PREPARED FOR THE U.S. DEPARTMENT OF ENERGY,
UNDER CONTRACT DE-AC02-76CH03073

PPPL-3482
UC-70

PPPL-3482

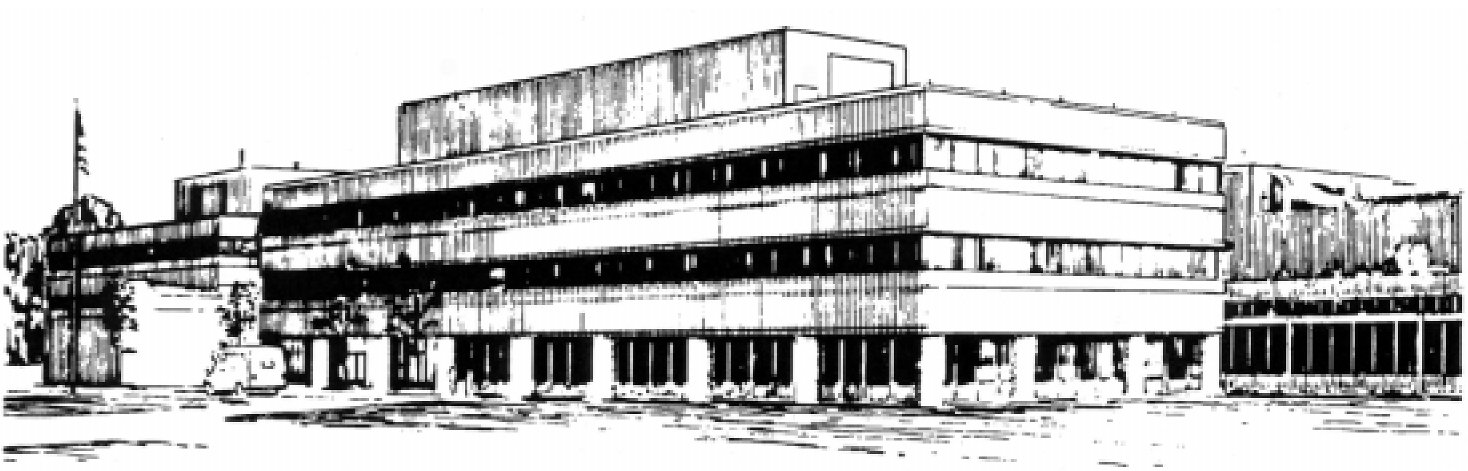
Gyrocenter-Gauge Kinetic Theory

by

H. Qin, W.M. Tang, and W.W. Lee

August 2000

PPPL
PRINCETON PLASMA PHYSICS LABORATORY



PRINCETON PLASMA PHYSICS LABORATORY
PRINCETON UNIVERSITY, PRINCETON, NEW JERSEY

Gyrocenter-Gauge Kinetic Theory

H. Qin, W. M. Tang, and W. W. Lee

Princeton Plasma Physics Laboratory

Princeton University, Princeton, NJ 08543

Abstract

Gyrocenter-gauge kinetic theory is developed as an extension of the existing gyrokinetic theories. In essence, the formalism introduced here is a kinetic description of magnetized plasmas in the gyrocenter coordinates which is fully equivalent to the Vlasov-Maxwell system in the particle coordinates. In particular, provided the gyroradius is smaller than the scale-length of the magnetic field, it can treat high frequency range as well as the usual low frequency range normally associated with gyrokinetic approaches. A significant advantage of this formalism is that it enables the direct particle-in-cell simulations of compressional Alfvén waves for MHD applications and of RF waves relevant to plasma heating in space and laboratory plasmas. The gyrocenter-gauge kinetic susceptibility for arbitrary wavelength and arbitrary frequency electromagnetic perturbations in a homogeneous magnetized plasma is shown to recover exactly the classical result obtained by integrating the Vlasov-Maxwell system in the particle coordinates. This demonstrates that all the waves supported by the Vlasov-Maxwell system can be studied using the gyrocenter-gauge kinetic model in the gyrocenter coordinates. This theoretical approach is so named to distinguish it from the existing gyrokinetic theory, which has been successfully developed and applied to many important low-frequency and long parallel wavelength problems, where the conventional meaning of "gyrokinetic" has been standardized. Besides the usual gyrokinetic distribution function, the gyrocenter-gauge kinetic theory emphasizes as well the gyrocenter-gauge distribution function, which sometimes contains all the physics of the problems being studied, and whose importance has not been realized previously. The gyrocenter-gauge distribution function enters Maxwell's equations through the pull-back transformation of the gyrocenter transformation, which depends on the perturbed fields. The efficacy of the gyrocenter-gauge kinetic approach is largely due to the fact that it directly decouples particle's gyromotion from its gyrocenter motion in the gyrocenter coordinates. As in the case of kinetic theories using guiding center coordinates, obtaining solutions for this kinetic system involves only following particles along their gyrocenter orbits. However, an added advantage here is that unlike the guiding center formalism, the gyrocenter coordinates used in this theory involves both the equilibrium and the perturbed components of the electromagnetic field. In terms of solving the kinetic system using particle simulation methods, the gyrocenter-gauge kinetic approach enables the reduction of computational complexity without the loss of important physical content.

I. INTRODUCTION

Most of the interesting plasmas in the laboratory and space are magnetized plasmas. Particle's motion in magnetized equilibrium plasmas consist of the fast gyromotion and the slow guiding center motion. Fast gyromotion puts a restrict constrain on the time step if particle simulations in the particle coordinates are used to simulate the magnetized plasmas. In the past twenty years, gyrokinetic theory has been developed to remove the fast gyromotion from the kinetic system for low frequency and long parallel wavelength phenomena.[1−15,23,24] Gyrokinetic particle simulations, which use much larger time step than the time scale of gyromotion,[4,25−31] have been successfully applied to the transport problem of fusion plasmas. Recently, gyrokinetic perpendicular dynamics[14,15] is identified and developed as an important component of the kinetic theory in the gyrocenter coordinates. The gyrokinetic perpendicular dynamics, which has not been systematically considered in the conventional gyrokinetic theories,[1–5,8–12] enables us to elegantly recover the compressional Aflvén wave, which is missing in the previous gyrokinetic description for waves with characteristic frequencies smaller than the gyrofrequency. Introducing the gyrokinetic perpendicular dynamics also extends the gyrokinetic model to arbitrary frequency modes. Since novel mathematical techniques, Lie perturbation and pull-back transformation, are utilized, the analytical formalism is much more general and transparent compared with previous attempts of gyrokinetic model for high frequency modes.[6,7]

In this paper, we further extend the gyrokinetic perpendicular dynamics into a kinetic description in the gyrocenter coordinates which includes all the magnetized plasma responses that are contained in the Vlasov-Maxwell system in the particle coordinates. In essence, the formalism introduced here is a kinetic description of magnetized plasmas in the gyrocenter coordinates which is fully equivalent to the Vlasov-Maxwell system in the particle coordinates. In particular, provided the gyroradius is smaller than the scale-length of the magnetic field, it can treat high frequency range as well as the usual low frequency range normally associated with gyrokinetic approaches. A significant advantage of this formalism is that it enables the direct particle-in-cell simulations of compressional Alfvén waves for MHD applications and of RF waves relevant to plasma heating in space and laboratory plasmas. The gyrocenter-gauge kinetic susceptibility for arbitrary wavelength and arbitrary frequency electromagnetic perturbations in a homogeneous magnetized plasma is shown to recover exactly the classical result obtained by integrating the Vlasov-Maxwell system in the particle coordinates. This demonstrates that all the waves supported by the Vlasov-Maxwell system can be studied using the gyrocenter-gauge kinetic model in the gyrocenter coordinates. We will refer this formalism as gyrocenter-gauge kinetic theory to distinguish it from the existing gyrokinetic theory, which has been successfully developed and applied to many important low-frequency and long parallel wavelength problems,[4,25–31] where the conventional meaning of "gyrokinetic" has been standardized. In this new theoretical approach, besides the usual gyrokinetic distribution function f, another indispensable distribution function S on the phase space and the corresponding governing equation is introduced. As shown in Sec. II, S sometimes plays an even more important role. The word "gyrocenter-gauge kinetic" reflects the fact that S is actually a gauge function associated with the symplectic gyrocenter transformation.

Before formally introducing the mathematical formalism, let's look at the basic concepts of the gyrocenter-gauge kinetic theory. As pointed out in Ref. 15, the absence of the compressional Aflvén wave and the difficulties of treating arbitrary frequency modes in the previous gyrokinetic models are fundamentally due to the lack of a systematic treatment for the plasma perpendicular response in these models. For a kinetic system, the kinetic equation can be viewed as a theoretical description for the response of the plasma, in terms of charge and current densities, to the electromagnetic field. It is not necessary to determine charge density independently, because we can solve for it from the continuity equation after knowing the current density. We can therefore infer that the reason that the compressional Aflvén wave is not recoverable from the previous gyrokinetic models must be the lack of complete information about the plasma response provided in these models. In the gyrocenter-gauge kinetic theory, all the information about the magnetized plasma response contained in the Vlasov-Maxwell system is kept by a complete description of the gyrocenter-gauge distribution function. The special features that particularly distinguish the gyrocenter-gauge kinetic theory in the gyrocenter coordinates from other gyrokinetic theories are the systematic treatment of the gyrocenter-gauge distribution function and the pull-back transformation. Since the construction of the gyrocenter coordinates involves the perturbed fields, the pull-back transformations of functions from the gyrocenter coordinates back to the particle coordinates must depend on the perturbed fields. This dependence shows up directly through the perturbed potential ϕ_1 and $\boldsymbol{A}_1$, as well as indirectly through the gyrocenter-gauge distribution function. The spirit of gyrocenter-gauge kinetic simplification is to decouple the gyromotion (the gyration due to the Lorentz force) from the gyrocenter motion (the orbit motion of gyrocenter due to the inhomogeneity of the magnetic field), instead of averaging out the gyromotion. This procedure can only be done rigorously and systematically using the Lie perturbation method. What gyrocenter-gauge kinetic theory offers is a simplified version of the Vlasov-Maxwell system by utilizing the fact that the particle's gyroradius is much smaller than the scale length of the magnetic field: $\epsilon_{B_0+B_1} \equiv \left|\rho/L_{B_0+B_1}\right| \ll 1$. As long as $\epsilon_{B_0+B_1}$ is small, we are able to construct a gyrocenter coordinate system in which the particle's gyromotion is decoupled from the rest of the particle dynamics. It is important to notice that the existence of the gyrocenter coordinates does not depend on the mode frequency directly. Therefore even when the mode frequency is comparable to or larger than the cyclotron frequency, we can still take advantage of the gyrocenter coordinates and simplify the kinetic system.[14,15] Three different coordinate systems appear in our formalism. $(\boldsymbol{x}, \boldsymbol{v})$ is the particle 'physical' coordinate system. $\boldsymbol{Z} = (\boldsymbol{X}, V_\parallel, \mu, \xi, w, t)$ is the (extended) 'guiding center' coordinate system in an equilibrium magnetic field. When the time-dependent electromagnetic field are introduced, we use the 'gyrocenter' coordinate system $\bar{\boldsymbol{Z}} = (\bar{\boldsymbol{X}}, \bar{V}_\parallel, \bar{\mu}, \bar{\xi}, \bar{w}, \bar{t})$ to describe the gyrocenter motion. Among other things, the most well-known difference between the guiding center motion and the gyrocenter motion is the polarization drift motion due to the time-dependent electrical perturbation, responsible for the finite Larmor radius correction to drift waves[4] and the compressional Alfvén wave.[15] We are following Brizard[11] and recent conventions[15] in using the terms 'gyrocenter' and 'guiding center' to distinguish these two different coordinate systems.

Recasting the Vlasov-Maxwell equations in the gyrocenter coordinates should not lose any physics content of the original system, if the mathematical procedure is carried out correctly while the simplification is achieved. In the gyrocenter-gauge kinetic theory, the information of the kinetic system is split into two parts, the usual gyrokinetic distribution f and the gyrocenter-gauge distribution function S. While f is gyrophase independent and mainly responsible for the shear Aflvén wave and drift waves, S is gyrophase dependent and solely responsible for the compressional Aflvén wave. We note that f and S is not a simple algebraic split of the full distribution function in the particle coordinates, but rather a geometric split of the information carried by it. In the gyrocenter-gauge kinetic system, the dynamics of f and S are governed by different kinetic equations in the gyrocenter phase space. Physics on the phase space should not depend on the choice of coordinate system. The guiding center coordinate system $\boldsymbol{Z} = (\boldsymbol{X}, V_{\parallel}, \mu, \xi)$ and the gyrocenter coordinate system $\bar{\boldsymbol{Z}} = (\bar{\boldsymbol{X}}, \bar{V}_{\parallel}, \bar{\mu}, \bar{\xi})$ are equivalent to the usual particle coordinate system $\boldsymbol{z} = (\boldsymbol{x}, \boldsymbol{v})$ in terms of describing the physics contained in the Vlasov-Maxwell equation system. We will show in Sec. III that the magnetized plasma linear response, expressed in the susceptibility, from the the gyrocenter-gauge kinetic theory recovers exactly the conventional magnetized plasma susceptibility derived from the Vlasov-Maxwell equations in the particle coordinate system. Recovering the classical plasma susceptibility completely from the gyrocenter-gauge kinetic theory guarantees the recovery, in the gyrocenter coordinate system, all the interesting waves that we have known from the classical theory, including the compressional Alfvén wave and the Bernstein wave, previously recovered by the gyrokinetic perpendicular dynamics[15].

Even though all coordinate systems are geometrically equivalent, the computational complexity involved and are different depending on the specific problems under investigation. For applications in magnetized plasmas, the advantage of the gyrocenter coordinate system lies at the fact that in and only in this coordinate system the fast time scale gyromotion is decoupled from the particle's gyrocenter orbit dynamics. For low frequency electrostatic and the shear Alfvén modes, the gyromotion is not important and is naturally decoupled from the system as if it is completely "averaged out". On the other hand, general frequency modes and the compressional Alfvén mode can be easily recovered by including the decoupled gyrocenter-gauge kinetic equation in the gyrocenter coordinate system, since the gyrocenter orbit motion is independent of the gyromotion. The current numerical codes and particle simulation codes based on gyrocenter orbit integration for low frequency electrostatic and shear Alvén modes can be extended to general frequency by appropriately adding in the gyrocenter-gauge component.

An interesting fact seldom discussed before is that the classical magnetized plasma susceptibility is actually gyro-phase independent. All the physics contained does not depend on the distribution over gyrophase. It is therefore natural and straightforward to work in the gyrocenter coordinates. As we will see later, it does not take too much calculation to obtain the plasma susceptibility after the basic formalism is rigorously set up.

The paper is organized as follows. In Sec. II, we introduce the basic analytical formalism of the gyrocenter-gauge kinetic theory. Then, the susceptibility of a magnetized plasma is derived from the gyrocenter-gauge kinetic theory in Sec. III. We show that this gyrocenter-gauge kinetic susceptibility recovers exactly the classical one. In the last section, we discuss the particle simulation method for the gyrocenter-gauge kinetic model and several related issues.

II. BASIC FORMALISM

A. Littlejohn's Standard Guiding Center Coordinates

We assume the equilibrium plasma is magnetostatic and magnetized, which means, by definition,

$$\epsilon_{B_0} \equiv |\frac{\boldsymbol{\rho}}{L_{B_0}}| \ll 1. \tag{1}$$

Here, $\boldsymbol{\rho} \equiv -\boldsymbol{v} \times \boldsymbol{b_0}/\Omega$ is the gyroradius, and $L_{B_0} \equiv |\boldsymbol{B}_0/\nabla\boldsymbol{B}_0|$ is the scale length of the equilibrium magnetic field $\boldsymbol{B}_0$. For magnetized plasmas, we can construct a set of non-canonical phase space coordinates in which the gyromotion is decoupled from the rest of the particle dynamics to any order in ϵ_{B_0}. This special set of coordinates is called "standard guiding center variables" by Littlejohn.[3] The underlying method is to look at the perturbation of the phase space Lagrangian when ϵ_{B_0} is small, and introduce a near identity coordinate transformation such that, in the new coordinate system, the gyromotion is decoupled. The guiding center transformation $T_{GC} : \boldsymbol{z} = (\boldsymbol{x}, \boldsymbol{v}) \longmapsto \boldsymbol{Z} = (\boldsymbol{X}, V_\|, \mu, \xi)$, which transfers particle coordinates $\boldsymbol{z} = (\boldsymbol{x}, \boldsymbol{v})$ into the standard guiding center coordinates $\boldsymbol{Z} = (\boldsymbol{X}, V_\|, \mu, \xi)$ can be found in Refs. 1–3, 11. Here, $\boldsymbol{X}$ is the configuration component of the guiding center coordinate, $V_\|$ is the parallel velocity, μ is the magnetic moment, and ξ is the gyrophase angle. For the present purpose, we do not need to display the expression except for the familiar

$$\boldsymbol{X} = \boldsymbol{x} - \boldsymbol{\rho}_0. \tag{2}$$

The regular phase space is extended to include the time coordinate t and its conjugate coordinate energy w such that time-dependent Hamiltonians can be treated on an equal footing with the time-independent ones. In the extended guiding center coordinates $(\boldsymbol{X}, V_\parallel, \mu, \xi, w, t)$, the extended phase space Lagrangian is[2,3,11,12]

$$
\begin{aligned}
\gamma_E =& \widehat{\gamma}_E - H_E d\tau \\
=& (\frac{e}{c}\boldsymbol{A} + mV_\parallel \boldsymbol{b} - \mu\frac{mc}{e}\boldsymbol{W}) \cdot d\boldsymbol{X} + \frac{mc}{e}\mu d\xi - w dt - (H - w)d\tau,
\end{aligned} \tag{3}
$$

where species subscripts are temporarily suppressed, and

$$
\boldsymbol{W} = \boldsymbol{R} + \frac{\boldsymbol{b}}{2}(\boldsymbol{b}\cdot\nabla\times\boldsymbol{b}), \qquad \boldsymbol{R} = (\nabla \boldsymbol{e}_1)\cdot \boldsymbol{e}_2, \qquad \boldsymbol{b} = \boldsymbol{B}/B. \tag{4}
$$

$\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ are unit vectors in two arbitrarily chosen perpendicular directions, and $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ are perpendicular to each other. All quantities are evaluated in the guiding center coordinates now. $\widehat{\gamma}_E$ gives the extended symplectic structure, $H_E = H - w$ is the extended Hamiltonian, and H is the regular Hamiltonian defined as

$$
H = \frac{mV_\parallel^2}{2} + \mu B.
$$

The corresponding Poisson bracket is obtained by inverting the matrix $\widehat{\gamma}_{Eij}$, which is the coefficient of the differential of the symplectic structure $d\widehat{\gamma}_E = \widehat{\gamma}_{Eij} dZ^i dZ^j$,[2,3,11]

$$
\begin{aligned}
\{F, G\} =& \frac{e}{mc}(\frac{\partial F}{\partial \xi}\frac{\partial G}{\partial \mu} - \frac{\partial F}{\partial \mu}\frac{\partial G}{\partial \xi}) - \frac{c\boldsymbol{b}}{eB_\parallel^*}\cdot[(\nabla F + \boldsymbol{W}\frac{\partial F}{\partial \xi}) \times (\nabla G + \boldsymbol{W}\frac{\partial G}{\partial \xi})] \\
&+ \frac{\boldsymbol{B}^*}{mB_\parallel^*}\cdot[(\nabla F + \boldsymbol{W}\frac{\partial F}{\partial \xi})\frac{\partial G}{\partial V_\parallel} - (\nabla G + \boldsymbol{W}\frac{\partial G}{\partial \xi})\frac{\partial F}{\partial V_\parallel}] + (\frac{\partial F}{\partial w}\frac{\partial G}{\partial t} - \frac{\partial F}{\partial t}\frac{\partial G}{\partial w}),
\end{aligned} \tag{5}
$$

where

$$
\boldsymbol{B}^* = \boldsymbol{B} + \frac{cmV_\parallel}{e}\nabla\times\boldsymbol{b}, \quad B_\parallel^* = \boldsymbol{b}\cdot\boldsymbol{B}^*. \tag{6}
$$

B. Symplectic Gyrocenter Transformation

When the time dependent perturbed electromagnetic field is introduced, the extended phase space Lagrangian still gives the dynamics of particles. However, it is perturbed accordingly,[10,11]

$$\begin{aligned} \gamma_E &= \gamma_{E0} + \gamma_{E1}, \\ \gamma_{E1} &= [\frac{e}{c}\boldsymbol{A}_1(T_{GC}^{-1}\boldsymbol{X}, t) \cdot d(T_{GC}^{-1}\boldsymbol{X})] - e\phi_1(T_{GC}^{-1}\boldsymbol{X}, t)d\tau, \end{aligned} \tag{7}$$

where T_{GC}^{-1} is the inverse of the guiding center transformation,

$$T_{GC}^{-1}\boldsymbol{X} = \boldsymbol{X} + \boldsymbol{\rho}_0 + O(\epsilon_B). \tag{8}$$

Expanding $d(T_{GC}^{-1}\boldsymbol{X})$, we obtain:

$$\gamma_{E1} = \frac{e}{c}\boldsymbol{A}_1(\boldsymbol{X} + \boldsymbol{\rho}_0, t) \cdot [(1 + \nabla\boldsymbol{\rho}_0) \cdot d\boldsymbol{X} + \frac{\partial \boldsymbol{\rho}_0}{\partial \mu} d\mu + \frac{\partial \boldsymbol{\rho}_0}{\partial \xi} d\xi] - e\phi_1(\boldsymbol{X} + \boldsymbol{\rho}_0, t)d\tau. \tag{9}$$

The essence of the Lie perturbation method is to introduce a near identity transformation from the equilibrium guiding center coordinates $\boldsymbol{Z} = (\boldsymbol{X}, V_\parallel, \mu, \xi, w, t)$ to the gyrocenter coordinates $\bar{\boldsymbol{Z}} = (\bar{\boldsymbol{X}}, \bar{V_\parallel}, \bar{\mu}, \bar{\xi}, \bar{w}, \bar{t})$ when the perturbed field is present such that the transformed extended phase space Lagrangian $\bar{\gamma}$ can be gyrophase independent.

For the transformation

$$\bar{Z}^i = (\ \mathrm{e}^{\boldsymbol{G}}\ \boldsymbol{Z})^i \approx Z^i + G^i(\boldsymbol{Z}), \tag{10}$$

the leading order transformed extended phase space Lagrangian is:

$$\bar{\gamma}_{E1} = \gamma_{E1} - i_{\boldsymbol{G}}\omega_{E0} + dS = \hat{\bar{\gamma}}_{E1} - \bar{H}_{E1}d\tau, \tag{11}$$

where $\omega_{E0} = d\gamma_{E0}$, S is an arbitrary gauge function, and $i_{\boldsymbol{G}}\omega_{E0}$ is the interior product between the vector field $\boldsymbol{G}$ and the two-form ω_{E0}. The fact that dS is a gauge transformation was pointed out by Littlejohn in Ref. 16, where the Lie perturbation method for Hamiltonian system in noncanonical coordinates was systematically introduced. It was also pointed out by Hahm in Ref. 9, where this method was first applied to the gyrokinetic theory. This point of view was subsequently adopted by Brizard.[11,12,17] In this paper, we refer S as the gyrocenter-gauge to reflect the fact that dS is the gauge transformation in the process of constructing the gyrocenter coordinates from the equilibrium guiding center coordinates and the perturbed fields.[18] We note that the Hamiltonian Lie perturbation procedure in noncanonical coordinates is different from the conventional canonical coordinate transformation, which can be characterized as those transformations $(\boldsymbol{q}, \boldsymbol{p}) \longrightarrow (\boldsymbol{Q}, \boldsymbol{P})$ which satisfy $\boldsymbol{p}d\boldsymbol{q} = \boldsymbol{P}d\boldsymbol{Q} + dS$ for some scalar S.[19,20,16,17] In the canonical limit, S serves as the scalar generating function which generates the canonical transformation. However, in the noncanonical cases, it is the vector field $\boldsymbol{G}$ that directly gives the transformation. The extra freedom associated with S allows us to pick the gauge which is computationally or analytically beneficial. There are several different ways to make $\widehat{\bar{\gamma}}_E$ and $\bar{H}_E d\tau$ gyrophase independent. We will choose $\boldsymbol{G}$ and S such that the transformation is symplectic, that is, there is no perturbation on the symplectic structure,

$$\widehat{\bar{\gamma}}_{E1} = 0. \tag{12}$$

Other non-symplectic transformations are also possible. Generally non-symplectic transformations are more algebraically involved.

Since we choose not to change the time variable t, $G^t = 0$. Other components of $\boldsymbol{G}$ are solved for from $\widehat{\widehat{\gamma}}_{E1} = 0$.

$$
\begin{aligned}
\boldsymbol{G}^{\boldsymbol{X}} &= -\frac{c}{eB_\parallel^*}\boldsymbol{b} \times (\frac{e}{c}\boldsymbol{A}_1 + \nabla S) - \frac{\boldsymbol{B}^*}{mB_\parallel^*}\frac{\partial S}{\partial V_\parallel} + O(\epsilon_B), \\
G^{V_\parallel} &= \frac{\boldsymbol{B}^*}{mB_\parallel^*} \cdot (\frac{e}{c}\boldsymbol{A}_1 + \nabla S) + O(\epsilon_B), \\
G^\mu &= \frac{e}{mc}(\frac{e}{c}\boldsymbol{A}_1 \cdot \frac{\partial \boldsymbol{\rho}_0}{\partial \xi} + \frac{\partial S}{\partial \xi}), \\
G^\xi &= -\frac{e}{mc}(\frac{e}{c}\boldsymbol{A}_1 \cdot \frac{\partial \boldsymbol{\rho}_0}{\partial \mu} + \frac{\partial S}{\partial \mu}) + O(\epsilon_B), \\
G^w &= -\frac{\partial S}{\partial t}.
\end{aligned}
\tag{13}
$$

The transformed Hamiltonian is thus uniquely determined by the choice of $\widehat{\widehat{\gamma}}_{E1} = 0$.

$$
\bar{H}_{E1} = H_{E1} - G^i\frac{\partial H_{E0}}{\partial x^i} + G^w = e\phi_1(\bar{\boldsymbol{X}} + \boldsymbol{\rho}_0, t) - \frac{e}{c}\boldsymbol{A}_1(\bar{\boldsymbol{X}} + \boldsymbol{\rho}_0, t) \cdot \{\bar{\boldsymbol{X}} + \boldsymbol{\rho}_0, H_{E0}\} - \{S, H_{E0}\},
\tag{14}
$$

in which

$$
\{\bar{\boldsymbol{X}} + \boldsymbol{\rho}_0, H_{E0}\} = \bar{\boldsymbol{V}} + \boldsymbol{v}_d,
\tag{15}
$$

where

$$
\bar{\boldsymbol{V}} = \bar{\boldsymbol{V}}_\perp + \bar{V}_\parallel \boldsymbol{b}, \qquad \bar{\boldsymbol{V}}_\perp = \{\boldsymbol{\rho}_0, H_{E0}\}.
\tag{16}
$$

In the calculation related to the gyrocenter transformation, we will only keep the lowest order in terms of ϵ_B, because the background FLR effects normally are not important.

$\bar{H}_{E1}$ has to be gyrophase independent as well. There is another freedom here. We choose

$$
\bar{H}_{E1} = \langle e\phi_1(\bar{\boldsymbol{X}} + \boldsymbol{\rho}_0, t) - \bar{\boldsymbol{V}} \cdot \frac{e}{c}\boldsymbol{A}_1(\bar{\boldsymbol{X}} + \boldsymbol{\rho}_0, t)\rangle,
\tag{17}
$$

where $\langle\ \rangle \equiv 1/2\pi \int_0^{2\pi} d\xi$ represents the gyrophase averaging operation. This leads to the equation determining the gauge function S:

$$
\begin{aligned}
\{S, H_{E0}\} &= \Omega\frac{\partial S}{\partial \xi} + \frac{\partial S}{\partial t} + \frac{\partial S}{\partial \bar{\boldsymbol{X}}} \cdot \{\boldsymbol{X}, H_{E0}\} + \frac{\partial S}{\partial V_\parallel}\{V_\parallel, H_{E0}\} \\
&= e\widetilde{\phi_1}(\bar{\boldsymbol{X}} + \boldsymbol{\rho}_0, t) - \frac{e}{c}\widetilde{\bar{\boldsymbol{V}} \cdot \boldsymbol{A}_1}(\bar{\boldsymbol{X}} + \boldsymbol{\rho}_0, t),
\end{aligned}
\tag{18}
$$

where $\widetilde{\phi}_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho}_0,t)$ and $\widetilde{\bar{\boldsymbol{V}}\cdot\boldsymbol{A}}_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho}_0,t)$ are the gyrophase dependent parts of $\phi_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho}_0,t)$ and $\bar{\boldsymbol{V}}\cdot\boldsymbol{A}_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho}_0,t)$ respectively.

$$\begin{aligned}\widetilde{\phi}_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho}_0,t) &= \phi_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho}_0,t)-\langle\phi_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho}_0,t)\rangle\\ \widetilde{\bar{\boldsymbol{V}}\cdot\boldsymbol{A}}_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho}_0,t) &= \bar{\boldsymbol{V}}\cdot\boldsymbol{A}_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho}_0,t)-\langle\bar{\boldsymbol{V}}\cdot\boldsymbol{A}_1(\bar{\boldsymbol{X}}+\boldsymbol{\rho}_0,t)\rangle.\end{aligned} \tag{19}$$

Here, we only carry out the analysis to the first order, we therefore study linear theory in this paper. Second order nonlinear theory is readily available by carrying out the analysis to the second order, but the algebra is somewhat tedious.

Since the transformation we have chosen is symplectic, $\widehat{\widetilde{\gamma}}_{E1}=0$, the Poisson bracket in the gyrocenter coordinates is the same as that in the guiding center coordinates, which is given by Eq. (5). After obtaining the desired gyrocenter coordinates, we will "push forward" objects on the original particle coordinates onto the new coordinates. The objects of physical interest here are Maxwell's equations and the Vlasov equation.

We will use $\boldsymbol{A}$ and ϕ to notate the perturbed field hereafter; the subscript "1" will be dropped. Unless clarity requires us to use the barred notation, we will also drop the bars for the gyrocenter coordinates hereafter.

C. Kinetic Equations, Pull-Back, and Push-Forward

In its geometric (coordinate independent) form, the Vlasov equation is $\{F,H_E\}=0$. In the gyrocenter coordinates, $\bar{F}$ and $\widetilde{F}$ can be decoupled because $\{\ \}$ and H_E are gyrophase independent.

$$\{\bar{F},H_E\}=0,\qquad\{\widetilde{F},H_E\}=0, \tag{20}$$

where $\bar{F} = \langle F \rangle$, and $\widetilde{F} = F - \bar{F}$. Let $F = F_0 + f$, where F_0 is the equilibrium distribution, and f is the perturbed distribution, we have

$$
\begin{aligned}
&\frac{\partial \bar{f}}{\partial t} + \dot{\boldsymbol{X}} \frac{\partial \bar{f}}{\partial \boldsymbol{X}} + \dot{V}_{\parallel} \frac{\partial \bar{f}}{\partial V_{\parallel}} = -\{\bar{F}_0, H_1\}, \\
&\frac{\partial \widetilde{f}}{\partial t} + \dot{\boldsymbol{X}} \frac{\partial \widetilde{f}}{\partial \boldsymbol{X}} + \dot{V}_{\parallel} \frac{\partial \widetilde{f}}{\partial V_{\parallel}} + \dot{\xi} \frac{\partial \widetilde{f}}{\partial \xi} = -\{\widetilde{F}_0, H_1\}.
\end{aligned}
\tag{21}
$$

However, $\bar{f}$ and $\widetilde{f}$ can not provide all the information about the distribution function in the phase space. The third kinetic equation in the gyrocenter-gauge kinetic theory is

$$
\frac{\partial S}{\partial t} + \dot{\boldsymbol{X}} \frac{\partial S}{\partial \boldsymbol{X}} + \dot{V}_{\parallel} \frac{\partial S}{\partial V_{\parallel}} + \dot{\xi} \frac{\partial S}{\partial \xi} = e\widetilde{\phi}(\boldsymbol{X} + \boldsymbol{\rho}_0, t) - \frac{e}{c} \widetilde{\boldsymbol{V} \cdot \boldsymbol{A}}(\boldsymbol{X} + \boldsymbol{\rho}_0, t). \tag{22}
$$

In the gyrocenter-gauge kinetic theory, the gyrocenter-gauge function S plays a significant role. S is not only a gauge, but more importantly, S is identified as the distribution function over the phase space which carries valuable physical information about the kinetic system. In many applications, such as the compressional Alfvén wave and the Bernstein wave, all the physics is hidden in S instead of the gyrokinetic distribution function f. Eq. (22) may look similar in form and dimension to Eq. (7) for a scalar field S_1 in Ref. 21 in the context of free energy method.[22] The scalar field S_1 in Ref. 21 is a first order generating function, which generates a canonical coordinate transformation, and therefore induces a transformation from the perturbed particle distribution to the unperturbed particle distribution f. Clearly, our gyro-center gauge function and Eq. (22) are different from the generating function S_1 and Eq. (7) in Ref. 21. First of all, S in our formalism is the gauge function for the noncanonical gyrocenter coordinate transformation, while S_1 in Ref. 21 is a generating function for a canonical transformation. Secondly, S_1 in Ref. 21 in the context of free energy method exists before the construction of gyrocenter coordinates or even when the gyrocenter coordinate system does not exist at all. Of course, after the gyrocenter coordinates are constructed, one can try to express S_1 and Eq. (7) in Ref. 21 in the gyrocenter coordinates with the purpose of developing a free energy method for the low-frequency gyrokinetic system,[21] a goal different from ours. As usual, Maxwell's equations are used to complete the gyrocenter-gauge kinetic system. It is not clear how to write Maxwell's equations directly in the gyrocenter coordinates. But the straightforward solution is to write Maxwell's equations in the particle coordinates first, then relate the charge and current densities to the distribution functions in the gyrocenter coordinates, *i.e.*, $\bar{f}$, $\widetilde{f}$, and S.

The Poisson equation is

$$-\nabla^2\phi(\boldsymbol{r},t) = 4\pi\sum_j e\int d^3\boldsymbol{v}f(\boldsymbol{r},\boldsymbol{v},t) + \frac{1}{c}\frac{\partial}{\partial t}\nabla\cdot\boldsymbol{A}(\boldsymbol{r},t), \tag{23}$$

where

$$\int d^3\boldsymbol{v}\ f(\boldsymbol{r},\boldsymbol{v},t) = \int d^6\boldsymbol{Z}\ [T^*_{GY}f](\boldsymbol{Z},t)\delta(T^{-1}_{GC}\boldsymbol{X}-\boldsymbol{r}). \tag{24}$$

Ampere's law is

$$\nabla\times(\nabla\times\boldsymbol{A}(\boldsymbol{r},t)) = \frac{4\pi}{c}\sum_j e\int d^3\boldsymbol{v}\ \boldsymbol{v}f(\boldsymbol{r},\boldsymbol{v},t), \tag{25}$$

where

$$\int d^3\boldsymbol{v}\boldsymbol{v}f(\boldsymbol{r},\boldsymbol{v},t) = \int d^6\boldsymbol{Z}\ \boldsymbol{V}_{GC}(\boldsymbol{Z})[T^*_{GY}f](\boldsymbol{Z},t)\delta(T^{-1}_{GC}\boldsymbol{X}-\boldsymbol{r}). \tag{26}$$

In the above equations, $d^6\boldsymbol{Z}$ is understood to be $(B^*_\parallel/m)\ d^3\boldsymbol{X}dV_\parallel d\mu d\xi$. T^*_{GY} is the pull-back transformation, which transforms the perturbed distribution f in the gyrocenter coordinates into that in the guiding center coordinates. T^{-1}_{GC} is the inverse of T_{GC} that transforms the particle physical coordinates $(\boldsymbol{r},\boldsymbol{v},t)$ into the guiding center coordinates. We assume the guiding center transformation T_{GC} and the corresponding pull-back transformation T^*_{GC}, and the gyrocenter transformation T_{GY} and the corresponding pull-back transformation T^*_{GY}

are one-one onto (bijective). Generally, for a macroscopic quantity $Q(\boldsymbol{r})$ in the particle coordinates and its microscopic counterpart in phase space $q(\boldsymbol{r},\boldsymbol{v})$, we have[5,9–11]

$$Q(\boldsymbol{r}) = \int q(\boldsymbol{r},\boldsymbol{v}) f_P(\boldsymbol{r},\boldsymbol{v},t)\ d^3\boldsymbol{v} = \int \delta(\boldsymbol{x}-\boldsymbol{r}) q(\boldsymbol{r},\boldsymbol{v}) f_P(\boldsymbol{z},t)\ d^6\boldsymbol{z}. \tag{27}$$

In the guiding center coordinates $\boldsymbol{Z}=(\boldsymbol{X},V_\parallel,\mu,\xi)$,

$$Q(\boldsymbol{r}) = \int [T_{GC}^{*}{}^{-1} q](\boldsymbol{Z}) f_{GC}(\boldsymbol{Z},t)\delta(T_{GC}^{-1}\boldsymbol{X}-\boldsymbol{r})\ d^6\boldsymbol{Z}. \tag{28}$$

Replacing $f_{GC}(\boldsymbol{Z},t)$ by its pull-back from the gyrocenter coordinates, we get,

$$Q(\boldsymbol{r}) = \int [T_{GC}^{*}{}^{-1} q](\boldsymbol{Z}) [T_{GY}^{*} f_{GY}](\boldsymbol{Z},t)\delta(T_{GC}^{-1}\boldsymbol{X}-\boldsymbol{r})\ d^6\boldsymbol{Z}. \tag{29}$$

The pull-back transformation from the gyrocenter coordinates to the guiding center coordinates is easily obtained from the expression for G given by Eq. (13),

$$\begin{aligned}
T_{GY}^{*}F &= F + L_{\boldsymbol{G}}F = F - \frac{\boldsymbol{b}}{B_\parallel^*}\times[\boldsymbol{A}(\boldsymbol{X}+\boldsymbol{\rho}_0,t)+\frac{c}{e}\nabla S]\cdot\nabla F - \frac{\boldsymbol{B}^*}{mB_\parallel^*}\frac{\partial S}{\partial V_\parallel}\cdot\nabla F \\
&+ \frac{e}{mc}\frac{\boldsymbol{B}^*}{B_\parallel^*}\cdot[\boldsymbol{A}(\boldsymbol{X}+\boldsymbol{\rho}_0,t)+\frac{c}{e}\nabla S]\frac{\partial F}{\partial V_\parallel} + \frac{e}{mc}[\frac{e}{c}\boldsymbol{A}(\boldsymbol{X}+\boldsymbol{\rho}_0,t)\cdot\frac{\partial \boldsymbol{\rho}_0}{\partial\xi}+\frac{\partial S}{\partial\xi}]\frac{\partial F}{\partial\mu} \\
&- [\frac{e}{c}\boldsymbol{A}(\boldsymbol{X}+\boldsymbol{\rho}_0,t)\cdot\frac{\partial \boldsymbol{\rho}_0}{\partial\mu}+\frac{\partial S}{\partial\mu}]\frac{\partial F}{\partial\xi} + O(\epsilon_B),
\end{aligned} \tag{30}$$

where $L_{\boldsymbol{G}}F$ represents the Lie derivative of F with respect to the vector field $\boldsymbol{G}$. As we will see in the next section, the pull-back transformation T_{GY}^{*} and therefore the gyrocenter-gauge distribution S lie at the center of the gyrocenter-gauge kinetic theory.

After the pull-back of f into the particle coordinates, the configuration variable $\boldsymbol{r}$ of the particle coordinates in Maxwell's equations can be viewed as a dummy variable, and can be replaced by the configuration variable $\boldsymbol{X}$ of the gyrocenter coordinates. As a result, we effectively obtain the push-forward of Maxwell's equation on the gyrocenter coordinates.

$$-\nabla^2\phi(\boldsymbol{X}',t) = 4\pi\sum_j e\int d^6\boldsymbol{Z}\ [T_{GY}^{*}f](\boldsymbol{Z},t)\delta(T_{GC}^{-1}\boldsymbol{X}-\boldsymbol{X}') + \frac{1}{c}\frac{\partial}{\partial t}\nabla\cdot\boldsymbol{A}(\boldsymbol{X}',t), \tag{31}$$

$$\nabla \times (\nabla \times \boldsymbol{A}(\boldsymbol{X}', t)) = \frac{4\pi}{c} \sum_j e \int d^6 \boldsymbol{Z}\, \boldsymbol{V}_{GC}(\boldsymbol{Z}) [T_{GY}^* f](\boldsymbol{Z}, t) \delta(T_{GC}^{-1} \boldsymbol{X} - \boldsymbol{X}'). \tag{32}$$

III. SUSCEPTIBILITY

As an electromagnetic medium, a plasma can be faithfully characterized by its susceptibility. For example, all the waves supported by plasmas can be derived from the plasma susceptibility. To a large degree, a theoretical model for plasmas can be characterized by the susceptibility it predicts. In this section, we derive the susceptibility for a magnetized plasma from the gyrocenter-gauge kinetic model, and prove that it recovers exactly the well-known result derived from the Vlasov-Maxwell system in the particle coordinates. By this recovery, we show that gyrocenter-gauge kinetic theory, as an extension of the gyrokinetic theory, includes all the physics that can be described by the Vlasov-Maxwell system in the particle coordinates.

We consider a homogeneous magnetized plasma with a constant magnetic field in the $\boldsymbol{e}_z$ direction. For a linear perturbation of the form $\exp(i\boldsymbol{k} \cdot \boldsymbol{r} - i\omega t)$, we can always choose the coordinate system such that $k_y = 0$, $k_\perp = k_x$, and $k_\parallel = k_z$. By definition,

$$\boldsymbol{j} = -\frac{i\omega}{4\pi} \chi \cdot \boldsymbol{E}. \tag{33}$$

To find out χ, we only need to express $\boldsymbol{j}$ in terms of $\boldsymbol{E}$. Our starting point is the pull-back formula

$$\boldsymbol{j}_1 = \left\{ e \int (\boldsymbol{V}_\perp + V_\parallel \boldsymbol{b}) [T_{GY}^* (F_0 + f)](\boldsymbol{Z}) \delta(\boldsymbol{X} + \boldsymbol{\rho}_0 - \boldsymbol{r})\, d^6 \boldsymbol{Z} \right\}_1 , \tag{34}$$

where $\boldsymbol{j}_1$ is the first order current flow in the laboratory frame. $\{\ \ \}_1$ represents the first order of the expression inside $\{\ \ \}$. In addition, we have used the relationship

$$T_{GC}^{*-1} \boldsymbol{v} = \boldsymbol{V}_\perp + V_\parallel \boldsymbol{b}. \tag{35}$$

As usual, it is reasonable to assume $\widetilde{F}_0 = 0$. Then, the kinetic equation for $\widetilde{f}$ is homogeneous and does not depend on the perturbed field.

$$\frac{\partial \widetilde{f}}{\partial t} + \dot{\boldsymbol{X}} \frac{\partial \widetilde{f}}{\partial \boldsymbol{X}} + \dot{V}_{\parallel} \frac{\partial \widetilde{f}}{\partial V_{\parallel}} + \dot{\xi} \frac{\partial \widetilde{f}}{\partial \xi} = 0. \tag{36}$$

For the initial value problem, $\widetilde{f}$ is purely a residual of the gyrophase dependent part of the initial $\widetilde{f}$. If we assume $\widetilde{f} = 0$ initially, then $\widetilde{f}$ vanishes all the time. The physics for the linear susceptibility does not depend on initial condition. We can therefore let $\widetilde{f} = 0$ and $f = \bar{f}$ for the current purpose. Useful information about the gyrophase dependent part of the distribution function is carried by S.

The integral in Eq. (34) is related to F_0, f, and S through:

$$[T^*_{GY}(F_0 + f)]_1 = f + \frac{e}{mc} \boldsymbol{b} \cdot [\boldsymbol{A}(\boldsymbol{X} + \boldsymbol{\rho}_0) + \frac{c}{e} \nabla S] \frac{\partial F_0}{\partial V_{\parallel}} + \frac{e}{mc} [\frac{e}{c} \boldsymbol{A}(\boldsymbol{X} + \boldsymbol{\rho}_0) \cdot \frac{\partial \boldsymbol{\rho}_0}{\partial \xi} + \frac{\partial S}{\partial \xi}] \frac{\partial F_0}{\partial \mu}. \tag{37}$$

To solve for f and S, we first calculate the linear drive H_1 and $e\widetilde{\phi}(\boldsymbol{X} + \boldsymbol{\rho}_0, t) - e/c\widetilde{\boldsymbol{V} \cdot \boldsymbol{A}}(\boldsymbol{X} + \boldsymbol{\rho}_0, t)$. Choosing the coordinate system for ξ such that

$$\begin{aligned} \boldsymbol{V}_{\perp} &= -V_{\perp} [\boldsymbol{e}_x \sin \xi + \boldsymbol{e}_y \cos \xi], \\ \boldsymbol{\rho}_0 &= \frac{V_{\perp}}{\Omega} [\boldsymbol{e}_x \cos \xi - \boldsymbol{e}_y \sin \xi], \end{aligned} \tag{38}$$

We have:

$$\begin{aligned} \phi(\boldsymbol{X} + \boldsymbol{\rho}_0) &= \mathrm{e}^{\boldsymbol{\rho}_0 \cdot \nabla} \phi(\boldsymbol{X}), \\ \langle \phi(\boldsymbol{X} + \boldsymbol{\rho}_0) \rangle &= \langle \mathrm{e}^{\boldsymbol{\rho}_0 \cdot \nabla} \rangle \phi(\boldsymbol{X}) = J_0 \phi(\boldsymbol{X}), \\ \widetilde{\phi}(\boldsymbol{X} + \boldsymbol{\rho}_0) &= (\mathrm{e}^{\boldsymbol{\rho}_0 \cdot \nabla} - J_0) \phi(\boldsymbol{X}), \\ J_0 &= J_0(\frac{V_{\perp} k_{\perp}}{\Omega}). \end{aligned} \tag{39}$$

Similarly,

$$
\begin{aligned}
\boldsymbol{V}\cdot\boldsymbol{A}(\boldsymbol{X}+\boldsymbol{\rho}_0) &= V_\parallel A_\parallel(\boldsymbol{X}+\boldsymbol{\rho}_0)+\boldsymbol{V}_\perp\cdot\boldsymbol{A}_\perp(\boldsymbol{X}+\boldsymbol{\rho}_0),\\
\langle\boldsymbol{V}_\perp\cdot\boldsymbol{A}_\perp(\boldsymbol{X}+\boldsymbol{\rho}_0)\rangle &= -V_\perp J_1 A_y,\\
\widetilde{\boldsymbol{V}_\perp\cdot\boldsymbol{A}}_\perp(\boldsymbol{X}+\boldsymbol{\rho}_0) &= -\ \mathrm{e}^{i\lambda\cos\xi}\ V_\perp\sin\xi A_x-(\ \mathrm{e}^{i\lambda\cos\xi}\ \cos\xi+J_1)V_\perp A_y,\\
J_1 &= J_1(\frac{V_\perp k_\perp}{\Omega}),\\
\lambda &= \rho_0 k_\perp=\rho_0 k_x.
\end{aligned}
\tag{40}
$$

The expression for H_1 and $e\widetilde{\phi}(\boldsymbol{X}+\boldsymbol{\rho}_0)-e/c\widetilde{\boldsymbol{V}\cdot\boldsymbol{A}}(\boldsymbol{X}+\boldsymbol{\rho}_0)$ are

$$
\begin{aligned}
H_1 &= e[J_0(\phi-\frac{V_\parallel}{c}A_\parallel)+\frac{V_\perp}{c}J_1A_y],\\
e\widetilde{\phi}(\boldsymbol{X}+\boldsymbol{\rho}_0)-\frac{e}{c}\widetilde{\boldsymbol{V}\cdot\boldsymbol{A}}(\boldsymbol{X}+\boldsymbol{\rho}_0) &= e[(\ \mathrm{e}^{i\lambda\cos\xi}\ -J_0)(\phi-\frac{V_\parallel}{c}A_z)+\ \mathrm{e}^{i\lambda\cos\xi}\ \sin\xi\frac{V_\perp}{c}A_x\\
&\quad+(\ \mathrm{e}^{i\lambda\cos\xi}\ \cos\xi+J_1)\frac{V_\perp}{c}A_y].
\end{aligned}
\tag{41}
$$

From the kinetic equation for f

$$\frac{\partial f}{\partial t}+V_\parallel\boldsymbol{b}\cdot\nabla f=\frac{1}{m}\boldsymbol{b}\cdot\nabla H_1\frac{\partial F_0}{\partial V_\parallel}, \tag{42}$$

we easily know the solution for f,

$$f=\frac{-ek_z}{m(\omega-k_zV_\parallel)}[J_0(\phi-V_\parallel/cA_\parallel)+\frac{V_\perp}{c}J_1A_y]\frac{\partial F_0}{\partial V_\parallel}. \tag{43}$$

One quickly notices that for those modes with $k_z=0$, such as the compressional Alfvén wave and the Bernstein wave, $f=0$, all the physics must be inside the gyrocenter-gauge distribution function S.

Introducing S^* defined by

$$S^*=S-\frac{eH_1}{i(\omega-k_zV_\parallel)}, \tag{44}$$

we have the kinetic equation for S^* from that for S,

$$\begin{aligned}\Omega\frac{\partial S^*}{\partial\xi}-i(\omega-k_zV_\parallel)S^* &= e\ \mathrm{e}^{i\lambda\cos\xi}\ [\phi(\boldsymbol{X})-\frac{1}{c}\boldsymbol{V}\cdot\boldsymbol{A}(\boldsymbol{X})]\\ &= e\ \mathrm{e}^{i\lambda\cos\xi}\ [(\phi-\frac{1}{c}V_zA_z)+\frac{V_\perp}{c}A_x\sin\xi+\frac{V_\perp}{c}A_y\cos\xi].\end{aligned}\tag{45}$$

The pull-back transformation depends only on S^* due to a cancellation,

$$\begin{aligned}[T^*_{GY}(F_0+f)]_1 &= \frac{e}{mc}A_z(\boldsymbol{X}+\boldsymbol{\rho}_0)\frac{\partial F_0}{\partial V_\parallel}+\frac{ik_z}{m}S^*\frac{\partial F_0}{\partial V_\parallel}\\ &+\frac{e}{mc}[\frac{e}{c}\boldsymbol{A}(\boldsymbol{X}+\boldsymbol{\rho}_0)\cdot\frac{\partial\boldsymbol{\rho}_0}{\partial\xi}+\frac{\partial S^*}{\partial\xi}]\frac{\partial F_0}{\partial\mu}.\end{aligned}\tag{46}$$

As a consequence, all the physics is included in S^*. Using (A9) to expand $\exp(i\lambda\cos\xi)$, we can solve Eq. (45) for S^*,

$$\begin{aligned}S^* = \frac{e}{\Omega}\sum_{n=-\infty}^{n=\infty}\{&\frac{I_n(i\lambda)\ \mathrm{e}^{in\xi}}{i(n-\overline{\omega}+\overline{k_zV_\parallel})}(\phi-\frac{V_\parallel}{c}A_z)\\ &+[\frac{I_n(i\lambda)\ \mathrm{e}^{i(n+1)\xi}}{-2(n+1-\overline{\omega}+\overline{k_zV_\parallel})}-\frac{I_n(i\lambda)\ \mathrm{e}^{i(n-1)\xi}}{-2(n-1-\overline{\omega}+\overline{k_zV_\parallel})}]\frac{V_\perp}{c}A_x\\ &+[\frac{I_n(i\lambda)\ \mathrm{e}^{i(n+1)\xi}}{2i(n+1-\overline{\omega}+\overline{k_zV_\parallel})}+\frac{I_n(i\lambda)\ \mathrm{e}^{i(n-1)\xi}}{2i(n-1-\overline{\omega}+\overline{k_zV_\parallel})}]\frac{V_\perp}{c}A_y\},\end{aligned}\tag{47}$$

where $\overline{\omega}=\omega/\Omega$ and $\overline{k_zV_\parallel}=k_zV_\parallel/\Omega$. We can re-arrange $\sum_{n=-\infty}^{n=\infty}$ using (A16), (A17), and (A18) to get

$$\begin{aligned}S^* = \frac{e}{\Omega}\sum_{n=-\infty}^{n=\infty}\{&\frac{I_n(\lambda)\ \mathrm{e}^{in\xi}}{i(n-\overline{\omega}+\overline{k_zV_\parallel})}(\phi-\frac{V_\parallel}{c}A_z)\\ &+\frac{nI_n(i\lambda)\ \mathrm{e}^{in\xi}}{-i\lambda(n-\overline{\omega}+\overline{k_zV_\parallel})}\frac{V_\perp}{c}A_x+\frac{I_n'(i\lambda)\ \mathrm{e}^{in\xi}}{i(n-\overline{\omega}+\overline{k_zV_\parallel})}\frac{V_\perp}{c}A_y\},\end{aligned}\tag{48}$$

$$\begin{aligned}\frac{\partial S^*}{\partial\xi} = \frac{e}{\Omega}\sum_{n=-\infty}^{n=\infty}\{&\frac{nI_n(\lambda)\ \mathrm{e}^{in\xi}}{i(n-\overline{\omega}+\overline{k_zV_\parallel})}(\phi-\frac{V_\parallel}{c}A_z)\\ &+\frac{n^2I_n(i\lambda)\ \mathrm{e}^{in\xi}}{-i\lambda(n-\overline{\omega}+\overline{k_zV_\parallel})}\frac{V_\perp}{c}A_x+\frac{nI_n'(i\lambda)\ \mathrm{e}^{in\xi}}{i(n-\overline{\omega}+\overline{k_zV_\parallel})}\frac{V_\perp}{c}A_y\}.\end{aligned}\tag{49}$$

Our strategy here is to calculate the current through the pull-back formula in terms of potentials ϕ and $\boldsymbol{A}$, and find out the susceptibility $\boldsymbol{\chi}_p$ defined by

$$\boldsymbol{j} = -\frac{i\omega}{4\pi}\boldsymbol{\chi}_p \cdot \begin{pmatrix} A_x \\ A_y \\ A_z \\ \phi \end{pmatrix}. \tag{50}$$

Here, subscript "p" refers to the fact that $\boldsymbol{\chi}_p$ is the susceptibility matrix connecting $\boldsymbol{j}$ and potentials ϕ and $\boldsymbol{A}$, while $\boldsymbol{\chi}$ is reserved for the susceptibility connecting $\boldsymbol{j}$ and electric field $\boldsymbol{E}$. $\boldsymbol{\chi}_p$ and $\boldsymbol{\chi}$ are simply related by

$$\boldsymbol{\chi}_p = i\boldsymbol{\chi} \cdot \begin{pmatrix} \frac{\omega}{c} & 0 & 0 & -k_x \\ 0 & \frac{\omega}{c} & 0 & 0 \\ 0 & 0 & \frac{\omega}{c} & -k_z \end{pmatrix}. \tag{51}$$

Furthermore, to break the expression into manageable pieces, we split $\boldsymbol{\chi}_p$ into two parts,

$$\boldsymbol{\chi}_p = \boldsymbol{\chi}_p^{\parallel} + \boldsymbol{\chi}_p^{\perp}, \tag{52}$$

where $\boldsymbol{\chi}_p^{\parallel}$ and $\boldsymbol{\chi}_p^{\perp}$ are terms proportional to $\partial F_0/\partial V_\parallel$ and $\partial F_0/\partial V_\perp$ respectively.

Let's start from the complete expression for $\boldsymbol{j}_1$,

$$\begin{aligned} \boldsymbol{j}_1(\boldsymbol{r}) &= e\int \delta(\boldsymbol{X}+\boldsymbol{\rho}_0-\boldsymbol{r})\ d^6\boldsymbol{Z}(\boldsymbol{V}_\perp + V_\parallel \boldsymbol{b})\{\frac{e}{mc}A_z(\boldsymbol{X}+\boldsymbol{\rho}_0)\frac{\partial F_0}{\partial V_\parallel} + \frac{ik_z}{m}S^*\frac{\partial F_0}{\partial V_\parallel} \\ &+ \frac{e}{mc}[\frac{e}{c}\boldsymbol{A}(\boldsymbol{X}+\boldsymbol{\rho}_0)\cdot\frac{\partial \boldsymbol{\rho}_0}{\partial \xi} + \frac{\partial S^*}{\partial \xi}]\frac{\partial F_0}{\partial \mu}\} \\ &= e\int d^3\boldsymbol{V}\ \mathrm{e}^{-\boldsymbol{\rho}_0\cdot\nabla}\ (\boldsymbol{V}_\perp + V_\parallel \boldsymbol{b})\{\frac{e}{mc}A_z(\boldsymbol{X}+\boldsymbol{\rho}_0)\frac{\partial F_0}{\partial V_\parallel} + \frac{ik_z}{m}S^*\frac{\partial F_0}{\partial V_\parallel} \\ &+ \frac{e}{mc}[\frac{e}{c}\boldsymbol{A}(\boldsymbol{X}+\boldsymbol{\rho}_0)\cdot\frac{\partial \boldsymbol{\rho}_0}{\partial \xi} + \frac{\partial S^*}{\partial \xi}]\frac{\partial F_0}{\partial \mu}\}\Big|_{\boldsymbol{X}\longmapsto\boldsymbol{r}}, \end{aligned} \tag{53}$$

where "$|_{\boldsymbol{X}\longmapsto\boldsymbol{r}}$" means replacing $\boldsymbol{X}$ by $\boldsymbol{r}$ after the velocity integral is finished. The expressions for S^* and $\partial S^*/\partial\xi$, Eqs. (48) and (49), can be substituted into the above equation to express $\boldsymbol{j}_1(\boldsymbol{r})$ in terms of the potentials ϕ and $\boldsymbol{A}$ exclusively.

First, we look at the $\partial F_0/\partial V_\parallel$ term in $\boldsymbol{j}_1(\boldsymbol{r})\cdot\boldsymbol{b}$,

$$\begin{aligned}
& e\int_0^\infty V_\perp dV_\perp \int_{-\infty}^\infty dV_\parallel \int_0^{2\pi} d\xi V_\parallel \frac{\partial F_0}{\partial V_\parallel} \sum_{n'=-\infty}^{n'=\infty} I_{n'}(-i\lambda)\ \mathrm{e}^{in'\xi} \\
& \{\ \mathrm{e}^{i\lambda\cos\xi}\ \frac{e}{mc}A_z + \frac{ik_z e}{m\Omega}\sum_{n=-\infty}^{n=\infty}[\frac{I_n(i\lambda)\ \mathrm{e}^{in\xi}}{i(n-\overline{\omega}+\overline{k_zV_\parallel})}(\phi-\frac{V_\parallel}{c}A_z) \\
& + \frac{n\ \mathrm{e}^{in\xi}\ I_n(i\lambda)}{-i\lambda(n-\overline{\omega}+\overline{k_zV_\parallel})}\frac{V_\perp}{c}A_x + \frac{\mathrm{e}^{in\xi}\ I_n'(i\lambda)}{i(n-\overline{\omega}+\overline{k_zV_\parallel})}\frac{V_\perp}{c}A_y]\} \\
& = e2\pi\int_0^\infty V_\perp dV_\perp \int_{-\infty}^\infty dV_\parallel V_\parallel \frac{\partial F_0}{\partial V_\parallel}\sum_{n=-\infty}^{n=\infty}[\frac{nJ_n^2(\lambda)}{-(n-\overline{\omega}+\overline{k_zV_\parallel})}\frac{k_z e}{k_x mc}A_x \\
& + \frac{iJ_n'(\lambda)J_n(\lambda)}{-(n-\overline{\omega}+\overline{k_zV_\parallel})}\frac{k_zV_\perp e}{\Omega mc}A_y + \frac{J_n^2(\lambda)(n-\overline{\omega})}{(n-\overline{\omega}+\overline{k_zV_\parallel})}\frac{e}{mc}A_z + \frac{J_n^2(\lambda)}{(n-\overline{\omega}+\overline{k_zV_\parallel})}\frac{k_z e}{m\Omega}\phi].
\end{aligned} \tag{54}$$

In the above derivation, we have used identity (A12) for the A_x and ϕ terms, (A13) for the A_y term, and (A5) for the A_z term. This equation gives the third row of $\boldsymbol{\chi}_p^\parallel$.

For the $\partial F_0/\partial V_\perp$ term in $\boldsymbol{j}_1(\boldsymbol{r})\cdot\boldsymbol{b}$, we have

$$\begin{aligned}
& e\int_0^\infty V_\perp dV_\perp \int_{-\infty}^\infty dV_\parallel \int_0^{2\pi} d\xi V_\parallel \frac{\partial F_0}{\partial V_\perp}\frac{B}{mV_\perp}\sum_{n'=-\infty}^{n'=\infty} I_{n'}(-i\lambda)\ \mathrm{e}^{in'\xi}\ \frac{e}{mc} \\
& \{\frac{e}{c}\ \mathrm{e}^{i\lambda\cos\xi}\ \frac{V_\perp}{\Omega}(-A_x\sin\xi - A_y\cos\xi) + \frac{e}{\Omega}\sum_{n=-\infty}^{n=\infty}[\frac{nI_n(i\lambda)\ \mathrm{e}^{in\xi}}{(n-\overline{\omega}+\overline{k_zV_\parallel})}(\phi-\frac{V_\parallel}{c}A_z) \\
& + \frac{n^2 I_n(i\lambda)\ \mathrm{e}^{in\xi}}{-\lambda(n-\overline{\omega}+\overline{k_zV_\parallel})}\frac{V_\perp}{c}A_x + \frac{nI_n'(i\lambda)\ \mathrm{e}^{in\xi}}{(n-\overline{\omega}+\overline{k_zV_\parallel})}\frac{V_\perp}{c}A_y]\} \\
& = \frac{e^2}{m}2\pi\int_0^\infty V_\perp dV_\perp \int_{-\infty}^\infty dV_\parallel \frac{\partial F_0}{\partial V_\perp}\frac{V_\parallel}{V_\perp}\sum_{n=-\infty}^{n=\infty}[\frac{J_n^2(\lambda)}{(n-\overline{\omega}+\overline{k_zV_\parallel})}(\phi-\frac{V_\parallel}{c}A_z) \\
& + \frac{\overline{\omega}-\overline{k_zV_\parallel}}{-(n-\overline{\omega}+\overline{k_zV_\parallel})}\frac{nJ_n^2(\lambda)V_\perp}{\lambda c}A_x + \frac{i(\overline{\omega}-\overline{k_zV_\parallel})}{-(n-\overline{\omega}+\overline{k_zV_\parallel})}\frac{J_n(\lambda)J_n'(\lambda)V_\perp}{c}A_y],
\end{aligned} \tag{55}$$

where we have used identity (A12) for the ϕ and A_z terms, (A12) and (A3) for the A_x term, and (A13) and (A4) for the A_y term. The third row of $\boldsymbol{\chi}_p^{\perp}$ can be easily read off from the above equation.

The algebra involved for the perpendicular component of $\boldsymbol{j}_1$ is a little bit more complicated. For the $\partial F_0/\partial V_\parallel$ term in $\boldsymbol{j}_1(\boldsymbol{r})\cdot\boldsymbol{e}_x$, we have

$$
\begin{aligned}
& e\int_0^\infty V_\perp dV_\perp \int_{-\infty}^{\infty} dV_\parallel \int_0^{2\pi} d\xi V_\parallel \frac{\partial F_0}{\partial V_\perp} \sum_{n'=-\infty}^{n'=\infty} -I_{n'}(-i\lambda)\ \mathrm{e}^{in'\xi}\ \sin\xi \\
& \{\ \mathrm{e}^{i\lambda\cos\xi}\ \frac{e}{mc}A_z + \frac{ik_z e}{m\Omega}\sum_{n=-\infty}^{n=\infty}[\frac{I_n(i\lambda)\ \mathrm{e}^{in\xi}}{i(n-\overline{\omega}+\overline{k_z V_\parallel})}(\phi - \frac{V_\parallel}{c}A_z) \\
& + \frac{n\ \mathrm{e}^{in\xi}\ I_n(i\lambda)}{-i\lambda(n-\overline{\omega}+\overline{k_z V_\parallel})}\frac{V_\perp}{c}A_x + \frac{\mathrm{e}^{in\xi}\ I_n'(i\lambda)}{i(n-\overline{\omega}+\overline{k_z V_\parallel})}\frac{V_\perp}{c}A_y]\} \\
& = e2\pi\int_0^\infty V_\perp dV_\perp \int_{-\infty}^{\infty} dV_\parallel V_\perp \frac{\partial F_0}{\partial V_\parallel}\sum_{n=-\infty}^{n=\infty}[\frac{n^2 J_n^2(\lambda)}{-(n-\overline{\omega}+\overline{k_z V_\parallel})}\frac{k_z e\Omega}{k_x^2 V_\perp mc}A_x \\
& + \frac{inJ_n'(\lambda)J_n(\lambda)}{-(n-\overline{\omega}+\overline{k_z V_\parallel})}\frac{k_z e}{k_x mc}A_y + \frac{nJ_n^2(\lambda)(n-\overline{\omega})}{(n-\overline{\omega}+\overline{k_z V_\parallel})}\frac{e\Omega}{k_x V_\perp mc}A_z + \frac{nJ_n^2(\lambda)}{(n-\overline{\omega}+\overline{k_z V_\parallel})}\frac{k_z e}{k_x V_\perp m}\phi].
\end{aligned}
\tag{56}
$$

To derive this expression, we have first used the following identity

$$
\sum_{n=-\infty}^{n=\infty} I_n(-i\lambda)\ \mathrm{e}^{in\xi}\ \sin\xi = \sum_{n=-\infty}^{n=\infty}\frac{nI_n(-i\lambda)}{\lambda}\ \mathrm{e}^{in\xi}\ , \tag{57}
$$

and then (A12) for the A_x term, (A13) for the A_y term, (A12) and (A3) for the A_z term, and (A13) for the ϕ term. This result gives the first row of $\boldsymbol{\chi}_p^{\parallel}$.

For the $\partial F_0/\partial V_\perp$ component in $\boldsymbol{j}_1(\boldsymbol{r})\cdot\boldsymbol{e}_x$, we have

$$
\begin{aligned}
& e\int_0^\infty V_\perp dV_\perp \int_{-\infty}^{\infty} dV_\parallel \int_0^{2\pi} d\xi V_\perp \frac{\partial F_0}{\partial V_\perp}\frac{B}{mV_\perp}\sum_{n'=-\infty}^{n'=\infty}\frac{-in' I_{n'}(-i\lambda)}{\lambda}\ \mathrm{e}^{in'\xi}\ \frac{e}{mc} \\
& \{\frac{e}{c}\ \mathrm{e}^{i\lambda\cos\xi}\ \frac{V_\perp}{\Omega}(-A_x\sin\xi - A_y\cos\xi) + \frac{e}{\Omega}\sum_{n=-\infty}^{n=\infty}[\frac{nI_n(i\lambda)\ \mathrm{e}^{in\xi}}{(n-\overline{\omega}+\overline{k_z V_\parallel})}(\phi - \frac{V_\parallel}{c}A_z) \\
& + \frac{n^2 I_n(i\lambda)\ \mathrm{e}^{in\xi}}{-\lambda(n-\overline{\omega}+\overline{k_z V_\parallel})}\frac{V_\perp}{c}A_x + \frac{nI_n'(i\lambda)\ \mathrm{e}^{in\xi}}{(n-\overline{\omega}+\overline{k_z V_\parallel})}\frac{V_\perp}{c}A_y]\} \\
& = \frac{e^2}{m}2\pi\int_0^\infty V_\perp dV_\perp \int_{-\infty}^{\infty} dV_\parallel \frac{\partial F_0}{\partial V_\perp}\sum_{n=-\infty}^{n=\infty}[\frac{n^2 J_n^2(\lambda)}{\lambda(n-\overline{\omega}+\overline{k_z V_\parallel})}(\phi - \frac{V_\parallel}{c}A_z) \\
& + \frac{\overline{\omega}-\overline{k_z V_\parallel}}{-(n-\overline{\omega}+\overline{k_z V_\parallel})}\frac{n^2 J_n^2(\lambda)V_\perp}{\lambda^2 c}A_x + \frac{i(\overline{\omega}-\overline{k_z V_\parallel})}{-(n-\overline{\omega}+\overline{k_z V_\parallel})}\frac{nJ_n(\lambda)J_n'(\lambda)V_\perp}{\lambda c}A_y],
\end{aligned}
\tag{58}
$$

where we have used (A12) and (A6) for the A_x term, (A13) and (A4) for the A_y term, (A12) for the ϕ and A_z terms. What we get from this equation is the first row of $\boldsymbol{\chi}_p^\perp$.

To obtain the equation for $\boldsymbol{j}_1\cdot\boldsymbol{e}_y$, we first invoke

$$
\mathrm{e}^{-i\lambda\cos\xi}\ \cos\xi = \sum_{n=-\infty}^{n=\infty} I_n'(-i\lambda)\ \mathrm{e}^{in\xi}\ . \tag{59}
$$

Then, for the $\partial F_0/\partial V_\parallel$ term,

$$
\begin{aligned}
& e\int_0^\infty V_\perp dV_\perp \int_{-\infty}^{\infty} dV_\parallel \int_0^{2\pi} d\xi V_\perp \frac{\partial F_0}{\partial V_\parallel}\sum_{n'=-\infty}^{n'=\infty} -I_{n'}(-i\lambda)\ \mathrm{e}^{in'\xi} \\
& \{\ \mathrm{e}^{i\lambda\cos\xi}\ \frac{e}{mc}A_z + \frac{ik_z e}{m\Omega}\sum_{n=-\infty}^{n=\infty}[\frac{I_n(i\lambda)\ \mathrm{e}^{in\xi}}{i(n-\overline{\omega}+\overline{k_z V_\parallel})}(\phi - \frac{V_\parallel}{c}A_z) \\
& + \frac{n\ \mathrm{e}^{in\xi}\ I_n(i\lambda)}{-i\lambda(n-\overline{\omega}+\overline{k_z V_\parallel})}\frac{V_\perp}{c}A_x + \frac{\mathrm{e}^{in\xi}\ I_n'(i\lambda)}{i(n-\overline{\omega}+\overline{k_z V_\parallel})}\frac{V_\perp}{c}A_y]\} \\
& = e2\pi\int_0^\infty V_\perp dV_\perp \int_{-\infty}^{\infty} dV_\parallel V_\perp \frac{\partial F_0}{\partial V_\parallel}\sum_{n=-\infty}^{n=\infty}[\frac{inJ_n(\lambda)J_n'(\lambda)}{(n-\overline{\omega}+\overline{k_z V_\parallel})}\frac{k_z V_\perp e}{\lambda\Omega mc}A_x \\
& - \frac{J_n'^2(\lambda)}{(n-\overline{\omega}+\overline{k_z V_\parallel})}\frac{k_z V_\perp e}{\Omega mc}A_y - \frac{iJ_n(\lambda)J_n'(\lambda)(n-\overline{\omega})}{(n-\overline{\omega}+\overline{k_z V_\parallel})}\frac{e}{mc}A_z - \frac{iJ_n(\lambda)J_n'(\lambda)}{(n-\overline{\omega}+\overline{k_z V_\parallel})}\frac{k_z e}{\Omega m}\phi],
\end{aligned}
\tag{60}
$$

where we have used (A14) for the ϕ and A_x terms, (A15) for the A_y term, (A14) and (A4) for the A_z term. This result gives the second row of $\boldsymbol{\chi}_p^{\parallel}$.

For the $\partial F_0/\partial V_\perp$ component,

$$
\begin{aligned}
& e\int_0^\infty V_\perp dV_\perp \int_{-\infty}^{\infty} dV_\parallel \int_0^{2\pi} d\xi V_\perp \frac{\partial F_0}{\partial V_\perp}\frac{B}{mV_\perp}\sum_{n'=-\infty}^{n'=\infty} -n' I_{n'}(-i\lambda)\ \mathrm{e}^{in'\xi}\ \frac{e}{mc} \\
& \{\frac{e}{c}\ \mathrm{e}^{i\lambda\cos\xi}\ \frac{V_\perp}{\Omega}(-A_x\sin\xi - A_y\cos\xi) + \frac{e}{\Omega}\sum_{n=-\infty}^{n=\infty}[\frac{nI_n(i\lambda)\ \mathrm{e}^{in\xi}}{(n-\overline{\omega}+\overline{k_zV_\parallel})}(\phi - \frac{V_\parallel}{c}A_z) \\
& + \frac{n^2 I_n(i\lambda)\ \mathrm{e}^{in\xi}}{-\lambda(n-\overline{\omega}+\overline{k_zV_\parallel})}\frac{V_\perp}{c}A_x + \frac{nI_n'(i\lambda)\ \mathrm{e}^{in\xi}}{(n-\overline{\omega}+\overline{k_zV_\parallel})}\frac{V_\perp}{c}A_y]\} \\
& = \frac{e^2}{m}2\pi\int_0^\infty V_\perp dV_\perp \int_{-\infty}^{\infty} dV_\parallel \frac{\partial F_0}{\partial V_\perp}\sum_{n=-\infty}^{n=\infty}[-\frac{inJ_n(\lambda)J_n'(\lambda)}{(n-\overline{\omega}+\overline{k_zV_\parallel})}(\phi - \frac{V_\parallel}{c}A_z) \\
& + \frac{i(\overline{\omega}-\overline{k_zV_\parallel})}{(n-\overline{\omega}+\overline{k_zV_\parallel})}\frac{nJ_n(\lambda)J_n'(\lambda)V_\perp}{\lambda c}A_x - \frac{\overline{\omega}-\overline{k_zV_\parallel}}{(n-\overline{\omega}+\overline{k_zV_\parallel})}\frac{J_n'^2(\lambda)V_\perp}{c}A_y],
\end{aligned}
\tag{61}
$$

where we have used (A14) and (A7) for the A_x term, (A15) and (A8) for the A_y term, (A14) for the ϕ and A_z terms. What we get from this equation is the second row of $\boldsymbol{\chi}_p^{\perp}$.

Assembling the above results together, we obtain the following result for the susceptibility in the gyrocenter-gauge kinetic theory.

$$
\boldsymbol{\chi}_p = \boldsymbol{\chi}_p^{\parallel} + \boldsymbol{\chi}_p^{\perp}, \tag{62}
$$

$$
\boldsymbol{\chi}_p^{\parallel} = \frac{4\pi e^2}{i\omega m\Omega} \sum_{n=-\infty}^{n=\infty} 2\pi \int_0^{\infty} V_\perp dV_\perp \int_{-\infty}^{\infty} dV_\parallel \frac{1}{(n-\overline{\omega}+\overline{k_z V_\parallel})} \frac{\partial F_0}{\partial V_\parallel} \times
$$

$$
\begin{pmatrix}
\frac{n^2 J_n^2(\lambda) k_z}{\lambda^2 c} V_\perp^2 & \frac{in J_n(\lambda) J_n'(\lambda) k_z}{\lambda c} V_\perp^2 & \frac{-n J_n^2(\lambda)(n-\overline{\omega})\Omega}{\lambda c} V_\perp & \frac{-n J_n^2(\lambda) k_z \Omega}{k_x} \\
-\frac{in J_n(\lambda) J_n'(\lambda) k_z}{\lambda c} V_\perp^2 & \frac{J_n'^2(\lambda) k_z}{c} V_\perp^2 & \frac{J_n(\lambda) J_n'(\lambda)(n-\overline{\omega})\Omega}{c} V_\perp & i J_n(\lambda) J_n'(\lambda) k_z V_\perp \\
\frac{n J_n^2(\lambda) k_z}{\lambda c} V_\perp V_\parallel & \frac{i J_n(\lambda) J_n'(\lambda) k_z}{c} V_\perp V_\parallel & \frac{-J_n^2(\lambda)(n-\overline{\omega})\Omega}{c} V_\parallel & -J_n^2(\lambda) k_z V_\parallel
\end{pmatrix},
\tag{63}
$$

$$
\boldsymbol{\chi}_p^{\perp} = \frac{4\pi e^2}{i\omega m\Omega} \sum_{n=-\infty}^{n=\infty} 2\pi \int_0^{\infty} V_\perp dV_\perp \int_{-\infty}^{\infty} dV_\parallel \frac{1}{(n-\overline{\omega}+\overline{k_z V_\parallel})} \frac{\partial F_0}{\partial V_\perp} \times
$$

$$
\begin{pmatrix}
\frac{n^2 J_n^2(\lambda)(\omega - k_z V_\parallel)}{\lambda^2 c} V_\perp & \frac{in J_n(\lambda) J_n'(\lambda)(\omega - k_z V_\parallel)}{\lambda c} V_\perp & \frac{n^2 J_n^2(\lambda)\Omega}{\lambda c} V_\parallel & \frac{-n^2 J_n^2(\lambda)\Omega}{\lambda} \\
\frac{-in J_n(\lambda) J_n'(\lambda)(\omega - k_z V_\parallel)}{\lambda c} V_\perp & \frac{J_n'^2(\lambda)(\omega - k_z V_\parallel)}{c} V_\perp & \frac{-in J_n(\lambda) J_n'(\lambda)\Omega}{c} V_\parallel & in J_n(\lambda) J_n'(\lambda)\Omega \\
\frac{n J_n^2(\lambda)(\omega - k_z V_\parallel)}{\lambda c} V_\parallel & \frac{i J_n(\lambda) J_n'(\lambda)(\omega - k_z V_\parallel)}{c} V_\parallel & \frac{n J_n^2(\lambda)\Omega V_\parallel^2}{c V_\perp} & \frac{-n J_n^2(\lambda)\Omega V_\parallel}{V_\perp}
\end{pmatrix}.
\tag{64}
$$

Finally, Eq. (62) recovers the classical result derived by integrating the Vlasov-Maxwell equations in the particle coordinates along unperturbed orbit. To see this, we take the result for $\boldsymbol{\chi}_s$ from the Eq. (10.45) of Ref. 32, and transform it into $\boldsymbol{\chi}_{s,p}$ by

$$\boldsymbol{\chi}_{s,p} = i\boldsymbol{\chi}_s \cdot \begin{pmatrix} \frac{\omega}{c} & 0 & 0 & -k_x \\ 0 & \frac{\omega}{c} & 0 & 0 \\ 0 & 0 & \frac{\omega}{c} & -k_z \end{pmatrix}. \tag{65}$$

This $\boldsymbol{\chi}_{s,p}$ is exactly the same as the result we have obtained in Eq. (62) from the gyrocenter-gauge kinetic theory.

IV. DISCUSSION

Gyrocenter-gauge kinetic theory is developed as a kinetic theory in the gyrocenter coordinates, fully equivalent to the Vlasov-Maxwell system in the particle coordinates. Taking advantage of the existence of the gyrocenter coordinates in magnetized plasmas, the gyrocenter-gauge kinetic theory simplifies the Vlasov equation by geometrically decoupling the gyrophase-independent part of the distribution function from the gyrophase-dependent part. Maxwell's equations in the particle coordinates can be easily pushed forward onto the gyrocenter coordinates by using the pull-back formula, which relates the charge and current densities to the distribution functions in the gyrocenter coordinates. As an extension of previous gyrokinetic models, the gyrocenter-gauge kinetic theory emphasizes the decoupling of the gyrophase dependent and independent informations, and the importance of the gyrocenter-gauge distribution function. Gyrocenter-gauge kinetic susceptibility is derived for homogeneous magnetized plasmas, and it recovers exactly the classical result derived by integrating the Vlasov-Maxwell equations in the particle coordinates along unperturbed orbit.

Even though only the susceptibility for homogeneous magnetized plasmas is derived here, the equation system in Sec. II is valid in general geometry. We expect that the gyrocenter-gauge kinetic equation system to bring substantial simplification compared with the usual Vlasov-Maxwell approach in treating inhomogeneous magnetized plasmas, while all the physics are kept intact. This is because the unperturbed orbit in the gyrocenter-gauge kinetic system is much simpler. It consists of two components, the gyromotion and the decoupled gyrocenter motion. The fact that gyrocenter motion is decoupled from the gyromotion enable us to eliminate the gyrophase variable ξ in the kinetic equations for f and S_n. In this sense, the gyrocenter-gauge kinetic model enjoys the same simplification and benefit as the conventional low frequency gyrokinetic models do, and further more, extends this benefit and simplification to arbitrary frequency modes.

For example, let's consider the case where $\partial/\partial y = 0$ for the perturbed field. We use this example to illustrate the basic feature of particle simulation method for the gyrocenter-gauge kinetic system. In the paper, we do not intent to give a comprehensive account on the gyrocenter-gauge particle simulation method, which will be the subject of future publications. For the current case, the kinetic equation for S is

$$\begin{aligned}&\frac{\partial S}{\partial t} + \dot{\boldsymbol{X}}\frac{\partial S}{\partial \boldsymbol{X}} + \dot{V}_\parallel \frac{\partial S}{\partial V_\parallel} + \Omega \frac{\partial S}{\partial \xi} = \\ &e[(\mathrm{e}^{i\lambda\cos\xi} - J_0)(\phi - \frac{V_\parallel}{c}A_z) + \mathrm{e}^{i\lambda\cos\xi}\sin\xi\frac{V_\perp}{c}A_x + (\mathrm{e}^{i\lambda\cos\xi}\cos\xi + J_1)\frac{V_\perp}{c}A_y],\end{aligned} \tag{66}$$

where k_x is understood to be $-i\partial/\partial x$. Since in (and only in) the gyrocenter coordinates $\dot{\boldsymbol{X}}$ and $\dot{V}_\parallel$ are gyrophase independent, different gyrophase harmonics for S are decoupled. Let

$$S = \sum_{n=-\infty}^{n=\infty} S_n \,\mathrm{e}^{in\xi} \,. \tag{67}$$

Using

$$
\begin{aligned}
\mathrm{e}^{i\lambda\cos\xi} &= \sum_{n=-\infty}^{n=\infty} I_n(i\lambda)\ \mathrm{e}^{in\xi}\ , \\
\mathrm{e}^{i\lambda\cos\xi}\ \sin\xi &= \sum_{n=-\infty}^{n=\infty} \frac{-nI_n(i\lambda)}{\lambda}\ \mathrm{e}^{in\xi}\ , \\
\mathrm{e}^{i\lambda\cos\xi}\ \cos\xi &= \sum_{n=-\infty}^{n=\infty} I_n'(i\lambda)\ \mathrm{e}^{in\xi}\ ,
\end{aligned}
\tag{68}
$$

we easily have the decoupled equations for S_n

$$
\begin{aligned}
& S_0 = 0, \\
& \frac{dS_n}{dt} + in\Omega S_n = e[I_n(i\lambda)(\phi - \frac{V_\parallel}{c}A_z) - \frac{nI_n(i\lambda)}{\lambda}\frac{V_\perp}{c}A_x + I_n'(i\lambda)\frac{V_\perp}{c}A_y], \quad n \neq 0, \\
& \frac{d}{dt} \equiv \frac{\partial}{\partial t} + \dot{\boldsymbol{X}}\frac{\partial}{\partial \boldsymbol{X}} + \dot{V}_\parallel \frac{\partial}{\partial V_\parallel}.
\end{aligned}
\tag{69}
$$

The above kinetic equations for S_n do not involve the gyrophase variable ξ, and the characteristics of the equations are particles' gyrocenter orbits. However, to solve these kinetic equations using particle simulation method, the time step ΔT for advancing S_n has to satisfy $\Delta T < 1/n\Omega$, even though the gyrocenter orbit motions are slower and satisfy

$$
\dot{\boldsymbol{X}}\frac{\partial}{\partial \boldsymbol{X}} + \dot{V}_\parallel \frac{\partial}{\partial V_\parallel} \ll n\Omega. \tag{70}
$$

This is because term $in\Omega S_n$ and the terms depending on ϕ and A_z are fast varying. Then, in terms of particle simulation for arbitrary frequency modes, what is the simplification brought by the gyrocenter-gauge kinetic system compared with the Vlasov-Maxwell system in the particle coordinates? To solve the kinetic equations for f and S_n, we truncate the equation system for S_n and keep those important harmonics for the problem under investigation. Along its gyrocenter orbit, each particle carries those S_n kept in the system, as well as the usual distribution f. For high frequency mode ($\omega \sim n\Omega$, for some integer n), we have to use small time step ($\Delta T < 1/n\Omega$) to advance f and S_n along particles' gyrocenter orbits. Since the gyrocenter motions themselves are slower motions with larger scale length, it is not necessary to use small time step to advance particles' gyrocenters in the gyrocenter phase space. Particularly, we can us an *adiabatic gyrocenter pusher*, which advances particles' phase-space coordinates in larger gyrocenter time step, and between gyrocenter time steps, f and S_n are advanced many time steps in smaller gyrofrequency time step while particles' phase space coordinates are kept constant. The slower gyrocenter time step is determined by the gyrocenter orbit motion, whereas the faster gyrofrequency time step is determined by the harmonics number n. In principle, we can use different gyrofrequency time steps for different harmonics S_n. In each gyrofrequency time step, Maxwell's equations in the gyrocenter coordinates has to be solved to update the field. f and S_n enter Maxwell's equations through the pull-back formula, which can be numerically implemented by the well-known multi-point averaging technique.[25] The computational simplification brought by the gyrocenter-gauge kinetic system is twofold. First, the gyrophase coordinate ξ is explicitly removed from the dynamic equation for particles. The gyrophase-dependent information is efficiently described by the harmonics S_n kept in the system, without increasing the number of simulation particles. If using the straightforward particle simulation for Vlasov-Maxwell system in the particle coordinates, we have to increase the number of simulation particles many times to achieve desired resolution in the gyrophase coordinate ξ. Obviously, the gyrocenter-gauge kinetic particle simulation requires less memory usage and computing time. Secondly, the gyrocenter-gauge kinetic particle simulation only advances particles' phase space coordinates along their gyrocenter motions, which are much slower motions with larger scale length compared with particles' motions in the particle coordinates, which each simulation particle has to follow if the simulation is carried out for the Vlasov-Maxwell system in the particle coordinates.

Therefore, gyrocenter-gauge kinetic particle simulation requires much less computing time to advance simulation particles.

The formalism presented in this manuscript can be easily extended to nonlinear case by carrying out the transformation between the (equilibrium) guiding center coordinates and the (perturbed) gyrocenter coordinates to the 2nd or higher order. The basic procedure is similar to those in Ref 9–12. In fact, the non-canonical Lie perturbation methods used here was originally introduced as an efficient and systematic approach for the nonlinear gyrokinetic systems. In the nonlinear case, the kinetic equations and the push-forward of Maxwell's equations keep the same forms, except that in the pull-back of distribution function, nonlinear perturbed fields appear. This is a direct result of the construction of the gyrocenter coordinates up to the 2nd or higher order.

So far, we have not considered collisions in our system. The gyrocenter-gauge kinetic system in the gyrocenter coordinates developed here is thus parallel to the collisionless Vlasov-Maxwell system in the particle coordinates. For many problems of wave-particle interactions and instabilities, collisions are not important, especially for the high frequency range. However, for applications such as neoclassical transport, it is necessary to include collisions in the gyrocenter-gauge kinetic system. The exact expressions of collision operators in the gyrocenter coordinates should be rigorously derived by *pushing forward* the corresponding collision operators in the particle coordinates. Compared with the collision operators in the particle coordinates, one distinguish feature of the collision operators in the gyrocenter coordinates is their explicit dependences on the perturbed fields and background inhomogeneities through the pull-back transformation. Since the collision operators normally involve high order differential in the phase space, the construction of the gyrocenter-gauge collision operators will be in the high order jet space. In terms of particle simulation, once the expression of the collision operators are obtained, they can be simulated by the usual Monte Carlo method.[33–35] Work in this direction will be reported in the future publications.

ACKNOWLEDGMENTS

This work is supported by U.S. DOE No. DE-AC02-76-CH0-3073. We thank Drs. G. Rewoldt, Z. H. Lin, T. S. Hahm, J. A. Krommes, and L. Chen for constructive comments and suggestions.

APPENDIX A: IDENTITIES FOR J_N AND I_N

The following identities for J_n and I_n are used.

For J_n,

$$J_{-n}(x) = J_n(-x) = (-1)^n J_n(x), \tag{A1}$$

$$J_n'(-x) = (-1)^{n+1} J_n'(x), \tag{A2}$$

$$\sum_{n=-\infty}^{n=\infty} n J_n^2 = 0, \tag{A3}$$

$$\sum_{n=-\infty}^{n=\infty} J_n J_n' = 0, \tag{A4}$$

$$\sum_{n=-\infty}^{n=\infty} J_n^2 = 1, \tag{A5}$$

$$\sum_{n=-\infty}^{n=\infty} n^2 J_n^2(x) = \frac{x^2}{2}, \tag{A6}$$

$$\sum_{n=-\infty}^{n=\infty} n J_n J_n' = 0, \tag{A7}$$

$$\sum_{n=-\infty}^{n=\infty} J_n'^2 = \frac{1}{2}, \tag{A8}$$

For I_n,

$$e^{\lambda \cos \xi} = \sum_{n=-\infty}^{n=\infty} I_n(\lambda)\, e^{in\xi} , \tag{A9}$$

$$I_n(x) = i^{-n} J_n(ix), \tag{A10}$$

$$I_n'(x) = i^{-n+1} J_n'(ix), \tag{A11}$$

$$I_n(ix) I_{-n}(-ix) = J_n^2(x), \tag{A12}$$

$$I_n'(ix) I_{-n}(-ix) = -i J_n(x) J_n'(x), \tag{A13}$$

$$I_{-n}'(-ix) I_n(ix) = i J_n(x) J_n'(x), \tag{A14}$$

$$I_n'(-ix) I_n'(ix) = J_n'^2(x), \tag{A15}$$

$$I_{n-1}(x) - I_{n+1}(x) = \frac{2n}{x} I_n(x), \tag{A16}$$

$$I_{n-1}(x) + I_{n+1}(x) = 2I_n'(x), \tag{A17}$$

$$I_n'(x) = I_{n-1}(x) - \frac{n}{x} I_n(x) = I_{n+1}(x) + \frac{n}{2} I_n(x). \tag{A18}$$

REFERENCES

[1] R. G. Littlejohn, J. Math. Phys. **20**, 2445 (1979).

[2] R. G. Littlejohn, Phys. Fluids **24**, 1730 (1981).

[3] R. G. Littlejohn, J. Plasma Phys. **29**, 111 (1983).

[4] W. W. Lee, Phys. Fluids **26**, 556 (1983).

[5] D. H. E. Dubin, J. A. Krommes, C. Oberman, and W. W. Lee, Phys. Fluids **26**, 3524 (1983).

[6] L. Chen and S. T. Tsai, Phys. Fluids **26**, 141 (1983).

[7] L. Chen and S. T. Tsai, Plasma Phys. **25**, 349 (1983).

[8] S. C. Yang and D. I. Choi, Phys. Lett. **108 A**, 25 (1985).

[9] T. S. Hahm, Phys. Fluids **31**, 2670 (1988).

[10] T. S. Hahm, W. W. Lee, and A. Brizard, Phys. Fluids **31**, 1940 (1988).

[11] A. J. Brizard, J. Plasma Phys. **41**, 541 (1989).

[12] A. J. Brizard, Phys. Fluids B **1**, 1381 (1989).

[13] H. Qin, W. M. Tang, and G. Rewoldt, Phys. Plasmas **5**, 1035 (1998).

[14] H. Qin, Ph. D. Dissertation (Princeton University, 1998).

[15] H. Qin, W. M. Tang, W. W. Lee, and G. Rewoldt, Phys. Plasmas **6**, 1575 (1999).

[16] R. G. Littlejohn, J. Math. Phys. **23**, 742(1982).

[17] A. J. Brizard, Ph. D. Dissertation (Princeton University, 1990), pp. 31.

[18] T. S. Hahm, private communication.

[19] H. Goldstein, *Classical Mechanics* (2nd ed., Addision-Wesley, Reading MA, 1980).

[20] V. I. Arnold, *Mathematical Methods of Classical Mechanics* (2nd ed., Springer-Verlag, New York, 1989), pp. 258.

[21] A. Brizard, Phys. Plasmas **1**, 2473 (1994).

[22] P. J. Morrison and D. Pfirsh, Phys. Fluid B **2**, 1105 (1990).

[23] H. Qin, W. M. Tang, and G. Rewoldt, Phys. Plasmas **6**, 2544 (1999).

[24] H. Sugama, Phys. Plasmas **7**, 466 (2000).

[25] W. W. Lee, J. Comput. Phys. **72**, 243 (1987).

[26] W. W. Lee and W. M. Tang, Phys. Fluids **31**, 612 (1988).

[27] S. E. Parker, W. W. Lee, and R. A. Santoro, Phys. Rev. Lett. **71**, 2042 (1993).

[28] J. C. Cummings, Ph. D. Dissertation (Princeton University, 1995).

[29] A. M. Dimits, T. J. Williams, J. A. Byers, and B. I. Cohen, Phys. Rev. Lett. **77**, 71 (1996).

[30] R. D. Sydora, V. K. Decyk, J. M. Dawson, Plasma Phys. Control. Fusion **38**, A281 (1996).

[31] Z. Lin, T. S. Hahm, W. W. Lee, W. M. Tang, and R. B. White, Science **281**, 1835 (1998).

[32] T. H. Stix, *Waves in Plasmas*, (AIP, New York, 1992), pp. 247 - 262.

[33] X. Xu and M. N. Rosenbluth, Phys. Fluid B **3**, 627 (1991)

[34] Z. H. Lin, W. M. Tang, and W. W. Lee, Phys. Rev. Lett. **78**, 456 (1997).

[35] Y. Chen and R. B. White, Phys. Plasmas **4**, 3591 (1997).

The Princeton Plasma Physics Laboratory is operated
by Princeton University under contract
with the U.S. Department of Energy.

Information Services
Princeton Plasma Physics Laboratory
P.O. Box 451
Princeton, NJ 08543

Phone: 609-243-2750
Fax: 609-243-2751
e-mail: pppl_info@pppl.gov
Internet Address: http://www.pppl.gov