Analysis of the polarization effects in the gyrokinetic theory of magnetized plasmas

Debing Zhang, Limin Yu, Erbing Xue, and Xianmei Zhang[*]
Department of Physics, East China University of Science and Technology, Shanghai, 200237, China

(Received 14 October 2020; accepted 9 January 2021; published 25 January 2021)

By adopting the hybrid coordinates, in which the nonlinearity of polarization displacement is included in the configuration space variables compared to the conventional gyrocenter coordinates, the polarization effects are analyzed by using the modern gyrokinetic (GK) theory of magnetized plasmas. Based on the invariant property, the velocity transformation between the gyrocenter and hybrid coordinates is calculated, and the phase-space velocity in terms of the hybrid coordinates is obtained. The linear and nonlinear polarization distribution functions are defined, and the evolutions for the polarization distribution functions are derived. It is well known that the polarization density is important in the GK calculation of particle density. Analogously, it is shown that the polarization current should be considered in the GK calculation of current density. In the case with electrostatic fluctuations, the roles of the polarization current are illustrated in the derivations of the Hasegawa-Mima equation and the dispersion relation for geodesic acoustic mode. In the case with magnetic fluctuations, the procedure for the GK calculation of perpendicular current is clarified, the dispersion relation for compressional Alfvén wave is derived, in which the effect of polarization current is discussed.

DOI: 10.1103/PhysRevE.103.013212

I. INTRODUCTION

Anomalous transport is a key issue in magnetized fusion plasmas. Kinetic description of the self-consistent anomalous transport includes the evolution of particle phase-space distribution function governed by the particle Vlasov equation and the evolution of electromagnetic fluctuations governed by the Maxwell's equations. Since the motion of charged particles appeared in the Vlasov equation depends on the electromagnetic fields, the charge density and current density appeared in the Maxwell's equations depend on the particle distribution, the Vlasov-Maxwell system is nonlinear [1,2].

Gyrokinetic (GK) theory is a powerful tool to study the anomalous transport both theoretically [3–9] and numerically [10–17]. Instead of the guiding center (GC) coordinates, the aim of the GK theory is to seek a new coordinate system, called as gyrocenter (GY) coordinates, in which the GY magnetic moment is constant, to decouple the fast gyromotion from the slow GY motion. The modern GK theory is developed by the Lie-transform perturbation method [18–20], where the phase-space transformation from the GC coordinates to the GY coordinates is a Lie transform. In terms of GY coordinates, the equations of motion are greatly simplified. However, the GY orbit does not follow the real particle orbit; the excursion is called as "polarization effects" [21].

The polarization effects play important roles in the GK theory. Lee first proposed the definition for polarization density, which denotes the excursion between real particle density and GY density [21]; the polarization density has been successfully introduced in the GK Poisson equation in amounts of GK simulation codes [10–17]. In Ref. [22], a new definition of GY is introduced to explicitly include the polarization drift in the GY equations of motion, in this way, the proximity of the GY to the actual particle position can be ensured for a longer time; alternatively, the polarization effect is included in the modification of the new GY phase-space volume, which leads that the GK Poisson equation and the energy invariant are not changed. Qin identified and developed the GK perpendicular dynamics; the perpendicular current, including the linear polarization current, was calculated in the GC picture; it is found that the physics of compressional Alfvén wave (CAW) can be recovered from the viewpoint of polarization current [23]. Later, Lee argued that the unique treatment of polarization effects in the GK theory is the key to add and suppress shear and compressional Alfvén waves without resorting to additional geometrical simplifications [24]. By considering the nonlinear terms in the polarization drift, the turbulent poloidal Reynolds stress can be formally added into the GK equation, which can be used to study the effects of poloidal Reynolds stress on zonal flows [25]. Brizard revealed that the exact energy conservation for the nonlinear GK Vlasov-Maxwell system relies on the connection between GC ponderomotive Hamiltonian and GY polarization and magnetization effects [26].

In our previous work [25], we introduced the hybrid (HY) coordinates, where the velocity space variables are the same as those in GY coordinates, and the configuration space variables explicitly include the polarization excursions, as is considered in Ref. [22]. In this work, we analyze the polarization effects in the GK theory with the help of HY coordinates. The transformation of distribution function and particle velocity between GY and HY coordinates is illustrated. The linear and nonlinear polarization distribution function and polarization drift are distinguished, and the evolutions for the polarization distribution functions are illustrated. The GK derivation of the

[*]zhangxm@ecust.edu.cn

dispersion relation for CAW in the slab geometry is shown, in which the polarization current is clarified.

The remaining part of this paper is organized as follows. In Sec. II, a brief review of the fundamentals in the GK theory is presented, and the HY coordinates are introduced; in Sec. III, the phase-space velocity in terms of the HY coordinates is illustrated, and the evolutions for the polarization distribution functions are discussed; in Sec. IV, the dispersion relation for CAW in the slab geometry is derived, in which the polarization current is discussed; Sec. V is the conclusions and discussions.

II. FUNDAMENTALS

In this section, we first present a brief review of the GK theory (more details can be found in Refs. [8,27]); we then present a review for the HY coordinates (more details can be found in Ref. [25]).

A. Review of the modern GK theory

In terms of the particle coordinates $\boldsymbol{z} = (\boldsymbol{x}, \boldsymbol{v})$, with $\boldsymbol{x}$ and $\boldsymbol{v}$ the particle position and velocity, respectively, the particle Vlasov equation is written as

$$\partial_t f + \boldsymbol{v}\cdot\partial_{\boldsymbol{x}} f + \frac{e}{m}(\boldsymbol{E} + \boldsymbol{v}\times\boldsymbol{B})\cdot\partial_{\boldsymbol{v}} f = 0. \tag{1}$$

Here, $f(\boldsymbol{z}, t)$ is the particle distribution function, e and m are the particle charge and mass, respectively; $\boldsymbol{E}$ and $\boldsymbol{B}$ are the electric and magnetic field, respectively. Generally, it is convenient to carry out the analysis in the GC coordinates $\boldsymbol{Z} = (\boldsymbol{X}, v_\parallel, \mu, \xi)$ with the coordinate transformation

$$\boldsymbol{x} = \boldsymbol{X} + \boldsymbol{\rho}_0(\mu, \xi), \tag{2a}$$

$$\boldsymbol{v} = \boldsymbol{v}(v_\parallel, \mu, \xi). \tag{2b}$$

Here, $\boldsymbol{X}$ is the GC position, $\boldsymbol{\rho}_0$ is the gyroradius, $v_\parallel$ is the GC parallel velocity, μ is the GC magnetic moment, and ξ is the gyroangle. In Eq. (2a), the spatial dependence in $\boldsymbol{\rho}_0$ is ignored. In the case without fluctuations, $\dot{\mu} = 0$, the fast gyromotion is decoupled from the slow GC motion.

We introduce the GC Hamiltonian theory in the unperturbed (UP) electromagnetic field. The UP fundamental one-form (the GC extended Lagrangian) written in terms of the extended GC coordinates is $\hat{\Gamma}_0 = \Gamma_0 - H_0 d\tau$, with the extended UP symplectic part and Hamiltonian

$$\Gamma_0 = (mv_\parallel \boldsymbol{b}_0 + e\boldsymbol{A}_0)\cdot d\boldsymbol{X} + \frac{m}{e}\mu\, d\xi - U\, dt, \tag{3a}$$

$$H_0 = \frac{1}{2}mv_\parallel^2 + \mu B_0 - U. \tag{3b}$$

Here, $(-U, t)$ are the conjugate energy-time coordinates, τ is an independent parameter. $\boldsymbol{b}_0 = \boldsymbol{B}_0/B_0$ with $\boldsymbol{B}_0 = \nabla\times\boldsymbol{A}_0$ the UP magnetic field. The UP Lagrange two-form is $\hat{\omega}_0 = \omega_0 - dH_0\wedge d\tau$,

$$\omega_0 = \frac{1}{2}\epsilon_{ijk} e B_0^{*,k} dX^i\wedge dX^j + mb_{0i} dv_\parallel\wedge dX^i + \frac{m}{e} d\mu\wedge d\xi + d(-U)\wedge dt, \tag{4}$$

with $\boldsymbol{B}_0^* = \boldsymbol{B}_0 + \frac{m}{e}v_\parallel\nabla\times\boldsymbol{b}_0$, $\wedge$ the exterior product, ϵ_{ijk} the permutation tensor, b_{0i} the component of $\boldsymbol{b}_0$. The components of the inverse of Lagrange tensor $\mathcal{J}_0 = \omega_0^{-1}$, known as the UP Poisson tensor, are the fundamental Poisson brackets $\mathcal{J}_0^{ij} = \{Z^i, Z^j\}$. The UP GC equation of motion is $\dot{Z}_0^i = \{Z^i, H_0\}$, with the Poisson bracket given by

$$\{f, g\} = \frac{\boldsymbol{B}^*}{mB_\parallel^*}\cdot\left(\nabla f\frac{\partial g}{\partial v_\parallel} - \nabla g\frac{\partial f}{\partial v_\parallel}\right) - \frac{\boldsymbol{b}_0}{eB_\parallel^*}\cdot(\nabla f\times\nabla g) + \frac{e}{m}\left(\frac{\partial f}{\partial\xi}\frac{\partial g}{\partial\mu} - \frac{\partial g}{\partial\xi}\frac{\partial f}{\partial\mu}\right) + \frac{\partial f}{\partial t}\frac{\partial g}{\partial U} - \frac{\partial g}{\partial t}\frac{\partial f}{\partial U}. \tag{5}$$

Here, $B_\parallel^* = \boldsymbol{B}_0^*\cdot\boldsymbol{b}_0$ denotes the Jacobian of the GC coordinates.

We then discuss the the Hamiltonian theory with the perturbed electric potential $\delta\phi$ and vector potential $\delta\boldsymbol{A}$, which are the functions of real space $\boldsymbol{r}$ and time t. The fundamental one-form is separated into the UP part $\hat{\Gamma}_0$ and the perturbed part $\hat{\Gamma}_1$,

$$\hat{\Gamma}_1 \equiv \Gamma_{1i} dZ^i = \delta\boldsymbol{A}\cdot d(\boldsymbol{X} + \boldsymbol{\rho}_0) - \delta\phi\, dt, \tag{6}$$

and the first-order Lagrange two-form is

$$\omega_1 = (\partial_i\Gamma_{1j} - \partial_j\Gamma_{1i}) dZ^i\wedge dZ^j \tag{7a}$$

$$= e\delta E^i\big(dX^i\wedge dt + \partial_\mu\rho_0^i d\mu\wedge dt + \partial_\xi\rho_0^i d\xi\wedge dt\big) + \epsilon_{ijk} e\delta B^k\left(\partial_\mu\rho_0^i d\mu\wedge dX^j + \partial_\xi\rho_0^i d\xi\wedge dX^j + \frac{1}{2}dX^i\wedge dX^j\right) - \frac{e\rho_0^2}{2\mu}\boldsymbol{b}_0\cdot\delta\boldsymbol{B}\, d\mu\wedge d\xi, \tag{7b}$$

where $\delta\boldsymbol{E} = -\nabla\delta\phi - \partial_t\delta\boldsymbol{A}$, $\delta\boldsymbol{B} = \nabla\times\delta\boldsymbol{A}$.

The Lie transform [18–20] is a method to simplify the equations of motion by transforming the GC coordinates $\boldsymbol{Z}$ to the GY coordinates $\bar{Z} = (\bar{X}, \bar{v}_\parallel, \bar{\mu}, \bar{\xi})$, in which $\bar{\mu}$ is a constant. We introduce the ordering parameter $\epsilon_\delta \sim \frac{e\delta\phi}{T} \sim \frac{e\boldsymbol{v}\cdot\delta\boldsymbol{A}}{T}$, where T is the UP temperature. To $O(\epsilon_\delta^2)$, the transformations between GC and GY coordinates are

$$\bar{Z} = \boldsymbol{Z} + \boldsymbol{G}_1 + \boldsymbol{G}_2 + \tfrac{1}{2}\boldsymbol{G}_1\cdot\nabla_6\boldsymbol{G}_1(\boldsymbol{Z}), \tag{8a}$$

$$\boldsymbol{Z} = \bar{Z} - \boldsymbol{G}_1 - \boldsymbol{G}_2 + \tfrac{1}{2}\boldsymbol{G}_1\cdot\nabla_6\boldsymbol{G}_1(\bar{Z}). \tag{8b}$$

Here, $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$ are the first- and second-order generating vector fields, respectively; note that $G_n^t = 0$, which indicates that the time variable is not changed in the Lie transform. The symbol "$\boldsymbol{\nabla_6}$" denotes the partial derivative with respect to the six phase-space variables. From Eqs. (2a) and (12), it is easy to derive the transformation between the configuration space coordinates $\boldsymbol{x}$ and the GY coordinates

$$\boldsymbol{x} = \bar{X} + \bar{\rho} - G_1^{\boldsymbol{r}} - G_2^{\boldsymbol{r}} + \tfrac{1}{2}\boldsymbol{G}_1\cdot\nabla_6 G_1^{\boldsymbol{r}}(\bar{Z}). \tag{8c}$$

Here, $\bar{\rho} \equiv \boldsymbol{\rho}_0(\bar{\mu}, \bar{\xi})$, $G_n^{\boldsymbol{r}} = G_n^X + G_n^\rho$ with $G_n^\rho \equiv G_n^\mu\partial_{\bar{\mu}}\bar{\rho} + G_n^\xi\partial_{\bar{\xi}}\bar{\rho}$. From Eq. (8c), it is obvious that except for the gyroradius $\bar{\rho}$, the excursions between particle position $\boldsymbol{x}$ and GY position $\bar{X}$ relate to the polarization vectors $G_n^{\boldsymbol{r}}$, namely, $G_n^{\boldsymbol{r}}$ represent the polarization effects.

The push-forward transformation of the fundamental one-form reads as [8,19] $\hat{\bar{\Gamma}} \equiv \bar{\Gamma} - \bar{H}\, d\tau = \bar{\Gamma}_i d\bar{Z}^i - \bar{H}\, d\tau$, with $\bar{\Gamma} = \bar{\Gamma}_0 + \bar{\Gamma}_1 + \bar{\Gamma}_2$, $\bar{H} = \bar{H}_0 + \bar{H}_1 + \bar{H}_2$. For the usual GK symplectic transformation, the transformed two-form is formally the same as the UP one, that is, $\bar{\Gamma}_0 = \Gamma_0$, $\bar{\Gamma}_1 = \bar{\Gamma}_2 = 0$;

on the other hand, $\bar{H}_n$'s are set to be independent of the gyroangle $\bar{\xi}$ to decouple the fast gyromotion from the GY motion, then the generating vector fields are explicitly written in a generic form [27]

$$G_n^i = G_{n,A}^i + G_{n,Y}^i, \tag{9}$$

with

$$G_{n,A}^{\boldsymbol{X}} = -\frac{\boldsymbol{b}_0}{B_\parallel^*} \times \delta\boldsymbol{\mathcal{A}}_n, \tag{10a}$$

$$G_{n,A}^{v_\parallel} = \frac{e\boldsymbol{B}_0^*}{mB_\parallel^*} \cdot \delta\boldsymbol{\mathcal{A}}_n, \tag{10b}$$

$$G_{n,A}^{\mu} = \frac{e^2}{m}\delta\boldsymbol{\mathcal{A}}_n \cdot \partial_{\bar{\xi}}\bar{\rho}, \tag{10c}$$

$$G_{n,A}^{\xi} = -\frac{e^2}{m}\delta\boldsymbol{\mathcal{A}}_n \cdot \partial_{\bar{\mu}}\bar{\rho}, \tag{10d}$$

$$G_{n,A}^{U} = 0; \tag{10e}$$

$$G_{n,Y}^{\boldsymbol{X}} = -\frac{\boldsymbol{b}_0}{eB_\parallel^*} \times \nabla S_n - \frac{\boldsymbol{B}_0^*}{mB_\parallel^*}\frac{\partial S_n}{\partial \bar{v}_\parallel}, \tag{11a}$$

$$G_{n,Y}^{v_\parallel} = \frac{\boldsymbol{B}_0^*}{mB_\parallel^*} \cdot \nabla S_n, \tag{11b}$$

$$G_{n,Y}^{\mu} = \frac{e}{m}\partial_{\bar{\xi}} S_n, \tag{11c}$$

$$G_{n,Y}^{\xi} = -\frac{e}{m}\partial_{\bar{\mu}} S_n, \tag{11d}$$

$$G_{n,Y}^{U} = \delta\psi_n - \partial_t S_n. \tag{11e}$$

Here, $\delta\psi_1 = \delta\phi$, $\delta\psi_2 = \frac{1}{2}G_1^r \cdot \delta\boldsymbol{E}$; $\delta\boldsymbol{\mathcal{A}}_1 = \delta\boldsymbol{A}$, $\delta\boldsymbol{\mathcal{A}}_2 = \frac{1}{2}G_1^r \times \delta\boldsymbol{B}$. S_n's are the nth-order gauge function used for removing the gyroangle dependence of the extended Lagrangian, which satisfies

$$\left(\frac{d}{dt}\right)_0 S_n = e\widetilde{\delta\Psi_n}. \tag{12}$$

Here, $(d/dt)_0$ means differential along the UP orbit, $\widetilde{\delta\Psi_n} = \delta\Psi_n - \langle\delta\Psi_n\rangle$ with $\langle\cdot\rangle$ the gyroaverage, and the effective potential $\delta\Psi_n$ is defined as

$$\delta\Psi_n = \delta\psi_n - (\dot{\bar{X}}_0 + \dot{\bar{\rho}}_0) \cdot \delta\boldsymbol{\mathcal{A}}_n, \tag{13}$$

with $\dot{\bar{\rho}}_0 = \Omega\partial_{\bar{\xi}}\bar{\rho}$, $\Omega = eB_0/m$. Note that in Eq. (12), S_n and $\delta\Psi_n$ are evaluated in the GY coordinates. The Hamiltonian is [8,28]

$$\bar{H}_0(\bar{Z}) = \tfrac{1}{2}m\bar{v}_\parallel^2 + \bar{\mu}B_0 - \bar{U}, \tag{14a}$$

$$\delta\bar{H}_1(\bar{Z}) = e\langle\delta\Psi_1\rangle, \tag{14b}$$

$$\delta\bar{H}_2(\bar{Z}) = -\tfrac{1}{2}\big\langle G_1^r \cdot \nabla\delta\bar{H}_1 + e\delta\boldsymbol{A} \cdot \{\boldsymbol{r}, \delta\bar{H}_1\}\big\rangle. \tag{14c}$$

In the GY coordinates, the GK Vlasov equation is written as

$$\partial_t\bar{F} + \nabla_6 \cdot (\dot{\bar{Z}}\bar{F}) = 0. \tag{15}$$

Here, the divergence is read as $\nabla_{\boldsymbol{6}} \cdot \boldsymbol{a} = \frac{1}{B_\parallel^*}\partial_i(B_\parallel^* a^i)$, $\bar{F}(\bar{Z}, t)$ is the GY distribution function, $\dot{\bar{Z}}$ are the phase-space velocity calculated as

$$\dot{\bar{Z}} = \{\bar{Z}, \bar{H}_0 + \delta\bar{H}_1 + \delta\bar{H}_2\}. \tag{16}$$

B. Review of the HY coordinates

We begin with the velocity moment integral in the particle coordinates, which is written as

$$\mathfrak{n}[\lambda](\boldsymbol{r}) = \int d^6z\, \delta(z; \boldsymbol{r}) f(z)\lambda(z). \tag{17}$$

Here, the four factors in the integral represent the volume element, the "location" function, the distribution function, and the moment function, respectively. As is known, $d^6z = d^3\boldsymbol{x}\, d^3\boldsymbol{v}$, $\delta(z; \boldsymbol{r}) = \delta_D(\boldsymbol{x} - \boldsymbol{r})$ with δ_D the standard Dirac function, which represents that the velocity moment should be evaluated at the field position $\boldsymbol{r}$; note that Eq. (17) is the standard definition for the velocity moment. The particle and current density are obtained by taking $\lambda = 1$ and $\dot{\boldsymbol{x}}$, respectively.

Analogous to Eq. (17), the velocity moment integral in the GY coordinates can be formally written as

$$\mathfrak{n}[\lambda](\boldsymbol{r}) = \int J_{\bar{Z}}(\bar{Z})d^6\bar{Z}\, \bar{\Delta}(\bar{Z}; \boldsymbol{r})\bar{F}(\bar{Z})\bar{\Lambda}(\bar{Z}). \tag{18}$$

The transformation for the four factors in Eqs. (17) and (18) can be obtained based on the scalar invariance property. In detail, the volume element in the GY coordinate is written as [4]

$$J_{\bar{Z}}d^6\bar{Z} = B_\parallel^* d^3\bar{X}\, d\bar{v}_\parallel d\bar{\mu}\, d\bar{\xi},$$

and the "location" function is [8,22,25]

$$\begin{aligned}\bar{\Delta}(\bar{Z}; \boldsymbol{r}) =& \delta_D(z; \boldsymbol{r}) \\ =& \delta_D - G_1^i\partial_i\delta_D - G_2^i\partial_i\delta_D + \tfrac{1}{2}G_1^j\partial_j\big(G_1^i\partial_i\delta_D\big),\end{aligned}$$

where the argument of the Dirac function is $\bar{X} + \bar{\rho} - \boldsymbol{r}$. After some straightforward integration by parts, Eq. (18) is formally rewritten as [25]

$$\begin{aligned}\mathfrak{n}[\lambda](\boldsymbol{r}) = \int d^6\bar{Z}\, \delta_D(\bar{X} + \bar{\rho} - \boldsymbol{r}) \times \Big\{ & B_\parallel^*\bar{F}\,\bar{\Lambda} \\ & + \nabla \cdot \Big[G_1^r B_\parallel^*\bar{F}\,\bar{\Lambda} + \frac{1}{2}\nabla \cdot \big(G_1^r G_1^r B_\parallel^*\bar{F}\,\bar{\Lambda}\big) \\ & + \Big(G_2^r - \frac{1}{2}G_1^j\partial_j G_1^r\Big)B_\parallel^*\bar{F}\,\bar{\Lambda}\Big]\Big\}.\end{aligned} \tag{19}$$

It should be noted that "∇" denotes the divergence in the configuration space, not in the phase space.

We further introduce the HY coordinates $\boldsymbol{Z}_h = (\boldsymbol{X}_h, \bar{v}_{\parallel,h}, \bar{\mu}_h, \bar{\xi}_h)$, the transformation between the GY and HY coordinates are

$$\boldsymbol{Z}_h = \bar{Z} - \boldsymbol{G}_{1h} - \boldsymbol{G}_{2h} + \tfrac{1}{2}\boldsymbol{G}_1 \cdot \nabla_6\boldsymbol{G}_{1h}(\bar{Z}), \tag{20a}$$

$$\bar{Z} = \boldsymbol{Z}_h + \boldsymbol{G}_{1h} + \boldsymbol{G}_{2h} - \tfrac{1}{2}\boldsymbol{G}_1 \cdot \nabla_6\boldsymbol{G}_{1h} + \boldsymbol{G}_{1h} \cdot \nabla_6\boldsymbol{G}_{1h}(\boldsymbol{Z}_h), \tag{20b}$$

where $\boldsymbol{G}_{nh} = (\boldsymbol{G}_n^r, \boldsymbol{0})$, that is, the velocity space variables in $\boldsymbol{Z}_h$ coordinates are the same as those in $\bar{Z}$ coordinates, from which one obtains that the gyroradius in $\boldsymbol{Z}_h$ coordinates is the same as that in $\bar{Z}$ coordinates. According to Eqs. (8c) and

(20a), it is obvious that $\boldsymbol{x} = \boldsymbol{X}_h + \bar{\rho}$. Since G_n^r are absorbed into $\boldsymbol{X}_h$, $\boldsymbol{X}_h$ can be regarded as the true GC that follows the particle orbit. We mention that the concept is the same as that in Ref. [22]. However, compared to the coordinates used in Ref. [22], the nonlinearity of polarization displacement is included in Eq. (20a); besides, the HY coordinates can be used in the cases with electrostatic and electromagnetic fluctuations. In the following, we denote $\boldsymbol{G}_{11h} \equiv -\frac{1}{2}\boldsymbol{G}_1 \cdot \nabla_6 \boldsymbol{G}_{1h}$ to simplify the notation.

Similarly, the velocity moment integral in the HY coordinates can be formally written as

$$\mathfrak{n}[\lambda](\boldsymbol{r}) = \int J_h d^6\boldsymbol{Z}_h \Delta_h(\boldsymbol{Z}_h;\boldsymbol{r}) F_h(\boldsymbol{Z}_h)\Lambda_h(\boldsymbol{Z}_h), \tag{21}$$

where $d^6\boldsymbol{Z}_h = d^3\boldsymbol{X}_h d\bar{v}_\parallel d\bar{\mu}\, d\bar{\xi}$. According to the scalar invariance, the volume element is determined by

$$J_h d^6\boldsymbol{Z}_h(\boldsymbol{Z}_h) = B_\parallel^* d^6\bar{Z}[\bar{Z}(\boldsymbol{Z}_h)], \tag{22}$$

where the Jacobian is calculated as [25]

$$\begin{aligned} J_h &= B_\parallel^*[\bar{Z}(\boldsymbol{Z}_h)]\left|\frac{\partial \bar{Z}}{\partial \boldsymbol{Z}_h}\right|[\bar{Z}(\boldsymbol{Z}_h)] \\ &= B_\parallel^*\bigg\{1 + \nabla_6 \cdot (\boldsymbol{G}_{1h} + \boldsymbol{G}_{2h} + \boldsymbol{G}_{11h}) \\ &\quad + \frac{1}{2}\nabla_6 \cdot [\nabla_6 \cdot (\boldsymbol{G}_{1h}\boldsymbol{G}_{1h})]\bigg\}. \end{aligned} \tag{23}$$

Obviously, the Jacobian is time dependent, and includes the polarization effects. The "location" function is obtained as

$$\Delta_h(\boldsymbol{Z}_h;\boldsymbol{r}) = \delta[z(\boldsymbol{Z}_h);\boldsymbol{r}] = \delta_D(\boldsymbol{X}_h + \bar{\rho} - \boldsymbol{r}), \tag{24}$$

and the distribution function is obtained as

$$\begin{aligned} F_h(\boldsymbol{Z}_h) &= \bar{F}[\bar{Z}(\boldsymbol{Z}_h)] \\ &= \bar{F} + (\boldsymbol{G}_{1h} + \boldsymbol{G}_{2h} + \boldsymbol{G}_{11h}) \cdot \nabla_6 \bar{F} \\ &\quad + \tfrac{1}{2}(\boldsymbol{G}_{1h} \cdot \nabla_6 \boldsymbol{G}_{1h}) \cdot \nabla_6 \bar{F} \\ &\quad + \tfrac{1}{2}(\boldsymbol{G}_{1h} \cdot \nabla_6)(\boldsymbol{G}_{1h} \cdot \nabla_6 \bar{F}). \end{aligned} \tag{25}$$

Combining Eqs. (23) and (25) yields

$$J_h F_h = B_\parallel^*(F_{h,0} + F_{h,1} + F_{h,2}), \tag{26}$$

where

$$F_{h,0} = \bar{F}, \tag{27a}$$

$$F_{h,1} = \nabla_6 \cdot (\boldsymbol{G}_{1h}\bar{F}), \tag{27b}$$

$$F_{h,2} = F_{h,2}^{(1)} + F_{h,2}^{(2)} + F_{h,2}^{(3)}, \tag{27c}$$

with

$$F_{h,2}^{(1)} = \nabla_6 \cdot (\boldsymbol{G}_{2h}\bar{F}), \tag{28a}$$

$$F_{h,2}^{(2)} = \nabla_6 \cdot (\boldsymbol{G}_{11h}\bar{F}), \tag{28b}$$

$$F_{h,2}^{(3)} = \frac{1}{2}\nabla_6 \cdot \nabla_6 \cdot (\boldsymbol{G}_{1h}\boldsymbol{G}_{1h}\bar{F}). \tag{28c}$$

Note that $\boldsymbol{G}_{nh}$ represents the polarization effects, $F_{h,1}$ and $F_{h,2}$ can be defined as the linear and nonlinear polarization distribution function, respectively.

By combining Eqs. (23)–(25), Eq. (21) can be rewritten as

$$\mathfrak{n}[\lambda](\boldsymbol{r}) = \int B_\parallel^* d^6\boldsymbol{Z}_h \delta_D(\boldsymbol{X}_h + \bar{\rho} - \boldsymbol{r})(F_{h,0} + F_{h,1} + F_{h,2})\Lambda_h. \tag{29}$$

Note that in Eq. (29), the distribution function and the moment function are separated, the form of the moment integral is similar to that in Eq. (17). We mention that Eq. (29) is more convenient to calculate the velocity moments compared to Eq. (19), especially for the first- and second-order velocity moments. Additionally, the distribution function (accompanied with the Jacobian) has been divided in terms of the polarization effects, when the moment function Λ_h is divided analogous to the distribution function, the polarization effects in the velocity moments can be easily distinguished. We will discuss this issue in Sec. III C in detail.

III. VELOCITY TRANSFORMATION BETWEEN GY AND HY COORDINATES

The velocity transformation between two coordinates can be obtained based on the scalar invariance, which obeys [19,29]

$$\dot{Z}_1^i(\boldsymbol{Z}_1) = \dot{Z}_2^j \frac{\partial Z_1^i}{\partial Z_2^j}[\boldsymbol{Z}_2(\boldsymbol{Z}_1)], \tag{30}$$

which should be understood as follows. For a vector field with the basic vectors fixed, the component of this vector field can be written in $\boldsymbol{Z}_1$ coordinates or in $\boldsymbol{Z}_2$ coordinates. In $\boldsymbol{Z}_1$ coordinates, the function for the component is $\dot{Z}_1^i$, which is evaluated at $\boldsymbol{Z}_1$; in $\boldsymbol{Z}_2$ coordinates, the function for the component is $\dot{Z}_2^j \frac{\partial Z_1^i}{\partial Z_2^j}$, which is evaluated at $\boldsymbol{Z}_2$. $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ represent the same point in the phase space, which leads that the values of component calculated in $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ coordinates are the same.

Thus, in practical calculation, there are two steps to obtain $\dot{Z}_1^i$ in $\boldsymbol{Z}_1$ coordinates: (i) calculate the expression for $\dot{Z}_2^j \frac{\partial Z_1^i}{\partial Z_2^j}$, which is evaluated at $\boldsymbol{Z}_2$, this step can be regarded as the push-forward of the function; (ii) take the Taylor expansion of $\dot{Z}_2^j \frac{\partial Z_1^i}{\partial Z_2^j}$ around $\boldsymbol{Z}_1$, this step can be regarded as the pull-back of the coordinates. Both in these two steps, the coordinate transformation between $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ is necessary.

A. Velocity transformation between GC and GY coordinates

We begin with the brief case about the velocity transformation between GC and GY coordinates to clarify the manipulation of Eq. (30). In detail, the first step is to obtain the GC velocity in terms of the $\bar{Z}$ coordinates, which is

$$\dot{\boldsymbol{Z}}(\bar{Z}) \equiv \dot{\bar{Z}} \cdot \nabla_7 \boldsymbol{Z} \tag{31a}$$

$$= \dot{\bar{Z}} - \dot{\bar{Z}} \cdot \nabla_7\big(\boldsymbol{G}_1 + \boldsymbol{G}_2 + \tfrac{1}{2}\boldsymbol{G}_1 \cdot \nabla_7 \boldsymbol{G}_1\big). \tag{31b}$$

Here and in the following, we define $\nabla_7 = \nabla_6 + \partial_t$; correspondingly, we add an artificial dimension, which corresponds to the time t, to the vectors such as $\dot{\bar{Z}}$ and $\boldsymbol{G}_n$, when they are written to dot product with ∇_7. Recall that $G_n^t = 0$, $\dot{t} = \dot{\bar{t}} = 1$, it is obvious that $\dot{\bar{Z}} \cdot \nabla_7 = \dot{\bar{Z}} \cdot \nabla_6 + \partial_t$, $\boldsymbol{G}_1 \cdot \nabla_7 = \boldsymbol{G}_1 \cdot \nabla_6$.

The existence of the terms with time derivative is due to the fact that the coordinate transformation between $\boldsymbol{Z}$ and $\bar{Z}$ is time dependent.

The second step is to take the Taylor expansion of Eq. (31b) around $\boldsymbol{Z}$ according to the coordinate transformation shown in Eq. (8b). Note that

$$\dot{\bar{Z}}(\bar{Z}) = \dot{\bar{Z}}(\boldsymbol{Z}) + (\boldsymbol{G}_1 + \boldsymbol{G}_2)\cdot\nabla_7\dot{\bar{Z}} + \tfrac{1}{2}\boldsymbol{G}_1\cdot\nabla_7(\boldsymbol{G}_1\cdot\nabla_7\dot{\bar{Z}}),$$

$$\dot{\bar{Z}}\cdot\nabla_7\boldsymbol{G}_n(\bar{Z}) = (1+\boldsymbol{G}_1\cdot\nabla_7)(\dot{\bar{Z}}\cdot\nabla_7\boldsymbol{G}_n)(\boldsymbol{Z}),$$

then the GC velocity in $\boldsymbol{Z}$ coordinates, expressed by the phase-space velocity in $\bar{Z}$ coordinates, is obtained as

$$\begin{aligned}\dot{\boldsymbol{Z}}(\boldsymbol{Z}) = {} & \dot{\bar{Z}} + \boldsymbol{G}_1\cdot\nabla_7\dot{\bar{Z}} - \dot{\bar{Z}}\cdot\nabla_7\boldsymbol{G}_1 \\ & + \boldsymbol{G}_2\cdot\nabla_7\dot{\bar{Z}} - \dot{\bar{Z}}\cdot\nabla_7\boldsymbol{G}_2 + \tfrac{1}{2}\boldsymbol{G}_1\cdot\nabla_7(\boldsymbol{G}_1\cdot\nabla_7\dot{\bar{Z}}) \\ & + \tfrac{1}{2}\dot{\bar{Z}}\cdot\nabla_7(\boldsymbol{G}_1\cdot\nabla_7\boldsymbol{G}_1) - \boldsymbol{G}_1\cdot\nabla_7(\dot{\bar{Z}}\cdot\nabla_7\boldsymbol{G}_1).\end{aligned} \tag{32}$$

By using $\nabla_7\cdot\dot{\bar{Z}} \equiv \frac{1}{B_\parallel^*}\partial_i(B_\parallel^*\dot{\bar{Z}}^i) = 0$ and $\nabla_7\cdot\boldsymbol{G}_n \equiv \frac{1}{B_\parallel^*}\partial_i(B_\parallel^* G_n^i) = 0$ [8,27], Eq. (32) can be rewritten in a concise form

$$\begin{aligned}\dot{\boldsymbol{Z}}(\boldsymbol{Z}) = {} & \dot{\bar{Z}} + \nabla_7\cdot(\mathbf{T}_1+\mathbf{T}_2) \\ & + \tfrac{1}{2}\nabla_7\cdot(\boldsymbol{G}_1\nabla_7\cdot\mathbf{T}_1 - \mathbf{T}_1\cdot\nabla_7\boldsymbol{G}_1),\end{aligned} \tag{33}$$

where we have defined $\mathbf{T}_n = \boldsymbol{G}_n\dot{\bar{Z}} - \dot{\bar{Z}}\boldsymbol{G}_n$, which is an antisymmetric tensor. Further, by using the identity

$$\nabla_7\cdot(\mathbf{T}_1\cdot\nabla_7\boldsymbol{G}_1)] = \nabla_7\cdot[(\nabla_7\cdot\mathbf{T}_1)\boldsymbol{G}_1],$$

Eq. (33) can be rewritten as

$$\dot{\boldsymbol{Z}}(\boldsymbol{Z}) = \dot{\bar{Z}} + \nabla_7\cdot\mathbf{T}_1 + \nabla_7\cdot\mathbf{T}_2 + \nabla_7\cdot\mathbf{T}_{11}. \tag{34}$$

Here, $\mathbf{T}_{11} = \frac{1}{2}\boldsymbol{G}_1(\nabla_7\cdot\mathbf{T}_1) - \frac{1}{2}(\nabla_7\cdot\mathbf{T}_1)\boldsymbol{G}_1$, which is also an antisymmetric tensor. Note that $\nabla_7\cdot\dot{\bar{Z}} = 0$, it is obvious $\nabla_7\cdot\dot{\boldsymbol{Z}} = 0$.

B. Velocity transformation between GY and HY coordinates

We then deal with the velocity transformation between GY and HY coordinates. For simplicity, we define

$$\boldsymbol{G}_0 \equiv \boldsymbol{G}_{1h} + \boldsymbol{G}_{2h} + \boldsymbol{G}_{11h},$$

then Eqs. (20a) and (20b) are rewritten as

$$\boldsymbol{Z}_h = \bar{Z} - \boldsymbol{G}_0(\bar{Z}), \tag{35a}$$

$$\bar{Z} = \boldsymbol{Z}_h + \boldsymbol{G}_0 + \boldsymbol{G}_0\cdot\nabla_6\boldsymbol{G}_0(\boldsymbol{Z}_h). \tag{35b}$$

Note that to $O(\epsilon_\delta^2)$, $\boldsymbol{G}_0\cdot\nabla_6\boldsymbol{G}_0 = \boldsymbol{G}_{1h}\cdot\nabla_6\boldsymbol{G}_{1h}$; also note that $G_0^t = 0$. In the following, we denote $\boldsymbol{A}_{00} \equiv \boldsymbol{G}_0\cdot\nabla_6\boldsymbol{G}_0$ to simplify the notation.

The first step is to push forward the function $\dot{\boldsymbol{Z}}_h$ to the $\bar{Z}$ coordinates, which is

$$\dot{\boldsymbol{Z}}_h(\bar{Z}) = \dot{\bar{Z}}\cdot\nabla_7\boldsymbol{Z}_h(\bar{Z}) = \dot{\bar{Z}}(\bar{Z}) - \dot{\bar{Z}}\cdot\nabla_7\boldsymbol{G}_0(\bar{Z}).$$

The second step is to take the Taylor expansion of $\dot{\boldsymbol{Z}}_h(\bar{Z})$ around $\boldsymbol{Z}_h$ by using Eq. (35b) to pull back the coordinates. Note that

$$\dot{\bar{Z}}(\bar{Z}) = \dot{\bar{Z}}(\boldsymbol{Z}_h) + \boldsymbol{G}_0\cdot\nabla_7\dot{\bar{Z}} + \boldsymbol{A}_{00}\cdot\nabla_7\dot{\bar{Z}} + \tfrac{1}{2}\boldsymbol{G}_0\boldsymbol{G}_0 : \nabla_7\nabla_7\dot{\bar{Z}},$$

$$\dot{\bar{Z}}\cdot\nabla_7\boldsymbol{G}_0(\bar{Z}) = (1+\boldsymbol{G}_0\cdot\nabla_7)(\dot{\bar{Z}}\cdot\nabla_7\boldsymbol{G}_0)(\boldsymbol{Z}_h),$$

then we obtain

$$\begin{aligned}\dot{\boldsymbol{Z}}_h(\boldsymbol{Z}_h) = {} & \dot{\bar{Z}} + \boldsymbol{G}_0\cdot\nabla_7\dot{\bar{Z}} - \dot{\bar{Z}}\cdot\nabla_7\boldsymbol{G}_0 \\ & + \boldsymbol{A}_{00}\cdot\nabla_7\dot{\bar{Z}} + \tfrac{1}{2}\boldsymbol{G}_0\boldsymbol{G}_0 : \nabla_7\nabla_7\dot{\bar{Z}} \\ & - \boldsymbol{G}_0\cdot\nabla_7(\dot{\bar{Z}}\cdot\nabla_7\boldsymbol{G}_0).\end{aligned} \tag{36}$$

Equation (36) can be manipulated similar to the calculation from Eq. (32) to (33), however, one should take notice of $\nabla_7\cdot\boldsymbol{G}_0 \neq 0$. After some straightforward calculation, Eq. (36) is finally calculated as

$$\begin{aligned}\dot{\boldsymbol{Z}}_h(\boldsymbol{Z}_h) = {} & \{1 - g_0 - \tfrac{1}{2}\nabla_7\cdot\nabla_7\cdot(\boldsymbol{G}_0\boldsymbol{G}_0) + g_0^2\}\dot{\bar{Z}} \\ & + (1-g_0)\boldsymbol{t}_0 + \tfrac{1}{2}\nabla_7\cdot(\boldsymbol{G}_0\boldsymbol{t}_0 - \boldsymbol{t}_0\boldsymbol{G}_0) \\ & + \tfrac{1}{2}\nabla_7\cdot(\boldsymbol{A}_{00}\dot{\bar{Z}} - \dot{\bar{Z}}\boldsymbol{A}_{00}),\end{aligned} \tag{37}$$

where $g_0 \equiv \nabla_7\cdot\boldsymbol{G}_0$, $\boldsymbol{t}_0 \equiv \nabla_7\cdot\mathbf{T}_0$ with $\mathbf{T}_0 \equiv \boldsymbol{G}_0\dot{\bar{Z}} - \dot{\bar{Z}}\boldsymbol{G}_0$. Here, we mention that by replacing ∇_7 with $\nabla_6 + \partial_t$, the time derivative in Eq. (37) can be explicitly expressed, which is $-\partial_t\boldsymbol{G}_0 - \boldsymbol{G}_0\cdot\nabla_6(\partial_t\boldsymbol{G}_0)$; this term can be regarded as the Taylor expansion of $-\partial_t\boldsymbol{G}_0(\bar{Z})$ around $\boldsymbol{Z}_h$ by keeping up to $O(\epsilon_\delta^2)$.

Combined with the Jacobian shown in Eq. (23), Eq. (37) can be rewritten as

$$\begin{aligned}\dot{\boldsymbol{Z}}_h(\boldsymbol{Z}_h) = {} & J_{h0}^{-1}\{\dot{\bar{Z}} + \boldsymbol{t}_0 + \tfrac{1}{2}\nabla_7\cdot(\boldsymbol{G}_0\boldsymbol{t}_0 - \boldsymbol{t}_0\boldsymbol{G}_0) \\ & + \tfrac{1}{2}\nabla_7\cdot(\boldsymbol{A}_{00}\dot{\bar{Z}} - \dot{\bar{Z}}\boldsymbol{A}_{00})\},\end{aligned} \tag{38}$$

where

$$J_{h0} = \frac{1}{B_\parallel^*}J_h = 1 + g_0 + \frac{1}{2}\nabla_7\cdot\nabla_7\cdot(\boldsymbol{G}_0\boldsymbol{G}_0).$$

By using $\nabla_7\cdot\dot{\bar{Z}} = 0$, it is easy to find that

$$\nabla_7\cdot(J_{h0}\dot{\boldsymbol{Z}}_h) = 0, \tag{39a}$$

which is alternatively written as

$$\partial_t J_h + \partial_i\big(J_h\dot{Z}_h^i\big) = 0, \tag{39b}$$

with i summed over the six phase-space variables. By neglecting the nonlinear parts in the Jacobian J_h and the phase-space velocity $\dot{Z}_h^i$, the Liouville theorem shown in Eq. (39b) reduces to Eq. (19) in Ref. [22] in the electrostatic case. Note that in Ref. [22], the Liouville theorem is derived based on the Lie transform for the phase-space coordinates with polarization drift; in our method, the Liouville theorem is derived based on the transformation between GY and HY coordinates. The equivalent results are due to the fact that the physical meaning of the coordinates adopted in [22] and our work are the same.

To end this section, we show that, mathematically speaking, Eq. (34) is similar to Eq. (38). For further explanation, we take $\boldsymbol{G}_{nh} = \boldsymbol{G}_n$, the HY coordinates $\boldsymbol{Z}_h$ are reduced back to the GC coordinates $\boldsymbol{Z}$. Note that

$$g_0 = \nabla_6\cdot\big(\boldsymbol{G}_1 + \boldsymbol{G}_2 - \tfrac{1}{2}\boldsymbol{G}_1\cdot\nabla_6\boldsymbol{G}_1\big) = -\tfrac{1}{2}\nabla_6\cdot\nabla_6\cdot(\boldsymbol{G}_1\boldsymbol{G}_1),$$

then according to Eq. (23), $J_{h0} = 1$. Additionally,

$$\boldsymbol{t}_0 = \nabla_7 \cdot (\mathbf{T}_1 + \mathbf{T}_2) - \tfrac{1}{2}\nabla_7 \cdot [(\boldsymbol{G}_1 \cdot \nabla_7 \boldsymbol{G}_1)\dot{\bar{Z}} - \dot{\bar{Z}}(\boldsymbol{G}_1 \cdot \nabla_7 \boldsymbol{G}_1)],$$

in which the last term is canceled out with the last term in Eq. (38) by keeping up to $O(\epsilon_\delta^2)$, then one can easily find that Eq. (38) is reduced back to Eq. (34).

C. The GK equation in the HY coordinates

Generally, the form of the Vlasov equation in $\mathcal{Z}$ coordinates is written as [18]

$$\partial_t(\mathcal{J}_{\mathcal{Z}}\mathcal{F}_{\mathcal{Z}}) + \partial_i(\dot{\mathcal{Z}}^i\mathcal{J}_{\mathcal{Z}}\mathcal{F}_{\mathcal{Z}}) = 0, \tag{40}$$

where $\mathcal{J}_{\mathcal{Z}}$, $\mathcal{F}_{\mathcal{Z}}$, and $\dot{\mathcal{Z}}$ denote the Jacobian, the distribution function, and the phase-space velocity in $\mathcal{Z}$ coordinates. We adopt $\mathcal{Z}$ as the $\boldsymbol{Z}_h$ coordinates. The distribution function (accompanied with the Jacobian) is shown in Eq. (26); similarly, the phase-space velocity in Eq. (38) is rewritten as

$$\dot{\boldsymbol{Z}}_h = \dot{\boldsymbol{Z}}_{h,0} + \dot{\boldsymbol{Z}}_{h,1} + \dot{\boldsymbol{Z}}_{h,2}, \tag{41}$$

$$\dot{\boldsymbol{Z}}_{h,0} = \dot{\bar{Z}}, \tag{42a}$$

$$\dot{\boldsymbol{Z}}_{h,1} = -\partial_t \boldsymbol{G}_{1h} - g_{1h}\dot{\bar{Z}} + \boldsymbol{t}_{1h}, \tag{42b}$$

$$\dot{\boldsymbol{Z}}_{h,2} = \dot{\boldsymbol{Z}}^{(1)}_{h,2} + \dot{\boldsymbol{Z}}^{(2)}_{h,2} + \dot{\boldsymbol{Z}}^{(3)}_{h,2}, \tag{42c}$$

with

$$\dot{\boldsymbol{Z}}^{(1)}_{h,2} = -\partial_t \boldsymbol{G}_{2h} - g_{2h}\dot{\bar{Z}} + \boldsymbol{t}_{2h}, \tag{43a}$$

$$\dot{\boldsymbol{Z}}^{(2)}_{h,2} = -\partial_t \boldsymbol{G}_{11h} - g_{11h}\dot{\bar{Z}} + \boldsymbol{t}_{11h}, \tag{43b}$$

$$\begin{aligned}\dot{\boldsymbol{Z}}^{(3)}_{h,2} = {} & -\boldsymbol{G}_{1h} \cdot \nabla_6(\partial_t \boldsymbol{G}_{1h}) + g_{1h}^2\dot{\bar{Z}} \\ & - \tfrac{1}{2}\nabla_6 \cdot \nabla_6 \cdot (\boldsymbol{G}_{1h}\boldsymbol{G}_{1h})\dot{\bar{Z}} - g_{1h}\boldsymbol{t}_{1h} \\ & + \tfrac{1}{2}\nabla_6 \cdot (\boldsymbol{G}_{1h}\boldsymbol{t}_{1h} - \boldsymbol{t}_{1h}\boldsymbol{G}_{1h}) \\ & + \tfrac{1}{2}\nabla_6 \cdot (\boldsymbol{G}_{1h} \cdot \nabla_6 \boldsymbol{G}_{1h}\dot{\bar{Z}} - \dot{\bar{Z}}\boldsymbol{G}_{1h} \cdot \nabla_6 \boldsymbol{G}_{1h}).\end{aligned} \tag{43c}$$

Here, $g_{nh} \equiv \nabla_6 \cdot \boldsymbol{G}_{nh}$, $\boldsymbol{t}_{nh} = \nabla_6 \cdot \mathbf{T}_{nh}$ with $\mathbf{T}_{nh} \equiv \boldsymbol{G}_{nh}\dot{\bar{Z}} - \dot{\bar{Z}}\boldsymbol{G}_{nh}$. Obviously, the linear and nonlinear polarization effects are included in $\dot{\boldsymbol{Z}}_{h,1}$ and $\dot{\boldsymbol{Z}}_{h,2}$, respectively.

We argue that Eq. (40) can be naturally separated according to the ordering of polarization effects. In detail, the ordering of $O(\boldsymbol{G}^0_{1h})$ reads as

$$\partial_t \bar{F} + \nabla_6 \cdot (\dot{\bar{Z}}\bar{F}) = 0. \tag{44}$$

The form is exactly the same as the GK equation shown in Eq. (15); note that the functions in Eq. (44) are evaluated at $\boldsymbol{Z}_h$. Obviously, we can conclude that no any polarization effect is included in Eq. (15). The ordering of $O(\boldsymbol{G}^1_{1h})$ reads as

$$\partial_t F_{h,1} + \nabla_6 \cdot (\dot{\bar{Z}}F_{h,1} + \dot{\boldsymbol{Z}}_{h,1}\bar{F}) = 0, \tag{45}$$

which shows the evolution of the linear polarization distribution. The ordering of $O(\boldsymbol{G}^2_{1h})$ reads as

$$\partial_t F^{(1)}_{h,2} + \nabla_6 \cdot \left(\dot{\bar{Z}}F^{(1)}_{h,2} + \dot{\boldsymbol{Z}}^{(1)}_{h,2}\bar{F}\right) = 0, \tag{46a}$$

$$\partial_t F^{(2)}_{h,2} + \nabla_6 \cdot \left(\dot{\bar{Z}}F^{(2)}_{h,2} + \dot{\boldsymbol{Z}}^{(2)}_{h,2}\bar{F}\right) = 0, \tag{46b}$$

$$\partial_t F^{(3)}_{h,2} + \nabla_6 \cdot \left(\dot{\bar{Z}}F^{(3)}_{h,2} + \dot{\boldsymbol{Z}}_{h,1}F_{h,1} + \dot{\boldsymbol{Z}}^{(3)}_{h,2}\bar{F}\right) = 0, \tag{46c}$$

which shows the evolution of the nonlinear polarization distribution. Note that the forms of Eqs. (46a) and (46b) are the same as that of Eq. (45) since the roles of $\boldsymbol{G}_{2h}$ and $\boldsymbol{G}_{11h}$ are the same as that of $\boldsymbol{G}_{1h}$ in the coordinates transformation from $\bar{Z}$ to $\boldsymbol{Z}_h$, as are shown in Eqs. (20a) and (20b).

Here, we make some remarks about Eqs. (44)–(46):

(i) By using Eq. (44) and the definition of $F_{h,1}$ shown in Eq. (27b), Eq. (45) can be obtained by directly calculating $\partial_t F_{h,1}$, as well as the illustration for Eqs. (46a) and (46b); with the help of Eqs. (44) and (45) and the definition of $F^{(3)}_{h,2}$ shown in Eq. (28c), Eq. (46c) can be obtained by directly calculating $\partial_t F^{(3)}_{h,2}$; more details are shown in Appendices A and B. In other words, Eqs. (44)–(46) exactly equal to Eq. (44) combined with the definitions of $F_{h,1}$ and $F_{h,2}$.

(ii) As is known, in terms of particle coordinates, the self-consistent evolution of electromagnetic turbulence is governed by the particle Vlasov equation and the Maxwell equations. In terms of the $\boldsymbol{Z}_h$ coordinates, the distribution function can be obtained by solving Eq. (44) to obtain $\bar{F}$ and by introducing the definitions of $F_{h,1}$ and $F_{h,2}$; the velocity moments can be obtained according to Eq. (29).

(iii) The particle density is calculated as

$$n(\boldsymbol{r}) = \int B^*_\parallel d^6\boldsymbol{Z}_h \delta_D(\boldsymbol{X}_h + \bar{\rho} - \boldsymbol{r})(\bar{F} + F_{h,1} + F_{h,2}). \tag{47}$$

Here, the integrals for $\bar{F}$, $F_{h,1}$, and $F_{h,2}$ denote GY, linear, and nonlinear polarization density, respectively. For the electrostatic case, by adopting the long-wave-limit approximation $\boldsymbol{G}_{1h} \approx \boldsymbol{G}^{\boldsymbol{\rho}}_1 = \nabla_\perp \delta\phi / B_0\Omega$, then the linear polarization density is calculated as $\nabla \cdot (n_g \nabla_\perp \delta\phi / B_0\Omega)$ with n_g the GY density, which reproduces the result in [21]. In Ref. [30], it is demonstrated that in the orbit computation, by considering the $\boldsymbol{G}_1\boldsymbol{G}_1$ terms, the short-time I-transform method which involves integrating only along the UP orbit agrees with the conventional method which integrates along the full orbit. We mention that consideration of the nonlinear polarization density in the GK Poisson equation may be beneficial for the nonlinear evolution of turbulence.

(iv) Analogously, the contributions from $\dot{\boldsymbol{Z}}_{h,1}$ and $\dot{\boldsymbol{Z}}_{h,2}$ should be added in the calculation of the current density, namely,

$$\begin{aligned}\boldsymbol{j}(\boldsymbol{r}) = {} & \int B^*_\parallel d^6\boldsymbol{Z}_h \delta_D(\boldsymbol{X}_h + \bar{\rho} - \boldsymbol{r}) \\ & \times (\dot{\bar{Z}} + \dot{\boldsymbol{Z}}_{h,1} + \dot{\boldsymbol{Z}}_{h,2})(\bar{F} + F_{h,1} + F_{h,2}).\end{aligned} \tag{48}$$

It is shown that the velocity and distribution function are both divided in terms of the polarization orderings, then the polarization effects in the current can be easily clarified. Obviously, the integral for $\dot{\bar{Z}}\bar{F}$ denotes the GY current density, the linear polarization current density contains the contribution from the part of polarization drift ($\dot{\boldsymbol{Z}}_{h,1}\bar{F}$) and the part of polarization distribution ($\dot{\bar{Z}}F_{h,1}$); the rest contributes to the nonlinear polarization current density.

Here, we give two examples to clarify the polarization effects in the case with electrostatic fluctuations. The first example is about the Hasegawa-Mima equation [31], which is a popular model to describe drift waves. For ions, by adopting the long-wave-limit approximation, the continuity equation is obtained by taking the velocity integral of Eqs. (44) and (45), which is

$$\partial_t \delta\hat{n}_i - \rho_s^2 \partial_t \nabla_\perp^2 \hat{\phi} + \rho_s^3 c_s \nabla\nabla_\perp^2 \hat{\phi} \cdot \nabla\hat{\phi} \times \boldsymbol{b}_0 - \rho_s c_s \nabla\hat{\phi} \times \boldsymbol{b}_0 \cdot \nabla \ln n_0 = 0 \tag{49}$$

Here, the equilibrium distribution function is adopted as the Maxwellian distribution. $\delta\hat{n}_i$ is normalized by the equilibrium density n_0, $\hat{\phi} = e\delta\phi/T$, $c_s = \sqrt{T/m}$ and $\rho_s = c_s/\Omega$. For electrons, the adiabatic assumption is adopted, that is, $\delta\hat{n}_e = \hat{\phi}$. Then, by taking the quasi-neutrality condition, Eq. (49) is rewritten as

$$\partial_t \left(1 - \rho_s^2 \nabla_\perp^2\right)\hat{\phi} - \rho_s^3 c_s \boldsymbol{b}_0 \times \nabla\hat{\phi} \cdot \nabla\nabla_\perp^2 \hat{\phi} - \rho_s c_s \nabla\hat{\phi} \times \boldsymbol{b}_0 \cdot \nabla \ln n_0 = 0, \tag{50}$$

which is the Hasegawa-Mima equation [31]. It is strange at the first sight that there exists a nonlinear term ($\sim\hat{\phi}^2$) in Eq. (50). In fact, Eq. (50) can be read from the viewpoint of polarization, that is, the term with $\partial_t \nabla_\perp^2 \hat{\phi}$ comes from $\partial_t \boldsymbol{G}_{1h}$ in $\dot{\boldsymbol{Z}}_{h,1}$, and the nonlinear term ($\sim\hat{\phi}^2$) comes from $g_{1h}\dot{\bar{Z}}$ in $\dot{\boldsymbol{Z}}_{h,1}$; both of these two terms relate to the linear polarization effect.

The second example is about the geodesic acoustic mode (GAM), which has been extensively studied in the past decades [32–36]. The dispersion relation for GAM is governed by $\nabla \cdot \boldsymbol{j} = \langle j^r \rangle_F = 0$, with $\langle \cdot \rangle_F$ denoting the magnetic surface average. From Eq. (48), the perturbed current density consists of two parts: one part is the perturbed GY diamagnetic current due to $\dot{\bar{Z}}_0 \bar{F}_1$, with $\bar{F}_1$ the perturbed GY distribution function; the other part is the perturbed polarization current due to $\dot{\bar{Z}}_0 F_{h,1} + \dot{\boldsymbol{Z}}_{h,1}\bar{F}_0$. Thus, the dispersion relation for GAM is

$$\langle v_{dr}\bar{F}_1 + v_{dr}F_{h,1} + u_p \bar{F}_0 \rangle_F = 0. \tag{51}$$

Here, v_{dr} denotes the radial diamagnetic drift velocity, $u_p = -\partial_t G_{1h}^r$ denotes the radial linear polarization drift velocity.

At last, we mention that the long-wave-limit approximation is not necessary in the GK calculation based on the HY coordinates, the HY coordinates may be used in the issues with short wavelength, such as the short-wavelength effects on the residual zonal flow [37,38] and the polarization effects on energetic particles [39] or impurities [40]. We leave these issues as a future work.

IV. POLARIZATION EFFECTS IN THE DISPERSION RELATION FOR CAW

Effects of the perpendicular magnetic potential fluctuation $\delta\boldsymbol{A}_\perp$ (or the parallel magnetic perturbation $\delta B_\parallel$) are usually neglected in a low-β plasma [41], with β the ratio of kinetic pressure to magnetic pressure. However, in the finite-β case, the $\delta\boldsymbol{A}_\perp$ effects are important. For example, the GTC simulation results show that $\delta\boldsymbol{A}_\perp$ can play an important role in drift-Alfvén instabilities in tokamaks, especially as β increases [42]. The theoretical analysis indicates that neglect of $\delta\boldsymbol{A}_\perp$ can cause enhanced stability on MHD instabilities [43]. In a high-β plasma in National Spherical Torus Experiment (NSTX), the CAWs are often observed during neutral beam injection [44,45]. The evolution of $\delta\boldsymbol{A}_\perp$ fluctuation is governed by the perpendicular Ampère's law, in which the current relates to the perpendicular velocity moment. In this section, we derive the dispersion relation for CAWs to (i) show the application of the perpendicular GK theory in the case with magnetic fluctuations and (ii) discuss the polarization effects in the perpendicular current.

For simplicity, we adopt the slab geometry, that is, (x, y, z) are used to denote the radial, poloidal, and toroidal directions, respectively; the GC and GY coordinates are labeled as (X, Y, Z_g) and $(\bar{X}, \bar{Y}, \bar{Z}_g)$, respectively. The background magnetic field $\boldsymbol{B}_0 = B_0 \boldsymbol{e}_z$ is spatially uniform, where $(\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z)$ are the unit vectors. The UP Hamiltonian is $\bar{H}_0 = \frac{1}{2}m\bar{v}_\parallel^2 + \bar{\mu}B_0$, and the UP phase-space velocities are $\dot{\bar{X}}_0 = \bar{v}_\parallel \boldsymbol{e}_z$, $\dot{\bar{\xi}} = \Omega$.

We assume the perturbed magnetic potential as $\delta\boldsymbol{A}_\perp = \boldsymbol{e}_x A_x(y, t)$. The Fourier transformation is

$$A_x(y, t) = \sum_{\boldsymbol{k}} A_{\boldsymbol{k}} e^{i(ky - \omega t)} \tag{52a}$$

$$\equiv \sum_{\boldsymbol{k}} e^{i\mathcal{M}} A_{\boldsymbol{k}} \sum_n J_n(\alpha) e^{-in\xi}. \tag{52b}$$

Here, Eqs. (52a) and (52b) denote the Fourier decomposition in the particle and GC coordinates, respectively. $\mathcal{M} = kY - \omega t$, the wave vector is $\boldsymbol{k} = k\boldsymbol{e}_y$, $\boldsymbol{\rho}_0 = \frac{v_\perp}{\Omega}(\boldsymbol{e}_x \cos\xi - \boldsymbol{e}_y \sin\xi)$, J_n is the nth Bessel function with the argument $\alpha = k\rho_0$. The first-order effective potential is

$$\delta\Psi_1 = -\dot{\boldsymbol{\rho}}_0 \cdot \delta\boldsymbol{A}_\perp\big|_{Z=\bar{Z}} \equiv \sum_{\boldsymbol{k}} e^{i\bar{\mathcal{M}}}\Psi_{\boldsymbol{k}}, \tag{53a}$$

$$\Psi_{\boldsymbol{k}} = i\bar{v}_\perp A_{\boldsymbol{k}} \sum_n J_n' e^{-in\bar{\xi}}. \tag{53b}$$

In Eq. (53a), the symbol "$|_{Z=\bar{Z}}$" denotes that the function is evaluated at $\bar{Z}$. $\bar{\mathcal{M}} = k\bar{Y} - \omega t$, $\bar{v}_\perp = \sqrt{2\bar{\mu}B_0/m}$, the "prime" superscript denotes the derivative to α.

The first-order Hamiltonian is

$$\delta\bar{H}_1 = -ie\bar{v}_\perp \sum_{\boldsymbol{k}} e^{i\bar{\mathcal{M}}} A_{\boldsymbol{k}} J_1, \tag{54}$$

which gives the first-order drift velocity as

$$\delta\dot{\bar{X}}_1 = -\frac{\bar{v}_\perp}{B_0} \sum_{\boldsymbol{k}} e^{i\bar{\mathcal{M}}} k A_{\boldsymbol{k}} J_1 \boldsymbol{e}_x. \tag{55}$$

According to Eq. (12), the first-order gauge function is solved as

$$S_1 \equiv \sum_{\boldsymbol{k}} e^{i\bar{\mathcal{M}}} S_{\boldsymbol{k}}, \tag{56a}$$

$$S_{\boldsymbol{k}} = -\frac{e\bar{v}_\perp A_{\boldsymbol{k}}}{\Omega} \sum_{n\neq 0} \frac{e^{-in\bar{\xi}}}{\hat{\omega} + n} J_n', \tag{56b}$$

where $\hat{\omega} = \omega/\Omega$.

From $\{\boldsymbol{r}, \boldsymbol{r}\} = 0$, it is obvious that $G_{1A}^r = 0$, then we only need to deal with G_{1Y}^r. According to Eqs. (11a)–(11d), it is found that

$$G_{1Y,\boldsymbol{k}}^X = -\frac{A_{\boldsymbol{k}}}{B_0} \sum_{n\neq 0} \frac{e^{-in\bar{\xi}}}{n + \hat{\omega}} i\alpha J_n', \tag{57a}$$

$$G_{1Y,\boldsymbol{k}}^Y = 0, \tag{57b}$$

$$G^{\mu}_{1Y,\boldsymbol{k}} = \frac{e\bar{v}_\perp A_{\boldsymbol{k}}}{B_0}\sum_{n\neq 0}\frac{e^{-in\bar{\xi}}}{n+\hat{\omega}}inJ_n', \tag{57c}$$

$$G^{\xi}_{1Y,\boldsymbol{k}} = \frac{eA_{\boldsymbol{k}}}{m\bar{v}_\perp}\sum_{n\neq 0}\frac{e^{-in\bar{\xi}}}{n+\hat{\omega}}(\alpha J_n')', \tag{57d}$$

from which we obtain

$$G^{\boldsymbol{\rho}}_{1Y,\boldsymbol{k}} = (a_1\boldsymbol{e}_x + a_2\boldsymbol{e}_y)\frac{A_{\boldsymbol{k}}}{B_0}, \tag{58a}$$

$$a_1 = \frac{i}{2}\sum_{n\neq 0}\frac{e^{-in\bar{\xi}}}{n+\hat{\omega}}\left[(\alpha J_{n+1})'e^{-i\bar{\xi}} + (\alpha J_{n-1})'e^{i\bar{\xi}}\right], \tag{58b}$$

$$a_2 = \frac{1}{2}\sum_{n\neq 0}\frac{e^{-in\bar{\xi}}}{n+\hat{\omega}}\left[(\alpha J_{n+1})'e^{-i\bar{\xi}} - (\alpha J_{n-1})'e^{i\bar{\xi}}\right]. \tag{58c}$$

Note that

$$(d/dt)_0 G^X_{1Y} = -\frac{\bar{v}_\perp}{B_0}\sum_{\boldsymbol{k}} e^{i\bar{\mathcal{M}}}kA_{\boldsymbol{k}}\sum_{n\neq 0}J_n' e^{-in\bar{\xi}}, \tag{59}$$

then we obtain

$$\delta\dot{\boldsymbol{x}}_1 \equiv \delta\dot{\bar{X}}_1 - (d/dt)_0 G^X_{1Y}\boldsymbol{e}_x \tag{60a}$$

$$= \frac{-i\bar{v}_\perp \sin\bar{\xi}}{B_0}\sum_{\boldsymbol{k}} e^{ik(\bar{Y}+\bar{\rho}_y)-i\omega t}kA_{\boldsymbol{k}}\boldsymbol{e}_x, \tag{60b}$$

that is, the perturbed potential $\delta\boldsymbol{A}_\perp$ appeared in $\delta\dot{\boldsymbol{x}}_1$ should be evaluated at $\bar{Y}+\bar{\rho}_y$.

The polarization drift velocity is

$$\boldsymbol{u}_p = -(\partial_t + \Omega\partial_{\bar{\xi}})G^{\boldsymbol{\rho}}_{1Y} \equiv \sum_{\boldsymbol{k}} e^{i\bar{\mathcal{M}}}\boldsymbol{u}_{p,\boldsymbol{k}}, \tag{61a}$$

$$u_{p,x,\boldsymbol{k}} = -\frac{eA_{\boldsymbol{k}}}{2m}\sum_{n\neq 0}\frac{e^{-in\bar{\xi}}}{n+\hat{\omega}}\left[(\hat{\omega}+n+1)(\alpha J_{n+1})'e^{-i\bar{\xi}} + (\hat{\omega}+n-1)(\alpha J_{n-1})'e^{i\bar{\xi}}\right], \tag{61b}$$

$$u_{p,y,\boldsymbol{k}} = \frac{ieA_{\boldsymbol{k}}}{2m}\sum_{n\neq 0}\frac{e^{-in\bar{\xi}}}{n+\hat{\omega}}\left[(\hat{\omega}+n+1)(\alpha J_{n+1})'e^{-i\bar{\xi}} - (\hat{\omega}+n-1)(\alpha J_{n-1})'e^{i\bar{\xi}}\right]. \tag{61c}$$

For the case $\hat{\omega} \ll 1$, by keeping up to the leading order of $k\rho_0$, Eqs. (61b) and (61c) are reduced to

$$u_{p,x,\boldsymbol{k}} = \frac{eA_{\boldsymbol{k}}}{m}\hat{\omega}^2, \tag{62a}$$

$$u_{p,y,\boldsymbol{k}} = -\frac{ieA_{\boldsymbol{k}}}{m}\hat{\omega}, \tag{62b}$$

which are transformed back to the real space as

$$\boldsymbol{u}_p = \frac{-\partial_t\delta\boldsymbol{A}_\perp\times\boldsymbol{b}_0}{B_0} - \frac{\partial_t^2\delta\boldsymbol{A}_\perp}{B_0\Omega}. \tag{62c}$$

Note that the perturbed electric field is $\delta\boldsymbol{E}_\perp = -\partial_t\delta\boldsymbol{A}_\perp$, thus in Eq. (62c), the first term represents the $\boldsymbol{E}\times\boldsymbol{B}$ drift, which is equal to ions and electrons; the second term represents the polarization drift, which is proportional to $1/\Omega$, thus, ions dominate this term.

The linear dispersion relation for CAW is governed by the perpendicular Ampère's law

$$-\nabla^2 A_x(\boldsymbol{r}) = \mu_0 e\int d^6\boldsymbol{Z}_h\delta_D(\boldsymbol{X}_h+\bar{\rho}-\boldsymbol{r}) \times(\dot{\bar{X}}\bar{F} + \dot{\bar{X}}F_{h,1} + \dot{X}_{h,1}\bar{F}), \tag{63}$$

with μ_0 the permeability of vacuum. In the following, the equilibrium distribution function is adopted as

$$\bar{F}_0 = \frac{n_0}{\pi^{3/2}v_{th}^3}e^{-\hat{v}_\parallel^2-\hat{v}_\perp^2}, \tag{64}$$

where $v_{th} = \sqrt{2T/m}$ is the thermal velocity, $\hat{v}_\parallel$ and $\hat{v}_\perp$ are the parallel and perpendicular velocity normalized to v_{th}, respectively. By taking the linearization of Eq. (15), one can easily find that $\bar{F}_1 = 0$, then $\bar{F} = \bar{F}_0$.

In Eq. (63), since the fluctuations relate to y, the three-dimensional Dirac function is reduced to $\delta_D(Y_h+\bar{\rho}_y-y)$, then, after the integrals in the configuration space, the factor e^{ikY_h} in the fluctuations should be replaced by $e^{iky}e^{-ik\bar{\rho}_y}$. According to Eq. (41), the radial velocity is $\boldsymbol{e}_x\cdot(\dot{\bar{\rho}}+\delta\dot{\boldsymbol{x}}_1+\boldsymbol{u}_p)$ by keeping up to $O(\epsilon_\delta)$, from which the perturbed radial current density is calculated as

$$j_x = \sum_{\boldsymbol{k}} e^{i(ky-\omega t)}(j_{D,x,\boldsymbol{k}} + j_{L,x,\boldsymbol{k}} + j_{p,x,\boldsymbol{k}}). \tag{65}$$

Here,

$$j_{D,x,\boldsymbol{k}} = e\int d^3\bar{\boldsymbol{v}}\,\dot{x}_{1,\boldsymbol{k}}\bar{F}_0 \tag{66}$$

is the current contributed from the first-order drift velocity; $\int d^3\bar{\boldsymbol{v}} = \int B_0 d\bar{v}_\parallel d\bar{\mu}\,d\bar{\xi}$. According to Eq. (60b), $\dot{x}_{1,\boldsymbol{k}}$ is periodic of $\bar{\xi}$, thus $j_{D,x,\boldsymbol{k}} = 0$ by taking the $\bar{\xi}$ integral,

$$j_{L,x,\boldsymbol{k}} = e\int d^3\bar{\boldsymbol{v}}\,\dot{\bar{\rho}}_x ikG^{\rho_y}_{1Y,\boldsymbol{k}}e^{-ik\bar{\rho}_y}\bar{F}_0 \tag{67a}$$

is the current contributed from the Larmor motion and the polarization distribution. Note the facts that

$$\dot{\bar{\rho}}_x e^{-ik\bar{\rho}_y} = i\bar{v}_\perp\sum_l J_l'(\alpha)e^{il\bar{\xi}},$$

$$\frac{1}{2\pi}\int d\bar{\xi}\,J_l e^{i(l-n)\bar{\xi}} = J_n,$$

$$\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}d\hat{v}_\parallel e^{-\hat{v}_\parallel^2} = 1,$$

Eq. (67a) is calculated as

$$j_{L,x,\boldsymbol{k}} = \frac{e^2 n_0}{m}A_{\boldsymbol{k}}\sum_{n\neq 0}\frac{2n}{\hat{\omega}^2-n^2}\mathcal{I}^{(1)}_{n+1}. \tag{67b}$$

Here, $\mathcal{I}^{(1)}_{n+1}(\alpha_0) \equiv \int_0^\infty dx\,x^2 e^{-x^2}\frac{d(xJ_{n+1})}{dx}\frac{dJ_{n+1}}{dx}$, where the argument of the Bessel function is $\alpha_0 x$ with $\alpha_0 = kv_{th}/\Omega$,

$$j_{p,x,\boldsymbol{k}} = e\int d^3\bar{\boldsymbol{v}}\,u_{p,x,\boldsymbol{k}}e^{-ik\bar{\rho}_y}\bar{F}_0 \tag{68a}$$

is the current contributed from the polarization drift. Similarly, Eq. (68a) is finally calculated as

$$j_{p,x,\boldsymbol{k}} = -A_{\boldsymbol{k}}\frac{e^2 n_0}{m}\sum_{n\neq 0}\frac{\hat{\omega}^2-n(n+1)}{\hat{\omega}^2-n^2}\mathcal{I}^{(2)}_{n+1}. \tag{68b}$$

Here, $\mathcal{I}^{(2)}_{n+1}(\alpha_0) = 2\int dx\, x^3 e^{-x^2} J^2_{n+1}$, where the argument of the Bessel function is $\alpha_0 x$.

Thus, the perturbed current in the linear Ampère's law contains two parts, both of which relate to the polarization effects: one part is from the contribution of polarization distribution function ($j_{L,x}$), and the other part is from the contribution of polarization drift velocity ($j_{p,x}$). Note that the polarization effects are proportional to $1/\Omega$, the perturbed current due to the contribution of electrons can be neglected, then the dispersion relation governed by the linear Ampère's law is

$$1 = \frac{\Omega^2}{k^2 v_A^2}\sum_{n\neq 0}\frac{1}{\hat{\omega}^2 - n^2}\left\{2n\mathcal{I}^{(1)}_{n+1} - [\hat{\omega}^2 - n(n+1)]\mathcal{I}^{(2)}_{n+1}\right\}, \tag{69}$$

with $v_A = \sqrt{B_0^2/\mu_0 n_0 m}$ is the Alfvén velocity.

We analyze Eq. (69) in the long-wave-limit approximation. For $n=-1$, $\mathcal{I}^{(1)}_0 = -\frac{1}{4}\alpha_0^2$, $\mathcal{I}^{(2)}_0 = 1-\alpha_0^2$; for $n=-2$, $\mathcal{I}^{(1)}_{-1} = \frac{1}{4}\alpha_0^2$, $\mathcal{I}^{(2)}_{-1} = \frac{1}{2}\alpha_0^2$; the other nth terms are the modulation of $O(\alpha_0^3)$. By adopting $\hat{\omega} \ll 1$ and by keeping up to $O(\alpha_0^2)$, Eq. (69) is reduced to be

$$\omega^2 = k^2\left(v_A^2 + c_s^2\right). \tag{70}$$

Note that the Alfvén velocity term is from the contribution of $j_{p,x}$, one-half of the acoustic velocity term is from the contribution of $j_{p,x}$, and the other one-half is from the contribution of $j_{L,x}$.

At last, we mention that the perturbed current can also be derived by taking the velocity integral in the GC coordinates, as is illustrated in Ref. [23]. In our opinion, each of these two methods has its own advantages. The method in Ref. [23] avoids the calculation of polarization vector; the manipulation is relatively simpler. Our method is more convenient to analyze the polarization effects from the polarization vector qualitatively and quantitatively. Both of these two methods reveal the fact that the polarization effects should be included in the calculation of perpendicular current. It should be noted that all the quantities in our method can be obtained from the standard modern GK theory [8]; our method can be directly used to compute the fluxes in GK simulations.

V. CONCLUSIONS AND DISCUSSIONS

In conclusion, the polarization effects in the GK theory are analyzed. The velocity transformation between two coordinate systems obeys the scalar invariance property. We first calculate the velocity transformation between the GY and GC coordinates to illustrate the key points in the procedure of velocity transformation. Then, the velocity transformation is manipulated between the GY and HY coordinates. The velocity in the HY coordinates is obtained, and is divided in terms of the polarization vector $\boldsymbol{G}_{nh}$, as is shown in Eq. (41). In the HY coordinates, we define the linear and nonlinear polarization distribution functions, which relate to $\boldsymbol{G}_{nh}$; the evolutions for the polarization distribution functions are derived. As is known, the polarization particle density should be considered in the GK calculation of particle density. Analogously, the polarization current should be considered in the GK calculation of current density. Two examples, derivations of the Hasegawa-Mima equation and the dispersion relation for GAM, are shown to address the importance of polarization current in the electrostatic case.

In the case with perpendicular magnetic potential fluctuations, the importance of polarization current is illustrated in the derivation of the dispersion relation for CAW. First, the linear polarization vector induced by the perpendicular magnetic potential is derived, from which the polarization drift is calculated; in the low-frequency limit, the polarization drift velocity is found to be the same as that from the GC theory [23]. Then, the perpendicular current is calculated; it is found that the perpendicular current consists two parts, both of which relate to the polarization effects: one part relates to the contribution of polarization distribution function, the other part relates to the contribution of polarization drift velocity. Finally, the dispersion relation for CAW is derived, as is shown in Eq. (69), which is analyzed by adopting the long-wave-limit approximation in the low-frequency case. It is found that the current from the contribution of polarization drift velocity contributes to the Alfvén velocity term and one-half of the acoustic velocity term, and the current from the contribution of polarization distribution function contributes to the other one-half of the acoustic velocity term.

At last, we mention that the nonlinear polarization effects in the distribution function and phase-space velocity are theoretically illustrated. It is demonstrated that, by keeping up to $O(\epsilon_\delta^2)$, the GK equation [Eq. (44)] combined with the definitions of $F_{h,1}$ and $F_{h,2}$ equal to the particle Vlasov equation, consideration of the nonlinear polarization effects in the GK simulation may be beneficial for the nonlinear evolution of turbulence. Since these nonlinear terms are expressed in terms of the quantities from the standard GK theory [8], they can be easily revealed in the GK simulation in principle. We note that the $\boldsymbol{G}_1\boldsymbol{G}_1$ term has been demonstrated to be beneficial to the GC orbit calculation [30]; this part has been added into the orbit module of NLT code [16,17]. Based on this experience, we will try to add the contribution of the $F_{h,2}$ term in the NLT code to numerically evaluate the influence of $F_{h,2}$ on the nonlinear behavior of turbulence in the future work.

ACKNOWLEDGMENTS

One of the authors, D. Zhang, thanks Professor S. Wang for his helpful discussion about the velocity transformation between the GC and GY coordinates. This work was supported by the National Natural Science Foundation of China under Grants No. 12005063, No. 11875254, No. 11775265, and No. GG2030040320.

APPENDIX A: DERIVATION OF EQ. (45) FROM $\partial_t F_{h,1}$

In this Appendix, we show an alternative way to obtain Eq. (45) by directly calculating $\partial_t F_{h,1}$. Note that $F_{h,1} = \nabla_6\cdot(\boldsymbol{G}_{1h}\bar{F})$, then

$$\partial_t F_{h,1} = \nabla_6\cdot(\partial_t\boldsymbol{G}_{1h}\bar{F}) + \nabla_6\cdot(\boldsymbol{G}_{1h}\partial_t\bar{F}). \tag{A1}$$

With the help of Eq. (44), we obtain

$$\nabla_6\cdot(\boldsymbol{G}_{1h}\partial_t\bar{F}) = -\nabla_6\cdot[\boldsymbol{G}_{1h}\nabla_6\cdot(\dot{\bar{Z}}\bar{F})] \tag{A2a}$$

$$= -\nabla_6\cdot[\nabla_6\cdot(\dot{\bar{Z}}\boldsymbol{G}_{1h}\bar{F}) - \dot{\bar{Z}}\bar{F}\cdot\nabla_6\boldsymbol{G}_{1h}]. \tag{A2b}$$

By using the identity $\nabla_6 \cdot \nabla_6 \cdot (\boldsymbol{F}\boldsymbol{G}) = \nabla_6 \cdot \nabla_6 \cdot (\boldsymbol{G}\boldsymbol{F})$, Eq. (A2) is rewritten as

$$\nabla_6 \cdot (\boldsymbol{G}_{1h}\partial_t \bar{F}) = -\nabla_6 \cdot [\nabla_6 \cdot (\boldsymbol{G}_{1h}\bar{F}\dot{\bar{Z}}) - \dot{\bar{Z}}\bar{F} \cdot \nabla_6 \boldsymbol{G}_{1h}] \tag{A3a}$$

$$= -\nabla_6 \cdot [F_{h,1}\dot{\bar{Z}} + \bar{F}(\boldsymbol{G}_{1h} \cdot \nabla_6 \dot{\bar{Z}} - \dot{\bar{Z}} \cdot \nabla_6 \boldsymbol{G}_{1h})] \tag{A3b}$$

$$= -\nabla_6 \cdot \left[F_{h,1}\dot{\bar{Z}} + \bar{F}(\boldsymbol{t}_{1h} - g_{1h}\dot{\bar{Z}})\right]. \tag{A3c}$$

Substituting Eq. (A3c) into Eq. (A1) yields

$$\partial_t F_{h,1} + \nabla_6 \cdot (\dot{\bar{Z}} F_{h,1} + \dot{\boldsymbol{Z}}_{h,1}\bar{F}) = 0, \tag{A4}$$

which is exactly the same as Eq. (45).

APPENDIX B: DERIVATION OF EQ. (46) FROM $\partial_t F_{h,2}$

In this Appendix, we show an alternative way to obtain Eq. (46) by directly calculating $\partial_t F_{h,2}$. Equations (46a) and (46b) can be easily illustrated following the manipulation shown in Appendix A. In the following, we show the calculation for Eq. (46c). Note that

$$F_{h,2}^{(3)} = \tfrac{1}{2}\nabla_6 \cdot \nabla_6 \cdot (\boldsymbol{G}_{1h}\boldsymbol{G}_{1h}\bar{F}) = \tfrac{1}{2}\nabla_6 \cdot (F_{h,1}\boldsymbol{G}_{1h} + \bar{F}\boldsymbol{A}_{11}),$$

where $\boldsymbol{A}_{11} = \boldsymbol{G}_{1h} \cdot \nabla_6 \boldsymbol{G}_{1h}$. In the following, we denote $F_{h,2}^{(3),1} \equiv \nabla_6 \cdot (\bar{F}\boldsymbol{A}_{11})$, $F_{h,2}^{(3),2} \equiv \nabla_6 \cdot (F_{h,1}\boldsymbol{G}_{1h})$. Similar to the derivations in Eqs. (A3a)–(A3c), it is found that

$$\nabla_6 \cdot (\boldsymbol{A}_{11}\partial_t \bar{F}) = -\nabla_6 \cdot \left[F_{h,2}^{(3),1}\dot{\bar{Z}} + \bar{F}(\boldsymbol{t}_{11} - g_{11}\dot{\bar{Z}})\right], \tag{B1}$$

where $\boldsymbol{t}_{11} = \nabla_6 \cdot (\boldsymbol{A}_{11}\dot{\bar{Z}} - \dot{\bar{Z}}\boldsymbol{A}_{11})$, $g_{11} = \nabla_6 \cdot \boldsymbol{A}_{11}$. Then

$$\partial_t F_{h,2}^{(3),1} + \nabla_6 \cdot \left[F_{h,2}^{(3),1}\dot{\bar{Z}} + \bar{F}(\boldsymbol{t}_{11} - g_{11}\dot{\bar{Z}} - \partial_t \boldsymbol{A}_{11})\right] = 0. \tag{B2}$$

Note that

$$\partial_t F_{h,2}^{(3),2} = \nabla_6 \cdot (\partial_t F_{h,1}\boldsymbol{G}_{1h}) + \nabla_6 \cdot (F_{h,1}\partial_t \boldsymbol{G}_{1h}),$$

by using Eq. (45), the first term can be calculated as

$$\begin{aligned}
&\nabla_6 \cdot (\partial_t F_{h,1}\boldsymbol{G}_{1h}) \\
&\quad = -\nabla_6 \cdot [\nabla_6 \cdot (\dot{\bar{Z}} F_{h,1} + \dot{\boldsymbol{Z}}_{h,1}\bar{F})\boldsymbol{G}_{1h}] && \text{(B3a)} \\
&\quad = -\nabla_6 \cdot [\nabla_6 \cdot (F_{h,1}\boldsymbol{G}_{1h}\dot{\bar{Z}} + \bar{F}\boldsymbol{G}_{1h}\dot{\boldsymbol{Z}}_{h,1})] \\
&\qquad + \nabla_6 \cdot (F_{h,1}\dot{\bar{Z}} \cdot \nabla_6 \boldsymbol{G}_{1h} + \bar{F}\dot{\boldsymbol{Z}}_{h,1} \cdot \nabla_6 \boldsymbol{G}_{1h}) && \text{(B3b)} \\
&\quad = -\nabla_6 \cdot [F_{h,2}^{(3),2}\dot{\bar{Z}} + F_{h,1}(\boldsymbol{t}_{1h} - g_{1h}\dot{\bar{Z}} + \dot{\boldsymbol{Z}}_{h,1})] \\
&\qquad - \nabla_6 \cdot \left\{\bar{F}[\boldsymbol{t}_2 - g_{1h}\dot{\boldsymbol{Z}}_{h,1} + (\nabla_6 \cdot \dot{\boldsymbol{Z}}_{h,1})\boldsymbol{G}_{1h}]\right\}, && \text{(B3c)}
\end{aligned}$$

with $\boldsymbol{t}_2 = \nabla_6 \cdot (\boldsymbol{G}_{1h}\dot{\boldsymbol{Z}}_{h,1} - \dot{\boldsymbol{Z}}_{h,1}\boldsymbol{G}_{1h})$. Then, we obtain

$$\begin{aligned}
&\partial_t F_{h,2}^{(3),2} + \nabla_6 \cdot \left\{F_{h,2}^{(3),2}\dot{\bar{Z}} + 2F_{h,1}\dot{\boldsymbol{Z}}_{h,1}\right. \\
&\qquad \left. - \bar{F}[\boldsymbol{t}_2 - g_{1h}\dot{\boldsymbol{Z}}_{h,1} + (\nabla_6 \cdot \dot{\boldsymbol{Z}}_{h,1})\boldsymbol{G}_{1h}]\right\} = 0.
\end{aligned} \tag{B4}$$

Combining Eqs. (B2) and (B4) yields

$$\partial_t F_{h,2}^{(3)} + \nabla_6 \cdot \left(\dot{\bar{Z}} F_{h,2}^{(3)} + \dot{\boldsymbol{Z}}_{h,1}F_{h,1} + \dot{\boldsymbol{Z}}_2\bar{F}\right) = 0, \tag{B5}$$

where

$$\begin{aligned}
\dot{\boldsymbol{Z}}_2 &= \tfrac{1}{2}\big[\boldsymbol{t}_{11} - g_{11}\dot{\bar{Z}} - \partial_t \boldsymbol{A}_{11} \\
&\qquad + \boldsymbol{t}_2 - g_{1h}\dot{\boldsymbol{Z}}_{h,1} + (\nabla_6 \cdot \dot{\boldsymbol{Z}}_{h,1})\boldsymbol{G}_{1h}\big] \\
&= \left\{-\tfrac{1}{2}\nabla_6 \cdot (g_{1h}\boldsymbol{G}_{1h} + \boldsymbol{A}_{11}) + g_{1h}^2\right\}\dot{\bar{Z}} \\
&\qquad - g_{1h}\boldsymbol{t}_{1h} + \tfrac{1}{2}\nabla_6 \cdot (\boldsymbol{G}_{1h}\boldsymbol{t}_{1h} - \boldsymbol{t}_{1h}\boldsymbol{G}_{1h}) \\
&\qquad + \tfrac{1}{2}\nabla_6 \cdot (\boldsymbol{A}_{11}\dot{\bar{Z}} - \dot{\bar{Z}}\boldsymbol{A}_{11}) - \boldsymbol{G}_{1h} \cdot \nabla_6(\partial_t \boldsymbol{G}_{1h}),
\end{aligned}$$

which is exactly the same as $\dot{\boldsymbol{Z}}_{h,2}^{(3)}$ shown in Eq. (43c). Thus, we have proved that Eq. (B5) is exactly the same as Eq. (46c).

[1] R. Balesc, *Aspects of Anomalous Transport in Plasmas* (Institute of Physics Publishing, Bristol, 2005).
[2] P. H. Diamond, S. I. Itoh, and K. Itoh, *Physical Kinetics of Turbulent Plasmas* (Cambridge University Press, New York, 2010).
[3] E. A. Frieman and L. Chen, Phys. Fluids **25**, 502 (1982).
[4] T. S. Hahm, Phys. Fluids **31**, 2670 (1988).
[5] T. S. Hahm, W. W. Lee, and A. Brizard, Phys. Fluids **31**, 1940 (1988).
[6] A. Brizard, J. Plasma Phys. **41**, 541 (1989).
[7] A. J. Brizard, Phys. Plasmas **2**, 459 (1995).
[8] A. J. Brizard and T. S. Hahm, Rev. Mod. Phys. **79**, 421 (2007).
[9] S. Wang, Phys. Rev. E **87**, 063103 (2013).
[10] Z. Lin, T. S. Hahm, W. W. Lee, W. M. Tang, and R. B. White, Science **281**, 1835 (1998).
[11] F. Jenko, W. Dorland, M. Kotschenreuther, and B. N. Rogers, Phys. Plasmas **7**, 1904 (2000).
[12] J. Candy and R. E. Waltz, J. Comput. Phys. **186**, 545 (2003).
[13] V. Grandgirard, Y. Sarazin, P. Angelino, A. Bottino, N. Crouseilles, G. Darmet, G. Dif-Pradalier, X. Garbet, P. Ghendrih, S. Jolliet *et al.*, Plasma Phys. Controlled Fusion **49**, B173 (2007).
[14] A. Peeters, Y. Camenen, F. Casson, W. Hornsby, A. Snodin, D. Strintzi, and G. Szepesi, Comput. Phys. Commun. **180**, 2650 (2009).
[15] X. Garbet, Y. Idomura, L. Villard, and T. H. Watanabe, Nucl. Fusion **50**, 043002 (2010).
[16] L. Ye, Y. Xu, X. Xiao, Z. Dai, and S. Wang, J. Comput. Phys. **316**, 180 (2016).
[17] Y. Xu, L. Ye, Z. Dai, X. Xiao, and S. Wang, Phys. Plasmas **24**, 082515 (2017).
[18] R. G. Littlejohn, Phys. Fluids **24**, 1730 (1981).
[19] R. G. Littlejohn, J. Math. Phys. **23**, 742 (1982).
[20] R. G. Littlejohn, J. Plasma Phys. **29**, 111 (1983).
[21] W. W. Lee, Phys. Fluids **26**, 556 (1983).
[22] L. Wang and T. S. Hahm, Phys. Plasmas **17**, 082304 (2010).
[23] H. Qin, W. M. Tang, W. W. Lee, and G. Rewoldt, Phys. Plasmas **6**, 1575 (1999).
[24] W. W. Lee and H. Qin, Phys. Plasmas **10**, 3196 (2003).
[25] D. Zhang, Y. Xu, Z. Dai, and S. Wang, Nucl. Fusion **60**, 046015 (2020).
[26] A. J. Brizard, Phys. Plasmas **17**, 042303 (2010).
[27] S. Wang, Phys. Plasmas **19**, 062504 (2012).

[28] Y. Xu and S. Wang, Plasma Phys. Controlled Fusion **55**, 015009 (2013).
[29] Y. Xu and S. Wang, Phys. Plasmas **20**, 012515 (2013).
[30] Y. Xu, Z. Dai, and S. Wang, Phys. Plasmas **21**, 042505 (2014).
[31] A. Hasegawa and K. Mima, Phys. Fluids **21**, 87 (1978).
[32] F. Zonca and L. Chen, Europhys. Lett. **83**, 35001 (2008).
[33] Z. Gao, K. Itoh, H. Sanuki, and J. Dong, Phys. Plasmas **15**, 072511 (2008).
[34] Z. Qiu, L. Chen, and F. Zonca, Plasma Sci. Technol. **20**, 094004 (2018).
[35] W. Huang, H. Ren, and X. Xu, Phys. Plasmas **26**, 022506 (2019).
[36] Z. Qiu, L. Chen, F. Zonca, and W. Chen, Nucl. Fusion **59**, 066031 (2019).
[37] Y. Xiao and P. J. Catto, Phys. Plasmas **13**, 102311 (2006).
[38] T. Hahm, L. Wang, W. Wang, E. Yoon, and F. Duthoit, Nucl. Fusion **53**, 072002 (2013).
[39] Z. Lu, X. Wang, P. Lauber, E. Fable, A. Bottino, W. Hornsby, T. Hayward-Schneider, F. Zonca, and C. Angioni, Plasma Phys. Controlled Fusion **61**, 044005 (2019).
[40] W. Guo, L. Wang, and G. Zhuang, Nucl. Fusion **57**, 056012 (2017).
[41] L. Chen and F. Zonca, Rev. Mod. Phys. **88**, 015008 (2016).
[42] G. Dong, J. Bao, A. Bhattacharjee, A. Brizard, Z. Lin, and P. Porazik, Phys. Plasmas **24**, 081205 (2017).
[43] J. P. Graves, D. Zullino, D. Brunetti, S. Lanthaler, and C. Wahlberg, Plasma Phys. Controlled Fusion **61**, 104003 (2019).
[44] E. D. Fredrickson, N. Gorelenkov, C. Z. Cheng, R. Bell, D. Darrow, D. Johnson, S. Kaye, B. LeBlanc, J. Menard, S. Kubota, and W. Peebles, Phys. Rev. Lett. **87**, 145001 (2001).
[45] E. V. Belova, N. N. Gorelenkov, E. D. Fredrickson, K. Tritz, and N. A. Crocker, Phys. Rev. Lett. **115**, 015001 (2015).