PHYSICS OF PLASMAS VOLUME 7, NUMBER 2 FEBRUARY 2000

Gyrokinetic field theory

H. Sugama[a)]
National Institute for Fusion Science, Toki 509-5292, Japan

(Received 2 August 1999; accepted 25 October 1999)

The Lagrangian formulation of the gyrokinetic theory is generalized in order to describe the particles' dynamics, as well as the self-consistent behavior of the electromagnetic fields. The gyrokinetic equation for the particle distribution function and the gyrokinetic Maxwell's equations, for the electromagnetic fields, are both derived from the variational principle for the Lagrangian consisting of the parts of particles, fields, and their interaction. In this generalized Lagrangian formulation, the energy conservation property for the total nonlinear gyrokinetic system of equations is directly shown from Noether's theorem. This formulation can be utilized in order to derive the nonlinear gyrokinetic system of equations and the rigorously conserved total energy for fluctuations with arbitrary frequencies. Simplified gyrokinetic systems of equations with the conserved energy are obtained from the Lagrangian with the small electron gyroradii, quasineutrality, and linear polarization–magnetization approximations. © *2000 American Institute of Physics.* [S1070-664X(00)02502-7]

I. INTRODUCTION

The gyrokinetic theory[1–9] is a basic framework to describe microinstabilities, turbulence, and resultant anomalous transport observed in magnetically confined plasmas. Basic equations for the gyrokinetic theory are the gyrokinetic equations for the particle distribution functions and Maxwell's equations for the electromagnetic fields. The gyrokinetic theory treats the fluctuations with perpendicular wavelengths on the order of the gyroradius ρ and frequencies on the order of the diamagnetic frequency $\omega_* \sim (\rho/L)\Omega$, and it employs the ratio ρ/L as the perturbation expansion parameter, where L is the equilibrium gradient scale length and Ω is the gyrofrequency.

Two types of methods to derive the gyrokinetic equation are known. One of them is the recursive technique,[1–5] which is also used for derivation of the drift kinetic equation.[10] The recursive method is combined with the ballooning representation,[11,12] and yields the gyrokinetic equation, in which the distribution function is separated into equilibrium and perturbed parts. Another modern derivation is based on the Hamiltonian and Lagrangian formulations.[6–9] The resultant gyrokinetic equation describes the total distribution function as an invariant along the particle motion. This formulation was first utilized by Littlejohn to derive the equation for the guiding center motion.[13–15] There, the motion equation is derived from the gyrophase-independent Hamiltonian, which automatically ensures the conservation of the phase space volume and the magnetic moment even in the approximate expressions obtained by truncating the perturbation expansion up to the finite order. Also, the Hamiltonian is regarded as the conserved energy for the particle in the static electromagnetic fields.

In the gyrokinetic theory, the particle Hamiltonian (or the particle energy) is not an invariant since the fluctuating electromagnetic fields are treated. Instead, the conserved quantity is the total energy of the system, which is given by the sum of the kinetic energy of the particles and the energy of the electromagnetic fields. However, the proof of the total energy conservation[16] is not trivial in the conventional formulation, where only the particle dynamics are described by the Hamiltonian or Lagrangian. Then, it seems natural that the formulation should be extended in order to derive governing equations for both the particles and the electromagnetic fields from the first principle. The purpose of the present work is to present such an extended formulation of the gyrokinetic theory.

In this paper, the gyrokinetic equation for the particle distribution function and the gyrokinetic Maxwell's equations for the electromagnetic fields are both derived from the variational principle using the Lagrangian, which consists of the parts of the particles, fields, and their interaction. This generalized Lagrangian includes the single-particle Lagrangian as a part, which has been used for the conventional Lagrangian derivation of the gyrokinetic equation. Since all the governing equations for the system are derived from the generalized Lagrangian, we can directly show the conservation of the total energy of the system with the help of Noether's theorem.[17] This seems to be the most natural and easiest way to prove the energy conservation. The Lagrangian formulation given in this work uses the technique of the classical field theory[17,18] and is closely related to several works on the variational (or action) principle for the Vlasov–Maxwell equations.[19–23] In these works, especially, Ye and Kaufman[23] treated the Vlasov–Maxwell system with high-frequency fluctuations and derived similar equations to those which are derived here. However, in Ye and Kaufman,[23] the governing equations for low-frequency gyrokinetic fluctuations are not explicitly given, and the resultant equations for high-frequency electromagnetic fluctuations are expressed in terms of the Fourier transform in space and time, which are not suitable for simulating the time evolution of the electro-

[a)]Electronic mail: sugama@nifs.ac.jp

magnetic fields. The gyrokinetic system of equations obtained here are represented in real space and time, and also their simplified versions in several limiting cases are shown in detail, which are considered to be significantly useful for numerical simulations of gyrokinetic turbulence in magnetically confined systems like tokamaks and stellarators.

The variational (or action) principle to yield the governing equations for the system considered here is written in the well-known form

$$\delta I \equiv \delta \int_{t_1}^{t_2} L\, dt = 0, \tag{1}$$

where I is called the action integral and δ represents the variation. The end points for the integral with respect to the time t are fixed to t_1 and t_2. The Lagrangian, to describe the Vlasov–Poisson–Ampère system, is written as

$$L \equiv \sum_a \int d^3\mathbf{x}_0 \int d^3\mathbf{v}_0 f_a(\mathbf{x}_0,\mathbf{v}_0,t_0) \times L_a[\mathbf{x}_a(\mathbf{x}_0,\mathbf{v}_0,t_0;t),\mathbf{v}_a(\mathbf{x}_0,\mathbf{v}_0,t_0;t), \dot{\mathbf{x}}_a(\mathbf{x}_0,\mathbf{v}_0,t_0;t)] + L_f. \tag{2}$$

Here, the single-particle Lagrangian L_a for species a is defined by

$$L_a(\mathbf{x}_a,\mathbf{v}_a,\dot{\mathbf{x}}_a) \equiv \left(m_a \mathbf{v}_a + \frac{e_a}{c}\mathbf{A}(\mathbf{x}_a,t)\right)\cdot \dot{\mathbf{x}}_a - \left(\frac{1}{2} m_a |\mathbf{v}_a|^2 + e_a \phi(\mathbf{x}_a,t)\right) \equiv \mathbf{p}_a \cdot \dot{\mathbf{x}}_a - H_a, \tag{3}$$

where $\mathbf{p}_a$ and H_a represents the canonical momentum and Hamiltonian for a single particle, respectively, and $\dot{\mathbf{x}}_a \equiv d\mathbf{x}_a/dt$. The field part L_f of the Lagrangian in Eq. (1) is defined by

$$L_f \equiv \int_V d^3\mathbf{x}\,\mathcal{L}_f \equiv \frac{1}{8\pi}\int_V d^3\mathbf{x}\left(|\nabla \phi(\mathbf{x},t)|^2 - |\nabla \times \mathbf{A}(\mathbf{x},t)|^2 + \frac{2}{c}\lambda(\mathbf{x},t)\nabla\cdot\mathbf{A}(\mathbf{x},t)\right), \tag{4}$$

which is slightly different from the one found in standard text books[17,18] in that $\partial\mathbf{A}/\partial t$ is not contained in Eq. (4). Consequently, the variational principle yields Ampère's law instead of Maxwell's equation with the displacement current. It implies that, in the present work as well as in the conventional gyrokinetic theory, we do not treat the electromagnetic waves with the speed of light. Also, $\lambda\nabla\cdot\mathbf{A}$ is included in Eq. (4) in order to derive the Coulomb (or transverse) gauge condition $\nabla\cdot\mathbf{A}=0$. The variational field λ plays the role of the Lagrange undetermined multipliers. In Eq. (1), $f_a(\mathbf{x}_0,\mathbf{v}_0,t_0)$ is the distribution function at an arbitrarily specified initial time t_0, $\mathbf{x}_a(\mathbf{x}_0,\mathbf{v}_0,t_0;t)$ and $\mathbf{v}_a(\mathbf{x}_0,\mathbf{v}_0,t_0;t)$ represents the position and velocity, respectively, of the particle at the time t, which satisfy the initial conditions

$$\mathbf{x}_a(\mathbf{x}_0,\mathbf{v}_0,t_0;t_0)=\mathbf{x}_0, \quad \mathbf{v}_a(\mathbf{x}_0,\mathbf{v}_0,t_0;t_0)=\mathbf{v}_0. \tag{5}$$

From $\delta I/\delta\mathbf{x}_a = \delta I/\delta \mathbf{v}_a = 0$, we obtain the nonrelativistic Newton's particle motion equations

$$\dot{\mathbf{x}}_a = \mathbf{v}_a, \quad m_a \dot{\mathbf{v}}_a = e_a\left[\mathbf{E}(\mathbf{x}_a,t) + \frac{1}{c}\mathbf{v}_a \times \mathbf{B}(\mathbf{x}_a,t)\right], \tag{6}$$

where $\mathbf{E} = -\nabla\phi - c^{-1}\partial\mathbf{A}/\partial t$ and $\mathbf{B}=\nabla\times\mathbf{A}$. Equation (6) determines the functional forms of $\mathbf{x}_a(\mathbf{x}_0,\mathbf{v}_0,t_0;t)$ and $\mathbf{v}_a(\mathbf{x}_0,\mathbf{v}_0,t_0;t)$. Once that they are obtained, the distribution function f_a for the time t is given by

$$f_a(\mathbf{x},\mathbf{v},t) = \int d^3\mathbf{x}_0 \int d^3\mathbf{v}_0\, \delta^3[\mathbf{x}-\mathbf{x}_a(\mathbf{x}_0,\mathbf{v}_0,t_0;t)] \times \delta^3[\mathbf{v}-\mathbf{v}_a(\mathbf{x}_0,\mathbf{v}_0,t_0;t)] f_a(\mathbf{x}_0,\mathbf{v}_0,t_0). \tag{7}$$

Then, we find from Eqs. (6) and (7) that the distribution function f_a satisfies the Vlasov equation

$$\left[\frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla + \frac{e_a}{m_a}\left\{ \mathbf{E}(\mathbf{x},t) + \frac{1}{c}\mathbf{v}\times\mathbf{B}(\mathbf{x},t)\right\}\cdot\frac{\partial}{\partial \mathbf{v}}\right] f_a(\mathbf{x},\mathbf{v},t) = 0. \tag{8}$$

The Coulomb gauge condition $\nabla\cdot\mathbf{A}=0$ is derived from $\delta I/\delta\lambda = 0$. From $\delta I/\delta\phi=0$ and $\delta I/\delta\mathbf{A}=0$, we obtain Poisson's equation

$$\nabla^2\phi(\mathbf{x},t) = -4\pi\sum_a e_a \int f_a(\mathbf{x},\mathbf{v},t)\, d^3\mathbf{v} \equiv -4\pi\sum_a e_a n_a \tag{9}$$

and

$$\nabla^2\mathbf{A}(\mathbf{x},t) - \frac{1}{c}\nabla\lambda(\mathbf{x},t) = -\frac{4\pi}{c}\sum_a e_a \int f_a(\mathbf{x},\mathbf{v},t)\mathbf{v}\, d^3\mathbf{v} \equiv -\frac{4\pi}{c}\mathbf{j}, \tag{10}$$

respectively, where n_a and $\mathbf{j}$ represent the particle number density for species a and the current density, respectively. The current density (or any vector field) can be written as $\mathbf{j} = \mathbf{j}_L + \mathbf{j}_T$, where $\mathbf{j}_L \equiv -(4\pi)^{-1}\nabla\int d^3\mathbf{x}'(\nabla'\cdot\mathbf{j})/|\mathbf{x}-\mathbf{x}'|$ and $\mathbf{j}_T \equiv (4\pi)^{-1}\nabla\times(\nabla\times\int d^3\mathbf{x}'\,\mathbf{j}/|\mathbf{x}-\mathbf{x}'|)$ represent the transverse (or solenoidal) and longitudinal (or irrotational), respectively.[24] Then, the transverse part of Eq. (10) is written as Ampère's law

$$\nabla^2\mathbf{A}(\mathbf{x},t) = -\frac{4\pi}{c}\mathbf{j}_T. \tag{11}$$

We can see that λ is unnecessary for determining the distribution function and the electromagnetic fields. Thus, Eqs. (8), (9), and (11) [instead of Eq. (10)] are regarded as the governing equations for the Vlasov–Poisson–Ampère system. Using the longitudinal part of Eq. (10), Poisson's equation [Eq. (9)], and the charge conservation law obtained from the Vlasov equation [Eq. (8)], it can be shown that $-\nabla\lambda = -4\pi\mathbf{j}_L = \partial\mathbf{E}_L/\partial t$, where $\mathbf{E}_L = -\nabla\phi$ is the longitudinal part of the electric field. Then, it is confirmed that the field equations [Eqs. (9)–(11)] are the same as those in the Darwin model.[25] It is shown from Noether's theorem (see Appendix) that the governing equations [Eqs. (8), (9), and (11)] conserve the total energy, which is written as

$$E_{\rm tot}=\sum_a \int d^3\mathbf{x}\int d^3\mathbf{v} f_a(\mathbf{x},\mathbf{v},t)\left[\frac{1}{2}m_a|\mathbf{v}|^2+e_a\phi(\mathbf{x},t)\right]-L_f$$
$$=\sum_a \int d^3\mathbf{x}\int d^3\mathbf{v} f_a(\mathbf{x},\mathbf{v},t)\frac{1}{2}m_a|\mathbf{v}|^2 + \frac{1}{8\pi}\int d^3\mathbf{x}(|\nabla\phi(\mathbf{x},t)|^2+|\nabla\times\mathbf{A}(\mathbf{x},t)|^2), \tag{12}$$

where the Poisson equation [Eq. (9)] is also used.

In Sec. II, the gyrocenter coordinates are introduced to represent the total Lagrangian in Eq. (2), from which the gyrokinetic Vlasov–Poisson–Ampère's equations are derived based on the conventional low-frequency assumption $\omega \ll \Omega$. Here ω is the characteristic fluctuation frequency and Ω is the particle gyrofrequency. Also, the gyrokinetic version of the rigorously conserved total energy is shown there.

The linear gyrokinetic theory, which can describe fluctuations with arbitrary frequencies including $\omega\sim\Omega$, was presented by Chen and Tsai[26,27] based on the recursive method. Also, recently, the Lagrangian formulation of the linear gyrokinetic theory for arbitrary-frequency fluctuations was given by H. Qin *et al.*[28–30] In Sec. III, we derive the nonlinear gyrokinetic system of equations with the rigorously conserved total energy for the fluctuations with arbitrary frequencies. A similar problem of the high-frequency ($\omega\sim\Omega$) fluctuations was also treated by Ye and Kaufman.[23] However, they treated the high-frequency fluctuations separately from the low-frequency fluctuations in the frequency-wave-number representation, while the resultant equations shown here describe the low-frequency and high-frequency fluctuations simultaneously in real space and time.

Several limiting cases, in which the gyrokinetic equations are simplified, are considered in Sec. IV. The small electron gyroradius limit, the quasineutrality, and the linear polarization-magnetization approximation are treated as examples. The simplified gyrokinetic system of equations are shown in detail, which can describe the high-frequency electrostatic plasma fluctuations (such as the ion Bernstein waves) in the uniform magnetic field.

Finally, conclusions are given in Sec. V. The Appendix gives brief explanation of the variational principle and Noether's theorem for systems including field variables.

II. LAGRANGIAN FORMULATION FOR THE GYROKINETIC VLASOV–POISSON–AMPÈRE SYSTEM

In this section, the gyrokinetic Vlasov equation, Poisson's equation, and Ampère's law are all derived by the Lagrangian formulation based on the conventional gyrokinetic low-frequency ($\omega\ll\Omega$) assumption. Thus, by applying Noether's theorem to the Lagrangian for the whole system, the total energy conservation for the gyrokinetic system can be proved more directly and easily than in the conventional gyrokinetic works, where only the gyrokinetic Vlasov equation is derived from the single-particle Lagrangian.[6–9,16]

Following Brizard's terminology,[9] we refer to the single-particle phase-space variables defined from the equilibrium and perturbed fields as the guiding-center and gyrocenter coordinates, respectively. The gyrocenter coordinates are used as independent variables of the particle distribution function in the gyrokinetic Vlasov equation. First, we consider the perturbation expansion of the single-particle Lagrangian in order to define these coordinates.

A. Perturbation expansion of the single-particle Lagrangian

In the gyrokinetic system, the electromagnetic fields and the corresponding scalar and vector potentials are assumed to consist of the equilibrium and perturbation parts,

$$\mathbf{E}=\mathbf{E}_0(\mathbf{x})+\Delta\mathbf{E}_1(\mathbf{x},t),\quad \mathbf{B}=\mathbf{B}_0(\mathbf{x})+\Delta\mathbf{B}_1(\mathbf{x},t),$$
$$\phi=\phi_0(\mathbf{x})+\Delta\phi_1(\mathbf{x},t),\quad \mathbf{A}=\mathbf{A}_0(\mathbf{x})+\Delta\mathbf{A}_1(\mathbf{x},t). \tag{13}$$

Here, Δ represents the order of the perturbation amplitude, which is used as an expansion parameter in the gyrokinetic theory. The canonical momentum of a single particle for species a is written as

$$\mathbf{p}_a\equiv m_a\mathbf{v}_a+\frac{e_a}{c}(\mathbf{A}_0+\Delta\mathbf{A}_1)\equiv m_a\mathbf{v}_{a0}+\frac{e_a}{c}\mathbf{A}_0, \tag{14}$$

where $\mathbf{A}_0$ and $\mathbf{A}_1$ are evaluated at the position $\mathbf{x}=\mathbf{x}_a$, and the zeroth-order particle velocity $\mathbf{v}_{a0}$ is defined in terms of the canonical momentum $\mathbf{p}_a$ and the zeroth-order vector field $\mathbf{A}_0$ as $\mathbf{v}_{a0}\equiv m_a^{-1}(\mathbf{p}_a-e_a\mathbf{A}_0/c)$. The (solenoidal) equilibrium current $\mathbf{j}_0$ is related to $\mathbf{A}_0$ by $\nabla^2\mathbf{A}_0=-(4\pi/c)\mathbf{j}_0$, where $\nabla\cdot\mathbf{A}_0=0$ is assumed. In the present work, we assume that the equilibrium $\mathbf{E}\times\mathbf{B}$ drift velocity is $\mathcal{O}(\epsilon v_T)$, where $\epsilon\sim\rho/L$ is the drift ordering parameter and v_T is the thermal velocity. Then, we put $\mathbf{E}_0=\phi_0=0$ and consider $\mathbf{E}_1$ and ϕ_1 to include the fluctuation part as well as the equilibrium part corresponding to the $\mathcal{O}(\epsilon v_T)$ $\mathbf{E}\times\mathbf{B}$ drift velocity. This causes no inconsistency in the results derived in this work. [The guiding-center and gyrocenter theories for the case of the $\mathcal{O}(v_T)$ $\mathbf{E}\times\mathbf{B}$ drift velocity are found in Refs. 14 and 32–36. The extension of the general Lagrangian formulation in the present work to this large $\mathbf{E}\times\mathbf{B}$ case is possible although it is not treated here for simplicity.] Here, we also neglect the induction field $-\partial\mathbf{A}_0/\partial t$, since it is $\mathcal{O}(\epsilon^2)$ according to the conventional transport ordering and its effect on the fluctuation dynamics is negligible.

Using Eqs. (13) and (14), the single-particle Lagrangian defined in Eq. (3) is rewritten as

$$L_a=L_{a0}+\Delta L_{a1}+\Delta^2 L_{a2}, \tag{15}$$

with

$$L_{a0}=\left(m_a\mathbf{v}_{a0}+\frac{e_a}{c}\mathbf{A}_0\right)\cdot\dot{\mathbf{x}}_a-\frac{1}{2}m_a|\mathbf{v}_{a0}|^2\equiv\mathbf{p}_a\cdot\dot{\mathbf{x}}_a-H_{a0}, \tag{16}$$

$$L_{a1}=-e_a\left(\phi_1-\frac{1}{c}\mathbf{v}_{a0}\cdot\mathbf{A}_1\right)\equiv-e_a\psi_a\equiv-H_{a1}, \tag{17}$$

and

$$L_{a2} = -\frac{e_a^2}{2m_a c^2}|\mathbf{A}_1|^2 \equiv -H_{a2}, \tag{18}$$

where L_{an} and H_{an} $(n=0,1,2)$ denote the nth order single-particle Lagrangian and Hamiltonian in Δ for species a, respectively. By using $\mathbf{v}_{a0}$ as the zeroth-order variable instead of $\mathbf{v}_a$, all the perturbation parts of the Lagrangian given by Eqs. (17) and (18) are confined in the Hamiltonian part, and they do not depend on $\dot{\mathbf{x}}_a$. This enables the variable transformation from the guiding-center to gyrocenter coordinates to be symplectic, as shown later. The velocity variable $\mathbf{v}_{a0}$ is not used in Brizard's formulation,[9] where the symplectic part of the Lagrangian is also perturbed. Therefore, definitions of the gyrocenter coordinates given in Sec. II C take slightly different forms from those in Brizard.[9] Velocity variables similar to $\mathbf{v}_{a0}$ are also employed by Hahm *et al.*[8] and by Ye and Kaufman.[23]

B. Guiding-center coordinates

The single-particle guiding-center coordinates $\mathbf{Z}_a = (Z_a^i)_{i=1,\ldots,6} = (\mathbf{X}_a, U_a, \mu_a, \xi_a)$ for species a are defined by taking account of the equilibrium electromagnetic fields. First, we consider the preliminary transformation

$$(\mathbf{x}_a, \mathbf{v}_{a0}) \rightarrow \mathbf{z}_a = (z_a^i)_{i=1,\ldots,6} = (\mathbf{x}_a, v_{a0\parallel}, \mu_{a0}, \theta_a). \tag{19}$$

Here, $v_{a0\parallel}$, μ_{a0}, and θ_a are defined by

$$v_{a0\parallel} = \mathbf{v}_{a0}\cdot\mathbf{b}, \quad \mu_{a0} = \frac{m_a v_{a0\perp}^2}{2B_0} \tag{20}$$

and

$$\mathbf{v}_{a0\perp} = \mathbf{v}_{a0} - v_{a0\parallel}\mathbf{b} = -v_{a0\perp}(\sin\theta_a\,\mathbf{e}_1 + \cos\theta_a\,\mathbf{e}_2), \tag{21}$$

respectively, where $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{b} \equiv \mathbf{B}_0/B_0)$ are unit vectors which form a right-handed orthogonal system at $\mathbf{x}_a$.

In order to remove the gyrophase dependence from the equilibrium part L_{a0} of the single-particle Lagrangian given by Eq. (16), we introduce the guiding-center transformation of the phase-space coordinates

$$\mathbf{z}_a = (\mathbf{x}_a, v_{a0\parallel}, \mu_{a0}, \theta_a) \rightarrow \mathbf{Z}_a = (\mathbf{X}_a, U_a, \mu_a, \xi_a), \tag{22}$$

This guiding-center transformation is the near-identity Lie transform,[14,31]

$$\begin{aligned} &\mathbf{X}_a = \mathbf{x}_a - \epsilon\boldsymbol{\rho}_{a0} + \mathcal{O}(\epsilon^2), \quad U_a = v_{a0\parallel} + \mathcal{O}(\epsilon), \\ &\mu_a = \mu_{a0} + \mathcal{O}(\epsilon), \quad \xi_a = \theta_a + \mathcal{O}(\epsilon), \end{aligned} \tag{23}$$

where $\boldsymbol{\rho}_{a0} \equiv \mathbf{b}\times\mathbf{v}_{a0}/\Omega_a$ and $\Omega_a \equiv e_a B_0/(m_a c)$. Detailed expressions for the $\mathcal{O}(\epsilon)$ and $\mathcal{O}(\epsilon^2)$ terms are found in Ref. 14. In terms of the guiding-center coordinates $(\mathbf{X}_a, U_a, \mu_a, \xi_a)$, the Lagrangian L_{a0} is written as

$$\begin{aligned} L_{a0} = {} & \epsilon^{-1}\frac{e_a}{c}\mathbf{A}_a^*(\mathbf{X}_a, U_a, \mu_a)\cdot\dot{\mathbf{X}}_a \\ & + \epsilon\frac{m_a c}{e_a}\mu_a\dot{\xi}_a - \overline{H}_{a0}(\mathbf{X}_a, U_a, \mu_a). \end{aligned} \tag{24}$$

Here, the definitions of $\overline{H}_{a0}$ and $\mathbf{A}_a^*$ are written, up to the third lowest order in ϵ, by

$$\begin{aligned} \overline{H}_{a0}(\mathbf{X}_a, U_a, \mu_a) &= \tfrac{1}{2}m_a|\mathbf{v}_{a0}(\mathbf{Z}_a)|^2 \\ &= \tfrac{1}{2}m_a U_a^2 + \mu_a B_0(\mathbf{X}_a) \end{aligned} \tag{25}$$

and

$$\begin{aligned} \mathbf{A}_a^*(\mathbf{X}_a, U_a, \mu_a) = {} & \mathbf{A}_0(\mathbf{X}_a) + \epsilon\frac{m_a c}{e_a}U_a\mathbf{b}(\mathbf{X}_a) \\ & - \epsilon^2\frac{m_a c^2}{e_a^2}\mu_a\mathbf{W}(\mathbf{X}_a), \end{aligned} \tag{26}$$

respectively, where

$$\begin{aligned} \mathbf{v}_{a0}(\mathbf{Z}_a) = {} & U_a\mathbf{b}(\mathbf{X}_a) - [2\mu_a B_0(\mathbf{X}_a)/m_a]^{1/2}[\sin\xi_a\,\mathbf{e}_1(\mathbf{X}_a) \\ & + \cos\xi_a\,\mathbf{e}_2(\mathbf{X}_a)] \end{aligned} \tag{27}$$

and

$$\begin{aligned} \mathbf{W}(\mathbf{X}_a) = {} & [\nabla\mathbf{e}_1(\mathbf{X}_a)]\cdot\mathbf{e}_2(\mathbf{X}_a) \\ & + \tfrac{1}{2}\mathbf{b}(\mathbf{X}_a)\mathbf{b}(\mathbf{X}_a)\cdot[\nabla\times\mathbf{b}(\mathbf{X}_a)]. \end{aligned} \tag{28}$$

The single-particle Lagrangian in Eq. (24) determines the symplectic structure, which is represented by the differential two-form ω, and the Hamiltonian flow in the single-particle phase space.[9,31,37] Taking the inverse of the matrix (ω_{ij}) with ω_{ij} being the components of the symplectic structure ω, the Poisson brackets for pairs of the guiding-center coordinates are obtained. Consequently, the nonvanishing Poisson brackets are given by

$$\{\mathbf{X}_a, \mathbf{X}_a\} = \epsilon\frac{c}{e_a B_{a\parallel}^*}\mathbf{b}\times\mathbf{I}, \tag{29}$$

$$\{\mathbf{X}_a, U_a\} = \frac{\mathbf{B}_a^*}{m_a B_{a\parallel}^*}, \tag{30}$$

$$\{\mathbf{X}_a, \xi_a\} = \epsilon\frac{c}{e_a B_{a\parallel}^*}\mathbf{b}\times\mathbf{W}, \tag{31}$$

$$\{U_a, \xi_a\} = -\frac{\mathbf{B}_a^*\cdot\mathbf{W}}{m_a B_{a\parallel}^*}, \tag{32}$$

$$\{\xi_a, \mu_a\} = \epsilon^{-1}\frac{e_a}{m_a c}, \tag{33}$$

where

$$\mathbf{B}_a^* \equiv \nabla\times\mathbf{A}_a^*, \quad B_{a\parallel}^* \equiv \mathbf{B}_a^*\cdot\mathbf{b}, \tag{34}$$

and $\mathbf{I} \equiv \mathbf{e}_1\mathbf{e}_1 + \mathbf{e}_2\mathbf{e}_2 + \mathbf{b}\mathbf{b}$ represents the unit dyadic. The quantities in the right-hand sides of Eqs. (29)–(32) are evaluated at $\mathbf{Z}_a = (\mathbf{X}_a, U_a, \mu_a, \xi_a)$, and $\nabla = \partial/\partial\mathbf{X}_a$ in Eq. (34). It should be noted that the Poisson bracket $\{\cdot,\cdot\}$ written here is relevant to the symplectic structure in the particle phase space only. The notation $\{\cdot,\cdot\}$ does not represent the Poisson bracket in the phase space for the total system, which requires to treat the electromagnetic fields as part of the phase-space coordinates.

We find from Eqs. (24), (25), and (29)–(33) that, in the guiding-center coordinates $\mathbf{Z}_a = (\mathbf{X}_a, U_a, \mu_a, \xi_a)$, dependence on the gyrophase ξ_a disappears from the equilibrium part of the single-particle Lagrangian L_{a0}, Hamiltonian $\overline{H}_{a0}$, and the Poisson brackets. Therefore, if there are no perturbed electromagnetic fields, the gyromotion is completely decoupled from the equations of motion, and the magnetic moment μ_a is a constant of motion. However, for the turbulent system, the gyrophase dependence appears through the perturbation part of the Lagrangian, which is removed by transformation from the guiding-center to gyrocenter coordinates, as shown in the next subsection.

C. Gyrocenter coordinates

As mentioned at the end of Sec. II. A, owing to the use of $\mathbf{v}_{a0}$, the perturbations given by Eqs. (17) and (18) change only the Hamiltonian part of the single-particle Lagrangian, although the other part (or the symplectic part) is not perturbed. As shown in Eq. (24), the symplectic part of L_{a0} has already taken a desired form in the guiding-center coordinates, which gives the gyrophase-independent Poisson brackets in Eqs. (29)–(33). Then, by the gyrocenter transformation from the guiding-center to gyrocenter coordinates

$$\mathbf{Z}_a=(\mathbf{X}_a,U_a,\mu_a,\xi_a)\rightarrow\overline{\mathbf{Z}}_a=(\overline{\mathbf{X}}_a,\overline{U}_a,\overline{\mu}_a,\overline{\xi}_a), \tag{35}$$

we remove the gyrophase dependence of the perturbed Hamiltonian without changing the symplectic structure or the form of the Poisson brackets for the guiding-center coordinates. This is done by the symplectic Lie (or canonical) transform,[31] which is associated with appropriate generating functions [see Eq. (41)]. The resultant expression for the single-particle Lagrangian in terms of the gyrocenter coordinates $\overline{\mathbf{Z}}_a=(\overline{Z}_a^i)_{i=1,\ldots,6}=(\overline{\mathbf{X}}_a,\overline{U}_a,\overline{\mu}_a,\overline{\xi}_a)$ is given by

$$\begin{aligned}L_a&=L_{a0}+\Delta L_{a1}+\Delta^2 L_{a2}\\&=\epsilon^{-1}\frac{e_a}{c}\mathbf{A}_a^*(\overline{\mathbf{X}}_a,\overline{U}_a,\overline{\mu}_a)\cdot\dot{\overline{\mathbf{X}}}_a\\&\quad+\epsilon\frac{m_a c}{e_a}\overline{\mu}_a\dot{\overline{\xi}}_a-\overline{H}_a(\overline{\mathbf{X}}_a,\overline{U}_a,\overline{\mu}_a,t),\end{aligned} \tag{36}$$

where $\mathbf{A}_a^*(\overline{\mathbf{X}}_a,\overline{U}_a,\overline{\mu}_a)$ is given by Eq. (26) with $(\mathbf{X}_a,U_a,\mu_a)$ replaced by $(\overline{\mathbf{X}}_a,\overline{U}_a,\overline{\mu}_a)$, and the gyrophase-independent Hamiltonian is written up to $\mathcal{O}(\Delta^2)$ as

$$\overline{H}_a(\overline{\mathbf{X}}_a,\overline{U}_a,\overline{\mu}_a,t)=\overline{H}_{a0}+\Delta\overline{H}_{a1}+\Delta^2\overline{H}_{a2}. \tag{37}$$

The zeroth-order Hamiltonian $\overline{H}_{a0}(\overline{\mathbf{X}}_a,\overline{U}_a,\overline{\mu}_a,t)$ is given by Eq. (25) with $(\mathbf{X}_a,U_a,\mu_a)$ replaced by $(\overline{\mathbf{X}}_a,\overline{U}_a,\overline{\mu}_a)$, and the first and second-order Hamiltonians are written as

$$\begin{aligned}&\overline{H}_{a1}(\overline{\mathbf{X}}_a,\overline{U}_a,\overline{\mu}_a,t)\\&\quad=e_a\langle\psi_a(\overline{\mathbf{Z}},t)\rangle_{\overline{\xi}_a}\\&\quad=e_a\left\langle\phi_1(\overline{\mathbf{X}}_a+\epsilon\overline{\boldsymbol{\rho}}_a,t)-\frac{1}{c}\mathbf{v}_{a0}(\overline{\mathbf{Z}}_a)\cdot\mathbf{A}_1(\overline{\mathbf{X}}_a+\epsilon\overline{\boldsymbol{\rho}}_a,t)\right\rangle_{\overline{\xi}_a},\end{aligned} \tag{38}$$

and

$$\begin{aligned}\overline{H}_{a2}(\overline{\mathbf{X}}_a,\overline{U}_a,\overline{\mu}_a,t)&=\frac{e_a^2}{2m_ac^2}\langle|\mathbf{A}_1(\overline{\mathbf{X}}_a+\epsilon\overline{\boldsymbol{\rho}}_a,t)|^2\rangle_{\overline{\xi}_a}\\&\quad-\frac{e_a}{2}\langle\{\tilde{S}_{a1}(\overline{\mathbf{Z}}_a,t),\tilde{\psi}_a(\overline{\mathbf{Z}}_a,t)\}\rangle_{\overline{\xi}_a},\end{aligned} \tag{39}$$

respectively, where $\overline{\boldsymbol{\rho}}_a=\boldsymbol{\rho}_{a0}(\overline{\mathbf{Z}}_a)=\mathbf{b}(\overline{\mathbf{X}}_a)\times\mathbf{v}_{a0}(\overline{\mathbf{Z}}_a)/\Omega_a(\overline{\mathbf{X}}_a)$. Here, the gyrophase-average and gyrophase-dependent parts of an arbitrary periodic gyrophase function $Q(\overline{\xi}_a)$ are defined by

$$\langle Q\rangle_{\overline{\xi}_a}\equiv\oint\frac{d\overline{\xi}_a}{2\pi}Q(\overline{\xi}_a)\quad\text{and}\quad\tilde{Q}\equiv Q-\langle Q\rangle_{\overline{\xi}_a}, \tag{40}$$

respectively. The Poisson brackets $\{\overline{Z^i},\overline{Z^j}\}$ for the gyrocenter coordinates have the same forms as those for the guiding-center coordinates, which are given by Eqs. (29)–(33) with $\mathbf{Z}_a=(\mathbf{X}_a,U_a,\mu_a,\xi_a)$ replaced by $\overline{\mathbf{Z}}_a=(\overline{\mathbf{X}}_a,\overline{U}_a,\overline{\mu}_a,\overline{\xi}_a)$.

The relations of the gyrocenter coordinates $\overline{\mathbf{Z}}_a=(\overline{\mathbf{X}}_a,\overline{U}_a,\overline{\mu}_a,\overline{\xi}_a)$ to the guiding-center coordinates $\mathbf{Z}_a=(\mathbf{X}_a,U_a,\mu_a,\xi_a)$ are written as

$$\overline{\mathbf{Z}}_a=\mathbf{Z}_a+\Delta\{\tilde{S}_{a1}(\mathbf{Z}_a,t),\mathbf{Z}_a\}+\mathcal{O}(\Delta^2). \tag{41}$$

Here, the first-order generating function $\tilde{S}_{a1}$ is determined as the solution of

$$\frac{\partial\tilde{S}_{a1}(\overline{\mathbf{Z}}_a,t)}{\partial t}+\{\tilde{S}_{a1}(\overline{\mathbf{Z}}_a,t),\overline{H}_{a0}(\overline{\mathbf{X}}_a,\overline{U}_a,\overline{\mu}_a,t)\}=e_a\tilde{\psi}_a(\overline{\mathbf{Z}}_a,t). \tag{42}$$

Using the conventional gyrokinetic assumption that $\Omega_a^{-1}\partial/\partial t=\mathcal{O}(\epsilon)$, and neglecting higher order terms in ϵ, Eq. (42) reduces to $\epsilon^{-1}\Omega_a\partial\tilde{S}_{a1}/\partial\overline{\xi}_a=e_a\tilde{\psi}_a$, the solution of which is given by

$$\tilde{S}_{a1}(\overline{\mathbf{Z}}_a,t)=\epsilon\frac{e_a}{\Omega_a(\overline{\mathbf{X}}_a)}\int\tilde{\psi}_a(\overline{\mathbf{Z}}_a,t)d\overline{\xi}_a, \tag{43}$$

where the integral constant is determined from the condition $\langle\tilde{S}_{a1}\rangle_{\overline{\xi}_a}=0$. [The case, in which the fluctuation frequencies are allowed to be on the order of the gyrofrequency, is considered in the next section.] Then, we find from Eqs. (41) and (43) that $\overline{\mathbf{X}}_a=\mathbf{X}_a+\mathcal{O}(\Delta\epsilon,\Delta^2)$. Following Brizard,[9] for the particle position $\mathbf{x}_a$ as the argument of the perturbation fields, we put $\mathbf{x}_a=\overline{\mathbf{X}}_a+\epsilon\overline{\boldsymbol{\rho}}_a$ by neglecting $\mathcal{O}(\Delta\epsilon,\Delta^2,\epsilon^2)$ terms. This approximation has already been used to evaluate the fluctuations ϕ_1 and $\mathbf{A}_1$ at the position $\mathbf{x}_a$ in Eqs. (38) and (39).

D. Gyrokinetic Vlasov–Poisson–Ampère system

Substituting Eq. (36) into Eq. (2), the Lagrangian for the gyrokinetic Vlasov–Poisson–Ampère system is given by

$$L=\sum_a \int d^6\bar{\mathbf{Z}}_0\, D_a(\bar{\mathbf{Z}}_0)F_a(\bar{\mathbf{Z}}_0,t_0)$$
$$\times L_a[\bar{\mathbf{Z}}_a(\bar{\mathbf{Z}}_0,t_0;t),\dot{\bar{\mathbf{Z}}}_a(\bar{\mathbf{Z}}_0,t_0;t),t]$$
$$+\frac{1}{8\pi}\int_V d^3\mathbf{x}\bigg(\Delta^2|\nabla\phi_1(\mathbf{x})|^2$$
$$-|\nabla\times[\mathbf{A}_0(\mathbf{x})+\Delta\mathbf{A}_1(\mathbf{x},t)]|^2+\Delta\frac{2}{c}\lambda(\mathbf{x},t)\nabla\cdot\mathbf{A}_1(\mathbf{x},t)\bigg), \tag{44}$$

where $\bar{\mathbf{Z}}_0\equiv(\bar{\mathbf{X}}_0,\bar{U}_0,\bar{\mu}_0,\bar{\xi}_0)$ and $\int d^6\bar{\mathbf{Z}}_0\equiv\int_V d^3\bar{\mathbf{X}}_0 \times\int_{-\infty}^{\infty}d\bar{U}_0\int_0^{\infty}d\bar{\mu}_0\int_0^{2\pi}d\bar{\xi}_0$. The functional form of the single-particle Lagrangian L_a for species a is defined by Eq. (36). Here, $D_a(\bar{\mathbf{Z}}_0)\equiv B^*_{a\parallel}(\bar{\mathbf{Z}}_0)/m_a$ is the Jacobian, $F_a(\bar{\mathbf{Z}}_0,t_0)$ denotes the distribution function for species a at an arbitrarily specified initial time t_0, and $\bar{\mathbf{Z}}_a(\bar{\mathbf{Z}}_0,t_0;t)$ represents the gyrocenter coordinates of the particle at the time t, which satisfy the initial condition

$$\bar{\mathbf{Z}}_a(\bar{\mathbf{Z}}_0,t_0;t_0)=\bar{\mathbf{Z}}_0. \tag{45}$$

Then, the distribution function $F_a(\bar{\mathbf{Z}},t)$ for the time t is determined by

$$D_a(\bar{\mathbf{Z}})F_a(\bar{\mathbf{Z}},t)=\int d^6\mathbf{Z}_0\, D_a(\bar{\mathbf{Z}}_0)F_a(\bar{\mathbf{Z}}_0,t_0)\,\delta^6[\bar{\mathbf{Z}}-\bar{\mathbf{Z}}_a(\bar{\mathbf{Z}}_0,t_0;t)], \tag{46}$$

where $\delta^6(\bar{\mathbf{Z}}-\bar{\mathbf{Z}}_a)=\delta^3(\bar{\mathbf{X}}-\bar{\mathbf{X}}_a)\,\delta(\bar{U}-\bar{U}_a)\,\delta(\bar{\mu}-\bar{\mu}_a)\,\delta[\bar{\xi}-\bar{\xi}_a(\mathrm{mod}\,2\pi)]$.

The gyrocenter motion equations are obtained from $\delta I/\delta\bar{\mathbf{Z}}_a=0$ as

$$\frac{d\bar{\mathbf{Z}}_a}{dt}=\{\bar{\mathbf{Z}}_a,\bar{H}_a(\bar{\mathbf{Z}}_a,t)\}, \tag{47}$$

which are rewritten as

$$\frac{d\bar{\mathbf{X}}_a}{dt}=\frac{1}{B^*_{a\parallel}}\Bigg[\left(\bar{U}_a+\Delta\frac{e_a}{m_a}\frac{\partial\Psi_a(\bar{\mathbf{Z}}_a)}{\partial\bar{U}_a}\right)\mathbf{B}^*_a+\epsilon c\mathbf{b}\times\left(\frac{\bar{\mu}_a}{e_a}\nabla B_0+\Delta\nabla\Psi_a(\bar{\mathbf{Z}}_a)\right)\Bigg], \tag{48}$$

$$\frac{d\bar{U}_a}{dt}=-\frac{\mathbf{B}^*_a}{m_aB^*_{a\parallel}}\cdot[\bar{\mu}_a\nabla B_0+\Delta e_a\nabla\Psi_a(\bar{\mathbf{Z}}_a)], \tag{49}$$

$$\frac{d\bar{\mu}_a}{dt}=0, \tag{50}$$

and

$$\frac{d\bar{\xi}_a}{dt}=\frac{\Omega_a}{\epsilon}+\mathbf{W}\cdot\frac{d\bar{\mathbf{X}}_a}{dt}+\frac{\Delta}{\epsilon}\frac{e_a^2}{m_ac}\frac{\partial\Psi_a(\bar{\mathbf{Z}}_a)}{\partial\bar{\mu}_a}, \tag{51}$$

where the effects of the fluctuating electromagnetic fields are included in the potential Ψ_a defined by

$$\Psi_a(\bar{\mathbf{Z}}_a)=\langle\psi_a(\bar{\mathbf{Z}},t)\rangle_{\bar{\xi}_a}+\Delta\left[\frac{e_a}{2m_ac^2}\langle|\mathbf{A}_1(\bar{\mathbf{X}}_a+\epsilon\bar{\boldsymbol{\rho}}_a,t)|^2\rangle_{\bar{\xi}_a}-\frac{1}{2}\langle\{\tilde{S}_{a1}(\bar{\mathbf{Z}}_a,t),\tilde{\psi}_a(\bar{\mathbf{Z}}_a,t)\}\rangle_{\bar{\xi}_a}\right]. \tag{52}$$

Here, the Poisson brackets in Eqs. (29)–(33) and the single-particle Hamiltonian defined by Eqs. (37)–(39) and (25) are used.

Since Eqs. (48)–(50) are independent of the gyrophase $\bar{\xi}_a$, it is easily found that $\bar{\mathbf{X}}_a(\bar{\mathbf{Z}}_0,t_0;t)$, $\bar{U}_a(\bar{\mathbf{Z}}_0,t_0;t)$, and $\bar{\mu}_a(\bar{\mathbf{Z}}_0,t_0;t)$ are independent of the initial gyrophase $\bar{\xi}_0$. The Jacobian D_a is also gyrophase-independent. Then, we find from Eq. (46) that, if F_a is initially gyrophase-independent, it is gyrophase-independent at any time. Hereafter, we assume without loss of generality that F_a is gyrophase-independent, $\partial F_a(\bar{\mathbf{Z}},t)/\partial\bar{\xi}=0$. We also obtain the gyrocenter phase-space conservation law

$$\frac{\partial}{\partial\bar{\mathbf{Z}}}\cdot[D_a(\bar{\mathbf{Z}})\{\bar{\mathbf{Z}},\bar{H}_a(\bar{\mathbf{Z}},t)\}]=0. \tag{53}$$

From Eqs. (46) and (47), we have the gyrokinetic Vlasov equation in the conservation form

$$\frac{\partial}{\partial t}[D_a(\bar{\mathbf{Z}})F_a(\bar{\mathbf{Z}},t)]+\frac{\partial}{\partial\bar{\mathbf{Z}}}\cdot[D_a(\bar{\mathbf{Z}})F_a(\bar{\mathbf{Z}},t)\{\bar{\mathbf{Z}},\bar{H}_a(\bar{\mathbf{Z}},t)\}]=0, \tag{54}$$

which is rewritten with the help of Eq. (53) in the convection form

$$\left[\frac{\partial}{\partial t}+\{\bar{\mathbf{Z}},\bar{H}_a(\bar{\mathbf{Z}},t)\}\cdot\frac{\partial}{\partial\bar{\mathbf{Z}}}\right]F_a(\bar{\mathbf{Z}},t)=0. \tag{55}$$

The Coulomb gauge condition $\nabla\cdot\mathbf{A}_1=0$ is derived from $\delta I/\delta\lambda=0$. From $\delta I/\delta\phi_1=\delta I/\delta\mathbf{A}_1=0$, the gyrokinetic Poisson's equation and the gyrokinetic Ampère's law are obtained as

$$\Delta\nabla^2\phi_1(\mathbf{x},t)=-4\pi\sum_a e_a\int d^6\bar{\mathbf{Z}}D_a(\bar{\mathbf{Z}})\times\delta^3[\bar{\mathbf{X}}+\epsilon\bar{\boldsymbol{\rho}}_{a0}(\bar{\mathbf{Z}})-\mathbf{x}][F_a(\bar{\mathbf{Z}},t)+\Delta\{\tilde{S}_{a1}(\bar{\mathbf{Z}},t),F_a(\bar{\mathbf{Z}},t)\}]\equiv-4\pi\sum_a e_an_{Ga} \tag{56}$$

and

$$\Delta\nabla^2\mathbf{A}_1(\mathbf{x},t)=-\frac{4\pi}{c}[(\mathbf{j}_G)_T(\mathbf{x},t)-\mathbf{j}_0(\mathbf{x},t)], \tag{57}$$

respectively, where $\mathbf{j}_0=-(c/4\pi)\nabla^2\mathbf{A}_0$, $\int d^6\bar{\mathbf{Z}}\equiv\int_V d^3\bar{\mathbf{X}}\int_{-\infty}^{\infty}d\bar{U}\int_0^{\infty}d\bar{\mu}\int_0^{2\pi}d\bar{\xi}$, and

$$\mathbf{j}_G \equiv \sum_a e_a \int d^6\overline{\mathbf{Z}} D_a(\overline{\mathbf{Z}}) \delta^3[\overline{\mathbf{X}} + \epsilon \overline{\boldsymbol{\rho}}_{a0}(\overline{\mathbf{Z}}) - \mathbf{x}] \times \left(\left[\mathbf{v}_{a0}(\overline{\mathbf{Z}}) - \Delta \frac{e_a}{m_a c} \mathbf{A}_1(\overline{\mathbf{X}} + \epsilon \overline{\boldsymbol{\rho}}_{a0}(\overline{\mathbf{Z}}), t) \right] F_a(\overline{\mathbf{Z}}, t) + \Delta \mathbf{v}_{a0}(\overline{\mathbf{Z}}) \{ \tilde{S}_{a1}(\overline{\mathbf{Z}}, t), F_a(\overline{\mathbf{Z}}, t) \} \right). \tag{58}$$

Here, n_{Ga} and $\mathbf{j}_G$ represent the gyrokinetic expressions of the density and current, respectively, and $(\mathbf{j}_G)_T \equiv (4\pi)^{-1}\nabla \times (\nabla \times \int d^3\mathbf{x}' \, \mathbf{j}_G / |\mathbf{x} - \mathbf{x}'|)$ represents the transverse part of $\mathbf{j}_G$. In the same way as in Eq. (11), the longitudinal part $\nabla\lambda = 4\pi(\mathbf{j}_G)_L$ is suppressed in Eq. (57) since λ is unnecessary for determining the distribution function and the electromagnetic fields. It should be noted that the distribution function F_a^{gc} in the guiding-center coordinates is related to the distribution function F_a in the gyrocenter coordinates by $F_a^{gc}(\overline{\mathbf{Z}}, t) = F_a(\overline{\mathbf{Z}}, t) + \Delta\{\tilde{S}_{a1}(\overline{\mathbf{Z}}, t), F_a(\overline{\mathbf{Z}}, t)\} + \mathcal{O}(\Delta^2)$. Thus, the right-hand sides of Eqs. (56) and (58) represent the velocity-space integrals of F_a^{gc} and $\mathbf{v}_a F_a^{gc}$, respectively, with $\mathcal{O}(\Delta^2)$ terms neglected. Then, we find that the gyrocenter motion equations [Eqs. (48)–(51)] with Eq. (52) are accurate up to $\mathcal{O}(\Delta^2)$ while the gyrokinetic Poisson–Ampère equations [Eqs. (56) and (57)] are accurate up to $\mathcal{O}(\Delta)$. This combination of unbalanced orders of accuracy is a direct result of the variational principle based on the Lagrangian (44) and is necessary for the existence of the invariant total energy. Thus, the orders of accuracy for all the governing equations are determined more systematically in the present formulation based on the Lagrangian for the whole system than in the conventional gyrokinetic theories based only on the single-particle Lagrangian.

Applying the Noether's theorem to the Lagrangian in Eq. (44) [see Eq. (A18) in the Appendix] and using Eq. (56), the conserved total energy is given by

$$\begin{aligned} E_{G\,\mathrm{tot}} &= \sum_a \int d^6\overline{\mathbf{Z}}_0 \, D_a(\overline{\mathbf{Z}}_0) F_a(\overline{\mathbf{Z}}_0, t_0) \dot{\overline{\mathbf{Z}}}_a \cdot \frac{\partial L_a(\overline{\mathbf{Z}}_a, \dot{\overline{\mathbf{Z}}}_a, t)}{\partial \dot{\overline{\mathbf{Z}}}_a} - L \\ &= \sum_a \int d^6\overline{\mathbf{Z}} \, D_a(\overline{\mathbf{Z}}) F_a(\overline{\mathbf{Z}}, t) \overline{H}_a(\overline{\mathbf{Z}}, t) - L_f \\ &= \sum_a \int d^6\overline{\mathbf{Z}} \, D_a(\overline{\mathbf{Z}}) F_a(\overline{\mathbf{Z}}, t) \left(\frac{1}{2} m_a \left[\mathbf{v}_{a0}(\overline{\mathbf{Z}}) - \Delta \frac{e_a}{m_a c} \mathbf{A}_1(\overline{\mathbf{X}} + \epsilon \overline{\boldsymbol{\rho}}_{a0}(\overline{\mathbf{Z}}), t) \right]^2 + \frac{e_a^2}{2\Omega_a(\overline{\mathbf{X}})} \Delta^2 \left[\left\{ \int (\tilde{\phi}_1)_a d\overline{\xi}, (\tilde{\phi}_1)_a \right\} \right. \right. \\ &\quad \left. \left. - \frac{1}{c^2} \left\{ \int (\widetilde{\mathbf{v}_0 \cdot \mathbf{A}_1})_a d\overline{\xi}, (\widetilde{\mathbf{v}_0 \cdot \mathbf{A}_1})_a \right\} \right] \right) + \frac{1}{8\pi} \int_V d^3\mathbf{x} (\Delta^2 |\nabla \phi(\mathbf{x}, t)|^2 + |\nabla \times [\mathbf{A}_0(\mathbf{x}) + \Delta \mathbf{A}_1(\mathbf{x}, t)]|^2), \end{aligned} \tag{59}$$

where $(\phi_1)_a = \phi_1(\overline{\mathbf{X}} + \epsilon \overline{\boldsymbol{\rho}}_{a0}(\overline{\mathbf{Z}}), t)$ and $(\mathbf{v}_0 \cdot \mathbf{A}_1)_a = \mathbf{v}_0(\overline{\mathbf{Z}}) \cdot \mathbf{A}_1(\overline{\mathbf{X}} + \epsilon \overline{\boldsymbol{\rho}}_{a0}(\overline{\mathbf{Z}}), t)$. The total energy $E_{G\,\mathrm{tot}}$ in Eq. (59) contains the $\mathcal{O}(\Delta^2)$ terms rewritten by

$$\begin{aligned} &\frac{e_a^2}{2\Omega_a(\overline{\mathbf{X}})} \Delta^2 \left[\left\{ \int (\tilde{\phi}_1)_a d\overline{\xi}, (\tilde{\phi}_1)_a \right\} - \frac{1}{c^2} \left\{ \int (\widetilde{\mathbf{v}_0 \cdot \mathbf{A}_1})_a d\overline{\xi}, (\widetilde{\mathbf{v}_0 \cdot \mathbf{A}_1})_a \right\} \right] \\ &= \frac{e_a}{2} \Delta^2 \left\{ \tilde{S}_{a1}(\overline{\mathbf{Z}}, t), (\tilde{\phi}_1)_a + \frac{1}{c} (\widetilde{\mathbf{v}_0 \cdot \mathbf{A}_1})_a \right\} \\ &= \Delta^2 \left\{ \tilde{S}_{a1}(\overline{\mathbf{Z}}, t), \frac{\Omega_a(\overline{\mathbf{X}})}{2\epsilon} \frac{\partial \tilde{S}_{a1}(\overline{\mathbf{Z}}, t)}{\partial \overline{\xi}} + \frac{e_a}{c} (\widetilde{\mathbf{v}_0 \cdot \mathbf{A}_1})_a \right\} \\ &\simeq \Delta^2 \left\{ \tilde{S}_{a1}(\overline{\mathbf{Z}}, t), \frac{e_a}{c} (\widetilde{\mathbf{v}_0 \cdot \mathbf{A}_1})_a + \frac{1}{2} \left\{ \tilde{S}_{a1}(\overline{\mathbf{Z}}, t), \frac{1}{2} m_a |\mathbf{v}_{a0}(\overline{\mathbf{Z}})|^2 \right\} \right\}, \end{aligned} \tag{60}$$

which coincide with the residual terms occurring in representing the particle kinetic energy in the gyrocenter coordinates $\overline{\mathbf{Z}}$ associated with the generating function $\tilde{S}_{a1}$.[6,7] The energy related to the ion polarization–magnetization is shown to be included in these terms. The conservation of the total energy (59) is a direct result of the Noether's theorem applied to the Lagrangian (44) while, in the conventional single-particle Lagrangian (or Hamiltonian) gyrokinetic theories, it is more troublesome to prove the energy conservation directly from the gyrokinetic Vlasov equation and the Poisson–Ampère equations.

III. GYROKINETIC THEORY FOR ARBITRARY FREQUENCIES

In this section, the gyrokinetic theory for arbitrary fluctuation frequencies is presented. The gyrokinetic system of equations derived here are applicable even for studying high-frequency fluctuations in the gyrofrequency range. Similar high-frequency Vlasov–Maxwell equations were derived by Ye and Kaufman.[23] However, they treated the low-frequency and high-frequency cases separately, and represented the fluctuating fields in the frequency-wave number space to derive the high-frequency wave equations (they did not show the low-frequency wave equations explicitly). Thus, it is difficult to obtain the conventional low-frequency gyrokinetic Maxwell (or Poisson–Ampère) equations from their high-frequency wave equations. The arbitrary-frequency gyrokinetic Vlasov–Poisson–Ampère equations obtained in this section can easily reproduce the results in the previous section in the low-frequency limit, and they are represented in real space and time, which is more suitable for simulating turbulence in magnetically confined plasmas.

For the case of arbitrary fluctuation frequencies, Eq. (43) is no longer valid, since the time derivative term in Eq. (42) can not be neglected. Then, the generating function $\widetilde{S}_{a1}$ is not determined by the fluctuation field $\widetilde{\psi}_a$ at the instant time t but takes the form of the time integral of $\widetilde{\psi}_a$. If $\widetilde{S}_{a1}$ in the Lagrangian L is regarded as the time integral of the fluctuation field, the action integral I contains the double time integral and the conventional variational principle is not applicable directly. Instead, we regard $\widetilde{S}_{a1}$ as an independent variational field and utilize the method of Lagrange undetermined multipliers to derive Eq. (42) as a result of the variational principle.

Now, let us write the total Lagrangian for the gyrokinetic Vlasov–Poisson–Ampère system with arbitrary-frequency fluctuations as

$$\begin{aligned}L=&\sum_a\int d^6\overline{\mathbf{Z}}_0\,D_a(\overline{\mathbf{Z}}_0)F_a(\overline{\mathbf{Z}}_0,t_0)L_a[\overline{\mathbf{Z}}_a(\overline{\mathbf{Z}}_0,t_0;t),\dot{\overline{\mathbf{Z}}}_a(\overline{\mathbf{Z}}_0,t_0;t),t]\\&+\frac{1}{8\pi}\int_V d^3\mathbf{x}\left(\Delta^2|\nabla\phi_1(\mathbf{x},t)|^2-|\nabla\times[\mathbf{A}_0(\mathbf{x})+\Delta\mathbf{A}_1(\mathbf{x},t)]|^2+\Delta\frac{2}{c}\lambda(\mathbf{x},t)\nabla\cdot\mathbf{A}_1(\mathbf{x},t)\right)\\&+\sum_a\int d^6\overline{\mathbf{Z}}_0\,D_a(\overline{\mathbf{Z}}_0)F_a(\overline{\mathbf{Z}}_0,t_0)\int d^6\overline{\mathbf{Z}}\Lambda_a[\overline{\mathbf{Z}};\overline{\mathbf{X}}_a(\overline{\mathbf{Z}}_0,t_0;t),\overline{U}_a(\overline{\mathbf{Z}}_0,t_0;t),\overline{\mu}_a(\overline{\mathbf{Z}}_0,t_0;t);t]\\&\times\left[\left(\frac{\partial}{\partial t}+\frac{\Omega_a(\overline{\mathbf{X}})}{\epsilon}\frac{\partial}{\partial\overline{\xi}}\right)\widetilde{S}_{a1}(\overline{\mathbf{Z}},t)-e_a\widetilde{\psi}_a(\overline{\mathbf{Z}},t)\right],\end{aligned}\tag{61}$$

where $L_a[\overline{\mathbf{Z}}_a,\dot{\overline{\mathbf{Z}}}_a,t]$ is the single-particle Lagrangian defined by Eq. (36). It should be noted that the Lagrangian in Eq. (61) contains the constraint part given by $L_c=\Sigma_a\int d^6\overline{\mathbf{Z}}_0D_a(\overline{\mathbf{Z}}_0)F_a(\overline{\mathbf{Z}}_0,t_0)\int d^6\overline{\mathbf{Z}}(\mathcal{L}_c)_a$ with $(\mathcal{L}_c)_a=\Lambda_a[\overline{\mathbf{Z}};\overline{\mathbf{X}}_a,\overline{U}_a,\overline{\mu}_a;t][(\partial/\partial t+\epsilon^{-1}\Omega_a(\overline{\mathbf{X}})\partial/\partial\overline{\xi})\widetilde{S}_{a1}(\overline{\mathbf{Z}},t)-e_a\widetilde{\psi}_a(\overline{\mathbf{Z}},t)]$. Here, $\widetilde{S}_{a1}$ and $\Lambda_a(\overline{\mathbf{Z}};\overline{\mathbf{X}}_a,\overline{U}_a,\overline{\mu}_a;t)$ are regarded as new independent variational fields for the variational principle $\delta I=\delta\int_{t_1}^{t_2}L\,dt=0$. The field Λ_a plays the role of Lagrange undetermined multipliers. Then, from $\delta I/\delta\Lambda_a=0$, the constraint on $\widetilde{S}_{a1}$ is obtained as

$$\left(\frac{\partial}{\partial t}+\frac{\Omega_a(\overline{\mathbf{X}})}{\epsilon}\frac{\partial}{\partial\overline{\xi}}\right)\widetilde{S}_{a1}(\overline{\mathbf{Z}},t)=e_a\widetilde{\psi}_a(\overline{\mathbf{Z}},t),\tag{62}$$

which corresponds to Eq. (42) with the time derivative term retained but higher ϵ order terms neglected. From $\delta I/\delta\widetilde{S}_{a1}=0$, we have

$$\begin{aligned}&\left(\frac{\partial}{\partial t}+\frac{\Omega_a(\overline{\mathbf{X}})}{\epsilon}\frac{\partial}{\partial\overline{\xi}}\right)\Lambda_a(\overline{\mathbf{Z}};\overline{\mathbf{X}}_a,\overline{U}_a,\overline{\mu}_a;t)\\&\quad=-\frac{e_a}{2}\langle\{\widetilde{\psi}_a(\overline{\mathbf{Z}}_a,t),\delta^3(\overline{\mathbf{X}}_a-\overline{\mathbf{X}})\,\delta(\overline{U}_a-\overline{U})\\&\qquad\times\delta(\overline{\mu}_a-\overline{\mu})\,\delta[\overline{\xi}_a-\overline{\xi}(\mathrm{mod}\,2\pi)]\}_{\overline{\mathbf{Z}}_a}\rangle_{\overline{\xi}_a}.\end{aligned}\tag{63}$$

We find from comparison between Eqs. (62) and (63) that Λ_a can be given by

$$\begin{aligned}&\Lambda_a(\overline{\mathbf{Z}};\overline{\mathbf{X}}_a,\overline{U}_a,\overline{\mu}_a;t)\\&\quad=-\tfrac{1}{2}\langle\{\widetilde{S}_{a1}(\overline{\mathbf{Z}}_a,t),\delta^3(\overline{\mathbf{X}}_a-\overline{\mathbf{X}})\\&\qquad\times\delta(\overline{U}_a-\overline{U})\,\delta(\overline{\mu}_a-\overline{\mu})\,\delta[\overline{\xi}_a-\overline{\xi}(\mathrm{mod}\,2\pi)]\}_{\overline{\mathbf{Z}}_a}\rangle_{\overline{\xi}_a},\end{aligned}\tag{64}$$

where $\{\cdot,\cdot\}_{\overline{\mathbf{Z}}_a}$ represents the Poisson bracket for functions of the phase-space coordinates $\overline{\mathbf{Z}}_a$ (not of $\overline{\mathbf{Z}}$).

The same form of motion equations as Eqs. (47) are obtained from $\delta I/\delta \overline{\mathbf{Z}}_a = 0$, and the same form of the gyrokinetic Vlasov equation as Eq. (55) is derived. Also, the same form of gyrokinetic Poisson's equation and Ampère's law as Eqs. (56) and (57) are derived from $\delta I/\delta \phi_1 = 0$ and $\delta I/\delta \mathbf{A}_1 = 0$, respectively, with the help of Eq. (64). Then, the gyrokinetic theory for the arbitrary-frequency fluctuations is given by the gyrokinetic Vlasov equation [Eq. (55)], the gyrokinetic Poisson's equation [Eq. (56)], the gyrokinetic Ampère's law [Eq. (57)], and the generating-function equation [Eq. (62)].

The conserved total energy derived from the total Lagrangian in Eq. (61) is written as

$$\begin{aligned} E_{G\,\mathrm{tot}} &= \sum_a \int d^6\overline{\mathbf{Z}}_0\, D_a(\overline{\mathbf{Z}}_0) F_a(\overline{\mathbf{Z}}_0, t_0) \left[\dot{\overline{\mathbf{Z}}}_a \cdot \frac{\partial L_a(\overline{\mathbf{Z}}_a, \dot{\overline{\mathbf{Z}}}_a, t)}{\partial \dot{\overline{\mathbf{Z}}}_a} + \int d^6\overline{\mathbf{Z}} \dot{\tilde{S}}_{a1} \frac{\partial (\mathcal{L}_c)_a}{\partial \dot{\tilde{S}}_{a1}} \right] - L \\ &= \sum_a \int d^6\overline{\mathbf{Z}}\, D_a(\overline{\mathbf{Z}}) F_a(\overline{\mathbf{Z}}, t) \left[\overline{H}_a(\overline{\mathbf{Z}}, t) - \frac{1}{2} \left\{ \tilde{S}_{a1}(\overline{\mathbf{Z}}, t), \frac{\partial \tilde{S}_{a1}(\overline{\mathbf{Z}}, t)}{\partial t} \right\} \right] - L_f \\ &= \sum_a \int d^6\overline{\mathbf{Z}}\, D_a(\overline{\mathbf{Z}}) F_a(\overline{\mathbf{Z}}, t) \left(\frac{1}{2} m_a \left[\mathbf{v}_{a0}(\overline{\mathbf{Z}}) - \Delta \frac{e_a}{m_a c} \mathbf{A}_1(\overline{\mathbf{X}}_a + \epsilon \overline{\boldsymbol{\rho}}_a, t) \right]^2 \right. \\ &\quad \left. + \Delta^2 \left\{ \tilde{S}_{a1}(\overline{\mathbf{Z}}, t), \frac{\Omega_a(\overline{\mathbf{X}})}{2\epsilon} \frac{\partial \tilde{S}_{a1}(\overline{\mathbf{Z}}, t)}{\partial \overline{\xi}} + \frac{e_a}{c} (\widetilde{\mathbf{v}_0 \cdot \mathbf{A}_1})_a \right\} \right) \\ &\quad + \frac{1}{8\pi} \int_V d^3\mathbf{x} (\Delta^2 |\nabla \phi_1(\mathbf{x}, t)|^2 + |\nabla \times [\mathbf{A}_0(\mathbf{x}) + \Delta \mathbf{A}_1(\mathbf{x}, t)]|^2), \end{aligned} \tag{65}$$

where Eqs. (56) and (62) are also used.

IV. LIMITING CASES

In this section, we consider several limiting cases, in which the gyrokinetic system of equations presented in the foregoing sections can be simplified. It is emphasized here that all simplifications or approximations should be done on the level of the original total Lagrangian. Once a simplified total Lagrangian is specified, a simplified gyrokinetic system of equations with the invariant total energy are straightforwardly derived from it. Therefore, in this section, we mainly show the ways of simplifying the total Lagrangian rather than the resultant simplified gyrokinetic system of equations. These simplified system of equations, which retain the rigorous energy conservation, are considered to be useful for numerical simulation of plasma turbulence and anomalous transport.

A. Neglect of W

Neglecting the $\mathcal{O}(\epsilon^2)$ term of $\mathbf{A}_a^*$ in Eq. (26) simplifies the motion equations [Eq. (48)–(51)]. This corresponds to putting $\mathbf{W} \to 0$, and gives $\{\mathbf{X}_a, \xi_a\} = \{U_a, \xi_a\} = 0$ in Eqs. (31) and (32).

In the case of the uniform equilibrium magnetic field $\mathbf{B}_0 = \mathrm{const}$, $\mathbf{W} = 0$ is rigorously obtained, and more simplifications of the motion equations are given from $\mathbf{B}_a^* = \mathbf{B}_0$ and $B_{a\|}^* = B_0$.

B. Small electron gyroradii

When the electron gyroradii are negligibly small compared to the fluctuation scale lengths, we can put

$$\boldsymbol{\rho}_e \to 0. \tag{66}$$

Then, the particle, guiding-center, and gyrocenter variables for electrons are regarded as equivalent to each other, $\mathbf{z}_e = \mathbf{Z}_e = \overline{\mathbf{Z}}_e$. The single-electron Lagrangian is given by

$$\begin{aligned} L_e = &- \frac{e}{c} \mathbf{A}_e^*(\overline{\mathbf{X}}_e, \overline{U}_e, \overline{\mu}_e) \cdot \dot{\overline{\mathbf{X}}}_e - \frac{m_e c}{e} \overline{\mu}_e \dot{\overline{\xi}}_e \\ &- \overline{H}_e(\overline{\mathbf{X}}_e, \overline{U}_e, \overline{\mu}_e), \end{aligned} \tag{67}$$

where

$$\mathbf{A}_e^*(\overline{\mathbf{X}}_e, \overline{U}_e, \overline{\mu}_e) = \mathbf{A}_0(\overline{\mathbf{X}}_e) - \frac{m_e c}{e} \overline{U}_e \mathbf{b}(\overline{\mathbf{X}}_e), \tag{68}$$

$$\begin{aligned} \overline{H}_e(\overline{\mathbf{X}}_e, \overline{U}_e, \overline{\mu}_e) = &\frac{1}{2} m_e \overline{U}_e^2 + \overline{\mu}_e B_0(\overline{\mathbf{X}}_e) \\ &- e \psi_e(\overline{\mathbf{X}}_e, \overline{U}_e, t) \\ &+ \frac{e^2}{2 m_e c^2} |\mathbf{A}_1(\overline{\mathbf{X}}_e, t)|^2, \end{aligned} \tag{69}$$

and

$$\psi_e(\overline{\mathbf{X}}_e,\overline{U}_e,t)=\phi_1(\overline{\mathbf{X}}_e,t)-\overline{U}_e A_{1\parallel}(\overline{\mathbf{X}}_e,t). \tag{70}$$

Here, the $\mathcal{O}(\epsilon^2)$ term in $\mathbf{A}_e^*$ is neglected. Here and hereafter, the drift-ordering parameter ϵ and the perturbation expansion parameter Δ are suppressed in the equations.

C. Quasineutrality

The quasineutrality approximation corresponds to putting $(1/8\pi)\int_V d^3\mathbf{x}|\nabla\phi(\mathbf{x})|^2\rightarrow 0$ in the field Lagrangian part. Then, we have

$$L_f=-\frac{1}{8\pi}\int_V d^3\mathbf{x}\Bigg(|\nabla\times[\mathbf{A}_0(\mathbf{x})+\mathbf{A}_1(\mathbf{x},t)]|^2 -\frac{2}{c}\lambda(\mathbf{x},t)\nabla\cdot\mathbf{A}_1(\mathbf{x},t)\Bigg). \tag{71}$$

Using this field Lagrangian, the left-hand-side term $\nabla^2\phi_1(\mathbf{x},t)$ in the gyrokinetic Poisson's equation [Eq. (56)] reduces to the quasineutrality condition $\Sigma_a e_a n_{Ga}=0$. Under this approximation, the electric field energy $(1/8\pi)\int_V d^3\mathbf{x}|\phi_1(\mathbf{x},t)|^2$ disappears from the total energy (59) [or Eq. (65)].

D. Linear polarization and magnetization

In this and next subsections, the distribution function $F_a(\overline{\mathbf{Z}},t)$ is assumed to be given by the sum of a time-independent equilibrium part $F_{a0}(\overline{\mathbf{Z}})$ and a small deviation from it. Here, $F_a(\overline{\mathbf{Z}},t)$ and $F_{a0}(\overline{\mathbf{Z}})$ are both independent of the gyrophase $\overline{\xi}$.

The right-hand sides of the gyrokinetic Poisson's equation [Eq. (56)] and the gyrokinetic Ampère's law (57) and (58) contain the nonlinear polarization-magnetization terms, which are given by the Poisson bracket between the generating function $\widetilde{S}_{a1}$ and the distribution function F_a. These nonlinear polarization-magnetization terms originate from the $\mathcal{O}(\Delta^2)$ terms in the Lagrangian

$$\sum_a \frac{e_a}{2}\int d^6\overline{\mathbf{Z}}_0 D_a(\overline{\mathbf{Z}}_0)F_a(\overline{\mathbf{Z}}_0,t_0) \times\langle\{\widetilde{S}_{a1}(\overline{\mathbf{Z}}_a(\overline{\mathbf{Z}}_0,t_0;t),t),\widetilde{\psi}_a(\overline{\mathbf{Z}}_a(\overline{\mathbf{Z}}_0,t_0;t),t)\}_{\overline{\mathbf{Z}}_a}\rangle_{\overline{\xi}_a}$$
$$\text{(for nonlinear polarization–magnetization)}. \tag{72}$$

The linear polarization—magnetization approximation is done by replacing the above terms in the Lagrangian (44) with

$$\sum_a \frac{e_a}{2}\int d^6\overline{\mathbf{Z}}_0 D_a(\overline{\mathbf{Z}}_0)F_{a0}(\overline{\mathbf{Z}}_0) \times\{\widetilde{S}_{a1}(\overline{\mathbf{Z}}_0,t),\widetilde{\psi}_a(\overline{\mathbf{Z}}_0,t)\}_{\overline{\mathbf{Z}}_0}$$
$$\text{(for linear polarization–magnetization)}. \tag{73}$$

In fact, from the variational principle using the Lagrangian with this replacement for the linear polarization–magnetization, the Poisson bracket terms $\{\widetilde{S}_{a1},F_{a0}\}$ appear instead of $\{\widetilde{S}_{a1},F_a\}$ as the polarization–magnetization terms in the resultant gyrokinetic Poisson's equation [Eq. (56)] and in the gyrokinetic Ampère's law (57) and (58). Also, the $\mathcal{O}(\Delta^2)$ part $-(e_a/2)\langle\{\widetilde{S}_{a1},\widetilde{\psi}_a\}\rangle_{\overline{\xi}_a}$ of the single-particle Hamiltonian is not involved in the resultant motion equation [Eq. (47)]. Thus, the terms associated with $\widetilde{S}_{a1}$ disappears from the gyrokinetic Vlasov equation [Eq. (55)]. Then, the $\mathcal{O}(\Delta^2)$ terms in Eq. (60) are connected not to F_a but to F_{a0} in the total energy (59).

The results described above are still valid when the linear polarization–magnetization approximation is applied to the the arbitrary-frequency case in Sec. III. For that case, the linear polarization–magnetization approximation is done for the Lagrangian (61) by the replacement of Eq. (72) to Eq. (73) and by replacing the constraint part L_c with

$$L_c^{\text{lpm}}=\sum_a\int d^6\overline{\mathbf{Z}}_0 D_a(\overline{\mathbf{Z}}_0)F_{a0}(\overline{\mathbf{Z}}_0) \times\int d^6\overline{\mathbf{Z}}\,\Lambda_a(\overline{\mathbf{Z}};\overline{\mathbf{X}}_0,\overline{U}_0,\overline{\mu}_0;t) \times\left[\left(\frac{\partial}{\partial t}+\Omega_a\frac{\partial}{\partial\overline{\xi}}\right)\widetilde{S}_{a1}(\overline{\mathbf{Z}},t)-e_a\widetilde{\psi}_a(\overline{\mathbf{Z}},t)\right]. \tag{74}$$

A detailed example of the linear polarization approximation is given in the next subsection for the electrostatic case.

E. High-frequency electrostatic waves in the uniform magnetic field

The approximations given in the foregoing subsections are applicable to the gyrokinetic theory for arbitrary-frequency fluctuations shown in Sec. III. In this subsection, we present a simplified gyrokinetic system of equations, which are valid even for high fluctuation frequencies in the ion-gyrofrequency range. Here, for simplicity, we consider only electrostatic fluctuations in the uniform magnetic field $\mathbf{B}_0=\text{const}$, although more general cases can be treated straightforwardly by the formulation given in Sec. III. The resultant equations can describe the ion Bernstein waves. [In fact, the rigorous linear dispersion relation for the ion Bernstein waves is immediately derived from the linearized version of Eqs. (78)–(80) and (76).[29,30]] We also take the small $\boldsymbol{\rho}_e$ limit for electrons, and use the linear polarization approximation for multispecies ions. Then, the total Lagrangian is written as

$$L=\int d^6\overline{\mathbf{Z}}_0\,D_eF_e(\overline{\mathbf{Z}}_0,t_0)\Bigg[\left(-\frac{e}{c}\mathbf{A}_0(\overline{\mathbf{X}}_e)+m_e\overline{U}_e\mathbf{b}\right)\cdot\dot{\overline{\mathbf{X}}}_e-\frac{m_ec}{e}\overline{\mu}_e\dot{\overline{\xi}}_e$$
$$-\left(\frac{1}{2}m_e\overline{U}_e^2+\overline{\mu}_eB_0-e\,\phi_1(\overline{\mathbf{X}}_e,t)\right)\Bigg]+\sum_{a(\text{ions})}\int d^6\overline{\mathbf{Z}}_0\,D_aF_a(\overline{\mathbf{Z}}_0,t_0)$$
$$\times\Bigg[\left(\frac{e_a}{c}\mathbf{A}_0(\overline{\mathbf{X}}_a)+m_a\overline{U}_a\mathbf{b}\right)\cdot\dot{\overline{\mathbf{X}}}_a+\frac{m_ac}{e_a}\overline{\mu}_a\dot{\overline{\xi}}_a-\left(\frac{1}{2}m_a\overline{U}_a^2+\overline{\mu}_aB_0+e_a\langle\phi_1(\overline{\mathbf{X}}_a+\overline{\boldsymbol{\rho}}_a,t)\rangle_{\overline{\xi}_a}\right)\Bigg]$$
$$-\sum_{a(\text{ions})}\frac{e_a^2}{2m_ac}\int d^6\overline{\mathbf{Z}}\,D_a\frac{\partial F_{a0}(\overline{\mathbf{Z}})}{\partial\mu}\frac{\partial\widetilde{S}_{a1}(\overline{\mathbf{X}},\overline{\mu},\overline{\xi},t)}{\partial\overline{\xi}}\phi_1(\overline{\mathbf{X}}+\overline{\boldsymbol{\rho}}_{a0}(\overline{\mathbf{Z}}),t)+\frac{1}{8\pi}\int_Vd^3\mathbf{x}|\nabla\phi_1(\mathbf{x})|^2$$
$$+\sum_{a(\text{ions})}\int d^6\overline{\mathbf{Z}}_0\,D_aF_{a0}(\overline{\mathbf{Z}}_0)\int d^3\overline{\mathbf{X}}d\overline{\mu}d\overline{\xi}\,\Lambda_a(\overline{\mathbf{X}},\overline{\mu},\overline{\xi};\mathbf{X}_0,\mu_0;t)$$
$$\times\left[\left(\frac{\partial}{\partial t}+\Omega_a\frac{\partial}{\partial\overline{\xi}}\right)\widetilde{S}_{a1}(\overline{\mathbf{X}},\overline{\mu},\overline{\xi},t)-e_a\widetilde{\phi}_a(\overline{\mathbf{X}}+\overline{\boldsymbol{\rho}}_{a0}(\overline{\mathbf{Z}}),t)\right],\qquad(75)$$

where $D_e=B_0/m_e$, $D_a=B_0/m_a$, $\overline{\boldsymbol{\rho}}_a=\boldsymbol{\rho}_{a0}(\overline{\mathbf{Z}}_a)$, and $\overline{\mathbf{Z}}_a=(\overline{\mathbf{X}}_a,\overline{U}_a,\overline{\mu}_a,\overline{\xi}_a)=\overline{\mathbf{Z}}_a(\overline{\mathbf{Z}}_0,t_0;t)$. Here, we have used the linear polarization part of the Lagrangian

$$-\Sigma_{a(\text{ions})}(e_a^2/2m_ac)\int d^6\overline{\mathbf{Z}}\,D_a[\partial F_{a0}(\overline{\mathbf{Z}})/\partial\mu]$$
$$\times[\partial\widetilde{S}_{a1}(\overline{\mathbf{X}},\overline{\mu},\overline{\xi},t)/\partial\overline{\xi}]\phi_1(\overline{\mathbf{X}}+\overline{\boldsymbol{\rho}}_{a0}(\overline{\mathbf{Z}}),t)$$

which is a electrostatic version of Eq. (73) with higher ϵ-order terms neglected.

From $\delta I/\delta\Lambda_a=0$, we obtain

$$\left(\frac{\partial}{\partial t}+\Omega_a\frac{\partial}{\partial\overline{\xi}}\right)\widetilde{S}_{a1}(\overline{\mathbf{X}},\overline{\mu},\overline{\xi},t)=e_a\widetilde{\phi}_1(\overline{\mathbf{X}}+\overline{\boldsymbol{\rho}}_{a0}(\overline{\mathbf{Z}}),t)\qquad(76)$$

The equation for Λ_a is derived from $\delta I/\delta\widetilde{S}_{a1}=0$, which is solved with the help of Eq. (76) to give

$$\Lambda_a(\overline{\mathbf{X}},\overline{\mu},\overline{\xi};\mathbf{X}_0,\mu_0;t)$$
$$=\frac{e_a}{4\pi m_ac}\delta^3(\overline{\mathbf{X}}-\overline{\mathbf{X}}_0)\frac{\partial\delta(\overline{\mu}-\overline{\mu}_0)}{\partial\mu}\frac{\partial\widetilde{S}_{a1}(\overline{\mathbf{X}},\overline{\mu},\overline{\xi},t)}{\partial\overline{\xi}}.\qquad(77)$$

It should be noted that, in the electrostatic case, $\widetilde{S}_{a1}$ and Λ_a are both independent of the parallel velocity, $\partial\widetilde{S}_{a1}/\partial\overline{U}=\partial\Lambda_a/\partial\overline{U}=\partial\Lambda_a/\partial\overline{U}_0=0$.

The motion equations for electrons and ions are derived from $\delta I/\delta\overline{\mathbf{Z}}_e=\delta I/\delta\overline{\mathbf{Z}}_a=0$, and the kinetic equations for the electron and ion distribution functions are given by

$$\Bigg[\frac{\partial}{\partial t}+\left(\overline{U}\mathbf{b}+\frac{c}{B_0}\mathbf{b}\times\nabla\phi_1(\overline{\mathbf{X}},t)\right)\cdot\nabla$$
$$+\frac{e}{m_e}\mathbf{b}\cdot\nabla\phi_1(\overline{\mathbf{X}},t)\frac{\partial}{\partial\overline{U}}\Bigg]F_e(\overline{\mathbf{X}},\overline{U},\overline{\mu},t)=0,\qquad(78)$$

and

$$\Bigg[\frac{\partial}{\partial t}+\left(\overline{U}\mathbf{b}+\frac{c}{B_0}\mathbf{b}\times\nabla\langle\phi_1(\overline{\mathbf{X}}+\overline{\boldsymbol{\rho}}_{a0}(\overline{\mathbf{Z}}),t)\rangle_{\overline{\xi}}\right)\cdot\nabla$$
$$-\frac{e_a}{m_a}\mathbf{b}\cdot\nabla\langle\phi_1(\overline{\mathbf{X}}+\overline{\boldsymbol{\rho}}_{a0}(\overline{\mathbf{Z}}),t)\rangle_{\overline{\xi}}\frac{\partial}{\partial\overline{U}}\Bigg]F_a(\overline{\mathbf{X}},\overline{U},\overline{\mu},t)=0,\qquad(79)$$

respectively. From $\delta I/\delta\phi_1=0$, we obtain the Poisson's equation

$$\nabla^2\phi(\mathbf{x})=4\pi e\int d^6\overline{\mathbf{Z}}\,D_e\delta^3(\overline{\mathbf{X}}-\mathbf{x})F_e(\overline{\mathbf{Z}},t)$$
$$-4\pi\sum_{a(\text{ions})}e_a\int d^6\overline{\mathbf{Z}}\,D_a\delta^3[\overline{\mathbf{X}}+\epsilon\overline{\boldsymbol{\rho}}_{a0}(\overline{\mathbf{Z}})-\mathbf{x}]$$
$$\times\left[F_a(\overline{\mathbf{Z}},t)+\frac{e_a}{m_ac}\frac{\partial F_{a0}(\overline{\mathbf{Z}})}{\partial\mu}\frac{\partial\widetilde{S}_{a1}(\overline{\mathbf{X}},\overline{\mu},\overline{\xi},t)}{\partial\overline{\xi}}\right].\qquad(80)$$

The closed nonlinear gyrokinetic system of equations (78)–(80) and (76) can describe both the low-frequency and high-frequency electrostatic plasma fluctuations in the uniform magnetic fields. They rigorously conserve the total energy, which is given by

$$E_{G\,\text{tot}}=\int d^6\overline{\mathbf{Z}}\,D_eF_e(\overline{\mathbf{Z}},t)\left(\frac{1}{2}m_e\overline{U}^2+\overline{\mu}B_0\right)$$
$$+\sum_{a(\text{ions})}\int d^6\overline{\mathbf{Z}}\,D_a\Bigg[F_a(\overline{\mathbf{Z}},t)\left(\frac{1}{2}m_a\overline{U}^2+\overline{\mu}B_0\right)$$
$$-\frac{\Omega_a^2}{2B_0}\frac{\partial F_{a0}(\overline{\mathbf{Z}})}{\partial\mu}\left(\frac{\partial\widetilde{S}_{a1}(\overline{\mathbf{X}},\overline{\mu},\overline{\xi},t)}{\partial\overline{\xi}}\right)^2\Bigg]$$
$$+\frac{1}{8\pi}\int_Vd^3\mathbf{x}|\nabla\phi(\mathbf{x})|^2.\qquad(81)$$

It should be noted that, in deriving the conservation of $E_{G\,\mathrm{tot}}$, the fluctuations on the boundary surface are assumed to make no contribution [see Eqs. (A13) and (A14) in Appendix]. If there are any external energy sources or sinks, $E_{G\mathrm{tot}}$ is not conserved. When F_{a0} is assumed to take the Maxwellian form in the velocity space perpendicular to the magnetic field, we can write $\partial F_{a0}(\overline{\mathbf{Z}})/\partial\overline{\mu} = -B_0 F_{a0}/T_a$ with the perpendicular temperature T_a.

As shown in Sec. IV C, from the variational principle using the Lagrangian (75) with the term $(1/8\pi)\int_V d^3\mathbf{x}|\nabla\phi(\mathbf{x})|^2$ neglected, the quasineutrality condition [Eq. (80) with the left-hand side term vanishing] is derived. Then, the electric field energy $(1/8\pi)\int_V d^3\mathbf{x}|\phi_1(\mathbf{x},t)|^2$ disappears from the total energy (81).

Further simplification is given by the adiabatic–electron approximation. That corresponds to the following replacement of the electron Lagrangian part in Eq. (75):

$$\int d^6\overline{\mathbf{Z}}_0\, D_e F_e(\overline{\mathbf{Z}}_0,t_0) L_e(\overline{\mathbf{Z}}_e,\dot{\overline{\mathbf{Z}}}_e,t) \rightarrow \int_V d^3\mathbf{x}\, e\,\phi_1(\mathbf{x},t) n_0(\mathbf{x})\left[1+\frac{e}{2T_e(\mathbf{x})}\phi_1(\mathbf{x},t)\right] \quad \text{(for adiabatic electrons)}, \tag{82}$$

where n_0 is the equilibrium electron density and T_e is the equilibrium electron temperature. In fact, it is easily confirmed that the variational principle for the Lagrangian using Eq. (82) makes changes in the gyrokinetic Poisson's equation [Eq. (80)] and in the conserved total energy (81), which are written as

$$\int d^6\overline{\mathbf{Z}}\, D_e \delta^3(\overline{\mathbf{X}}-\mathbf{x}) F_e(\overline{\mathbf{Z}},t) \rightarrow n_0(\mathbf{x})\left[1+\frac{e}{T_e(\mathbf{x})}\phi_1(\mathbf{x},t)\right] \quad \text{(adiabatic electron density)}, \tag{83}$$

and

$$\int d^6\overline{\mathbf{Z}}\, D_e F_e(\overline{\mathbf{Z}},t)\left(\frac{1}{2}m_e\overline{U}^2+\overline{\mu}B_0\right) \rightarrow \int_V d^3\mathbf{x}\, n_0(\mathbf{x})\frac{e^2}{2T_e(\mathbf{x})}|\phi_1(\mathbf{x},t)|^2, \tag{84}$$

respectively. In this approximation, no equation for the electron distribution function like Eq. (78) is derived or required for the closed system of equations.

The difference between the high-frequency gyrokinetic theory and the conventional low-frequency one is that, for the high-frequency case, the generating function $\tilde{S}_{a1}$ cannot be determined instantly from the fluctuation ϕ_1 due to the time derivative term retained in Eq. (76). Let us write the electrostatic potential in terms of the Fourier components with wave number vectors $\mathbf{k}$

$$\phi_1(\mathbf{x},t) = \sum_{\mathbf{k}} \phi_{\mathbf{k}}(t) e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{85}$$

Then, the solution of Eq. (76) is explicitly written as

$$\tilde{S}_{a1}(\overline{\mathbf{X}},\overline{\mu},\overline{\xi},t) = \sum_{\mathbf{k}}\sum_{n\neq 0} e^{i\mathbf{k}\cdot\overline{\mathbf{X}}} e^{in(\overline{\xi}-\alpha_{\mathbf{k}})} J_n(k_\perp\overline{\rho_a}) \times \int_{t_0}^{t} dt'\, e^{-in\Omega_a(t-t')}\phi_{\mathbf{k}}(t'), \tag{86}$$

where the initial condition $\tilde{S}_{a1}(\overline{\mathbf{X}},\overline{\mu},\overline{\xi},t_0)=0$ is used. Here, $\mathbf{k}=k_\parallel\mathbf{b}-k_\perp(\sin\alpha_{\mathbf{k}}\,\mathbf{e}_1+\cos\alpha_{\mathbf{k}}\,\mathbf{e}_2)$, $\overline{\rho_a}=(c/e_a)(2m_a\overline{\mu}/B_0)^{1/2}$, $n=\pm1,\pm2,\ldots$ $(n\neq 0)$, and J_n is the nth order Bessel function. In the low-frequency limit, $\int_{t_0}^{t} dt'\, e^{-in\Omega_a(t-t')}\phi_{\mathbf{k}}(t')$ in Eq. (86) is replaced by $(in\Omega_a)^{-1}\phi_{\mathbf{k}}(t)$, which reproduces the generating function given by Eq. (43).

Ye and Kaufman[23] transformed the guiding-center coordinates to the gyrocenter (oscillation-center in their terminology) coordinates only for nonresonant particles. That was done by using a window function in the frequency-wave-number representation to remove the resonance part from the fluctuating field in the right-hand side of the generating function equation [Eq. (76)]. In that case, since the gyrophase-dependent resonance part survives in the Hamiltonian for the motion equations, the magnetic moment is no longer an invariant for the resonant particles, and the gyrophase dependence of the distribution function must be taken into account in their Vlasov equation. In the arbitrary-frequency gyrokinetic theory given by the present work, the guiding-center to gyrocenter-coordinate transformation is done for all particles so that the gyrophase dependence does not appear in the resultant gyrokinetic Vlasov equation [Eq. (55)] where the magnetic moment is regarded as a constant parameter. As seen from Eq. (86), if the fluctuations have the resonance frequencies $\omega\simeq n\Omega_a$ $(n=\pm1,\pm2,\ldots)$, the amplitude of the generating function $\tilde{S}_{a1}$ grows, and the deviation of the gyrocenter coordinates from the guiding-center coordinates increases. For a turbulent system, where the fluctuations with the resonance frequencies exist only temporarily or randomly, the generating function given by Eq. (76) is not considered to increase divergently. Then, the use of the gyrocenter coordinates for all particles is effective. However, if the waves with the resonance frequencies are steadily produced like in the case of radio-frequency heating, the generating function diverges and the gyrocenter coordinates are not valid for the resonant particles. In such a case, the divergence at the resonance frequencies can be avoided in the real space and time representation by performing the following replacement in the generating function equation [Eq. (76)] and the Lagrangian (75)

$$\tilde{\phi}_1 \rightarrow \tilde{\phi}_1^{\text{non-res}} \equiv \tilde{\phi}_1 - \tilde{\phi}_1^{\text{res}} \equiv \tilde{\phi}_1\left[1-\left\langle w\left(\left(\frac{\partial}{\partial t}+\Omega_a\frac{\partial}{\partial\overline{\xi}}\right)\ln\tilde{\phi}_1\right)\right\rangle_{\overline{\xi}}\right], \tag{87}$$

where $\tilde{\phi}_1^{\mathrm{res}}$ and $\tilde{\phi}_1^{\mathrm{non-res}}$ represent the resonance and nonresonance parts of $\tilde{\phi}_1$, respectively. Here, w corresponds to the window function by Ye and Kaufman[23] although it is now used in the real space and time representation. The definition of w is as follows. In the region $|x|<\nu$, $w(x)\simeq 1$ and $w(0)=1$ while $w(x)\simeq 0$ for $|x|>\nu$, where $0<\nu\ll\Omega_a$. Accompanied with the replacement shown by Eq. (87), we must add the resonance potential $e_a\tilde{\phi}_1^{\mathrm{res}}$ to the single-particle Hamiltonian. Therefore, the Hamiltonian acquires the gyrophase dependence and the magnetic moment can change due to the resonance. In this case, using the Lagrangian with these changes in Eq. (75), the ion gyrokinetic equation [Eq. (79)] and the gyrokinetic Poisson's equation [Eq. (80)] are modified as well as the generating function equation (76). The treatment of the resonance shown here can be extended straightforwardly to the more general case in Sec. III.

V. CONCLUSIONS

In this work, the Lagrangian variational principle is presented to derive the nonlinear gyrokinetic Vlasov–Poisson–Ampère equations, and the rigorously conserved total energy for them is directly derived from the Noether's theorem. The nonlinear gyrokinetic system of equations and the conserved energy for the case of arbitrary fluctuation frequencies are also shown. The high-frequency properties of the fluctuations are included in the generating functions for the gyrocenter-variable transformation.

Several limiting cases are considered, in which the gyrokinetic equations are simplified and more easily tractable for numerical simulation. The small electron gyroradius limit, the quasineutrality, and the linear polarization–magnetization approximation are treated as examples. All the simplifications, which are applicable to the arbitrary fluctuation frequency case as well, are done on the level of the original Lagrangian. Then, the variational principle automatically yields the simplified gyrokinetic equations for the particles (or the distribution functions) and the fields, for which a corresponding conserved energy exists. The simplified gyrokinetic system of equations are written in detail to describe the high-frequency electrostatic plasma fluctuations in the uniform magnetic field. They are useful for studying the fluctuations in the ion-gyrofrequency range.

ACKNOWLEDGMENTS

Part of this work started during the author's visit to the Princeton Plasma Physics Laboratory as an exchange researcher of the Joint Institute for Fusion Theory, US-Japan Fusion Cooperation Program. He appreciates useful discussions with Dr. H. Qin and Dr. T. S. Hahm on the Lagrangian formulation of the gyrokinetic equations. He also thanks Dr. T.-H. Watanabe and Dr. S. Murakami for helpful discussions on the high-frequency electrostatic waves. This work is supported in part by a Grant-in-Aid from the Japanese Ministry of Education, Science, and Culture.

APPENDIX: VARIATIONAL PRINCIPLE AND NOETHER'S THEOREM

In this Appendix, the Lagrangian variational principle and the Noether's theorem are briefly explained in a partly modified way from the standard text books.[17,18] The action integral is given by

$$I=\int_{t_1}^{t_2} L\,dt. \tag{A1}$$

The total Lagrangians L considered in this work are all written in the form,

$$L=L[(\eta_\alpha),(\dot{\eta}_\alpha)], \tag{A2}$$

where the field variables η_α are functions of $(\mathbf{x}_\alpha,t)$, α is a label to specify the field, and $\cdot=\partial/\partial t$ is the time derivative. Here, $\mathbf{x}_\alpha$ denotes a l_α-dimensional vector variable, $\mathbf{x}_\alpha=(x_{\alpha 1},\ldots,x_{\alpha l_\alpha})$. When $l_\alpha=0$, η_α represents a function of the time t alone as seen in the case of the single-particle Lagrangian. The electromagnetic potential fields $\phi(\mathbf{x},t)$ and $\mathbf{A}(\mathbf{x},t)$ correspond to η_α with $l_\alpha=3$. Also, $l_\alpha=6$ is given for $\mathbf{x}_a(\mathbf{x}_0,\mathbf{v}_0,t_0;t)$ (where t_0 is a fixed parameter) in Eq. (2), and $l_\alpha=11$ for Λ_a in Eq. (61).

The Lagrangian L is a functional of the fields (or $\mathbf{x}_\alpha$-functions) η_α and $\dot{\eta}_\alpha$. We note that the part of the Lagrangian associated with η_α and $\dot{\eta}_\alpha$ for specified α are written in the form

$$L_\alpha(\eta_\alpha,\dot{\eta}_\alpha)=\int d^{l_\alpha}\mathbf{x}_\alpha \mathcal{L}_\alpha[\eta_\alpha(\mathbf{x}_\alpha,t),\dot{\eta}_\alpha(\mathbf{x}_\alpha,t),\nabla_\alpha\eta_\alpha(\mathbf{x}_\alpha,t),\ldots], \tag{A3}$$

where $\nabla_\alpha=\partial/\partial\mathbf{x}_\alpha$, and $\cdots$ represents possible dependencies on $\mathbf{x}_\alpha$ and on the other fields η_β $(\beta\neq\alpha)$. For example, in the case of Eq. (2), $\mathcal{L}_a(\mathbf{x}_a,\dot{\mathbf{x}}_a)=f_a(\mathbf{x}_0,\mathbf{v}_0,t_0)[(m_a\mathbf{v}_a-(e_a/c)\mathbf{A}(\mathbf{x}_a,t))\cdot\dot{\mathbf{x}}_a-e_a\phi(\mathbf{x}_a,t)]$ and $\mathcal{L}_\phi=(1/8\pi)|\nabla\phi(\mathbf{x},t)|^2-e_a\int d^3\mathbf{x}_0\int d^3\mathbf{v}_0 f_a(\mathbf{x}_0,\mathbf{v}_0,t_0)\phi(\mathbf{x},t)\,\delta^3(\mathbf{x}-\mathbf{x}_a)$. In this case, $\mathcal{L}_a$ and $\mathcal{L}_\phi$ share a part of each other. Thus, as shown by this example, we generally have $L\neq\Sigma_\alpha L_\alpha$.

The variational principle is written as

$$\delta I=\sum_\alpha \int_{t_1}^{t_2}dt\int d^{l_\alpha}\mathbf{x}_\alpha\left[\frac{\partial\mathcal{L}_\alpha}{\partial\eta_\alpha}-\frac{\partial}{\partial t}\left(\frac{\partial\mathcal{L}_\alpha}{\partial\dot{\eta}_\alpha}\right)-\nabla_\alpha\cdot\left(\frac{\partial\mathcal{L}_\alpha}{\partial\nabla_\alpha\eta_\alpha}\right)\right]\delta\eta_\alpha=0, \tag{A4}$$

where the variation $\delta\eta_\alpha(\mathbf{x}_\alpha,t)$ is taken to be zero at the temporal endpoints t_1 and t_2 as well as on the boundary surface of the integral $\int d^{l_\alpha}\mathbf{x}_\alpha$. We obtain from Eq. (A4) the Euler–Lagrange equations

$$\frac{\delta I}{\delta\eta_\alpha}\equiv\frac{\partial\mathcal{L}_\alpha}{\partial\eta_\alpha}-\frac{\partial}{\partial t}\left(\frac{\partial\mathcal{L}_\alpha}{\partial\dot{\eta}_\alpha}\right)-\nabla_\alpha\cdot\left(\frac{\partial\mathcal{L}_\alpha}{\partial\nabla_\alpha\eta_\alpha}\right)=0. \tag{A5}$$

Next, let us consider the following infinitesimal transformations of t, $\mathbf{x}_\alpha$, and $\eta_\alpha(\mathbf{x}_\alpha,t)$ simultaneously:

$$t \to t' = t + \delta t,$$

$$\mathbf{x}_\alpha \to \mathbf{x}'_\alpha = \mathbf{x}_\alpha + \delta \mathbf{x}_\alpha, \tag{A6}$$

$$\eta_\alpha(\mathbf{x}_\alpha, t) \to \eta'_\alpha(\mathbf{x}'_\alpha, t') = \eta_\alpha(\mathbf{x}_\alpha, t) + \delta\eta_\alpha(\mathbf{x}_\alpha, t).$$

Here, δt and $\delta \mathbf{x}_\alpha$ are generally functions of $(\mathbf{x}_\alpha, t)$, and $\delta\eta_\alpha(\mathbf{x}_\alpha, t)$ consists of the variations in the functional form of η_α and in the variables $(\mathbf{x}_\alpha, t)$,

$$\delta\eta_\alpha(\mathbf{x}_\alpha, t) = \overline{\delta}\eta_\alpha(\mathbf{x}_\alpha, t) + \delta t\, \dot{\eta}_\alpha + \delta\mathbf{x}_\alpha \cdot \nabla_\alpha \eta_\alpha, \tag{A7}$$

where $\mathcal{O}(\delta^2)$ terms are neglected, and $\overline{\delta}\eta_\alpha(\mathbf{x}_\alpha, t) = \eta'_\alpha(\mathbf{x}_\alpha, t) - \eta_\alpha(\mathbf{x}_\alpha, t)$. The infinitesimal transformations in Eq. (A6) also causes the variation in the action integral

$$I \to I' = \int_{t'_1}^{t'_2} L'\, dt', \tag{A8}$$

where, as in Eq. (A3), the part of L' associated with η'_α and $\dot{\eta}'_\alpha = \partial\eta'_\alpha/\partial t'$ for specified α is given by

$$L'_\alpha = \int d^{l_\alpha}\mathbf{x}'_\alpha \mathcal{L}_\alpha[\eta'_\alpha(\mathbf{x}'_\alpha, t'), \dot{\eta}'_\alpha(\mathbf{x}'_\alpha, t'), \nabla'_\alpha \eta'_\alpha(\mathbf{x}'_\alpha, t'), \ldots]. \tag{A9}$$

Using the Euler–Lagrange equation [Eq. (A5)], the variation in the action integral under the transformations in Eq. (A6) is written as

$$\delta I = I' - I = -\int_{t_1}^{t_2} dt \left[\frac{dG}{dt} + \sum_\alpha \int d^{l_\alpha}\mathbf{x}_\alpha \nabla_\alpha \cdot \mathbf{J}_\alpha\right], \tag{A10}$$

where

$$G = \delta t \left(\sum_\alpha \int d^{l_\alpha}\mathbf{x}_\alpha\, \dot{\eta}_\alpha \frac{\partial \mathcal{L}_\alpha}{\partial \dot{\eta}_\alpha} - L\right) + \sum_\alpha \int d^{l_\alpha}\mathbf{x}_\alpha \left(\delta\mathbf{x}_\alpha \cdot \nabla_\alpha \eta_\alpha \frac{\partial \mathcal{L}_\alpha}{\partial \dot{\eta}_\alpha} - \delta\eta_\alpha \frac{\partial \mathcal{L}_\alpha}{\partial \dot{\eta}_\alpha}\right) \tag{A11}$$

and

$$\mathbf{J}_\alpha = \delta t\, \dot{\eta}_\alpha \frac{\partial \mathcal{L}_\alpha}{\partial \nabla_\alpha \eta_\alpha} + \delta\mathbf{x}_\alpha \cdot \nabla_\alpha \eta_\alpha \frac{\partial \mathcal{L}_\alpha}{\partial \nabla_\alpha \eta_\alpha} - \delta\mathbf{x}_\alpha \mathcal{L}_\alpha - \delta\eta_\alpha \frac{\partial \mathcal{L}_\alpha}{\partial \nabla_\alpha \eta_\alpha}. \tag{A12}$$

When the action integral I is invariant under the transformations in Eq. (A6), we obtain from Eq. (A10) with arbitrariness of t_1 and t_2

$$\frac{dG}{dt} + \sum_\alpha \int d^{l_\alpha}\mathbf{x}_\alpha \nabla_\alpha \cdot \mathbf{J}_\alpha = 0, \tag{A13}$$

which is the main conclusion of the Noether's theorem. If $\mathbf{J}_\alpha$ vanish on the boundaries of the integral regions $\int d^{l_\alpha}\mathbf{x}_\alpha$, G is conserved

$$\frac{dG}{dt} = 0. \tag{A14}$$

The Noether's theorem is widely applicable to derivation of the conservation laws. For example, when $\mathcal{L}_\alpha$ is independent of η_α for $\alpha = \overline{\alpha}$, the action integral I is obviously invariant under the transform given by

$$\delta t = 0, \quad \delta\mathbf{x}_\alpha = 0, \quad \delta\eta_\alpha = \epsilon\, \delta_{\alpha\overline{\alpha}}, \tag{A15}$$

where ϵ is an infinitesimal constant parameter. Then, we find from Eqs. (A11) and (A14) that

$$\int d^{l_{\overline{\alpha}}}\mathbf{x}_{\overline{\alpha}} \frac{\partial \mathcal{L}_{\overline{\alpha}}}{\partial \dot{\eta}_{\overline{\alpha}}} = \text{const.} \tag{A16}$$

The conservation of the magnetic moment for the gyrophase-independent Lagrangian is regarded as a special case of this example.

The total Lagrangians considered in this work have no explicit time dependence, which means that their time dependencies are only through the functions $\eta_\alpha(\mathbf{x}_\alpha, t)$. Thus, the action integral I is invariant under the infinitesimal transformation given by

$$\delta t = \epsilon, \quad \delta\mathbf{x}_\alpha = 0, \quad \delta\eta_\alpha = 0. \tag{A17}$$

Then, from Eqs. (A11) and (A14) we immediately obtain the total energy conservation

$$E_{\text{tot}} \equiv \sum_\alpha \int d^{l_\alpha}\mathbf{x}_\alpha\, \dot{\eta}_\alpha \frac{\partial \mathcal{L}_\alpha}{\partial \dot{\eta}_\alpha} - L = \text{const.} \tag{A18}$$

[1] P. H. Rutherford and E. A. Frieman, Phys. Fluids **11**, 569 (1968).
[2] J. B. Taylor and R. J. Hastie, Phys. Plasmas **10**, 479 (1968).
[3] T. M. Antonsen, Jr. and B. Lane, Phys. Fluids **23**, 1205 (1980).
[4] P. J. Catto, W. M. Tang, and D. E. Baldwin, Phys. Plasmas **23**, 639 (1981).
[5] E. A. Frieman and L. Chen, Phys. Fluids **25**, 502 (1982).
[6] D. H. E. Dubin, J. A. Krommes, C. Oberman, and W. W. Lee, Phys. Fluids **26**, 3524 (1983).
[7] T. S. Hahm, W. W. Lee, and A. Brizard, Phys. Fluids **31**, 1940 (1988).
[8] T. S. Hahm, Phys. Fluids **31**, 2670 (1988).
[9] A. J. Brizard, J. Plasma Phys. **41**, 541 (1989).
[10] R. D. Hazeltine, Phys. Plasmas **15**, 77 (1973).
[11] R. D. Hezeltine and J. D. Meiss, *Plasma Confinement* (Addison-Wesley, Redwood City, California, 1992), p. 298.
[12] R. L. Dewar and A. H. Glasser, Phys. Fluids **26**, 3038 (1983).
[13] R. G. Littlejohn, J. Math. Phys. **20**, 2445 (1979).
[14] R. G. Littlejohn, Phys. Fluids **24**, 1730 (1981).
[15] R. G. Littlejohn, J. Plasma Phys. **29**, 111 (1983).
[16] A. J. Brizard, Phys. Rev. B **1**, 1381 (1989).
[17] H. Goldstein, *Classical Mechanics*, 2nd ed. (Addison-Wesley, Reading, Massachusetts, 1980), Chap. 12.
[18] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, 4th ed. (Pergamon Oxford, 1975), Chap. 4.
[19] F. E. Low, Proc. R. Soc. London, Ser. A **248**, 282 (1958).
[20] P. L. Similon, Phys. Lett. A **112**, 33 (1985).
[21] D. Pfirsch and P. J. Morrison, Phys. Rev. **32**, 1714 (1985).
[22] H. Ye and P. J. Morrison, Phys. Fluids B **4**, 771 (1992).
[23] H. Ye and A. N. Kaufman, Phys. Fluids B **4**, 1735 (1992).
[24] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1998), Sec. 6.3.
[25] A. N. Kaufman and P. S. Rostler, Phys. Fluids **14**, 446 (1971).
[26] L. Chen and S. T. Tsai, Phys. Fluids **26**, 141 (1983).
[27] L. Chen and S. T. Tsai, Phys. Plasmas **25**, 349 (1983).
[28] H. Qin, W. M. Tang, and G. Rewoldt, Phys. Plasmas **5**, 1035 (1998).
[29] H. Qin, Ph. D. thesis, Princeton University, 1998.
[30] H. Qin, W. M. Tang, W. W. Lee, and G. Rewoldt, Phys. Plasmas **6**, 1575 (1999).

[31] R. G. Littlejohn, J. Math. Phys. **23**, 742 (1982).
[32] M. Artun and W. M. Tang, Phys. Fluids B **4**, 1102 (1992).
[33] M. Artun and W. M. Tang, Phys. Plasmas **1**, 2682 (1994).
[34] A. J. Brizard, Phys. Plasmas **2**, 459 (1995).
[35] T. S. Hahm, Phys. Plasmas **3**, 4658 (1996).
[36] H. Sugama and W. Horton, Phys. Plasmas **5**, 2560 (1998).
[37] V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer, New York, 1978), Chap. 8.