boldsymbol{e}_y\cos\theta)$. $\boldsymbol{e}_x$ and $\boldsymbol{e}_y$ are two perpendicular directions in the configuration space, and $(\boldsymbol{e}_x,\boldsymbol{e}_y,\boldsymbol{b})$ is a right-handed orthogonal frame.

For the guiding center transformation $G:\boldsymbol{z}\mapsto\boldsymbol{Z}$, Eq. (26) becomes

$$
\begin{aligned}
q(\boldsymbol{r})&=\int q(\boldsymbol{r},\boldsymbol{v})f(\boldsymbol{r},\boldsymbol{v})d^3\boldsymbol{v}\\
&=\int q(\boldsymbol{z})f(\boldsymbol{z})\delta(\boldsymbol{x}-\boldsymbol{r})d^6\boldsymbol{z}\\
&=\int G^{-1}*[q(\boldsymbol{z})\delta(\boldsymbol{x}-\boldsymbol{r})]F(\boldsymbol{Z})d^6\boldsymbol{Z}\\
&=\int Q(\boldsymbol{Z})\delta[\boldsymbol{X}+\boldsymbol{\rho}-\boldsymbol{r}]F(\boldsymbol{Z})d^6\boldsymbol{Z},
\end{aligned}
\tag{29}
$$

where it is assumed that the guiding center transformation G is a diffeomorphism (one–one onto and smooth) almost everywhere, and

$$
\begin{aligned}
&d^6\boldsymbol{Z}\equiv B_{\parallel}^*/m\,d^3\boldsymbol{X}dV_{\parallel}d\mu d\xi,\\
&B_{\parallel}^*=\boldsymbol{b}\cdot\boldsymbol{B}^*,\quad \boldsymbol{B}^*=\boldsymbol{B}+\frac{cmV_{\parallel}}{e}\nabla\times\boldsymbol{b},\\
&Q(\boldsymbol{Z})=G^{-1}*q(\boldsymbol{z}),\\
&\boldsymbol{\rho}=G^{-1}*\boldsymbol{\rho}_0.
\end{aligned}
\tag{30}
$$

The physics encapsulated in the pullback formula Eq. (29) is that an observable $q(\boldsymbol{r})$ at a certain location $\boldsymbol{r}$ in the laboratory frame is the average of its microscopic counterpart expressed in the guiding center coordinates $Q(\boldsymbol{Z})$ over nearby guiding centers with $\boldsymbol{X}(\boldsymbol{Z})+\boldsymbol{\rho}(\boldsymbol{Z})=\boldsymbol{r}$. This is illustrated in Fig. 2, where three examples of such guiding centers are shown. For the number density in laboratory phase space coordinates, we use $q(\boldsymbol{z})=1$ and $G^{-1}*1=1$,

$$
\begin{aligned}
n(\boldsymbol{r})&=\int \delta(\boldsymbol{X}+\boldsymbol{\rho}-\boldsymbol{r})F(\boldsymbol{Z})d^6\boldsymbol{Z}\\
&=\int \delta(\boldsymbol{X}-\boldsymbol{r})F(\boldsymbol{Z})d^6\boldsymbol{Z}+O(\epsilon^2)\\
&=2\pi\int F(\boldsymbol{Z})B_{\parallel}^*/m\,dV_{\parallel}d\mu\Big|_{\boldsymbol{X}\mapsto\boldsymbol{r}}+O(\epsilon^2),
\end{aligned}
\tag{31}
$$

where "$|_{\boldsymbol{X}\mapsto\boldsymbol{r}}$" means replacing $\boldsymbol{X}$ by $\boldsymbol{r}$.

For the fluid velocity in laboratory phase space coordinates $\boldsymbol{u}(\boldsymbol{r})$, we have $q(\boldsymbol{z})=\boldsymbol{v}=\dot{\boldsymbol{x}}$, $G^{-1}*\boldsymbol{v}=\dot{\boldsymbol{X}}+\dot{\boldsymbol{\rho}}(\boldsymbol{X})+O(\epsilon^2)$, and

$$
\begin{aligned}
\boldsymbol{u}(\boldsymbol{r})&=\int (\dot{\boldsymbol{X}}+\dot{\boldsymbol{\rho}})\delta(\boldsymbol{X}+\boldsymbol{\rho}-\boldsymbol{r})F(\boldsymbol{Z})d^6\boldsymbol{Z}+O(\epsilon^2)\\
&=\int [V_{\parallel}\boldsymbol{b}+\boldsymbol{V}_{\boldsymbol{E}\times\boldsymbol{B}}+\boldsymbol{V}_d]\delta(\boldsymbol{X}-\boldsymbol{r})F(\boldsymbol{Z})d^6\boldsymbol{Z}\\
&\quad+\int \dot{\boldsymbol{\rho}}\delta(\boldsymbol{X}+\boldsymbol{\rho}-\boldsymbol{r})F(\boldsymbol{Z})d^6\boldsymbol{Z}+O(\epsilon^2).
\end{aligned}
\tag{32}
$$

The first term can be reduced to

$$
\begin{aligned}
&\int [V_{\parallel}\boldsymbol{b}+\boldsymbol{V}_{\boldsymbol{E}\times\boldsymbol{B}}+\boldsymbol{V}_d]\delta(\boldsymbol{X}-\boldsymbol{r})F(\boldsymbol{Z})d^6\boldsymbol{Z}\\
&\quad=\left[nU_{\parallel}\boldsymbol{b}+\frac{c}{B}n\boldsymbol{E}\times\boldsymbol{b}+\frac{c}{eB}\boldsymbol{b}\times\left(W_{\perp}\frac{\nabla B}{B}+W_{\parallel}\boldsymbol{b}\cdot\nabla\boldsymbol{b}\right)\right]\Bigg|_{\boldsymbol{X}\mapsto\boldsymbol{r}}\\
&\qquad+O(\epsilon^2),
\end{aligned}
$$

where

$$
U_{\parallel}\equiv\frac{2\pi}{n}\int V_{\parallel}B_{\parallel}^*/mF(\boldsymbol{Z})dV_{\parallel}d\mu,
\tag{33}
$$

$$
W_{\perp}\equiv 2\pi\int B\mu F(\boldsymbol{Z})B_{\parallel}^*/m\,dV_{\parallel}d\mu,
\tag{34}
$$

$$
W_{\parallel}\equiv 2\pi\int mV_{\parallel}^2F(\boldsymbol{Z})B_{\parallel}^*/m\,dV_{\parallel}d\mu.
\tag{35}
$$

The second term is the diamagnetic flow, which can be simplified in terms of $W_{\perp}$,

$$
\begin{aligned}
&\int \dot{\boldsymbol{\rho}}\delta(\boldsymbol{X}+\boldsymbol{\rho}-\boldsymbol{r})F(\boldsymbol{Z})d^6\boldsymbol{Z}\\
&\quad=\int \dot{\boldsymbol{\rho}}\boldsymbol{\rho}\cdot\nabla\delta(\boldsymbol{X}-\boldsymbol{r})F(\boldsymbol{Z})d^6\boldsymbol{Z}+O(\epsilon^2)\\
&\quad=-\int \nabla\cdot[\boldsymbol{\rho}\dot{\boldsymbol{\rho}}B_{\parallel}^*/mF(\boldsymbol{Z})]\delta(\boldsymbol{X}-\boldsymbol{r})dV_{\parallel}d\mu d\xi+O(\epsilon^2)\\
&\quad=-\frac{c}{e}\nabla\times\left(\boldsymbol{b}\frac{W_{\perp}}{B}\right)\Bigg|_{\boldsymbol{X}\mapsto\boldsymbol{r}}+O(\epsilon^2).
\end{aligned}
\tag{36}
$$

In Eq. (36), the following equations are used:

$$
\dot{\boldsymbol{\rho}}=\{\boldsymbol{\rho},H\}=\sqrt{\frac{2\mu B}{m}}\boldsymbol{e}_{\xi}+O(\epsilon),
\tag{37}
$$

$$\left(\int \dot{\boldsymbol{\rho}}\boldsymbol{\rho} d\xi \right)_{ij} = \frac{2\pi\mu c}{e} \epsilon_{ij\boldsymbol{b}} + O(\epsilon). \tag{38}$$

Here $\epsilon_{ij\boldsymbol{b}}$ is the Kronecker symbol, and the subscript $\boldsymbol{b}$ represents the dimension parallel to $\boldsymbol{B}$.

Overall,

$$n(\boldsymbol{r})\boldsymbol{u}(\boldsymbol{r}) = \left[nU_\parallel \boldsymbol{b} + \frac{c\boldsymbol{b}}{eB} \times (W_\perp \nabla B + W_\parallel \boldsymbol{b}\cdot\nabla\boldsymbol{b}) + \frac{cn}{B}\boldsymbol{E}\times\boldsymbol{b} + \frac{c}{e}\nabla\times\left(\frac{W_\perp}{B}\boldsymbol{b}\right)\right]\Bigg|_{\boldsymbol{X}\mapsto\boldsymbol{r}} + O(\epsilon^2). \tag{39}$$

It can be shown that the $n(\boldsymbol{r})$ and $n(\boldsymbol{r})\boldsymbol{u}(\boldsymbol{r})$ derived above satisfy the usual fluid equations derived from the Vlasov equation in the laboratory phase space coordinate system by taking the velocity moments.[27] As an example, the equilibrium force balance equation can be recovered here. First, it is shown that

$$W_\perp(\boldsymbol{r}) = 2\pi \int B\mu F(\boldsymbol{Z}) B_\parallel^*/m\, dV_\parallel d\mu \Big|_{\boldsymbol{X}\mapsto\boldsymbol{r}} = \int \frac{1}{2} m v_\perp^2 f d^3\boldsymbol{v} + O(\epsilon) = p_\perp(\boldsymbol{r}) + O(\epsilon), \tag{40}$$

$$\begin{aligned} W_\parallel(\boldsymbol{r}) &= 2\pi \int m V_\parallel^2 F(\boldsymbol{Z}) B_\parallel^*/m\, dV_\parallel d\mu \Big|_{\boldsymbol{X}\mapsto\boldsymbol{r}} \\ &= \left[mnU_\parallel^2 + 2\pi \int m(V_\parallel - U_\parallel)^2 F(\boldsymbol{Z}) B_\parallel^*/m\, dV_\parallel d\mu \right]\Bigg|_{\boldsymbol{X}\mapsto\boldsymbol{r}} \\ &= mnu_\parallel^2(\boldsymbol{r}) + p_\parallel(\boldsymbol{r}) + O(\epsilon). \end{aligned} \tag{41}$$

From Eq. (39),

$$\begin{aligned} n\boldsymbol{u}_\perp &= \left\{ \frac{c\boldsymbol{b}}{eB}\times(W_\perp\nabla B + W_\parallel \boldsymbol{b}\cdot\nabla\boldsymbol{b}) + \frac{cn}{B}\boldsymbol{E}\times\boldsymbol{b} + \frac{c}{q}\left[\nabla\times\left(\frac{W_\perp}{B}\boldsymbol{b}\right)\right]_\perp \right\}\Bigg|_{\boldsymbol{X}\mapsto\boldsymbol{r}} + O(\epsilon^2) \\ &= \frac{c}{e}\left\{ -\frac{\nabla p_\perp \times \boldsymbol{b}}{B} + p_\perp\left[\frac{\boldsymbol{b}\times\nabla B}{B^2} - \left(\nabla\times\frac{\boldsymbol{b}}{B}\right)_\perp\right] + p_\parallel \frac{(\nabla\times\boldsymbol{b})_\perp}{B} + mnu_\parallel^2 \frac{(\nabla\times\boldsymbol{b})_\perp}{B} \right\} \\ &\quad + n\frac{\boldsymbol{E}\times\boldsymbol{b}}{B}c + O(\epsilon^2). \end{aligned} \tag{42}$$

Using

$$(\boldsymbol{u}\cdot\nabla\boldsymbol{u})\times\boldsymbol{b} = -u_\parallel^2(\nabla\times\boldsymbol{b})_\perp + O(\epsilon^2), \tag{43}$$

$$\frac{\boldsymbol{b}\times\nabla B}{B^2} - \left(\nabla\times\frac{\boldsymbol{b}}{B}\right)_\perp = -\frac{(\nabla\times\boldsymbol{b})_\perp}{B}, \tag{44}$$

then leads to the result,

$$n\boldsymbol{u}_\perp = -\frac{c}{eB}[mn\boldsymbol{u}\cdot\nabla\boldsymbol{u}\times\boldsymbol{b} + \nabla_\perp p_\perp \times \boldsymbol{b} - (p_\parallel - p_\perp)(\nabla\times\boldsymbol{b})_\perp - en\boldsymbol{E}\times\boldsymbol{b}] + O(\epsilon^2). \tag{45}$$

The first term on the right-hand side of the above equation is smaller than the left-hand side by a factor of ρ/L, where ρ and L are the characteristic gyroradius and scale length of the equilibrium flow $\boldsymbol{u}$. In tokamak experiments, this term is often neglected when E_r is evaluated from the spectroscopic measurements of flows and pressure gradient.[29] For a neutral plasma

$$\boldsymbol{j}_\perp = \sum_s (en\boldsymbol{u})_s = \frac{c}{B}\left[\boldsymbol{b}\times\nabla\sum_s p_\perp + \left(\sum_s p_\parallel - \sum_s p_\perp\right)(\nabla\times\boldsymbol{b})_\perp\right]. \tag{46}$$

This is the transverse equilibrium force balance equation. In particular, when the distribution function F is isotropic, $\Sigma_s p_\parallel = \Sigma_s p_\perp = \Sigma_s p$, the familiar fluid result, $\boldsymbol{j}_\perp = c/B\boldsymbol{b}\times\nabla\Sigma_s p$, is recovered.

In the above derivation, the pullback formula has been used in the form of Eq. (26). The pullback formula in the form of Eq. (4) can be used to obtain the same results.

IV. PULLBACK ASSOCIATED WITH THE NONSYMPLECTIC GYROCENTER CENTER COORDINATE TRANSFORMATION

When time-dependent electromagnetic perturbations are introduced into a magnetized plasma, the guiding center coordinates used in Sec. III to study the gyrokinetic equilibrium will cease to be the "good" coordinate system where the gyromotion is decoupled from the rest of particle dynamics. To preserve the desirable decoupling of the gyromotion, a nonsymplectic gyrocenter transformation,

$$\mathrm{Gy}: \boldsymbol{Z} = (\boldsymbol{X}, U, \mu, \xi) \mapsto \overline{\boldsymbol{Z}} = (\overline{\boldsymbol{X}}, \overline{U}, \overline{\mu}, \overline{\xi}), \tag{47}$$

can be constructed using a Lie perturbation method such that the transformed symplectic structure has the same functional form as that in the unperturbed guiding center coordinates. Because the perturbation on the symplectic structure of the guiding center is "deperturbed away" by the perturbative gyrocenter coordinate transformation, particle dynamics in the gyrocenter coordinates are exactly the same as those in the guiding center coordinates except for a perturbation in the Hamiltonian. Since two consecutive coordinate transformations are involved, two pullback transformations are needed to relate the distribution function in the gyrocenter coordinates F_{Gy} to the macroscopic physical quantities in the laboratory coordinates $q(\boldsymbol{r})$. In the guiding center coordinate $\boldsymbol{Z} = (\boldsymbol{X}, U, \mu, \xi)$,

$$q(\boldsymbol{r}) = \int [G^{-1*}q](\boldsymbol{Z}) F(\boldsymbol{Z}) \delta(G^{-1}\boldsymbol{X} - \boldsymbol{r}) d^6\boldsymbol{Z}. \tag{48}$$

Using Eq. (26) again, we have

$$q(\boldsymbol{r}) = \int [G_y^{-1*} G^{-1*} q](\overline{\boldsymbol{Z}}) F(\overline{\boldsymbol{Z}}) \delta(G_y^{-1} G^{-1} \overline{\boldsymbol{X}} - \boldsymbol{r}) d^6\overline{\boldsymbol{Z}}. \tag{49}$$

Alternatively, we can use Eq. (4) to replace $F(\boldsymbol{Z})$ by its pullback from the gyrocenter coordinate to obtain

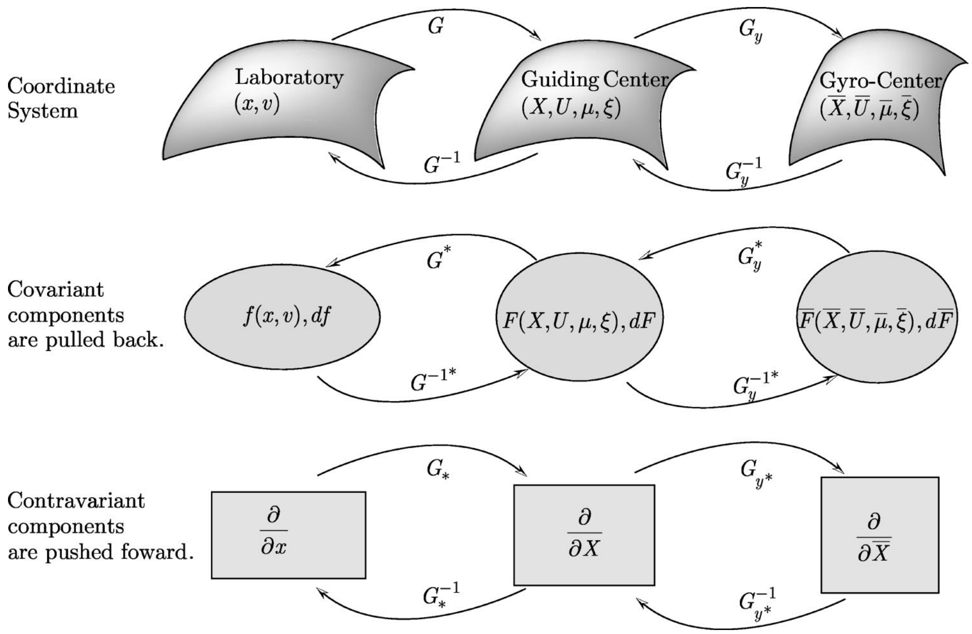

FIG. 3. Coordinate transformation and the associated pullback and push-forward.

$$q(\boldsymbol{r})=\int [G^{-1}{*}q](\boldsymbol{Z})[\mathrm{Gy}^{*}F_{\mathrm{Gy}}](\boldsymbol{Z})\,\delta(G^{-1}\boldsymbol{X}-\boldsymbol{r})\,d^6\boldsymbol{Z}. \tag{50}$$

In the above equations, $d^6\boldsymbol{Z}$ is understood to be $(B_\parallel^*/m)d^3\boldsymbol{X}dUd\mu d\xi$. Gy* is the pullback transformation associated with the gyrocenter transformation, which transforms the distribution function in the gyrocenter coordinates into that in the guiding center coordinates. G^{-1} is the inverse of G that transforms the laboratory phase space coordinate system into the guiding center coordinates. It is assumed that the guiding center transformation G and the gyrocenter transformation Gy are bijective and smooth. The relationship between the three coordinate systems involved and the objects defined on them are illustrated in Fig. 3. Covariant objects such as functions and forms are pulled back by the associated coordinate transformation, while contravariant objects such as vectors are pushed forward.

The pullback transformation from the gyrocenter coordinates to the guiding center coordinates is easily obtained from the expression for G given by Refs. 14, 23–25. Since the focus of this paper is not the gyrocenter coordinate transformation, the expression for the pullback transformation is displayed in terms of the perturbed fields $(\boldsymbol{A}_1,\phi_1)$ without derivation,

$$\begin{aligned}\mathrm{Gy}^*F=F+L_{\boldsymbol{G}}F&=F-\frac{\boldsymbol{b}}{B}\times\left(\boldsymbol{A}_1+\frac{c}{e}\nabla S\right)\cdot\nabla F\\&\quad+\frac{e}{mc}\boldsymbol{b}\cdot\left(\boldsymbol{A}_1+\frac{c}{e}\nabla S\right)\frac{\partial F}{\partial U}\\&\quad+\frac{e}{mc}\left[\frac{e}{c}\boldsymbol{A}_1\cdot\frac{\partial\boldsymbol{\rho}}{\partial\xi}+\frac{\partial S}{\partial\xi}\right]\frac{\partial F}{\partial\mu}+O(\epsilon_B),\end{aligned} \tag{51}$$

where the gauge function S satisfies

$$\begin{aligned}\{S,H_0\}&=\Omega\frac{\partial S}{\partial\bar{\xi}}+\frac{\partial S}{\partial t}+\frac{\partial S}{\partial\bar{\boldsymbol{X}}}\cdot\{\bar{\boldsymbol{X}},H_0\}+\frac{\partial S}{\partial\bar{U}}\{\bar{U},H_0\}\\&=e\tilde{\phi}_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho},t)-\frac{e}{c}\widetilde{\bar{\boldsymbol{V}}\cdot\boldsymbol{A}}_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho},t),\end{aligned} \tag{52}$$

in the coordinates $(\bar{\boldsymbol{X}},\bar{U},\bar{\mu},\bar{\xi})$. Here, $\tilde{\phi}_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho},t)$ and $\widetilde{\bar{\boldsymbol{V}}\cdot\boldsymbol{A}}_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho},t)$ are the gyrophase dependent parts of $\phi_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho},t)$ and $\bar{\boldsymbol{V}}\cdot\boldsymbol{A}_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho},t)$, respectively,

$$\tilde{\phi}_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho},t)=\phi_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho},t)-\langle\phi_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho},t)\rangle, \tag{53}$$

$$\widetilde{\bar{\boldsymbol{V}}\cdot\boldsymbol{A}}_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho}_0,t)=\bar{\boldsymbol{V}}\cdot\boldsymbol{A}_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho},t)-\langle\bar{\boldsymbol{V}}\cdot\boldsymbol{A}_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho},t)\rangle, \tag{54}$$

and H_0 is the unperturbed Hamiltonian

$$H_0=\frac{m\bar{U}^2}{2}+\bar{\mu}B.$$

In the coordinates $(\bar{\boldsymbol{X}},\bar{U},\bar{\mu},\bar{\xi})$, the linear gyrokinetic equation is

$$\begin{aligned}&\frac{\partial f}{\partial t}+(\bar{U}\boldsymbol{b}+\boldsymbol{v}_d)\cdot\nabla f-\frac{1}{m}\boldsymbol{b}\cdot\nabla H_0\frac{\partial f}{\partial\bar{U}}\\&\quad=\frac{c}{eB}\boldsymbol{b}\cdot(\nabla F_0\times\nabla H_1)-\frac{1}{m}\boldsymbol{b}\cdot\left(\nabla F_0\frac{\partial H_1}{\partial\bar{U}}-\nabla H_1\frac{\partial F_0}{\partial\bar{U}}\right),\end{aligned} \tag{55}$$

where

$$F=F_0+f, \tag{56}$$

$$H_1=\left\langle e\phi_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho},t)-\frac{e}{c}\boldsymbol{V}\cdot\boldsymbol{A}_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho},t)\right\rangle. \tag{57}$$

The importance of the pullback formula in Eqs. (50) and (51) is demonstrated in the following nontrivial examples. We note that Eqs. (50) and (51) are written in the guiding center coordinates, while Eqs. (52) and (55) are in the gyrocenter coordinates. In the following equations, $\boldsymbol{A}$ and ϕ will be used to represent the perturbed field. The subscript "1" will be dropped.

A. Gyrokinetic shear Alfvén wave

For shear Alfvén physics, we can only keep the parallel component of the vector potential $\boldsymbol{A}=A_\parallel \boldsymbol{b}$. To the leading order, Eq. (52) for shear Alfvén modes reduce to

$$\Omega\frac{\partial S}{\partial \bar{\xi}}=\frac{e}{\Omega}\left[\tilde{\phi}(\bar{\boldsymbol{X}}+\boldsymbol{\rho},t)-\frac{1}{c}\widetilde{UA}_\parallel(\bar{\boldsymbol{X}}+\boldsymbol{\rho},t)\right]$$

$$\approx\frac{e}{\Omega}\boldsymbol{\rho}_0\cdot\left[\nabla\phi(\bar{\boldsymbol{X}},t)-\frac{1}{c}U\nabla A_\parallel(\bar{\boldsymbol{X}},t)\right]. \tag{58}$$

Using Eq. (51), we have the pullback transformation for shear Alfvén modes,

$$\mathrm{Gy}^*F=F+\frac{e}{mc}A_\parallel(\boldsymbol{X}+\boldsymbol{\rho},t)\frac{\partial F}{\partial U}$$

$$+\frac{e}{B}\left[\tilde{\phi}(\boldsymbol{X}+\boldsymbol{\rho},t)-\frac{1}{c}\widetilde{UA}_\parallel(\boldsymbol{X}+\boldsymbol{\rho},t)\right]\frac{\partial F}{\partial\mu}. \tag{59}$$

The perturbed density, perturbed flow, and perturbed current can be derived from the general form of Eq. (50),

$$n_1(\boldsymbol{r})=\left\{\int[\mathrm{Gy}^*(F_0+f)](\boldsymbol{Z})\,\delta(\boldsymbol{X}+\boldsymbol{\rho}-\boldsymbol{r})d^6\boldsymbol{Z}\right\}_1$$

$$=\int f(\boldsymbol{Z})\,\delta(\boldsymbol{X}-\boldsymbol{r})d^6\boldsymbol{Z}+\int[\delta(\boldsymbol{X}+\boldsymbol{\rho}-\boldsymbol{r})$$

$$-\delta(\boldsymbol{X}-\boldsymbol{r})]f(\boldsymbol{Z})d^6\boldsymbol{Z}+\int\delta(\boldsymbol{X}+\boldsymbol{\rho}-\boldsymbol{r})$$

$$\times\left\{\frac{e}{mc}A_\parallel(\boldsymbol{X}+\boldsymbol{\rho},t)\frac{\partial F_0}{\partial U}+\frac{e}{B}\left[\tilde{\phi}(\boldsymbol{X}+\boldsymbol{\rho},t)\right.\right.$$

$$\left.\left.-\frac{1}{c}\widetilde{UA}_\parallel(\boldsymbol{X}+\boldsymbol{\rho},t)\right]\frac{\partial F_0}{\partial\mu}\right\}d^6\boldsymbol{Z}. \tag{60}$$

From Eq. (60), the perturbed density in laboratory coordinates consists of three parts corresponding to the three integrals on the right-hand side of the equation. The first integral is the perturbed density in gyrocenter coordinates, the second integral is the guiding center correction, and the third integral is the gyrocenter correction. After some lengthy algebra,

$$n_1(\boldsymbol{r},t)=\int J_0 f(\boldsymbol{r},U,\mu,t)d^3\boldsymbol{v}+\frac{e}{m}\nabla_\perp\frac{n_0}{\Omega^2}\nabla\phi(\boldsymbol{r},t)$$

$$+\frac{3}{4}\frac{ev_t^2n_0}{m\Omega^4}\nabla_\perp^4\phi(\boldsymbol{r},t), \tag{61}$$

where $d^3\boldsymbol{v}=2\pi(B/m)dUd\mu$, $J_0=J_0(v_\perp\nabla_\perp/i\Omega)$ is the zeroth order Bessel function of the first kind, and only terms up to $O(v_\perp^4\nabla_\perp^4/\Omega^4)$ for the L_GF_0 part of the pullback transformation Gy* have been retained. For the perturbed parallel flow,

$$n_0u_{\parallel 1}(\boldsymbol{r})=\left\{\int U[\mathrm{Gy}^*(F_0+f)](\boldsymbol{Z})\,\delta(\boldsymbol{X}+\boldsymbol{\rho}-\boldsymbol{r})d^6\boldsymbol{Z}\right\}_1$$

$$=\int Uf(\boldsymbol{Z})\,\delta(\boldsymbol{X}-\boldsymbol{r})d^6\boldsymbol{Z}+\int U[\delta(\boldsymbol{X}+\boldsymbol{\rho}-\boldsymbol{r})$$

$$-\delta(\boldsymbol{X}-\boldsymbol{r})]f(\boldsymbol{Z})d^6\boldsymbol{Z}+\int U\delta(\boldsymbol{X}+\boldsymbol{\rho}-\boldsymbol{r})$$

$$\times\left\{\frac{e}{mc}A_\parallel(\boldsymbol{X}+\boldsymbol{\rho},t)\frac{\partial F_0}{\partial U}+\frac{e}{B}\left[\tilde{\phi}(\boldsymbol{X}+\boldsymbol{\rho},t)\right.\right.$$

$$\left.\left.-\frac{1}{c}\widetilde{UA}_\parallel(\boldsymbol{X}+\boldsymbol{\rho},t)\right]\frac{\partial F_0}{\partial\mu}\right\}d^6\boldsymbol{Z}. \tag{62}$$

Again, the algebra here is straightforward but involved. The final result is

$$n_0u_{\parallel 1}(\boldsymbol{r},t)=\int J_0Uf(\boldsymbol{r},U,\mu,t)d^3\boldsymbol{v}+\int\frac{e}{mc}\langle UA_\parallel(\boldsymbol{r}+\boldsymbol{\rho}_0)\rangle$$

$$\times\frac{\partial F_0}{\partial U}2\pi\frac{B}{m}d\mu dU+\frac{en_0v_t^2}{2mc\Omega^2}\nabla_\perp^2A_\parallel, \tag{63}$$

where the first integral on the right-hand side is the perturbed parallel flow of the gyrocenter, and the second integral and the third term are the gyrocenter correction generated by the pullback transformation.

From Eq. (61), the quasineutrality condition is

$$\sum_j e\left[\int J_0fd^3\boldsymbol{v}+\frac{e}{m}\nabla_\perp\frac{n_0}{\Omega^2}\nabla_\perp\phi+\frac{3e}{4m}\frac{v_t^2}{\Omega^2}\frac{n_0}{\Omega^2}\nabla_\perp^4\phi\right]=0. \tag{64}$$

From Eq. (63), the parallel Ampère's law is

$$[\nabla\times\nabla\times\boldsymbol{A}]_\parallel=\frac{4\pi}{c}\sum_j e\int\left(UJ_0f+\frac{\partial F_0}{\partial U}\frac{e}{mc}\langle UA_\parallel\rangle\right)d^3\boldsymbol{v}$$

$$+\frac{4\pi}{c}\frac{e^2n_0v_t^2}{2mc\Omega^2}\nabla_\perp^2A_\parallel. \tag{65}$$

In Eqs. (64) and (65), the spatial variable is the laboratory coordinate $\boldsymbol{r}$. However $\boldsymbol{r}$ is a dummy variable. What matters is the functional forms. We can replace $\boldsymbol{r}$ by the spatial coordinate of the gyrocenter coordinates $\bar{\boldsymbol{Z}}$. Equations (64) and (65) will be referred to as the gyrokinetic quasineutrality condition and the gyrokinetic parallel Ampère's law, respectively.

As a simple application of these results, we derive the local dispersion relation in an unsheared slab geometry with $\boldsymbol{B}_0=B(x)\boldsymbol{e}_z$ and $n_0=n_0(x)$. For local perturbations

$$(\phi,\psi_\parallel)\sim e^{i(k_yy+k_\parallel z)}, \tag{66}$$

where $\psi_\parallel$ is defined through

$$A_\parallel=\frac{c}{i\omega}(\nabla\psi_\parallel)_\parallel. \tag{67}$$

The solution of the gyrokinetic equation, Eq. (55), for shear Alfvén waves in slab geometry is

$$f=-\frac{e}{T}F_0\left(\phi-\frac{k_\parallel U}{\omega}\psi_\parallel\right)+\frac{e}{T}\frac{\omega-\omega_*}{\omega-k_\parallel U}F_0\left(\phi-\frac{k_\parallel U}{\omega}\psi_\parallel\right), \tag{68}$$

where ω_* is the diamagnetic drift frequency defined by

$$\omega_{*j}\equiv\left(\frac{cTk_y}{L_n eB}\right)_j,\quad L_n\equiv-\left(\frac{d\ln n}{dx}\right)^{-1}, \tag{69}$$

and the temperature gradient has been neglected. Substituting f into the quasineutrality condition, we have,

$$\begin{aligned}-\sum_j \frac{e^2}{m}\nabla_\perp\frac{n_0}{\Omega^2}\nabla_\perp\phi &= -\sum_j\frac{e^2 n_0}{T}[1+\zeta Z(\zeta)](\phi-\psi_\parallel)\\ &\quad+\sum_j\frac{e^2n_0}{T}\zeta Z(\zeta)\frac{\omega_*}{\omega}(\phi-\psi_\parallel)-\sum_j\frac{e^2n_0}{T}\frac{\omega_*}{\omega}\psi_\parallel .\end{aligned} \tag{70}$$

$Z(\zeta)$ is the plasma dispersion function and $\zeta\equiv\omega/k_\parallel v_{\mathrm{th}}$. Straightforward algebra shows that the parallel Ampère's law reduces to

$$k_\parallel^2\psi_\parallel=\frac{\omega^2}{v_A^2}\phi, \tag{71}$$

or in terms of $\omega_A\equiv k_\parallel v_A$,

$$\psi_\parallel=\frac{\omega^2}{\omega_A^2}\phi. \tag{72}$$

Inserting this polarization property into the quasineutrality condition, we obtain the desired dispersion relation,

$$\begin{aligned}\sum_j\frac{e^2n_0}{m\Omega^2}k_y^2 &= -\sum_j\frac{e^2n_0}{T}[1+\zeta Z(\zeta)]\left(1-\frac{\omega^2}{\omega_A^2}\right)\\ &\quad+\sum_j\frac{e^2n_0}{T}\zeta Z(\zeta)\frac{\omega_*}{\omega}\left(1-\frac{\omega^2}{\omega_A^2}\right)\\ &\quad-\sum_j\frac{e^2n_0}{T}\frac{\omega_*}{\omega}\frac{\omega^2}{\omega_A^2}.\end{aligned} \tag{73}$$

It contains many interesting physics effects for various parametric regimes. Some, which are relevant to tokamak plasmas, will be highlighted in the following discussion.

The fluid results are generally recovered from kinetic theory by ignoring the kinetic resonances and assuming the so-called "hot electron, cold ion expansion," that is,

$$\zeta_e=\frac{\omega}{v_{\mathrm{the}}k_\parallel}\ll 1, \tag{74}$$

$$\zeta_i=\frac{\omega}{v_{\mathrm{thi}}k_\parallel}\gg 1. \tag{75}$$

Using the Taylor expansion and the asymptotic form for $Z(\zeta)$, gives

$$\begin{aligned}\frac{n_0m_ic^2}{B^2}k_y^2 &= \left[-\frac{e^2n_{e0}}{T_e}+\frac{e_i^2n_{i0}}{T_i}\left(\frac{v_{\mathrm{thi}}k_\parallel}{\omega}\right)^2\right]\left(1-\frac{\omega^2}{\omega_A^2}\right)\\ &\quad-\frac{e_i^2n_{0i}}{T_i}\frac{\omega_{*i}}{\omega}\left(1-\frac{\omega^2}{\omega_A^2}\right)\\ &\quad-\left[\frac{e^2n_{e0}^2}{T_e}\frac{\omega_{*e}}{\omega}+\frac{e_i^2n_{i0}^2}{T_i}\frac{\omega_{*i}}{\omega}\right]\frac{\omega^2}{\omega_A^2}\\ &= \frac{e^2n_e}{T_e}\left[-1+\frac{e_jT_e}{|e|T_i}\left(\frac{v_{\mathrm{thi}}k_\parallel}{\omega}\right)^2+\frac{\omega_{*e}}{\omega}\right]\left[1-\frac{\omega^2}{\omega_A^2}\right].\end{aligned} \tag{76}$$

Note the following relationship between ω_{*j} and ω_{*e} has been used:

$$\omega_{*j}=-\frac{|e|T_j}{e_jT_e}\omega_{*e}. \tag{77}$$

This then leads to the familiar fluid result,[30]

$$\frac{e_ic_s^2}{|e|\Omega_i^2}k_\perp^2=\left[\frac{e_ic_s^2}{|e|\omega^2}k_\parallel^2-1+\frac{\omega_{*e}}{\omega}\right]\left[1-\frac{\omega^2}{\omega_A^2}\right], \tag{78}$$

where

$$c_s^2\equiv\frac{T_e}{m_i}. \tag{79}$$

It is commonly believed that when the plasma β (ratio of plasma to magnetic pressure) approaches zero, the magnetic perturbations are not important. However, it is not a correct conclusion that, when β goes to zero, there are no magnetic perturbations. Very obvious examples are the well-known shear Alfvén wave and the compressional Alfvén wave in a homogeneous magnetized plasma. Even in a zero β magnetized plasma, the shear Alfvén wave and the compressional Alfvén wave are both mathematically and physically well-defined. The physical mechanism maintaining these two waves is the balance between plasma kinetic energy and the restoring force due to the bending or compression of the equilibrium magnetic field. Theoretically, they are characterized by the dispersion relations $\omega^2=k_\parallel^2v_A^2$ and $\omega^2=k^2v_A^2$, respectively. Their existence is independent of the plasma β.

This fact can also be verified from the dispersion relation, Eq. (70). When $\omega_A\gg\omega_{*e}$, there exists a solution in the range of ω_A. For this range of ω, the dispersion relation is reduced to

$$\frac{n_0m_ic^2}{B^2}k_y^2=-\sum_j\frac{e^2n_0}{T}[1+\zeta Z(\zeta)]\left(1-\frac{\omega^2}{\omega_A^2}\right). \tag{80}$$

Under normal condition, term $(n_0m_ic^2/B^2)k_y^2$ is smaller by $O(\rho_s^2k_\perp^2)$, compared with the other terms. To the leading order,

$$\omega^2=\omega_A^2. \tag{81}$$

The fact that there are no assumptions about ζ_e and ζ_i needed here to obtain this shear Alfvén wave is consistent with the basic physical picture of the shear Alfvén wave. The statement that when the plasma β is small, magnetic perturbations are not important is always relevant only for a special class of electrostatic modes. In particular, for the electrostatic drift waves, magnetic perturbations are decoupled from these electrostatic perturbations when β is small. This is evident from the fact that

$$\frac{\omega_{*e}}{\omega_A}\rightarrow 0 \quad \text{as} \quad \beta\rightarrow 0. \tag{82}$$

It is also evident from the polarization property $\psi_\parallel = (\omega/\omega_A)\,\phi$. For the electrostatic drift wave, $\omega\sim\omega_{*e}$,

$$\psi_\parallel = \frac{\omega_{*e}}{\omega_A}\phi\rightarrow 0 \quad \text{as} \quad \beta\rightarrow 0. \tag{83}$$

For the shear Alfvén branch, $\omega\sim\omega_A$,

$$\psi_\parallel\sim\phi \quad \text{independent of} \quad \beta. \tag{84}$$

The above facts also lead to a "β_{critical}" where there is strong coupling between the electrostatic drift branch and the electromagnetic shear Alfvén branch. The criterion is $\omega_{*e}\sim\omega_A$. In tokamak geometry, it is[30]

$$\sqrt{\beta_{\text{critical}}}\sim\frac{r}{R_0}\frac{r}{\rho_s}. \tag{85}$$

For standard tokamak parameters, β_{critical} is not a very small number. However, ω_A is geometry dependent in complex geometries. It can be reduced to be as small as ω_{*e} even at low β. Another interesting limit is where $k_\parallel$ approaches zero around mode rational surfaces, so that ω_A could be much smaller than its normal characteristic value. In both cases, there would be strong coupling between the shear Alfvén branch and the drift branch.

B. Compressional Alfvén wave

In this subsection, the simplest example of the compressional Alfvén wave in a homogeneous magnetized plasma is used to demonstrate the essence of the pullback formula in the perpendicular direction. $\boldsymbol{B}_0$ is assumed to be in the $\boldsymbol{e}_z$ direction, and for simplicity, we let $\boldsymbol{k}=k_y\boldsymbol{e}_y$. The MHD results for the compressional Alfvén wave indicate that the magnetic perturbation is in the parallel direction, the electrical perturbation, and current perturbation are in the $\boldsymbol{e}_x$ direction, and the plasma displacement is in the $\boldsymbol{e}_y$ direction. From the kinetic point of view, we can choose $\phi=0$ and $\boldsymbol{A}=A_x\boldsymbol{e}_x$. Assuming $A_x, f\propto e^{ik_y y - i\omega t}$, the linear gyrokinetic equation Eq. (55) leads to $f=0$.

Interesting physics is found in the gyrocenter pullback transformation, which requires the knowledge of S. Let $S=S^{(0)}+\epsilon_\omega S^{(1)}+\epsilon_\omega^2 S^{(2)}+O(\epsilon_\omega^3)$, where $\epsilon_\omega\equiv\omega/\Omega\ll 1$. To the third order, the solution of Eq. (52) for S is

$$\begin{aligned}\Omega\frac{\partial S}{\partial\bar{\xi}} &\approx \Omega\frac{\partial S^{(0)}+S^{(1)}+S^{(2)}}{\partial\bar{\xi}}+O(\epsilon_\omega^3)\\ &= e\left(\tilde{\phi}-\frac{1}{c}\widetilde{\boldsymbol{V}\cdot\boldsymbol{A}}\right)-\frac{e}{\Omega}\frac{\partial}{\partial t}\int\left(\tilde{\phi}-\frac{1}{c}\widetilde{\boldsymbol{V}\cdot\boldsymbol{A}}\right)d\bar{\xi}\\ &\quad+\frac{e}{\Omega^2}\frac{\partial^2}{\partial t^2}\int\int\left(\tilde{\phi}-\frac{1}{c}\widetilde{\boldsymbol{V}\cdot\boldsymbol{A}}\right)d\bar{\xi}d\bar{\xi}+O(\epsilon_\omega^3).\end{aligned} \tag{86}$$

From the general form of the gyrocenter pullback transformation we have

$$\begin{aligned}&[\mathrm{Gy}^*(F_0+f)]_1\\ &\quad= f+\frac{e}{mc}\frac{\partial F_0}{\partial\mu}\left\{-\frac{ev_\perp^2}{\Omega^2 c}B_\parallel-\frac{e}{\Omega^2}\frac{\partial}{\partial t}\int\left(\tilde{\phi}-\frac{1}{c}\widetilde{\boldsymbol{V}\cdot\boldsymbol{A}}\right)d\xi\right.\\ &\qquad\left.+\frac{e}{\Omega^3}\frac{\partial}{\partial t^2}\int\int\left(\tilde{\phi}-\frac{1}{c}\widetilde{\boldsymbol{V}\cdot\boldsymbol{A}}\right)d\xi d\xi\right\}.\end{aligned} \tag{87}$$

In above derivation, we have used the following expressions for the gyroaverage:

$$\begin{aligned}H_1&=\left\langle e\phi(\boldsymbol{X}+\boldsymbol{\rho},t)-\frac{e}{c}\boldsymbol{V}\cdot\boldsymbol{A}(\boldsymbol{X}+\boldsymbol{\rho},t)\right\rangle\\ &\approx e\left[\phi(\boldsymbol{X},t)-\frac{1}{c}UA_\parallel(\boldsymbol{X},t)+\frac{v_\perp^2}{2c\Omega}B_\parallel(\boldsymbol{X},t)\right].\end{aligned} \tag{88}$$

The perpendicular Ampère's law is needed to complete this system of equations. For this purpose, it is necessary to obtain the perturbed perpendicular current,

$$\begin{aligned}n_0\boldsymbol{u}_\perp&=\left\{\int\boldsymbol{V}_\perp[\mathrm{Gy}^*(F_0+f)](\boldsymbol{Z})\,\delta(\boldsymbol{X}+\boldsymbol{\rho}-\boldsymbol{r})d^6\boldsymbol{Z}\right\}_1\\ &=\int\boldsymbol{V}_\perp\,\delta(\boldsymbol{X}+\boldsymbol{\rho}-\boldsymbol{r})f(\boldsymbol{Z})d^6\boldsymbol{Z}\\ &\quad+\int\boldsymbol{V}_\perp\,\delta(\boldsymbol{X}+\boldsymbol{\rho}-\boldsymbol{r})\frac{e}{mc}\frac{\partial F_0}{\partial\mu}\left\{\frac{-ev_\perp^2}{\Omega^2 c}B_\parallel\right.\\ &\quad-\frac{e}{\Omega^2}\frac{\partial}{\partial t}\int\left(\tilde{\phi}-\frac{1}{c}\widetilde{\boldsymbol{V}\cdot\boldsymbol{A}}\right)d\xi\\ &\quad\left.+\frac{e}{\Omega^3}\frac{\partial}{\partial t^2}\int\int\left(\tilde{\phi}-\frac{1}{c}\widetilde{\boldsymbol{V}\cdot\boldsymbol{A}}\right)d\xi d\xi\right\}d^6\boldsymbol{Z}.\end{aligned} \tag{89}$$

FLR effects are ignored here, and use has been made of the following expression for the particle perpendicular velocity

$$\boldsymbol{V}_\perp=-V_\perp[\sin(\xi)\boldsymbol{e}_x+\cos(\xi)\boldsymbol{e}_y]. \tag{90}$$

Finally, the perturbed perpendicular flow is

$$n_0\boldsymbol{u}_\perp = \int \boldsymbol{V}_\perp \delta(\boldsymbol{X}-\boldsymbol{r}) \frac{e}{mc}\frac{\partial F_0}{\partial \mu}\left\{ \frac{e}{\Omega^2}\frac{\partial}{\partial t}\int \frac{1}{c}\boldsymbol{V}_\perp \cdot \boldsymbol{A}_\perp d\xi - \frac{e}{\Omega^3}\frac{\partial^2}{\partial t^2}\int\int \frac{1}{c}\boldsymbol{V}_\perp \cdot \boldsymbol{A}_\perp d\xi d\xi \right\} d^6\boldsymbol{Z}$$
$$= \frac{-i\omega n_0 e^2 B}{m^2c^2\Omega^2}A_x\boldsymbol{e}_y + \frac{\omega^2 n_0 e^2 B}{m^2c^2\Omega^3}A_x\boldsymbol{e}_x$$
$$= \frac{n_0 c}{B^2}\boldsymbol{E}\times\boldsymbol{B} + \frac{n_0 mc^2}{eB^2}\frac{\partial \boldsymbol{E}_\perp}{\partial t}. \tag{91}$$

It is obvious that the perpendicular current is generated by the ion polarization drift,

$$\boldsymbol{j} = \sum_j (en_0\boldsymbol{u}_\perp)_j \approx \frac{n_0 m_i c}{B^2}\omega^2 A_x \boldsymbol{e}_x. \tag{92}$$

The perpendicular Ampère's law $(\nabla\times\nabla\times\boldsymbol{A})_\perp = 4\pi/c\boldsymbol{j}_\perp$ gives

$$k_y^2 A_x = \frac{4\pi n_0 m_i}{B^2}\omega^2 A_x, \quad \text{or} \quad \omega^2 = k_y^2 v_A^2. \tag{93}$$

This is the dispersion relation for the compressional Alfvén wave. The key element in this gyrokinetic description of the compressional Alfvén wave is the perturbed perpendicular current. To $O(\epsilon_\omega)$, the perpendicular flow is the $\boldsymbol{E}\times\boldsymbol{B}$ flow, which gives no current. Therefore, it is necessary to go to $O(\epsilon_\omega^2)$. The current to this order is the current generated by the polarization drift in the perturbed electromagnetic field.

C. Bernstein wave

In this subsection, the Bernstein wave and is recovered, and the application of the pullback transformations to high frequency modes is demonstrated. We consider an electrostatic wave propagating in a homogeneous magnetized plasma with $\omega \sim \Omega$. Let $\boldsymbol{B}_0 = B\boldsymbol{e}_z$ and $\boldsymbol{k} = k\boldsymbol{e}_x$. The solution for the linear gyrokinetic equation give $f=0$ because $k_\parallel = 0$. As in the case of compressional Alfvén wave, f, the gyrophase independent part of the distribution function, does not play any role, and the only physics content is found in the pullback of the perturbed density, which requires expressing the gauge function S in terms of the perturbed fields. The equation for S is

$$\{S,H_0\} = \Omega\frac{\partial S}{\partial \bar{\xi}} + \frac{\partial S}{\partial t}$$
$$= e\tilde{\phi}(\bar{\boldsymbol{X}}+\boldsymbol{\rho}) = e\left[e^{\boldsymbol{\rho}\cdot\nabla} - J_0\left(\frac{\boldsymbol{\rho}\cdot\nabla}{i}\right)\right]\phi. \tag{94}$$

Using the identity $\exp(\lambda\cos\bar{\xi}) = \Sigma_{n=-\infty}^{\infty} I_n(\lambda)\exp(\text{in}\bar{\xi})$, we can easily solve Eq. (94) for S,

$$S = \frac{e}{\Omega i\bar{\omega}}J_0\phi + \frac{e}{\Omega}\sum_{n=-\infty}^{\infty}\frac{I_n(i\rho k)}{i(n-\bar{\omega})}e^{in\bar{\xi}}\phi, \tag{95}$$

where $\bar{\omega} = \omega/\Omega$. Since $f=0$, the density response comes only from the pullback transformation,

$$n_1 = \int J_0 f d^3\boldsymbol{v} + \int \delta(\boldsymbol{X}+\boldsymbol{\rho}-\boldsymbol{r})\frac{e}{mc}\frac{\partial S}{\partial \xi}\frac{\partial F_0}{\partial \mu}d^6\boldsymbol{Z}$$
$$= \int [e^{\boldsymbol{\rho}\cdot\nabla}\delta(\boldsymbol{X}-\boldsymbol{r})]\frac{-e}{T}F_0\sum_{n=-\infty}^{\infty}\frac{nI_n(i\rho k)}{(n-\bar{\omega})}e^{in\xi}\phi d^6\boldsymbol{Z}. \tag{96}$$

Using the facts that

$$\int [e^{\rho\cdot\nabla}\delta(\boldsymbol{X}-\boldsymbol{r})]Qd^6\boldsymbol{Z} = \int \delta(\boldsymbol{X}-\boldsymbol{r})e^{-\boldsymbol{\rho}\cdot\nabla}Qd^6\boldsymbol{Z}, \tag{97}$$

and

$$\int_0^{2\pi} e^{i(m+n)\xi}d\xi = \delta_{m,-n}2\pi, \tag{98}$$

we have

$$n_1 = \frac{2\pi}{(2\pi T/m)^{3/2}}\int \frac{-n_0 e\phi}{T}\exp\left(-\frac{v_\parallel^2+v_\perp^2}{2T/m}\right) \times \sum_{n=-\infty}^{\infty}\frac{nI_{-n}(-i\rho k)I_n(i\rho k)}{(n-\bar{\omega})}v_\perp dv_\parallel dv_\perp. \tag{99}$$

Carrying out the algebra with the help of some identities related to the Bessel functions,[25] we obtain

$$n_1 = n_0\frac{e\phi}{T}\sum_{n=1}^{\infty}\frac{2n^2}{\left(\frac{\omega}{\Omega}\right)^2 - n^2}\exp\left(-\frac{k^2T}{\Omega^2 m}\right)I_n\left(\frac{k^2T}{\Omega^2 m}\right). \tag{100}$$

Finally, the Poisson equation $-\nabla^2\phi = \Sigma_j 4\pi(en_1)_j$ gives the dispersion relation of the Bernstein wave,

$$1 = \sum_j \frac{4\pi n_0 e^2}{Tk^2}\sum_{n=1}^{\infty}\frac{2n^2}{\left(\frac{\omega}{\Omega}\right)^2 - n^2}\exp\left(-\frac{k^2T}{\Omega^2 m}\right)I_n\left(\frac{k^2T}{\Omega^2 m}\right). \tag{101}$$

V. CONCLUSIONS

The pullback transformations associated with the phase space coordinate system transformations have been developed here in the context of gyrokinetic theory. The necessity of such an approach arises from the existence of three different coordinate systems in the gyrokinetic theory. The familiar gyrocenter coordinate system, where the gyromotion is decoupled from the rest of particle's dynamics, is noncanonical and nonfibered. On the other hand, Maxwell's equations, which are needed to complete a kinetic system, are first only defined in the fibered laboratory phase space coordinate system. The pullback transformations are needed to connect the distribution functions in the gyrocenter coordinates and Maxwell's equations in the laboratory phase space coordinates. In order to gain a systematic understanding of the mathematical construction and physical implications of the pullback transformations, a geometric (coordinate independent) viewpoint has been for the moment integrals originally defined in the laboratory phase space coordinate system. The moment integrals in kinetic theories are geometrically interpreted as integrals of 3-forms over a 3-subset of the phase-space. Therefore, they are independent of the coordinate system used for the phase space. Starting from their representations in the laboratory phase space coordinate systems, one can "pullback" the distribution functions or the moment forms to express the moment integrals in an arbitrary new coordinate system, which can be noncanonical and nonfibered.

This general construction has been applied to the pullbacks of the guiding center transformation and the gyrocenter transformation. It has been demonstrated that the systematic treatment of the moment integrals provided by the pullback transformation is an essential component of the gyrokinetic theory itself. Without such a systematic approach, the gyrokinetic theory is actually incomplete and many important physics features, such as the gyrokinetic equilibrium and the compressional Alfvén wave, cannot be readily recovered. Illustrative examples have been discussed in Secs. III and IV.

ACKNOWLEDGMENTS

We thank Dr. W. M. Nevins, Dr. T. S. Hahm and Dr. H. Sugama for constructive discussion and comments.

This research was supported by the U.S. Department of Energy under Contract No. AC02-76CH03073.

[1] E. A. Friemann, R. C. Davidson, and B. Langdon, Phys. Fluids **9**, 1475 (1966).
[2] R. C. Davidson, Phys. Fluids **10**, 669 (1967).
[3] P. H. Rutherford and E. A. Friemann, Phys. Fluids **11**, 569 (1968).
[4] R. G. Littlejohn, J. Math. Phys. **20**, 2445 (1979).
[5] R. G. Littlejohn, Phys. Fluids **24**, 1730 (1981).
[6] R. G. Littlejohn, J. Plasma Phys. **29**, 111 (1983).
[7] W. W. Lee, Phys. Fluids **26**, 556 (1983).
[8] D. H. E. Dubin, J. A. Krommes, C. Oberman, and W. W. Lee, Phys. Fluids **26**, 3524 (1983).
[9] L. Chen and S. T. Tsai, Phys. Fluids **26**, 141 (1983).
[10] L. Chen and S. T. Tsai, Plasma Phys. **25**, 349 (1983).
[11] S. C. Yang and D. I. Choi, Phys. Lett. **108A**, 25 (1985).
[12] T. S. Hahm, Phys. Fluids **31**, 2670 (1988).
[13] T. S. Hahm, W. W. Lee, and A. Brizard, Phys. Fluids **31**, 1940 (1988).
[14] A. J. Brizard, J. Plasma Phys. **41**, 541 (1989).
[15] H. Sugama, Phys. Plasmas **7**, 466 (2000).
[16] W. W. Lee, J. Comput. Phys. **72**, 243 (1987).
[17] W. W. Lee and W. M. Tang, Phys. Fluids **31**, 612 (1988).
[18] S. E. Parker, W. W. Lee, and R. A. Santoro, Phys. Rev. Lett. **71**, 2042 (1993).
[19] J. C. Cummings, Ph.D. dissertation, Princeton University, 1995.
[20] A. M. Dimits, T. J. Williams, J. A. Byers, and B. I. Cohen, Phys. Rev. Lett. **77**, 71 (1996).
[21] R. D. Sydora, V. K. Decyk, and J. M. Dawson, Plasma Phys. Controlled Fusion **38**, A281 (1996).
[22] Z. Lin, T. S. Hahm, W. W. Lee, W. M. Tang, and R. B. White, Science **281**, 1835 (1998).
[23] H. Qin, W. M. Tang, and G. Rewoldt, Phys. Plasmas **5**, 1035 (1998).
[24] H. Qin, Ph.D. dissertation, Princeton University, 1998.
[25] H. Qin, W. M. Tang, W. W. Lee, and G. Rewoldt, Phys. Plasmas **6**, 1575 (1999).
[26] H. Qin, W. M. Tang, and G. Rewoldt, Phys. Plasmas **6**, 2544 (1999).
[27] H. Qin, W. M. Tang, G. Rewoldt, and W. W. Lee, Phys. Plasmas **7**, 991 (2000).
[28] H. Qin, W. M. Tang, and W. W. Lee, Phys. Plasmas **7**, 4433 (2000).
[29] T. S. Hahm (private communication, 2003).
[30] W. M. Tang, Nucl. Fusion **18**, 1089 (1978).
[31] Formally, the pullback of a k-form ω is defined as the following: Let $F:M\rightarrow N$ is a C^∞ mapping of manifolds, and $\Omega^k(N)$ is the C^∞ sections of the vector bundle of exterior k-forms on the tangent spaces of N. For $\omega \in \Omega^k(N)$, define a k-form on M, $F^*\omega:M\rightarrow\omega^k(M)$ by $F^*\omega(m) = (T_mF)^*\circ\omega\circ F(m)$. We say $F^*\omega$ is the pullback of ω by F. Especially, if g is a C^∞ smooth function of N, i.e., $g\in\Omega^0(N)$, we have $F^*g=g\circ F$ (see p. 112, R. Abraham and J. E. Marsden, *Foundation of Mechanics*, 2nd ed., Second printing with revisions, 1980). Also, if F is not diffeomorphism (assuming C^∞ and onto), the push-forward of F is not defined. To define push-forward, it is necessary that F is diffeomorphism (assuming C^∞ and onto).