Local momentum balance in electromagnetic gyrokinetic systems

H. Sugama[1, 2, a)]
[1)] National Institute for Fusion Science, Toki 509-5292, Japan
[2)] Department of Advanced Energy, University of Tokyo, Kashiwa 277-8561, Japan

(Dated: 25 March 2024)

The Eulerian variational formulation is presented to obtain governing equations of the electromagnetic turbulent gyrokinetic system. A local momentum balance in the system is derived from the invariance of the Lagrangian of the system under an arbitrary spatial coordinate transformation by extending the previous work [H. Sugama *et al*., Phys. Plasmas **28**, 022312 (2021)]. Polarization and magnetization due to finite gyroradii and electromagnetic microturbulence are correctly described by the gyrokinetic Poisson equation and Ampère's law which are derived from the variational principle. Also shown is how the momentum balance is influenced by including collisions and external sources. Momentum transport due to collisions and turbulence is represented by a symmetric pressure tensor which originates in a variational derivative of the Lagrangian with respect to the metric tensor. The relations of the axisymmetry and quasi-axisymmetry of the toroidal background magnetic field to a conservation form of the local momentum balance equation are clarified. In addition, an ensemble-averaged total momentum balance equation is shown to take the conservation form even in the background field with no symmetry when a constraint condition representing the macroscopic Ampère's law is imposed on the background field. Using the WKB representation, the ensemble-averaged pressure tensor due to the microturbulence is expressed in detail and it is verified to reproduce the toroidal momentum transport derived in previous works for axisymmetric systems. The local momentum balance equation and the pressure tensor obtained in this work present a useful reference for elaborate gyrokinetic simulation studies of momentum transport processes.

I. INTRODUCTION

Gyrokinetics[1–7] is a powerful theoretical framework based on which a large number of analytical and numerical studies on microinstabilities and turbulent processes in magnetized plasmas[8] have been done. The original (or classical) gyrokinetic theory[1–3,9] adopts the WKB approximation (or ballooning representation)[10] and treats the perturbed parts of particle distribution functions and electromagnetic fields with gyroradius-scale perpendicular wavelengths. This type of gyrokinetic theory is widely employed as the basic model for local flux-tube gyrokinetic simulations[11–15] to evaluate turbulent particle and heat fluxes. The other type of (or modern) gyrokinetic theory uses the Lie transformation method[16] to obtain gyrocenter coordinates which obey the Lagrangian and/or Hamiltonian dynamics derived from the variational formulation.[4,5] The modern theory guarantees favorable conservation properties[17–22] of gyrokinetic equations for total distribution functions (including both background and fluctuation parts), which are generally used for long-time global gyrokinetic simulations.[23–30] It is also noted here that classical gyrokinetic equations are shown to be consistently derived from modern ones by properly taking account of different phase-space coordinate systems used in the two type of theories.[31]

Over the years, momentum transport processes have been attracting much attention because they determine profiles of plasma flows such as background plasma rotations and $\mathbf{E}\times\mathbf{B}$ zonal flows, which are regarded as important factors for stabilizing or regulating instabilities and improving plasma confinement.[32] Also, there are a lot of activities in designing advanced magnetic configurations such as toroidal systems with quasi-symmetry,[33–36] in which reduction of neoclassical transport and increase of plasma flows are expected. In the present paper, local momentum balance equations which describe the momentum transport processes in electromagnetic gyrokinetic turbulence are derived by extending the previous work[37] on the momentum balance in electrostatic turbulence based on the Eulerian variational formulation,[38] which is also called the Euler-Poincaré reduction procedure.[39–44]

In conventional studies, momentum balance equations are obtained by taking the first-order velocity-space moment of a kinetic equation or from a variation in an action integral of the Lagrangian under infinitesimal translation or rotation. In the former derivation, it is unclear how the momentum balance in the direction perpendicular to the background magnetic field can be obtained from the gyrokinetic equation in the gyrophase-averaged form. In the latter, Noether's theorem can be applied to connect the symmetry condition of the system directly with the canonical momentum conservation equation,[45,46] in which, however, local momentum transport is represented by the asymmetric canonical pressure tensor because of the vector potential included in the canonical momentum. In this work as well as in the previous works,[37,44] the invariance of the Lagrangian under arbitrary infinitesimal transformations of general spatial coordinates is used to derive the local momentum balance equation which contains the symmetric pressure tensor obtained by taking the variational derivatives of the Lagrangian with respect to the 3×3 metric tensor components. This is analogous to the derivation of energy-momentum conservation laws from the invariance of an action integral under arbitrary transformations of spatiotemporal coordinates in the theory of general relativity.[47] The relations of the symmetry and quasi-symmetry properties of the background magnetic field to the momentum balance equation are investigated with the help of the symmetric pressure tensor. In addition, the effects of collisions and/or external momentum sources can be easily included in the local momentum balance equation, by which both collisional[48–50]

[a)] Corresponding author: sugama.hideo@nifs.ac.jp

and turbulent transport processes are described.

In extending the previous study for electrostatic turbulence[37] to the case of electromagnetic turbulence, one need to consider the average and fluctuating parts of the magnetic field and accordingly those parts of the magnetic potential. Then, as shown in Ref.[31] and this paper, the variational derivative of the gyrokinetic Lagrangian with respect to the fluctuating part of the vector potential is used to correctly represent both the average and fluctuating parts of the local particle flux and the current density which appears in Ampère's law. On the other hand, the variational derivative of the gyrokinetic Lagrangian with respect to the average part of the vector potential also takes a form similar to the particle flux and it appears in the momentum balance equation derived using the variational technique in the present study. Comparison between the above-mentioned two types of particle fluxes shows that their average parts coincides with each other to the leading order in the small gyroradius expansion although their fluctuating parts do not. It is also shown in the present work that, using the ensemble average of the latter particle flux to self-consistently determine the average part of the magnetic field from Ampère's law, the conservation form of the ensemble-averaged local momentum balance equation can be derived.

Currently, large-scale gyrokinetic simulations such as global ones solving phenomena from a device size to an ion gyroradius scale and cross-scale flux-tube simulations treating interactions between both ion and electron gyroradius scales are actively conducted. Huge simulations including all scales from the machine size to the electron gyroradius remain a challenging future task, for which global simulations need to treat full finite ion gyroradius effects at least as done in flux-tube simulations. In principle, the gyrokinetic model and the momentum balance equation presented in the present work contain all scales ranging from macroscopic equilibrium gradient lengths to microscopic turbulence wavelengths of the order of the electron gyroradius. The macroscopic behaviors of the momentum transport processes are described by the ensemble-averaged momentum balance equation, which is shown to take the conservation form under a condition to adjust the background field to the macroscopic Ampère's law. Furthermore, the WKB representation is used to explicitly express the full gyroradius effects of the electromagnetic turbulence on the symmetric pressure tensor, the ijth component of which represents the turbulent transport of the ith momentum component in the jth direction. These expressions can be applied to evaluation of the local momentum transport by the flux-tube simulations.

The rest of this paper is organized as follows. In Sec. II, equations of the gyrocenter motion in turbulent electromagnetic fields are derived as the Euler-Lagrange equations from the Lagrangian given as a function of the gyrocenter coordinates. In Sec. III, the Lagrangian for the whole gyrokinetic system consisting of particles of all species and electromagnetic fields is presented to derive gyrokinetic Vlasov equations for gyrocenter distribution functions and the gyrokinetic Poisson and Ampère equations for electrostatic and vector potentials based on the Eulerian (or Euler-Poincaré) variational formulation. Then, the gyrokinetic and field parts of the Lagrangian are all represented in terms of general spatial coordinates in Sec. IV and the invariance of the Lagrangian under an arbitrary infinitesimal transformation of spatial coordinates is used to derive the momentum balance equations for a single-particle-species system and for a system including all particle species and electromagnetic fields in Sec. V. In Sec. VI, axisymmetric, non-axisymmetric, and quasi-axisymmetric toroidal systems are investigated from the viewpoint of momentum balance, and Sec. VII presents the ensemble-averaged momentum balance equation, which is shown to take the conservation form when the background field is determined by the condition representing the macroscopic Ampère's law. The ensemble-averaged pressure tensor caused by the electromagnetic turbulence is expressed in detail using the WKB representation in Sec. VIII. Finally, conclusions are given in Sec. IX. In Appendix A, the potential field included in the gyrocenter Hamiltonian is represented by gyroradius expansion around the gyrocenter, which is used in Appendix B to expand the electromagnetic interaction part of Lagrangian density in terms of the electrostatic and vector potentials and their derivatives. In the same way as in Appendix B, charge and current densities are expanded in Appendices C and D, respectively, where the polarization and magnetization parts are identified. Energy balance equations in electromagnetic gyrokinetic turbulence are presented in Appendix E.

II. EQUATIONS OF GYROCENTER MOTION IN TURBULENT ELECTROMAGNETIC FIELDS

The Lagrangian for describing the gyrocenter motion of the charged particle is given by[4,5,31]

$$L_{GYa}(\mathbf{Z},\dot{\mathbf{Z}},t) \equiv \frac{e_a}{c}\mathbf{A}^*_a(\mathbf{X},U,t)\cdot\dot{\mathbf{X}} + \frac{m_a c}{e_a}\mu\dot{\vartheta} - H_{GYa}(\mathbf{Z},t), \tag{1}$$

where the modified vector potential $\mathbf{A}^*_a$ is defined by $\mathbf{A}^*_a(\mathbf{X},U,t) \equiv \mathbf{A}(\mathbf{X},t) + (m_a c/e_a)U\mathbf{b}(\mathbf{X},t)$, the subscript a represents the particle species with mass m_a and charge e_a, and $\dot{} \equiv d/dt$ represents the time derivative along the motion of the particle in phase space. The gyrocenter phase-space coordinates $\mathbf{X}$, U, $\mu \equiv mv_\perp^2/(2B)$, and ϑ denote the gyrocenter position, the velocity component parallel to the magnetic field, the magnetic moment, and the gyrophase angle, respectively, The vector potential and the unit vector parallel to the background magnetic field $\mathbf{B}$ are written by $\mathbf{A}$ and $\mathbf{b} \equiv \mathbf{B}/B$, respectively. Here, it is supposed that $\mathbf{A}$ can weakly depend on time t and accordingly the background magnetic field $\mathbf{B} = \nabla\times\mathbf{A}$ is allowed to slowly vary in time. Thus we can treat the inductive electric field $\mathbf{E}_T \equiv -c^{-1}\partial\mathbf{A}/\partial t$ which drives the ohmic current in tokamaks.

The gyrocenter Hamiltonian which appears on the right-hand side of Eq. (1) is given by

$$H_{GYa}(\mathbf{Z},t) \equiv \frac{1}{2}m_a U^2 + \mu B(\mathbf{X},t) + e_a\Psi_a(\mathbf{Z},t), \tag{2}$$

and the potential Ψ_a including effects of the turbulent electromagnetic fields is defined by[31]

$$e_a\Psi_a(\mathbf{Z},t) \equiv e_a\langle\psi_a(\mathbf{Z},t)\rangle_\vartheta - \frac{e_a}{c}\mathbf{v}_{Ba}\cdot\langle\widehat{\mathbf{A}}(\mathbf{X}+\boldsymbol{\rho}_a,t)\rangle_\vartheta + \frac{e_a^2}{2m_ac^2}\langle|\widehat{\mathbf{A}}(\mathbf{X}+\boldsymbol{\rho}_a,t)|^2\rangle_\vartheta - \frac{e_a^2}{2B}\frac{\partial}{\partial\mu}\langle(\widetilde{\psi}_a)^2\rangle_\vartheta, \quad (3)$$

where $\boldsymbol{\rho}_a$ is the gyroradius vector given by $\boldsymbol{\rho}_a \equiv \mathbf{b}\times\mathbf{v}/\Omega_a$, $\Omega_a \equiv e_aB/(m_ac)$ is the gyrofrequency, and

$$\psi_a(\mathbf{Z},t) \equiv \phi(\mathbf{X}+\boldsymbol{\rho}_a,t) - \frac{\mathbf{v}}{c}\cdot\widehat{\mathbf{A}}(\mathbf{X}+\boldsymbol{\rho}_a,t). \quad (4)$$

The gyrophase average $\langle Q\rangle_\vartheta$ and the gyrophase-dependent part $\widetilde{Q}$ of an arbitrary function Q of the gyrocenter phase-space coordinates $\mathbf{Z}\equiv(\mathbf{X},U,\mu,\vartheta)$ are represented by

$$\langle Q\rangle_\vartheta \equiv \frac{1}{2\pi}\oint Q\,d\vartheta \quad \text{and} \quad \widetilde{Q}\equiv Q-\langle Q\rangle_\vartheta, \quad (5)$$

respectively. The particle's velocity $\mathbf{v}$ is written as

$$\mathbf{v}\equiv U\mathbf{b}(\mathbf{X},t) - W[\sin\vartheta\,\mathbf{e}_1(\mathbf{X},t)+\cos\vartheta\,\mathbf{e}_2(\mathbf{X},t)], \quad (6)$$

where $(\mathbf{e}_1,\mathbf{e}_2,\mathbf{b})$ are unit vectors which form a right-handed orthogonal triad and are regarded as functions of $(\mathbf{X},t)$. The magnetic moment is given by $\mu\equiv m_aW^2/2B$, and

$$\mathbf{v}_{Ba}\equiv\frac{c}{e_aB}\mathbf{b}\times\left(m_aU^2\mathbf{b}\cdot\nabla\mathbf{b}+\mu\nabla B\right) \quad (7)$$

is the first-order drift velocity consisting of curvature drift and ∇B drift. Second-order terms retained in Eq. (3) are necessary for correctly deriving the gyrokinetic Poisson and Ampère equations from variational derivatives with ϕ and $\widehat{\mathbf{A}}$, respectively, as shown in Sec. III. Especially, it is shown in Ref.[31] that the second-order term $-(e_a/c)\mathbf{v}_{Ba}\cdot\langle\widehat{\mathbf{A}}\rangle_\vartheta$, which is often neglected in conventional studies, should be kept for the variational derivative to obtain the gyrokinetic Ampère's law including both equilibrium and turbulent parts accurately. In Appendix A, Ψ_a is expanded in the gyroradius and decomposed into several parts which have different dependences on electrostatic and magnetic fluctuations. It is noted in Ref.[51–53] that basic equations including terms of higher-order, which are not considered here, are required for accurately describing the flux-surface-averaged momentum balance along the symmetry direction in up-down symmetric tokamaks and stellarator-symmetric quasisymmetric stellarators where the low-flow ordering is assumed as in the present work.

Using the gyrocenter Lagrangian in Eq. (1), the Euler-Lagrangian equations are given by

$$\frac{d}{dt}\left(\frac{\partial L_{GYa}}{\partial\dot{\mathbf{Z}}}\right)-\frac{\partial L_{GYa}}{\partial\mathbf{Z}}=0. \quad (8)$$

from which the gyrocenter motion equations are obtained as

$$\frac{d\mathbf{Z}}{dt}=\{\mathbf{Z},H_{GYa}\}+\{\mathbf{Z},\mathbf{X}\}\cdot\frac{e_a}{c}\frac{\partial\mathbf{A}_a^*}{\partial t}. \quad (9)$$

with the Poisson brackets defined by

$$\{\mathbf{X},\mathbf{X}\}=\frac{c}{e_aB_{a\|}^*}\mathbf{b}\times\mathbf{I}, \quad \{\mathbf{X},U\}=\frac{\mathbf{B}_a^*}{m_aB_{a\|}^*},$$

$$\{\mathbf{X},\vartheta\}=0, \quad \{U,\vartheta\}=0, \quad \{\vartheta,\mu\}=\frac{e_a}{m_ac}. \quad (10)$$

Equations (9) are rewritten as

$$\frac{d\mathbf{X}}{dt} = \frac{1}{B_{a\|}^*}\left[\left(U+\frac{e_a}{m_a}\frac{\partial\Psi_a}{\partial U}\right)\mathbf{B}_a^* + c\mathbf{b}\times\left(\frac{\mu}{e_a}\nabla B+\nabla\Psi_a+\frac{1}{c}\frac{\partial\mathbf{A}_a^*}{\partial t}\right)\right], \quad (11)$$

$$\frac{dU}{dt}=-\frac{\mathbf{B}_a^*}{m_aB_{a\|}^*}\cdot\left(\mu\nabla B+e_a\nabla\Psi_a+\frac{e_a}{c}\frac{\partial\mathbf{A}_a^*}{\partial t}\right), \quad (12)$$

$$\frac{d\mu}{dt}=0, \quad (13)$$

and

$$\frac{d\vartheta}{dt}=\Omega_a+\frac{e_a^2}{m_ac}\frac{\partial\Psi_a}{\partial\mu}, \quad (14)$$

where $\mathbf{B}_a^*$ and $B_{a\|}^*$ are defined by

$$\mathbf{B}_a^*\equiv\nabla\times\mathbf{A}_a^* \quad \text{and} \quad B_{a\|}^*\equiv\mathbf{B}_a^*\cdot\mathbf{b}, \quad (15)$$

respectively. The gyrocenter motion given by Eqs. (11)–(14) satisfies Liouville's theorem, which is expressed as

$$\frac{\partial D_a(\mathbf{Z},t)}{\partial t}+\frac{\partial}{\partial\mathbf{Z}}\cdot\left(D_a(\mathbf{Z},t)\frac{d\mathbf{Z}}{dt}\right)=0, \quad (16)$$

where the Jacobian $D_a(\mathbf{Z},t)$ is given by $D_a(\mathbf{Z},t)=B_{a\|}^*/m_a$.

III. GYROKINETIC VLASOV-POISSON-AMPÈRE SYSTEM

The action integral for the gyrokinetic Vlasov-Poisson-Ampère system is given by

$$I\equiv\int_{t_1}^{t_2}dt\,L_{GKF}\equiv\int_{t_1}^{t_2}dt\,(L_{GK}+L_F), \quad (17)$$

where the gyrokinetic Lagrangian L_{GK} is defined by the phase-space integral of the gyrocenter distribution function F_a multiplied by the gyrocenter Lagrangian L_{GYa} [see Eq. (1)] as

$$L_{GK}\equiv\sum_a L_{GKa}\equiv\sum_a\int d^6Z\,F_a(\mathbf{Z},t)L_{GYa}(\mathbf{Z},\mathbf{u}_{aZ}(\mathbf{Z},t),t). \quad (18)$$

Here, based on the Eulerian picture, the temporal change rates of the gyrocenter coordinates are regarded as functions of $(\mathbf{Z},t)$ and they are represented by

$$\mathbf{u}_{aZ}(\mathbf{Z},t)=[\mathbf{u}_{aX}(\mathbf{Z},t),u_{aU}(\mathbf{Z},t),u_{a\mu}(\mathbf{Z},t),u_{a\vartheta}(\mathbf{Z},t)], \quad (19)$$

which are used in Eq.(18) to evaluate $L_{GYa}(\mathbf{Z},\mathbf{u}_{aZ}(\mathbf{Z},t),t)$. Then, the distribution function F_a satisfies

$$\frac{\partial F_a(\mathbf{Z},t)}{\partial t}+\frac{\partial}{\partial\mathbf{Z}}\cdot\left[F_a(\mathbf{Z},t)\,\mathbf{u}_{aZ}(\mathbf{Z},t)\right]=0, \quad (20)$$

where the functional form of $\mathbf{u}_{aZ}(\mathbf{Z},t)$ is determined later by the Euler-Poincaré variational principle.

Here, the Darwin approximation is made to remove electromagnetic waves propagating at light speed, and the Lagrangian L_F is defined by[5]

$$L_F \equiv \frac{1}{8\pi}\int_V d^3x \Big[|\mathbf{E}_L(\mathbf{x},t)|^2 - |\mathbf{B}(\mathbf{x},t)+\widehat{\mathbf{B}}(\mathbf{x},t)|^2 + \frac{2}{c}\lambda(\mathbf{x},t)\nabla\cdot\widehat{\mathbf{A}}(\mathbf{x},t)\Big]. \quad (21)$$

where V represents the spatial domain of the system, λ plays the role of a Lagrange undetermined multiplier to derive the Coulomb gauge condition

$$\nabla\cdot\widehat{\mathbf{A}} = 0, \quad (22)$$

from the variational condition $\delta I/\delta\lambda = 0$ (or $\delta L_{GKF}/\delta\lambda = \delta L_F/\delta\lambda = 0$), and $\mathbf{E}_L$ is the longitudinal (or irrotational) part of the electric field written in terms of electrostatic potential ϕ as

$$\mathbf{E}_L \equiv -\nabla\phi. \quad (23)$$

Now, the trajectories of particles in the phase space as well as the electrostatic potential and perturbed vector potential are virtually varied to derive the governing equations of the collisionless electromagnetic gyrokinetic turbulent system from the Eulerian variational principle. Following the same procedure as in Refs.[31,44], the infinitesimal variation in the phase-space trajectory is represented in the Eulerian picture by

$$\delta\mathbf{Z}_{aE}(\mathbf{Z},t) = [\delta\mathbf{X}_{aE}(\mathbf{Z},t), \delta U_{a\parallel E}(\mathbf{Z},t), \delta\mu_{aE}(\mathbf{Z},t), \delta\zeta_{aE}(\mathbf{Z},t)]. \quad (24)$$

Then, the variations in the functional forms of $\mathbf{u}_{aZ} = (\mathbf{u}_{aX}, u_{aU}, u_{a\mu}, u_{a\vartheta})$ and $F_a(\mathbf{Z},t)$ are written in the Eulerian picture as

$$\delta\mathbf{u}_{aZ}(\mathbf{Z},t) = \left(\frac{\partial}{\partial t} + \mathbf{u}_{aZ}(\mathbf{Z},t)\cdot\frac{\partial}{\partial\mathbf{Z}}\right)\delta\mathbf{Z}_{aE}(\mathbf{Z},t) - \delta\mathbf{Z}_{aE}(\mathbf{Z},t)\cdot\frac{\partial}{\partial\mathbf{Z}}\,\mathbf{u}_{aZ}(\mathbf{Z},t), \quad (25)$$

and

$$\delta F_a(\mathbf{Z},t) = -\frac{\partial}{\partial\mathbf{Z}}\cdot[F_a(\mathbf{Z},t)\delta\mathbf{Z}_{aE}(\mathbf{Z},t)], \quad (26)$$

respectively. The variations in the electrostatic potential and the perturbed vector potential are denoted by $\delta\phi$ and $\delta\widehat{\mathbf{A}}$, respectively. Using Eqs. (17), (18), (25), and (26), the variation in the action integral I_{GKF} is given by

$$\delta I_{GKF} = \sum_a \int_{t_1}^{t_2} dt \int d^6Z\, F_a \times \left[\left(\frac{\partial L_{GYa}}{\partial\mathbf{Z}}\right)_u - \left(\frac{d}{dt}\right)_a\left(\frac{\partial L_{GYa}}{\partial\mathbf{u}_{aZ}}\right)\right]\cdot\delta\mathbf{Z}_{aE} + \int_{t_1}^{t_2} dt \int_V d^3x \left[\frac{\delta L_{GKF}}{\delta\phi}\delta\phi + \frac{\delta L_{GKF}}{\delta\widehat{\mathbf{A}}}\cdot\delta\widehat{\mathbf{A}}\right] + \text{B.T.}. \quad (27)$$

Here, B.T. represents boundary terms that appear due to partial integrals and $(d/dt)_a$ denotes the time derivative along the phase-space trajectory defined by

$$\left(\frac{d}{dt}\right)_a \equiv \frac{\partial}{\partial t} + \mathbf{u}_{aZ}\cdot\frac{\partial}{\partial\mathbf{Z}}, \quad (28)$$

from which one obtains $(d/dt)_a\mathbf{Z} = \mathbf{u}_{aZ}$. The variational derivative $\delta\mathscr{F}[f]/\delta f$ of any functional $\mathscr{F}[f]$ of a function f in three-dimensional space is defined as a function in the space which satisfies

$$\int d^3x \frac{\delta\mathscr{F}[f]}{\delta f}(\mathbf{x})\varphi(\mathbf{x}) = \lim_{\varepsilon\to 0}\frac{\mathscr{F}[f+\varepsilon\varphi]-\mathscr{F}[f]}{\varepsilon}, \quad (29)$$

from which one can also write

$$\frac{\delta\mathscr{F}[f(\mathbf{y})]}{\delta f(\mathbf{x})} \equiv \frac{d}{d\varepsilon}\mathscr{F}[f(\mathbf{y})+\varepsilon\delta^3(\mathbf{y}-\mathbf{x})]\Big|_{\varepsilon=0}, \quad (30)$$

where $\delta^3(\mathbf{y}-\mathbf{x}) \equiv \delta(y^1-x^1)\delta(y^2-x^2)\delta(y^3-x^3)$. From Eq. (30), one has

$$\frac{\delta\phi(\mathbf{X}+\boldsymbol{\rho}_a)}{\delta\phi(\mathbf{x})} = \delta(\mathbf{X}+\boldsymbol{\rho}_a-\mathbf{x}). \quad (31)$$

Now, it is required that $\delta I_{GKF} = 0$ holds for any variations $\delta\mathbf{Z}_{aE}$, $\delta\phi$, and $\delta\widehat{\mathbf{A}}$ which vanish on the boundaries of the integral region. Then, it is found from Eq. (27) that

$$\left(\frac{d}{dt}\right)_a\left(\frac{\partial L_{GYa}}{\partial\mathbf{u}_{aZ}}\right) - \left(\frac{\partial L_{GYa}}{\partial\mathbf{Z}}\right)_u = 0, \quad (32)$$

$\delta L_{GKF}/\delta\phi = 0$, and $\delta L_{GKF}/\delta\widehat{\mathbf{A}} = 0$ need to be satisfied. Here, since Eq. (32) is equivalent to Eq. (8), one finds that $\mathbf{u}_{aZ}(\mathbf{Z},t)$ should be given by the right-hand side of Eq. (9). Thus, the gyrokinetic Vlasov equation is given by Eq. (20) with Eq. (9) substituted into $\mathbf{u}_{aZ}(\mathbf{Z},t)$.

The gyrokinetic Poisson equation is derived from $\delta L_{GKF}/\delta\phi = 0$ and written as

$$\nabla\cdot\mathbf{E}_L = 4\pi\rho_c, \quad (33)$$

where the charge density ρ_c is given by

$$\rho_c = -\frac{\delta L_{GK}}{\delta\phi} = -\sum_a\frac{\delta L_{GKa}}{\delta\phi} = \sum_a e_a\int d^6Z\,\delta^3(\mathbf{X}+\boldsymbol{\rho}_a-\mathbf{x})\left(F_a + \frac{e_a\widetilde{\psi}_a}{B}\frac{\partial F_a}{\partial\mu}\right), \quad (34)$$

and $\delta L_F/\delta\phi = (1/4\pi)\nabla\cdot\mathbf{E}_L$ is used.

The gyrokinetic Ampère's law is derived from $\delta L_{GKF}/\delta\widehat{\mathbf{A}} = 0$ as

$$\nabla\times(\mathbf{B}+\widehat{\mathbf{B}}) = \frac{4\pi}{c}\mathbf{j} - \frac{1}{c}\nabla\lambda, \quad (35)$$

where the electric current density $\mathbf{j}$ is given by

$$\mathbf{j} = c\frac{\delta L_{GK}}{\delta\widehat{\mathbf{A}}} = c\sum_a\frac{\delta L_{GKa}}{\delta\widehat{\mathbf{A}}} = \sum_a e_a\int d^6Z\,\delta^3(\mathbf{X}+\boldsymbol{\rho}_a-\mathbf{x}) \times\left[F_a(\mathbf{Z},t)\left(\mathbf{v} - \frac{e_a}{m_a c}\widehat{\mathbf{A}} + \mathbf{v}_{Ba}\right) + \frac{e_a\widetilde{\psi}_a}{B}\frac{\partial F_a}{\partial\mu}\mathbf{v}\right], \quad (36)$$

and $\delta L_F/\delta \widehat{\mathbf{A}} = -(1/4\pi)\nabla\times(\mathbf{B}+\widehat{\mathbf{B}}) - (1/4\pi c)\nabla\lambda$ is used. It is noted here that an arbitrary vector field $\mathbf{a}$ is written as $\mathbf{a} = \mathbf{a}_L + \mathbf{a}_T$ where the longitudinal (or irrotational) and transverse (or solenoidal) parts of $\mathbf{a}$ are given by $\mathbf{a}_L(\mathbf{x}) = -(4\pi)^{-1}\nabla\int d^3x'\,(\nabla'\cdot\mathbf{a}(\mathbf{x}'))/|\mathbf{x}-\mathbf{x}'|$ and $\mathbf{a}_T(\mathbf{x}) = (4\pi)^{-1}\nabla\times(\nabla\times\int d^3x'\,\mathbf{a}(\mathbf{x}')/|\mathbf{x}-\mathbf{x}'|)$, respectively.[54] Then, the longitudinal part of Eq. (35) gives

$$\nabla\lambda = 4\pi\mathbf{j}_L. \tag{37}$$

From the transverse part of Eq. (35), the gyrokinetic Ampère's law is written as

$$\nabla\times(\mathbf{B}+\widehat{\mathbf{B}}) = \frac{4\pi}{c}\mathbf{j}_T. \tag{38}$$

In Eqs. (37) and (38), $\mathbf{j}_L$ and $\mathbf{j}_T$ represent the longitudinal and transverse parts of $\mathbf{j}$, respectively.

IV. REPRESENTATION IN GENERAL SPATIAL COORDINATES

In this section, general spatial coordinates are used to represent the Lagrangian of the electromagnetic turbulent gyrokinetic system defined in Sec. III. The Lagrangian is given as the integral of the Lagrangian density with respect to the general spatial coordinates, and it is invariant under an arbitrary spatial coordinate transformation.

A. The Lagrangian density represented in general spatial coordinates

The action integral I_{GKF} in Eq. (17) is written here as

$$I_{GKF} \equiv \int_{t_1}^{t_2} dt\, L_{GKF} \equiv \int_{t_1}^{t_2} dt \int_V d^3x\,\mathscr{L}_{GKF}, \tag{39}$$

where the Lagrangian density $\mathscr{L}_{GKF}$ is given by

$$\begin{aligned}
\mathscr{L}_{GKF} &\equiv \mathscr{L}_{GK} + \mathscr{L}_F \\
\mathscr{L}_{GK} &\equiv \sum_a \int d^3v\, F_a(x,v,t) L_{GYa}(x,v,t) \\
\mathscr{L}_F &\equiv \frac{\sqrt{g(x)}}{8\pi}\Big[g^{ij}(x)(E_L)_i(x,t)(E_L)_j(x,t) \\
&\quad - g_{ij}(x)\{B^i(x,t)+\widehat{B}^i(x,t)\}\{B^j(x,t)+\widehat{B}^j(x,t)\} \\
&\quad + \frac{2}{c}\lambda(x,t)g^{ij}(x)\nabla_i\widehat{A}_j(x,t)\Big].
\end{aligned} \tag{40}$$

Here, $x\equiv(x^i)_{i=1,2,3}$, $v\equiv(U,\mu,\vartheta)$, $d^3x\equiv dx^1dx^2dx^3$, and $d^3v\equiv dUd\mu d\vartheta$ are used, and ∇_j represents a covariant derivative. In the equation for $\mathscr{L}_{GK}$ shown in Eq. (40), $x\equiv(x^i)_{i=1,2,3}$ represent the coordinates not of the position of the particle but that of the gyrocenter (denoted by $\mathbf{X}$ in Sec. II). It should be emphasized that in this section, $x\equiv(x^i)_{i=1,2,3}$ are general spatial coordinates which can be either Cartesian or any other curved coordinates. However, the spatial position vector $\mathbf{r}=\mathbf{r}(x)$ is assumed to be a function of only the spatial coordinates $x\equiv(x^i)_{i=1,2,3}$ and it is independent of time t. The gyrocenter distribution function in the (x,v)-space is denoted by F_a, and the number of particles of species a in the phase-space volume element $d^3xd^3v\equiv dx^1dx^2dx^3dUd\mu d\vartheta$ at time t is given by $F_a(x,v,t)d^3xd^3v$. This paper employs the summation convention that the same symbol used for a pair of upper and lower indices within a term [such as seen in Eq. (40) as well as in the equations shown below] indicates summation over the range $\{1,2,3\}$ of the symbol index. The contravariant metric tensor components g^{ij} in the general spatial coordinates $x\equiv(x^i)$ are related to the covariant components g_{ij} by $g^{ik}g_{kj}=\delta^i_j$, where δ^i_j represents the Kronecker delta. The determinant of the covariant metric tensor matrix is denoted by $g(x)\equiv\det[g_{ij}(x)]$. As the spatial position vector $\mathbf{r}$ is a function of only the spatial coordinates $x\equiv(x^i)$, $g_{ij}(x)$, $g^{ij}(x)$, and $g(x)$ are all independent of time t.

The gyrocenter Lagrangian L_{GYa}, which enters $\mathscr{L}_{GK}$ in Eq. (40), is represented in the Eulerian picture by

$$\begin{aligned}
L_{GYa} &\equiv \left(\frac{e_a}{c}A_j(x,t) + m_aUb^i(x,t)g_{ij}(x)\right)u^j_{ax}(x,v,t) \\
&\quad + \frac{m_ac}{e_a}\mu u_{a\vartheta}(x,v,t) - H_{GYa}(x,U,\mu,t),
\end{aligned} \tag{41}$$

where $b^i\equiv B^i/B$ is the ith contravariant component of the unit vector parallel to the background magnetic field and the background field strength is given by $B(x,t)\equiv\sqrt{g_{ij}(x)B^i(x,t)B^j(x,t)}$. The contravariant components of the background and perturbed magnetic fields are expressed in terms of the covariant components of the vector potentials as

$$B^i(x,t)\equiv\frac{\epsilon^{ijk}}{\sqrt{g(x)}}\frac{\partial A_k(x,t)}{\partial x^j}, \qquad \widehat{B}^i(x,t)\equiv\frac{\epsilon^{ijk}}{\sqrt{g(x)}}\frac{\partial \widehat{A}_k(x,t)}{\partial x^j}, \tag{42}$$

where the Levi-Civita symbol is denoted by

$$\begin{aligned}
\epsilon^{ijk} &\equiv \epsilon_{ijk} \\
&\equiv \begin{cases} 1 & ((i,j,k)=(1,2,3),(2,3,1),(3,1,2)) \\ -1 & ((i,j,k)=(1,3,2),(2,1,3),(3,2,1)) \\ 0 & (\text{otherwise}). \end{cases}
\end{aligned} \tag{43}$$

The gyrocenter Hamiltonian is written here as

$$H_{GYa}(x,U,\mu,t)\equiv\frac{1}{2}m_aU^2+\mu B(x,t)+e_a\Psi_a(x,U,\mu,t), \tag{44}$$

and the fluctuation part is given by

$$\begin{aligned}
\Psi_a(x,\mu,t) &\equiv \phi(x,t)+\Psi_{E1a}(x,\mu,t)+\Psi_{\widehat{A}1a}(x,U,\mu,t) \\
&\quad +\Psi_{E2a}(x,\mu,t)+\Psi_{E\widehat{A}a}(x,U,\mu,t) \\
&\quad +\Psi_{\widehat{A}2a}(x,U,\mu,t),
\end{aligned} \tag{45}$$

where

$$\begin{aligned}
\Psi_{E1a}(x,\mu,t) &= \sum_{n=1}^{\infty}\frac{\alpha_a^{j_1\cdots j_n}(x,\mu,t)}{n!}\nabla_{j_1}\cdots\nabla_{j_n}\phi(x,t) \\
&= -\sum_{n=1}^{\infty}\frac{\alpha_a^{j_1\cdots j_n}(x,\mu,t)}{n!}\nabla_{j_1}\cdots\nabla_{j_{n-1}}(E_L)_{j_n}(x,t),
\end{aligned} \tag{46}$$

and

$$\Psi_{\widehat{A}1a}(x,U,\mu,t) = -\frac{1}{c}\sum_{n=0}^{\infty}\frac{1}{n!}\big[\alpha_a^{j_1\cdots j_n}\,(Ub^i + v_{Ba}^i) + \Omega_a \times \sqrt{g}\epsilon_{klm}\,\alpha_a^{j_1\cdots j_n k}\,b^l\,g^{im}\big]\,\nabla_{j_1}\cdots\nabla_{j_n}\widehat{A}_i(x,t) \quad (47)$$

Here, $\alpha_a^{j_1\cdots j_n}(x,\mu,t)$ is defined by Eqs. (A9)–(A11) in Appendix A with using $h^{ij} \equiv g^{ij} - b^i b^j$. The second-order parts $\Psi_{E2a}(x,\mu,t)$, $\Psi_{E\widehat{A}a}(x,U,\mu,t)$, and $\Psi_{\widehat{A}2a}(x,U,\mu,t)$ on the right-hand side of Eq. (45) are obtained from Eqs. (A15), (A17), and (A18), respectively, with the partial derivative ∂_j replaced by the covariant derivative ∇_j.

The temporal change rates of the gyrocenter coordinates in Eq. (19) are denoted here by $u_{ax}^i(x,v,t)$, $u_{aU}(x,v,t)$, $u_{a\mu}(x,v,t)$, and $u_{a\vartheta}(x,v,t)$. Then, Eq. (20) for the distribution function $F_a(x,v,t)$ is written as

$$\frac{\partial F_a}{\partial t} + \frac{\partial}{\partial x^j}(F_a u_{ax}^j) + \frac{\partial}{\partial U}(F_a u_{aU}) + \frac{\partial}{\partial \mu}(F_a u_{a\mu}) + \frac{\partial}{\partial \vartheta}(F_a u_{a\vartheta}) = 0. \quad (48)$$

The gyrocenter Hamiltonian H_{GYa} given in Eq. (44) takes a functional form,

$$H_{GYa} = H_{GYa}\big[v, \partial_j A_i(x,t), \partial_{jk} A_i(x,t), \{\partial_J \phi(x,t)\}, \{\partial_J \widehat{A}_i(x,t)\}, \{\partial_J g_{ij}(x)\}\big], \quad (49)$$

which depends on the velocity space coordinates (except for ϑ) as well as the general spatial coordinates $x \equiv (x^i)_{i=1,2,3}$ through the field variables $[\partial_j A_i(x,t), \partial_{jk} A_i(x,t), \{\partial_J \phi(x,t)\}, \{\partial_J \widehat{A}_i(x,t)\}, \{\partial_J g_{ij}(x)\}]$. Here, the notation $J \equiv (j_1, j_2, \cdots, j_n)$ $(n = 0,1,2,\cdots; j_1, j_2, \cdots, j_n = 1,2,3)$ is used to write

$$\partial_J \mathscr{F} \equiv \begin{cases} \mathscr{F} & (n=0) \\ \partial_{j_1 j_2 \cdots j_n}\mathscr{F} \equiv \partial^n \mathscr{F}/\partial x^{j_1}\partial x^{j_2}\cdots\partial x^{j_n} & (n \geq 1) \end{cases} \quad (50)$$

where $\mathscr{F}$ is an arbitrary function of $x = (x^i)_{i=1,2,3}$. Then, $\{\partial_J \phi\} \equiv \{\phi, \partial_j \phi, \partial_{jk}\phi, \partial_{jkl}\phi, \cdots\}$, and the definitions of other compact notations $\{\partial_J \widehat{A}_i\}$, and $\{\partial_J g_{ij}(x)\}$ in Eq. (49) are understood in the same way. One obtains $(E_L)_i \equiv -\partial_i \phi$ from Eq. (23) which is used to replace $\{\partial_J \phi\}$ with $\{\phi, \{\partial_J (E_L)_i\}\}$ where $\{\partial_J (E_L)_i\} \equiv \{(E_L)_i, \partial_j (E_L)_i, \partial_{jk}(E_L)_i, \partial_{jkl}(E_L)_i, \cdots\}$. Note that high-order spatial derivative terms due to finite gyroradii enter the gyrocenter Hamiltonian H_{GYa} as seen in Eqs. (46) and (47) where the covariant derivatives contain the spatial derivatives of g_{ij} through the Christoffel symbols [see Eq. (A4) in Ref.[37]].

It is found from Eqs. (41) and (49) that the functional form of the gyrocenter Lagrangian L_{GYa} is written as

$$L_{GYa} = L_{GYa}\big[v, u_{ax}^i(x,v,t), u_{a\vartheta}(x,v,t), A_i(x,t), \partial_j A_i(x,t), \partial_{jk} A_i(x,t), \{\partial_J \phi(x,t)\}, \{\partial_J \widehat{A}_i(x,t)\}, \{\partial_J g_{ij}(x)\}\big], \quad (51)$$

where $u_{ax}^i(x,v,t)$, $u_{a\vartheta}(x,v,t)$, $\phi(x,t)$, and $\widehat{A}_i(x,t)$ are the functions, the governing equations of which are derived from the variation principle in Sec. III.C while the dependence of L_{GYa} on $A_i(x,t)$, $\partial_j A_i(x,t)$, $\partial_{jk} A_i(x,t)$, and $\partial_J g_{ij}(x,t)$ is also explicitly shown because their variations need to be taken into account to evaluate the variation of L_{GYa} in Sec. IV where the local momentum balance is derived using the general spatial coordinate transformation which causes the variations in the functional forms of both $(u_{ax}^i, u_{a\vartheta}, \phi, \widehat{A}_i)$ and (A_i, g_{ij}).

B. The Lagrangian density associated with the electromagnetic interaction

It is found from Eqs. (40), (41), and (44) that the part of the Lagrangian density including Ψ_a is given by

$$\mathscr{L}_\Psi \equiv \sum_a \mathscr{L}_{\Psi a} \equiv -\sum_a \int d^3v\, F_a\, e_a \Psi_a = -\rho_c^{(g)}\phi + \mathscr{L}_{E1} + \mathscr{L}_{\widehat{A}1} + \mathscr{L}_{E2} + \mathscr{L}_{E\widehat{A}} + \mathscr{L}_{\widehat{A}2}, \quad (52)$$

which determines the electromagnetic interaction of particles. Here, the gyrocenter charge density $\rho_c^{(g)}$ is defined by

$$\rho_c^{(g)} \equiv \sum_a e_a N_a^{(g)} \equiv \sum_a e_a \int d^3v\, F_a, \quad (53)$$

where $N_a^{(g)}$ represents the gyrocenter density. The terms $\mathscr{L}_{E1}$, $\mathscr{L}_{\widehat{A}1}$, $\mathscr{L}_{E2}$, $\mathscr{L}_{E\widehat{A}}$, and $\mathscr{L}_{\widehat{A}2}$ on the right-hand side of Eq. (52), are defined by

$$\mathscr{L}_{E1} \equiv \sum_a \mathscr{L}_{E1a} = \sum_{k=1}^{\infty} Q_0^{j_1\cdots j_{2k}} \nabla_{j_1}\cdots\nabla_{j_{2k-1}}(E_L)_{j_{2k}}, \quad (54)$$

$$\mathscr{L}_{\widehat{A}1} \equiv \sum_a \mathscr{L}_{\widehat{A}1a} = \sum_{n=1}^{\infty} R_0^{j_1\cdots j_n} \nabla_{j_1}\cdots\nabla_{j_{n-1}}\widehat{A}_{j_n}, \quad (55)$$

$$\mathscr{L}_{E2} \equiv \sum_a \mathscr{L}_{E2a} = \frac{1}{2}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \chi_E^{i_1\cdots i_m;j_1\cdots j_n} \nabla_{i_1}\cdots\nabla_{i_{m-1}}(E_L)_{i_m} \times \nabla_{j_1}\cdots\nabla_{j_{n-1}}(E_L)_{j_n}, \quad (56)$$

$$\mathscr{L}_{E\widehat{A}} \equiv \sum_a \mathscr{L}_{E\widehat{A}a} = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \chi_{E\widehat{A}}^{j_1\cdots j_m;k_1,\cdots k_n} \nabla_{j_1}\cdots\nabla_{j_{m-1}}(E_L)_{j_m} \times \nabla_{k_1}\cdots\nabla_{k_{n-1}}\widehat{A}_{k_n}, \quad (57)$$

and

$$\mathscr{L}_{\widehat{A}2} \equiv \sum_a \mathscr{L}_{\widehat{A}2a} = \frac{1}{2}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \chi_{\widehat{A}}^{j_1\cdots j_m;k_1,\cdots k_n} \nabla_{j_1}\cdots\nabla_{j_{m-1}}\widehat{A}_{j_m} \times \nabla_{k_1}\cdots\nabla_{k_{n-1}}\widehat{A}_{k_n}, \quad (58)$$

respectively. Here, $Q_0^{j_1\cdots j_{2k}}$, $R_0^{j_1\cdots j_n}$, $\chi_E^{i_1\cdots i_m;j_1\cdots j_n}$, $\chi_{E\widehat{A}}^{j_1\cdots j_m;k_1,\cdots k_n}$, and $\mathscr{L}_{\widehat{A}2}$ included in Eqs. (54)–(58) are given by

$$\left[Q_0^{j_1\cdots j_{2k}}, R_0^{j_1\cdots j_n}\right] \equiv \sum_a \left[Q_{0a}^{j_1\cdots j_{2k}}, R_{0a}^{j_1\cdots j_n}\right],$$

$$\left[\chi_E^{i_1\cdots i_m;j_1\cdots j_n}, \chi_{E\widehat{A}}^{j_1\cdots j_m;k_1,\cdots k_n}, \chi_{\widehat{A}}^{j_1\cdots j_m;k_1,\cdots k_n}\right] \equiv \sum_a \left[\chi_{Ea}^{i_1\cdots i_m;j_1\cdots j_n}, \chi_{E\widehat{A}a}^{j_1\cdots j_m;k_1,\cdots k_n}, \chi_{\widehat{A}a}^{j_1\cdots j_m;k_1,\cdots k_n}\right], \quad (59)$$

where $Q_{0a}^{j_1\cdots j_{2k}}$, $R_{0a}^{j_1\cdots j_n}$, $\chi_{Ea}^{i_1\cdots i_m;j_1\cdots j_n}$, $\chi_{E\widehat{A}a}^{j_1\cdots j_m;k_1,\cdots k_n}$, and $\chi_{\widehat{A}a}^{j_1\cdots j_m;k_1,\cdots k_n}$ are defined by Eqs. (B4), (B6), and (B12) in Appendix B. As seen in Appendices C and D, the charge and current densities in the gyrokinetic Poisson and Ampère equations are derived from $\mathscr{L}_\psi$, of which the components $\mathscr{L}_{E1}$, $\mathscr{L}_{\widehat{A}1}$, $\mathscr{L}_{E2}$, $\mathscr{L}_{E\widehat{A}}$, and $\mathscr{L}_{\widehat{A}2}$ are associated with the polarization charge and the magnetization current.

C. Derivation of gyrokinetic Vlasov-Poisson-Ampère equations in general spatial coordinates

Here, general spatial coordinates $x = (x^i)_{i=1,2,3}$ are used for the Eulerian variational derivation of the gyrokinetic Vlasov-Poisson-Ampère equations for the the electromagnetic gyrokinetic system. The virtual variations in the phase-space trajectory are now represented in the Eulerian picture by $\delta x^i_{aE}(x,v,t)$, $\delta U_{aE}(x,v,t)$, $\delta\mu_{aE}(x,v,t)$, and $\delta\vartheta_{aE}(x,v,t)$. Then, Eqs. (25), (26), and (27) are rewritten as

$$\begin{aligned}&[\delta u^i_{ax}, \delta u_{aU}, \delta u_{a\mu}, \delta u_{a\vartheta}]\\ &\quad = \left(\frac{\partial}{\partial t} + u^j_{ax}\frac{\partial}{\partial x^j} + u_{aU}\frac{\partial}{\partial U} + u_{a\mu}\frac{\partial}{\partial\mu}\right.\\ &\qquad \left. + u_{a\vartheta}\frac{\partial}{\partial\vartheta}\right)[\delta x^i_{aE}, \delta U_{aE}, \delta\mu_{aE}, \delta\vartheta_{aE}]\\ &\quad - \left(\delta x^j_{aE}\frac{\partial}{\partial x^j} + \delta U_{aE}\frac{\partial}{\partial U} + \delta\mu_{aE}\frac{\partial}{\partial\mu}\right.\\ &\qquad \left. + \delta\vartheta_{aE}\frac{\partial}{\partial\vartheta}\right)[u^i_{ax}, u_{aU}, u_{a\mu}, u_{a\vartheta}], \end{aligned} \tag{60}$$

$$\begin{aligned}\delta F_a &= -\frac{\partial}{\partial x^j}(F_a\delta x^j_{aE}) - \frac{\partial}{\partial U_a}(F_a\delta U_{aE}) - \frac{\partial}{\partial\mu}(F_a\delta\mu_{aE})\\ &\quad - \frac{\partial}{\partial\vartheta}(F_a\delta\vartheta_{aE}),\end{aligned} \tag{61}$$

and

$$\begin{aligned}\delta I_{GKF} &= \sum_a \int_{t_1}^{t_2} dt \int_V d^3x \int d^3v\, F_a\\ &\times \left[\left\{\left(\frac{\partial L_{GYa}}{\partial x^i}\right)_u - \left(\frac{d}{dt}\right)_a\left(\frac{\partial L_{GYa}}{\partial u^i_{ax}}\right)\right\}\delta x^i_{aE}\right.\\ &+ \left(\frac{\partial L_{GYa}}{\partial U}\right)_u \delta U_{aE} + \left(\frac{\partial L_{GYa}}{\partial\mu}\right)_u \delta\mu_{aE}\\ &\left. + \left\{\left(\frac{\partial L_{GYa}}{\partial\vartheta}\right)_u - \left(\frac{d}{dt}\right)_a\left(\frac{\partial L_{GYa}}{\partial u_{a\vartheta}}\right)\right\}\delta\vartheta_{aE}\right]\\ &+ \int_{t_1}^{t_2} dt \int_V d^3x \left(\frac{\delta L_{GK}}{\delta\phi}\delta\phi + \frac{\delta L_{GK}}{\delta\widehat{A}_i}\delta\widehat{A}_i + \frac{\delta L_{GK}}{\delta\lambda}\delta\lambda\right)\\ &+ \text{B.T.},\end{aligned} \tag{62}$$

respectively, where $(\partial L_{GYa}/\partial x^i)_u$, $(\partial L_{GYa}/\partial U)_u$, $(\partial L_{GYa}/\partial\mu)_u$, and $(\partial L_{GYa}/\partial\vartheta)_u$ denote the derivatives of L_{GYa} in x^i, U, μ, and ϑ, respectively, with $(u^i_{ax}, u_{a\vartheta})$ kept fixed in L_{GYa}, and the time derivative along the phase-space trajectory is represented by

$$\left(\frac{d}{dt}\right)_a \equiv \frac{\partial}{\partial t} + u^k_{ax}\frac{\partial}{\partial x^k} + u_{aU}\frac{\partial}{\partial U} + u_{a\mu}\frac{\partial}{\partial\mu} + u_{a\vartheta}\frac{\partial}{\partial\vartheta}. \tag{63}$$

Using Eq. (27) and $\delta I_{GKF} = 0$, one first obtains

$$\left(\frac{d}{dt}\right)_a p_{ai} = \left(\frac{\partial L_{GYa}}{\partial x^i}\right)_u, \tag{64}$$

where $p_{ai} \equiv \partial L_{GYa}/\partial u^i_{ax} = (e_a/c)A_i(x,t) + m_aUb_i(x,t) \equiv (e_a/c)A^*_{ai}(x,U,t)$ represents the covariant vector component of the canonical momentum. Equation (64) can be deformed to obtain

$$m_a u_{aU} b_i = e_a\left(-\frac{\partial\Psi_a}{\partial x^i} - \frac{1}{c}\frac{\partial A^*_{ai}}{\partial t} + \frac{1}{c}\sqrt{g}\varepsilon_{ijk}u^j_x B^{*k}\right) - \mu\frac{\partial B}{\partial x^i}, \tag{65}$$

where $B^{*i}_a \equiv (\varepsilon^{ijk}/\sqrt{g})(\partial A^*_{ak}/\partial x^j)$. Next, combining Eq. (65) with $(\partial L_{GYa}/\partial U)_u = m_a(u^i_{ax}b_i - U) = 0$, $(\partial L_{GYa}/\partial\mu)_u = (m_ac/e_a)u_{a\vartheta} - B - e_a\partial\Psi_a/\partial\mu = 0$, and $(d/dt)_a(\partial L_{GYa}/\partial u_{a\vartheta}) = (m_ac/e_a)u_{a\mu} = (\partial L_{GYa}/\partial\vartheta)_u = 0$, the gyrocenter motion equations are derived as

$$\begin{aligned}u^i_{ax} &= \frac{1}{B^*_{a\parallel}}\left[\left(U + \frac{e_a}{m_a}\frac{\partial\Psi_a}{\partial U}\right)B^{*i}_a\right.\\ &\qquad \left. + c\frac{\varepsilon^{ijk}}{\sqrt{g}}b_j\left(\frac{\mu}{e_a}\frac{\partial B}{\partial x^k} + \frac{\partial\Psi_a}{\partial x^k} + \frac{1}{c}\frac{\partial A^*_{ak}}{\partial t}\right)\right],\end{aligned} \tag{66}$$

$$m_a u_{aU} = -\frac{B^{*i}_a}{B^*_{a\parallel}}\left[\mu\frac{\partial B}{\partial x^i} + e_a\left(\frac{\partial\Psi_a}{\partial x^i} + \frac{1}{c}\frac{\partial A^*_{ai}}{\partial t}\right)\right], \tag{67}$$

$$u_{a\mu} = 0, \tag{68}$$

and

$$u_{a\vartheta} = \Omega_a + \frac{e_a^2}{m_ac}\frac{\partial\Psi_a}{\partial\mu}, \tag{69}$$

where $B^*_{a\parallel} \equiv B^{*i}_a b_i$. Substituting Eqs. (66)–(69) into Eqs. (48) and taking its average with respect to the gyrophase ϑ, the gyrokinetic Vlasov equation is derived as

$$\begin{aligned}&\frac{\partial\overline{F}_a}{\partial t} + \frac{\partial}{\partial x^i}\left[\overline{F}_a\frac{1}{B^*_{a\parallel}}\left\{\left(U + \frac{e_a}{m_a}\frac{\partial\Psi_a}{\partial U}\right)B^{*i}_a + c\frac{\varepsilon^{ijk}}{\sqrt{g}}b_j\right.\right.\\ &\qquad\qquad \left.\left.\times\left(\frac{\mu}{e_a}\frac{\partial B}{\partial x^k} + \frac{\partial\Psi_a}{\partial x^k} + \frac{1}{c}\frac{\partial A^*_{ak}}{\partial t}\right)\right\}\right]\\ &\quad + \frac{\partial}{\partial U}\left[\overline{F}_a\frac{B^{*i}_a}{m_aB^*_{a\parallel}}\left\{-e_a\left(\frac{\partial\Psi_a}{\partial x^i} + \frac{1}{c}\frac{\partial A^*_{ai}}{\partial t}\right) - \mu\frac{\partial B}{\partial x^i}\right\}\right]\\ &\quad = 0,\end{aligned} \tag{70}$$

where $\overline{F}_a \equiv \langle F_a\rangle_\vartheta \equiv \oint F_a d\vartheta/2\pi$ is the gyrophase-averaged distribution function.

The remaining conditions for $\delta I_{GKF} = 0$ are given by $\delta L_{GK}/\delta\phi = \delta L_{GK}/\delta\widehat{A}_i = \delta L_{GK}/\delta\lambda = 0$. The Coulomb gauge condition is obtained as $\delta L_{GKF}/\delta\lambda = (2/c)\nabla_i \widehat{A}^i = 0$. The gyrokinetic Poisson equation is given by

$$\frac{\delta L_{GKF}}{\delta\phi} = -\rho_c + \frac{1}{4\pi}\frac{\partial}{\partial x^i}[\sqrt{g}(E_L)^i] = 0, \quad (71)$$

where the charge density ρ_c is written as

$$\rho_c = -\frac{\delta L_{GK}}{\delta\phi} = \rho^{(gc)} - \nabla_i P_G^i, \quad (72)$$

with the generalized polarization vector density P_G^i defined by

$$P_G^i \equiv \sum_{n=0}^{\infty}(-1)^n \nabla_{i_1}\cdots\nabla_{i_n} Q^{i i_1\cdots i_n}. \quad (73)$$

Here, the multipole moments $Q^{i_1\cdots i_m}$ are given by Eq. (C7) in Appendix C. The gyrokinetic Ampère's law is derived from the condition $\delta L_{GK}/\delta\widehat{A}_i = 0$ which is written as

$$\frac{\delta L_{GKF}}{\delta\widehat{A}_i} = \frac{1}{c}j^i - \frac{1}{4\pi}\varepsilon^{ijk}\frac{\partial(B_k+\hat{B}_k)}{\partial x^j} - \frac{\sqrt{g}}{4\pi c}g^{ij}\frac{\partial\lambda}{\partial x^j} = 0. \quad (74)$$

Here, the current density is written as

$$j^l = (j^0)^l + c\nabla_k N^{kl}, \quad (75)$$

where $(j^0)^l$ is defined by Eq. (D9) in Appendix D and N^{kl} is given by

$$N^{kl} \equiv \sum_{n=0}^{\infty}(-1)^{n+1}\nabla_{j_1}\cdots\nabla_{j_n} R^{j_1\cdots j_n kl}, \quad (76)$$

with $R^{j_1\cdots j_n k}$ defined in Eq. (D5).

V. DERIVATION OF THE MOMENTUM BALANCE

In this section, the invariance of the Lagrangian under arbitrary infinitesimal transformations of spatial coordinates is used to derive the local momentum balance equations for the single-particle-species system and for the whole system including particles of multiple species and electromagnetic fields.

A. Invariance of Lagrangians under infinitesimal transformations of spatial coordinates

An arbitrary infinitesimal transformation of spatial coordinates from $x = (x^i)_{i=1,2,3}$ to $x' = (x'^i)_{i=1,2,3}$, is written as

$$x'^i = x^i + \xi^i(x), \quad (77)$$

where the infinitesimal variation in the spatial coordinate x^i is denoted by $\xi^i(x)$ which is an arbitrary function of only the spatial coordinates $x = (x^i)_{i=1,2,3}$ and independent of time t.

The gyrokinetic Lagrangian L_{GKa} is written as

$$L_{GKa} \equiv \int_V d^3x\,\mathscr{L}_{GKa} \equiv \int_V d^3x \int d^3v\, F_a L_{GYa} \quad (78)$$

where F_a and L_{GYa} defined in Eq. (1) behave as a scalar density field and a scalar field, respectively, under the transformation of the spatial coordinates. The variation in L_{GKa} under the infinitesimal spatial coordinate transformation in Eq. (77) is written as

$$\overline{\delta}L_{GKa} \equiv \int_V d^3x\left(\frac{\partial(\xi^i\mathscr{L}_{GKa})}{\partial x^i} + \overline{\delta}\mathscr{L}_{GKa}\right) \quad (79)$$

Here and hereafter, the notation $\overline{\delta}\cdots$ represents the variation caused by the infinitesimal coordinate transformation in Eq. (77) and it should be distinguished from the variation $\delta\cdots$ due to the virtual displacement used in Secs. III and IV C. The expression of the integral in Eq. (79) takes the form often found in conventional textbooks (see for example Ref.[55]) to give the change in the integral caused by the infinitesimal transformation. In the integrand in Eq. (79), the divergence term $\partial(\xi^i\mathscr{L}_{GKa})/\partial x^i$ is obtained using Gauss's theorem for the difference between the domains of integrations in $x = (x^i)_{i=1,2,3}$ and $x' = (x'^i)_{i=1,2,3}$ while $\overline{\delta}\mathscr{L}_{GKa}$ is written using the Leibniz rule for the derivative operation by $\overline{\delta}$ as

$$\overline{\delta}\mathscr{L}_{GKa} = \int d^3v\,\overline{\delta}(F_a L_{GYa}) = \int d^3v\,(\overline{\delta}F_a\cdot L_{GYa} + F_a\cdot\overline{\delta}L_{GYa}), \quad (80)$$

where $\overline{\delta}F_a$ and $\overline{\delta}L_{GYa}$ represent the variations in the spatial functional forms of F_a and L_{GYa} under the infinitesimal spatial coordinate transformation.

Then, applying the chain rule formula for the derivative operation $\overline{\delta}L_{GYa}[u^i_{ax}, u_{a\vartheta}, \{\partial_J A_i\}, \{\partial_J\widehat{A}_i\}, \{\partial_J\phi\}, \{\partial_J g_{ij}\}]$ yields

$$\begin{aligned}\overline{\delta}L_{GYa} &= \frac{\partial L_{GYa}}{\partial u^i_{ax}}\overline{\delta}u^i_{ax} + \frac{\partial L_{GYa}}{\partial u_{a\vartheta}}\overline{\delta}u_{a\vartheta} + \sum_J\frac{\partial L_{GYa}}{\partial(\partial_J A_i)}\overline{\delta}(\partial_J A_i) \\ &\quad + \sum_J\frac{\partial L_{GYa}}{\partial(\partial_J\widehat{A}_i)}\overline{\delta}(\partial_J\widehat{A}_i) + \sum_J\frac{\partial L_{GYa}}{\partial(\partial_J\phi)}\overline{\delta}(\partial_J\phi) \\ &\quad + \sum_J\frac{\partial L_{GYa}}{\partial(\partial_J g_{ij})}\overline{\delta}(\partial_J g_{ij}),\end{aligned} \quad (81)$$

where $\partial L_{GYa}/\partial(\partial_J A_i) = 0$ when the order n of $J = (j_1,\cdots,j_n)$ is greater than or equal to three [see Eq. (51)].

As shown in Ref.[37], the variation in the functional form under the infinitesimal spatial coordinate transformation can be represented by $\overline{\delta} = -L_\xi$, where L_ξ is the Lie derivative[56] associated with the vector field given by $(\xi^i)_{i=1,2,3}$ and it acts on an arbitrary tensor field (as well as an arbitrary tensor field density). Using the fact that $L_\xi\mathscr{L}_{GKa} = \partial(\xi^i\mathscr{L}_{GKa})/\partial x^i$ holds from the definition of the Lie derivative acting on a scalar density field, one finds that the integrand in Eq. (79) is written as $L_\xi\mathscr{L}_{GKa} - \overline{\delta}\mathscr{L}_{GKa}$ which is found to vanish from $\overline{\delta} = -L_\xi$. Then, the integral in Eq. (79) also vanishes and accordingly $\overline{\delta}L_{GKa} = 0$, which means that L_{GKa} is a scalar constant which is invariant under the coordinate transformation. Using Eqs. (79)–(81) and $\overline{\delta} = -L_\xi$, $\overline{\delta}L_{GKa} = 0$ can also be written as

$$
\begin{aligned}
\overline{\delta} L_{GKa} &= \int_V d^3x \int d^3v\, F_a \left(\xi^i \frac{\partial L_{GYa}}{\partial x^i} + \overline{\delta} L_{GYa} \right) \\
&= \int_V d^3x \int d^3v\, F_a \left(\xi^i \frac{\partial L_{GYa}}{\partial x^i} + \frac{\partial L_{GYa}}{\partial u^i_{ax}} \overline{\delta} u^i_{ax} + \frac{\partial L_{GYa}}{\partial u_{a\vartheta}} \overline{\delta} u_{a\vartheta} \right. \\
&\quad + \sum_J \frac{\partial L_{GYa}}{\partial(\partial_J A_i)} \overline{\delta}(\partial_J A_i) + \sum_J \frac{\partial L_{GYa}}{\partial(\partial_J \widehat{A}_i)} \overline{\delta}(\partial_J \widehat{A}_i) \\
&\quad \left. + \sum_J \frac{\partial L_{GYa}}{\partial(\partial_J \phi)} \overline{\delta}(\partial_J \phi) + \sum_J \frac{\partial L_{GYa}}{\partial(\partial_J g_{ij})} \overline{\delta}(\partial_J g_{ij}) \right) \\
&= 0. \qquad (82)
\end{aligned}
$$

Recall that L_ξ annihilate any scalar constant. Therefore, when L_{GKa} is a scalar constant, one can naturally write $\overline{\delta} L_{GKa} = -L_\xi L_{GKa} (= 0)$. Thus, one can confirm that the relation $\overline{\delta} = -L_\xi$ is consistent and useful when treating all tensor variables (including scalar fields and scalar constants) and deriving the invariance properties associated with the coordinate transformation. This relation $\overline{\delta} = -L_\xi$ can be applied under the condition that all quantities, in which the variations due to the coordinate transformation are considered, can be written in terms of tensor fields (or tensor field densities) on which the operation of the Lie derivative can be defined. This condition means that integrals using such tensor fields yield scalar constants which represent geometric quantities and take invariant values independent of the choice of the spatial coordinates. It should be stressed that $\overline{\delta} L_{GKa} = -L_\xi L_{GKa} = 0$ and Eq. (82), which are derived only from the above-mentioned invariance property under the spatial coordinate transformation, are valid whether the gyrokinetic equation derived from the variational principle associated with the virtual variation in the phase-space trajectory in Secs. II and III holds or not.

In the same way as in Eq. (79), the invariance of the Lagrangian L_{GKF} of the whole system under the infinitesimal spatial coordinate transformation can be written as

$$
\begin{aligned}
\overline{\delta} L_{GKF} &\equiv \int_V d^3x \left(\frac{\partial(\xi^i \mathscr{L}_{GKF})}{\partial x^i} + \overline{\delta} \mathscr{L}_{GKF} \right) \\
&= \sum_a \overline{\delta} L_{GKa} + \int_V d^3x \left(\frac{\partial(\xi^i \mathscr{L}_F)}{\partial x^i} + \overline{\delta} \mathscr{L}_F \right) \\
&= 0, \qquad (83)
\end{aligned}
$$

where L_{GKF} is defined by Eq. (39) with Eq. (40). The variation $\overline{\delta} \mathscr{L}_F$ of the field Lagrangian density $\mathscr{L}_F$ defined in Eq. (40) can be written as

$$
\begin{aligned}
\overline{\delta} \mathscr{L}_F &= -\frac{\partial(\xi^i \mathscr{L}_F)}{\partial x^i} \\
&= \frac{\partial \mathscr{L}_F}{\partial(\partial_j A_i)} \overline{\delta}(\partial_j A_i) + \sum_J \frac{\partial \mathscr{L}_F}{\partial(\partial_J \widehat{A}_i)} \overline{\delta}(\partial_J \widehat{A}_i) \\
&\quad + \frac{\partial \mathscr{L}_F}{\partial(\partial_i \phi)} \overline{\delta}(\partial_i \phi) + \sum_J \frac{\partial \mathscr{L}_F}{\partial(\partial_J g_{ij})} \overline{\delta}(\partial_J g_{ij}) + \frac{\partial \mathscr{L}_F}{\partial \lambda} \overline{\delta} \lambda, \quad (84)
\end{aligned}
$$

where $\overline{\delta} = -L_\xi$, $L_\xi \mathscr{L}_F = \partial_i(\xi^i \mathscr{L}_F)$, and the chain rule for $\overline{\delta} \mathscr{L}_F[\{\partial_j A_i\}, \{\partial_J \widehat{A}_i\}, \{\partial_i \phi\}, \{\partial_J g_{ij}\}, \lambda]$ are used. We now use Eqs. (82), and (84) to rewrite Eq. (83) as

$$
\begin{aligned}
\overline{\delta} L_{GKF} &= \int_V d^3x \left[\sum_a \int d^3v\, F_a \left(\xi^i \frac{\partial L_{GYa}}{\partial x^i} + \frac{\partial L_{GYa}}{\partial u^i_{ax}} \overline{\delta} u^i_{ax} \right. \right. \\
&\quad + \frac{\partial L_{GYa}}{\partial u_{a\vartheta}} \overline{\delta} u_{a\vartheta} + \sum_J \frac{\partial L_{GYa}}{\partial(\partial_J A_i)} \overline{\delta}(\partial_J A_i) + \sum_J \frac{\partial L_{GYa}}{\partial(\partial_J \widehat{A}_i)} \overline{\delta}(\partial_J \widehat{A}_i) \\
&\quad \left. + \sum_J \frac{\partial L_{GYa}}{\partial(\partial_J \phi)} \overline{\delta}(\partial_J \phi) + \sum_J \frac{\partial L_{GYa}}{\partial(\partial_J g_{ij})} \overline{\delta}(\partial_J g_{ij}) \right) + \frac{\partial(\xi^i \mathscr{L}_F)}{\partial x^i} \\
&\quad + \frac{\partial \mathscr{L}_F}{\partial(\partial_j A_i)} \overline{\delta}(\partial_j A_i) + \sum_J \frac{\partial \mathscr{L}_F}{\partial(\partial_J \widehat{A}_i)} \overline{\delta}(\partial_J \widehat{A}_i) + \frac{\partial \mathscr{L}_F}{\partial(\partial_i \phi)} \overline{\delta}(\partial_i \phi) \\
&\quad \left. + \sum_J \frac{\partial \mathscr{L}_F}{\partial(\partial_J g_{ij})} \overline{\delta}(\partial_J g_{ij}) + \frac{\partial \mathscr{L}_F}{\partial \lambda} \overline{\delta} \lambda \right] \\
&= 0. \qquad (85)
\end{aligned}
$$

As seen from Eqs. (81), (82), (84), and (85), the invariance of the scalar constants L_{GKa} and L_{GKF} under the infinitesimal spatial transformation can be confirmed using the chain rule formulas for the derivative operation $\overline{\delta} = -L_\xi$ acting on the scalar field $L_{GYa}[u^i_{ax}, u_{a\vartheta}, \{\partial_J A_i\}, \{\partial_J \widehat{A}_i\}, \{\partial_J \phi\}, \{\partial_J g_{ij}\}]$ and the scalar field density $\mathscr{L}_F[\{\partial_j A_i\}, \{\partial_J \widehat{A}_i\}, \{\partial_i \phi\}, \{\partial_J g_{ij}\}, \lambda]$ which are given as composite functions. In Sec.V.B and Sec. V.C, Eqs. (82) and (85) are used to derive the local momentum balance equation for the single-particle-species system and that for the whole system consisting of particles of all species and electromagnetic fields.

B. Momentum balance for a single particle species

We now use $\overline{\delta} = -L_\xi$, $L_\xi u^j_{ax} = \xi^i \partial_i u^j_{ax} - u^i_{ax} \partial_i \xi^j$, $L_\xi u_{a\vartheta} = \xi^i \partial_i u_{a\vartheta}$, and the Euler-Lagrange equations for gyrocenter motion [Eq. (64), $(\partial L_{GYa}/\partial U)_u = 0$, $(\partial L_{GYa}/\partial \mu)_u = 0$, and $(d/dt)_a(\partial L_{GYa}/\partial u_{a\vartheta}) = (\partial L_{GYa}/\partial \vartheta)_u = 0$] to write the first three terms in the integrand on the right-hand side of Eq. (82) as

$$
\begin{aligned}
&F \left(\xi^i \frac{\partial L_{GYa}}{\partial x^i} + \frac{\partial L_{GYa}}{\partial u^j_{ax}} \overline{\delta} u^j_{ax} + \frac{\partial L_{GYa}}{\partial u_{a\vartheta}} \overline{\delta} u_{a\vartheta} \right) \\
&= \xi^i \left[\frac{\partial}{\partial t} \left(F_a \frac{\partial L_{GYa}}{\partial u^i_{ax}} \right) + \frac{\partial}{\partial U} \left(F_a u_{aU} \frac{\partial L_{GYa}}{\partial u^i_{ax}} \right) \right. \\
&\quad - \frac{\partial L_{GYa}}{\partial u^i_{ax}} \left\{ \frac{\partial F_a}{\partial t} + \frac{\partial}{\partial x^j}(F_a u^j_{ax}) + \frac{\partial}{\partial U}(F_a u_{aU}) \right. \\
&\quad \left. \left. + \frac{\partial}{\partial \vartheta}(F_a u_{a\vartheta}) \right\} \right] + \frac{\partial}{\partial x^j} \left(\xi^i F_a u^j_{ax} \frac{\partial L_{GYa}}{\partial u^i_{ax}} \right). \qquad (86)
\end{aligned}
$$

Then, substituting Eq. (86) into Eq. (82), using $\overline{\delta}\partial_J = \partial_J\overline{\delta}$, and performing partial integrals yield

$$\begin{aligned}\overline{\delta}L_{GKa} &= \int_V d^3x \left[\xi^i \int d^3v \left\{ \frac{\partial}{\partial t}\left(F_a \frac{\partial L_{GYa}}{\partial u^i_{ax}}\right) - D_t F_a \frac{\partial L_{GYa}}{\partial u^i_{ax}} \right\} + \frac{\delta L_{GKa}}{\delta A_i}\overline{\delta}A_i + \frac{\delta L_{GKa}}{\delta \widehat{A}_i}\overline{\delta}\widehat{A}_i + \frac{\delta L_{GKa}}{\delta \phi}\overline{\delta}\phi + \frac{\delta L_{GKa}}{\delta g_{ij}}\overline{\delta}g_{ij}\right] + \text{B.T.} \\ &= 0, \end{aligned} \tag{87}$$

where

$$D_t F_a \equiv \frac{\partial F_a}{\partial t} + \frac{\partial}{\partial x^j}(F_a u^j_{ax}) + \frac{\partial}{\partial U}(F_a u_{aU}) + \frac{\partial}{\partial \mu}(F_a u_{a\mu}) + \frac{\partial}{\partial \vartheta}(F_a u_{a\vartheta}). \tag{88}$$

Furthermore, substituting $\overline{\delta}A_i = -\xi^j(\partial_j A_i) - (\partial_i \xi^j)A_j = -\xi^j(\nabla_j A_i) - (\nabla_i \xi^j)A_j$, $\overline{\delta}\widehat{A}_i = -\xi^j(\partial_j \widehat{A}_i) - (\partial_i \xi^j)\widehat{A}_j = -\xi^j(\nabla_j \widehat{A}_i) - (\nabla_i \xi^j)\widehat{A}_j$, $\overline{\delta}\phi = -\xi^j\partial_j\phi = -\xi^j\nabla_j\phi$, and $\overline{\delta}g_{ij} = -\nabla_i\xi_j - \nabla_j\xi_i$ into Eq. (87) and performing partial integrals, one obtains

$$\overline{\delta}L_{GKa} = \int_V d^3x\, \xi^j (J_{GKa})_j + \text{B.T.} = 0, \tag{89}$$

where

$$\begin{aligned}(J_{GKa})_j &\equiv \frac{\partial}{\partial t}\left(\int d^3v F_a \frac{\partial L_{GYa}}{\partial u^j_{ax}}\right) - \int d^3v D_t F_a \frac{\partial L_{GYa}}{\partial u^j_{ax}} \\ &\quad + 2\nabla_i\left(g_{jk}\frac{\delta L_{GKa}}{\delta g_{ik}}\right) - \frac{\delta L_{GKa}}{\delta \phi}\nabla_j\phi - \frac{\delta L_{GKa}}{\delta A_k}\nabla_j A_k \\ &\quad - \frac{\delta L_{GKa}}{\delta \widehat{A}_k}\nabla_j \widehat{A}_k + \nabla_k\left(\frac{\delta L_{GKa}}{\delta A_k}A_j + \frac{\delta L_{GKa}}{\delta \widehat{A}_k}\widehat{A}_j\right). \end{aligned} \tag{90}$$

Here, it should be noted that

$$(J_{GKa})_j = 0 \tag{91}$$

holds because Eq. (89) is valid for an arbitrary infinitesimal vector field represented by ξ^j which vanishes on the boundary of V.

Recalling that the canonical momentum of a single particle is given by $p_{aj} \equiv \partial L_{GYa}/\partial u^j_{ax}$, it is found that the first term on the right-hand side of Eq. (90) represents the rate of change in the momentum of particles per volume. The pressure tensor of the particle species species a is given in terms of the variational derivative $\delta L_{GKa}/\delta g_{ij}$ as

$$\begin{aligned}P^{ij}_a &\equiv 2\frac{\delta L_{GKa}}{\delta g_{ij}} \equiv 2\sum_J (-1)^{\#J}\partial_J\left(\int d^3v\, F_a \frac{\partial L_{GYa}}{\partial(\partial_J g_{ij})}\right) \\ &= P^{ij}_{\mathrm{CGL}a} + \pi^{ij}_{\wedge a} + \pi^{ij}_{\parallel \Psi a} + P^{ij}_{\Psi a}, \end{aligned} \tag{92}$$

where $\#J = n$ represents the order of $J \equiv (j_1, j_2, \cdots, j_n)$,

$$P^{ij}_{\mathrm{CGL}a} = \int d^3v F_a[m_a U^2 b^i b^j + \mu B(g^{ij} - b^i b^j)], \tag{93}$$

$$\pi^{ij}_{\wedge a} \equiv \int d^3v F_a m_a U[b^i (u_{ax})^j_\perp + (u_{ax})^i_\perp b^j], \tag{94}$$

$$\pi^{ij}_{\parallel \Psi a} \equiv e_a \int d^3v F_a U \frac{\partial \Psi_a}{\partial U} b^i b^j, \tag{95}$$

and

$$P^{ij}_{\Psi a} \equiv -2e_a \frac{\delta}{\delta g_{ij}}\left(\int d^3x \int d^3v\, F_a \Psi_a\right). \tag{96}$$

The pressure tensor $P^{ij}_{\mathrm{CGL}a}$ defined in Eq. (93) takes the Chew-Goldberger-Low (CGL) form[49] and it plays an essential role in the neoclassical transport. Effects of turbulent fluctuations on the momentum transport are included in $\pi^{ij}_{\wedge a}$, $\pi^{ij}_{\parallel \Psi a}$, and $P^{ij}_{\Psi a}$, which are detailedly investigated in Sec. VIII. It is found from Eqs. (3), (4), (6), (7), and (95) that $\pi^{ij}_{\parallel \Psi a}$ vanishes when $\widehat{\mathbf{A}} = 0$.

The particle density $N^{(p)}_a$ and the particle flux Γ^i_a of species a are given from the functional derivatives $\delta L_{GKa}/\delta\phi$ and $\delta L_{GKa}/\delta \widehat{A}_i$ by

$$\begin{aligned} e_a N^{(p)}_a &\equiv -\frac{\delta L_{GKa}}{\delta \phi} \equiv -\sum_J (-1)^{\#J}\partial_J\left(\int d^3v\, F_a \frac{\partial L_{GYa}}{\partial(\partial_J\phi)}\right) \\ &\equiv -\int d^3v\, F_a \frac{\partial L_{GYa}}{\partial \phi} + \sum_{n=1}^{\infty}(-1)^n \\ &\quad \times \partial_{j_1\cdots j_{n-1}}\left(\int d^3v\, F_a \frac{\partial L_{GYa}}{\partial(\partial_{j_1\cdots j_{n-1}}(E_L)_{j_n})}\right), \end{aligned} \tag{97}$$

and

$$\frac{e_a}{c}\Gamma^i_a \equiv \frac{\delta L_{GKa}}{\delta \widehat{A}_i} \equiv \sum_J (-1)^{\#J}\partial_J\left(\frac{\partial \mathscr{L}_{GKa}}{\partial(\partial_J \widehat{A}_i)}\right), \tag{98}$$

respectively. The electric current density defined by $j^i \equiv \sum_a e_a \Gamma^i_a$ with the particle flux Γ^i_a in Eq. (98) enters the gyrokinetic Ampère's law which is derived from the variational principle $\delta L_{GKF}/\delta \widehat{A}_i = 0$ in Eq. (74). On the other hand, the variational derivative $\delta L_{GKa}/\delta A_i$ gives

$$\begin{aligned}\frac{e_a}{c}\Gamma^i_{\#a} &\equiv \frac{\delta L_{GKa}}{\delta A_i} \equiv \sum_J (-1)^{\#J}\partial_J\left(\frac{\partial \mathscr{L}_{GKa}}{\partial(\partial_J A_i)}\right) \\ &= \frac{\partial \mathscr{L}_{GKa}}{\partial A_i} - \frac{\partial}{\partial x^j}\left(\frac{\partial \mathscr{L}_{GKa}}{\partial(\partial_j A_i)}\right) + \frac{\partial^2}{\partial x^j \partial x^k}\left(\frac{\partial \mathscr{L}_{GKa}}{\partial(\partial_j\partial_k A_i)}\right), \end{aligned} \tag{99}$$

from which another type of the particle flux $\Gamma^i_{\#a}$ of species a is derived as

$$\begin{aligned}\Gamma^i_{\#a} &\equiv \int d^3v F_a u^i_{ax} + \frac{c}{e_a}\epsilon^{ijk}\frac{\partial}{\partial x^j}\Bigg(\int d^3v \frac{F_a}{\sqrt{g}} \\ &\quad \times \Bigg[-\mu b_k + \frac{m_a U}{B}\left\{(u_{ax})_k - (u_{ax})_l b^l b_k\right\} \\ &\quad - e_a\left\{\frac{\partial \Psi_a}{\partial B^k} - \frac{1}{F_a}\frac{\partial}{\partial x^l}\left(F_a \frac{\partial \Psi_a}{\partial(\partial_l B^k)}\right)\right\}\Bigg]\Bigg). \end{aligned} \tag{100}$$

Here, one note that neither of the last two terms in the last line of Eq. (100) is seen to take the form of a vector component in the general spatial coordinate system. However, the sum of these two terms is shown to be rewritten by

$$\frac{\partial \Psi_a}{\partial B^k} - \frac{1}{F_a}\frac{\partial}{\partial x^l}\left(F_a \frac{\partial \Psi_a}{\partial(\partial_l B^k)}\right) = \left(\frac{\partial \Psi_a}{\partial B^k}\right)_{\nabla B} - \frac{1}{F_a}\nabla_l \left(F_a \frac{\partial \Psi_a}{\partial(\nabla_l B^k)}\right). \tag{101}$$

On the left-hand side of Eq. (101), derivatives are performed with regarding Ψ_a as taking a functional form of $\Psi_a(x^i, B^k(x^i), \partial_j B^k(x^i))$ while, on the right-hand side, a functional form of $\Psi_a(x^i, B^k(x^i), \nabla_j B^k(x^i))$ is considered. Then, each term of the right-hand side of Eq. (101) is found to have the form of a vector component, and the sum of the two terms on the left-hand side also becomes a vector component. Therefore, it is understood that $\Gamma^i_{\# a}$ give by Eq. (100) is transformed as a vector density under the general spatial coordinate transformation by noting that F_a and ε^{ijk} are a scalar density and a tensor density.

The two types of particle fluxes $\Gamma^i_a = (c/e_a)\delta L_{GKa}/\delta \widehat{A}_i$ and $\Gamma^i_{\# a} = (c/e_a)\delta L_{GKa}/\delta A_i$ in Eqs. (98) and (99) arise because the separation of the magnetic field into the average and fluctuating parts are done in the case of electromagnetic turbulence. As described in Ref.[31], it is Γ^i_a that accurately represents both the average and fluctuating parts of the particle flux and is used to evaluate the current density in Ampère's law as shown in Eqs. (36), (74), and (D1). It is found from comparing $\Gamma^i_{\# a}$ with Γ^i_a that the average part of $\Gamma^i_{\# a}$ equals that of Γ^i_a to the lowest order in $\delta = \rho/L$ while their fluctuating parts differ from each other.

It is emphasized here that $(J_{GKa})_j = 0$ in Eq. (91) is valid for $(J_{GKa})_j$ defined in Eq. (90) where the gyrocenter distribution function F_a can be arbitrarily chosen and it does not need to be a solution of the gyrokinetic Vlasov equation given by Eq. (48) or Eq. (70). It is recalled that the variation associated with the spatial coordinate transformation should be clearly distinguished from the variation (or virtual displacement) used for deriving the gyrokinetic Vlasov equation; the fact that the the former variation of the Lagrangian vanishes can be used to derive the momentum balance equation even in a more general case where the governing kinetic equation differ from the gyrokinetic Vlasov equation derived using the latter variational principle. We now assume F_a to satisfy not the gyrokinetic Vlasov equation, Eq. (48), but a more general one, that is, the gyrokinetic Boltzmann equation given by

$$\begin{aligned} D_t F_a &\equiv \frac{\partial F_a}{\partial t} + \frac{\partial}{\partial x^j}(F_a u^j_{ax}) + \frac{\partial}{\partial U}(F_a u_{aU}) + \frac{\partial}{\partial \mu}(F_a u_{a\mu}) \\ &\quad + \frac{\partial}{\partial \vartheta}(F_a u_{a\vartheta}) \\ &= \mathscr{K}_a, \end{aligned} \tag{102}$$

where $\mathscr{K}_a$ represents the rate of temporal change in F_a due to collisions and/or external sources for the species a. It is assumed in the present work that

$$\sum_a e_a \int d^3v\, \mathscr{K}_a = 0 \tag{103}$$

is satisfied by $\mathscr{K}_a$. Therefore, the charge density is not changed by $\mathscr{K}_a$. Using Eqs. (90), (92), (97)–(99), and (102), Eq. (91) implies that the solution of the gyrokinetic Boltzmann equation, Eq. (102), satisfies the canonical momentum balance equation,

$$\begin{aligned} &\frac{\partial}{\partial t}\left(\int d^3v\, F_a p_{aj}\right) - \int d^3v\, \mathscr{K}_a p_{aj} + \nabla_i (P_a)^i_j \\ &= -e_a N^{(p)}_a \nabla_j \phi + \frac{e_a}{c}\Big[\Gamma^k_{\# a}\nabla_j A_k + \Gamma^k_a \nabla_j \widehat{A}_k \\ &\quad - \nabla_k\left(\Gamma^k_{\# a} A_j + \Gamma^k_a \widehat{A}_j\right)\Big], \end{aligned} \tag{104}$$

which can be written in the conventional dyadic notation representing vectors and tensors in terms of boldface letters as

$$\begin{aligned} &\frac{\partial}{\partial t}\left(\int d^3v\, F_a \mathbf{p}_a\right) - \int d^3v\, \mathscr{K}_a \mathbf{p}_a + \nabla\cdot\mathbf{P}_a \\ &= -e_a N^{(p)}_a \nabla\phi + \frac{e_a}{c}\Big[(\nabla\mathbf{A})\cdot\mathbf{\Gamma}_{\# a} + (\nabla\widehat{\mathbf{A}})\cdot\mathbf{\Gamma}_a \\ &\quad - \nabla\cdot\left(\mathbf{\Gamma}_{\# a}\mathbf{A} + \mathbf{\Gamma}_a\widehat{\mathbf{A}}\right)\Big]. \end{aligned} \tag{105}$$

Here, as expected from Noether's theorem, one can confirm that Eq. (105) takes the conservation form of the canonical momentum in the direction of the constant vector $\mathbf{e}$ if the term including $\mathscr{K}_a$ vanishes and the electric and magnetic fields satisfy the symmetry conditions $\mathbf{e}\cdot\nabla\phi = 0$ and $\mathbf{e}\cdot\nabla\mathbf{A} = \mathbf{e}\cdot\nabla\widehat{\mathbf{A}} = 0$. In a case where the electric and magnetic fields are axisymmetric, conservation of the toroidal angular momentum can also be derived from Eq. (105) in the same manner as shown in Sec. VI.

The canonical momentum balance equation, Eq. (105), is also written as

$$\begin{aligned} &\frac{\partial}{\partial t}\left(\int d^3v\, F_a \mathbf{p}_a\right) - \int d^3v\, \mathscr{K}_a \mathbf{p}_a + \nabla\cdot\mathbf{P}_a \\ &= -e_a N^{(p)}_a \nabla\phi + \frac{e_a}{c}\Big[\left(\mathbf{\Gamma}_{\# a}\times\mathbf{B} + \mathbf{\Gamma}_a\times\widehat{\mathbf{B}}\right) \\ &\quad - \mathbf{A}(\nabla\cdot\mathbf{\Gamma}_{\# a}) - \widehat{\mathbf{A}}(\nabla\cdot\mathbf{\Gamma}_a)\Big]. \end{aligned} \tag{106}$$

Furthermore, Eq. (106) is deformed to

$$\begin{aligned} &\frac{\partial}{\partial t}\left(m_a N^{(g)}_a V_{ag\parallel}\mathbf{b}\right) - \int d^3v\, \mathscr{K}_a m_a U\mathbf{b} + \nabla\cdot\mathbf{P}_a \\ &= e_a\left(N^{(p)}_a \mathbf{E}_L + N^{(g)}_a \mathbf{E}_T\right) \\ &\quad + \frac{e_a}{c}\Big[\left(\mathbf{\Gamma}_{\# a}\times\mathbf{B} + \mathbf{\Gamma}_a\times\widehat{\mathbf{B}}\right) - \widehat{\mathbf{A}}(\nabla\cdot\mathbf{\Gamma}_a)\Big], \end{aligned} \tag{107}$$

where $\mathbf{E}_L = -\nabla\phi$ and $\mathbf{E}_T = -c^{-1}\partial\mathbf{A}/\partial t$ are used. Here,

$$N^{(g)}_a \equiv \int d^3v\, F_a \quad \text{and} \quad N^{(g)}_a V_{ag\parallel} \equiv \int d^3v\, F_a U \tag{108}$$

represent the density and the parallel flux of the gyrocenters, respectively. From Eqs. (100), (102), and (108), one can obtain

$$\frac{\partial N^{(g)}_a}{\partial t} + \nabla\cdot\mathbf{\Gamma}_{\# a} = \int d^3v\, \mathscr{K}_a, \tag{109}$$

which is used to derive Eq. (107) from Eq. (106). The first term on the left-hand side of Eq. (107) is the change rate of the density of the kinetic momentum which is obtained by extracting the vector potential term $\mathbf{p}_a \equiv (e_a/c)\mathbf{A} + m_a U \mathbf{b}$. The second and third terms on the left-hand side of Eq. (107) represent the effects of collisions (or external sources) and the pressure tensor, respectively, while the right-hand side contains Lorentz forces due to the electric and magnetic fields. The last term on the right hand side appears due to the perturbed vector potential and it is carried over from Eq. (106).

It is noted that the time derivative terms in Eqs. (105)–(107) are missing the perpendicular kinetic momentum part. This originates from the fact that the canonical momentum $\mathbf{p}_a$ associated with the gyrocenter Lagrangian does not include the perpendicular kinetic moment due to the perpendicular velocity $\mathbf{v}_\perp$ which depends on the gyrophase angle. The perpendicular part of the kinetic momentum density is given by $m_a \mathbf{\Gamma}_{a\perp}$ and its time derivative $m_a \partial \mathbf{\Gamma}_{a\perp}/\partial t$ is considered as neglected in the gyrokinetic momentum balance [Eqs. (105)–(107)] which the gyrokinetic Boltzmann equation, Eq. (102), satisfies. The leading order of the magnitude of terms in the perpendicular part of Eqs. (105)–(107) is given from the order of the Lorentz force terms and it is estimated to be $\mathscr{O}(e_a|\mathbf{\Gamma}_{a\perp}|B/c) = \mathscr{O}(m_a \Omega_a |\mathbf{\Gamma}_{a\perp}|)$ where $|\mathbf{\Gamma}_{\#a\perp}| \sim |\mathbf{\Gamma}_{a\perp}| \sim N_a^{(g)} c|\mathbf{E}_L|/B \sim (\rho_a/L) N_a^{(g)} v_{Ta}$ is regarded as valid for both the average and fluctuating parts. Thus, the neglected term $m_a \partial \mathbf{\Gamma}_{a\perp}/\partial t$ in the perpendicular momentum balance is smaller than the leading-order terms by a factor $\mathscr{O}(\Omega_a^{-1}\partial/\partial t)$. Here, the transport time scale ordering gives $\Omega_a^{-1}\partial/\partial t \sim (\rho_a/L)^3$ for the ensemble-averaged part while $\Omega_a^{-1}\partial/\partial t \sim \rho_a/L$ is obtained for the fluctuating part from the gyrokinetic ordering. Thus, neglecting $m_a \partial \mathbf{\Gamma}_{a\perp}/\partial t$ is not considered to give a significant influence on the perpendicular part of the local momentum balance although this higher-order term should be correctly included for accurately describing the flux-surface-averaged momentum balance along the symmetry direction in up-down symmetric tokamaks and stellarator-symmetric quasisymmetric stellarators[51–53].

C. Momentum balance for the whole system

One can follow the same procedures as used in deriving Eq. (87) to deform Eq. (85) to

$$\begin{aligned}\overline{\delta} L_{GKF} &= \int_V d^3x \left[\xi^i \sum_a \int d^3v \left\{ \frac{\partial}{\partial t}\left(F_a \frac{\partial L_{GYa}}{\partial u^i_{ax}} \right) - D_t F_a \frac{\partial L_{GYa}}{\partial u^i_{ax}} \right\} + \frac{\delta L_{GKF}}{\delta A_i} \overline{\delta} A_i + \frac{\delta L_{GKF}}{\delta \widehat{A}_i} \overline{\delta} \widehat{A}_i \right. \\ &\quad \left. + \frac{\delta L_{GKF}}{\delta \phi} \overline{\delta}\phi + \frac{\delta L_{GKF}}{\delta g_{ij}} \overline{\delta} g_{ij} \right] + \text{B.T.} \\ &= 0. \end{aligned} \tag{110}$$

Substituting $\overline{\delta} A_i = -\xi^j (\nabla_j A_i) - (\nabla_i \xi^j) A_j$, $\overline{\delta}\widehat{A}_i = -\xi^j (\nabla_j \widehat{A}_i) - (\nabla_i \xi^j)\widehat{A}_j$, $\overline{\delta}\phi = -\xi^j \nabla_j \phi$ into Eq. (110), and performing a partial integral, one obtains

$$\overline{\delta} L_{GKF} = \int_V d^3x\, \xi^j (J_{GKF})_j + \text{B.T.} = 0, \tag{111}$$

where

$$\begin{aligned}(J_{GKF})_j &\equiv \frac{\partial}{\partial t}\left(\sum_a \int d^3v\, F_a \frac{\partial L_{GYa}}{\partial u^i_{ax}} \right) - \sum_a \int d^3v\, D_t F_a \frac{\partial L_{GYa}}{\partial u^i_{ax}} \\ &\quad - \frac{\delta L_{GKF}}{\delta \phi} \nabla_j \phi - \frac{\delta L_{GKF}}{\delta A_i} \nabla_j A_i - \frac{\delta L_{GKF}}{\delta \widehat{A}_i} \nabla_j \widehat{A}_i \\ &\quad + \nabla_i \left(\frac{\delta L_{GKF}}{\delta A_i} A_j + \frac{\delta L_{GKF}}{\delta \widehat{A}_i} \widehat{A}_j \right) + 2\nabla_i \left(g_{jk} \frac{\delta L_{GKF}}{\delta g_{ik}} \right). \end{aligned} \tag{112}$$

Since Eq. (111) is valid for any ξ^j which vanishes on the boundary of V,

$$(J_{GKF})_j = 0 \tag{113}$$

holds for $(J_{GKF})_j$ defined in Eq. (112) where F_a, ϕ, A_i, and $\widehat{A}_i$ can be arbitrarily chosen and they do not need to be determined by any governing equations.

When F_a, ϕ, and $\widehat{A}_i$ satisfy the gyrokinetic Boltzmann equation shown in Eq. (102) and the gyrokinetic Poisson-Ampère equations given by $\delta L_{GKF}/\delta\phi = 0$ and $\delta L_{GKF}/\delta \widehat{A}_i = 0$, Eqs. (112) and (113) lead to the total canonical momentum balance equation,

$$\begin{aligned}&\frac{\partial}{\partial t}\left(\sum_a \int d^3v\, F_a p_{aj} \right) - \sum_a \int d^3v\, \mathscr{K}_a p_{aj} + \nabla_i \Theta^i_j \\ &\quad = \frac{\delta L_{GKF}}{\delta A_i} \nabla_j A_i - \nabla_i \Big(\frac{\delta L_{GKF}}{\delta A_i} A_j \Big), \end{aligned} \tag{114}$$

where $p_{aj} \equiv \partial L_{GYa}/\partial u^j_{ax} \equiv (e_a/c) A_j(x,t) + m_a U b_j(x,t)$ and the total pressure tensor density $\Theta^{ij} \equiv \Theta^i_k g^{jk}$ is defined by

$$\Theta^{ij} \equiv 2 \frac{\delta L_{GKF}}{\delta g_{ij}} \equiv \Theta^{ij}_{GK} + \Theta^{ij}_F. \tag{115}$$

The gyrokinetic part Θ^{ij}_{GK} of Θ^{ij} is written as

$$\Theta^{ij}_{GK} \equiv 2 \frac{\delta L_{GK}}{\delta g_{ij}} \equiv 2 \sum_a \frac{\delta L_{GKa}}{\delta g_{ij}} = P^{ij}_{\rm CGL} + \pi^{ij}_\wedge + \pi^{ij}_{\|\Psi} + P^{ij}_\Psi, \tag{116}$$

where $P^{ij}_{\rm CGL}$, $\pi^{ij}_\wedge$, $\pi^{ij}_{\|\Psi}$, and P^{ij}_Ψ are defined using Eqs. (93)–(96) as

$$\left[P^{ij}_{\rm CGL}, \pi^{ij}_\wedge, \pi^{ij}_{\|\Psi}, P^{ij}_\Psi \right] \equiv \sum_a \left[P^{ij}_{\rm CGLa}, \pi^{ij}_{\wedge a}, \pi^{ij}_{\|\Psi a}, P^{ij}_{\Psi a} \right]. \tag{117}$$

The field part Θ^{ij}_F of Θ^{ij} is given by

$$\begin{aligned}\Theta^{ij}_F &\equiv 2 \frac{\delta L_F}{\delta g_{ij}} \equiv 2 \sum_J (-1)^{\# J} \frac{\partial \mathscr{L}_F}{\partial(\partial_J g_{ij})} \\ &= \frac{\sqrt{g}}{4\pi} \left[\frac{g^{ij}}{2} \left\{ (E_L)^k (E_L)_k + (B^k + \widehat{B}^k)(B_k + \widehat{B}_k) \right\} \right. \\ &\quad - \left\{ (E_L)^i (E_L)^j + (B^i + \widehat{B}^i)(B^j + \widehat{B}^j) \right\} \\ &\quad \left. + \frac{1}{c} \left\{ -g^{ij} (\nabla_k \lambda) \widehat{A}^k + (\nabla^i \lambda) \widehat{A}^j + (\nabla^j \lambda) \widehat{A}^i \right\} \right], \end{aligned} \tag{118}$$

which contains the well-known Maxwell stress tensor and the additional terms in the same form as found in the Vlasov-Darwin model.[45]

Now, using Cartesian spatial coordinates and the conventional dyadic notation representing vectors and tensors in terms of boldface letters, Eq. (114) is rewritten as

$$\frac{\partial}{\partial t}\left(\sum_a \int d^3v\, F_a \mathbf{p}_a\right) - \sum_a \int d^3v\, \mathscr{K}_a \mathbf{p}_a + \nabla\cdot\boldsymbol{\Theta} = (\nabla\mathbf{A})\cdot\frac{\delta L_{GKF}}{\delta\mathbf{A}} - \nabla\cdot\left(\frac{\delta L_{GKF}}{\delta\mathbf{A}}\mathbf{A}\right). \tag{119}$$

Here, $\delta L_{GKF}/\delta\mathbf{A}$ is given by

$$\frac{\delta L_{GKF}}{\delta\mathbf{A}} = \frac{\mathbf{j}_\#}{c} - \frac{1}{4\pi}\nabla\times(\mathbf{B}+\widehat{\mathbf{B}}), \tag{120}$$

where $\mathbf{j}_\#$ is defined from $\boldsymbol{\Gamma}_{\#a}$ in Eq. (100) by

$$\mathbf{j}_\# \equiv c\frac{\delta L_{GK}}{\delta\mathbf{A}} \equiv \sum_a e_a \boldsymbol{\Gamma}_{\#a} \equiv \mathbf{j}^{(gc)} + c\nabla\times\mathbf{M}_\#. \tag{121}$$

Here, $\mathbf{j}^{(gc)}$ is the gyrocenter current defined by Eq. (C12) in Appendix C and $\mathbf{M}_\#$ is given by

$$\mathbf{M}_\# \equiv \sum_a e_a \int d^3v\, F_a\left[-\mu\mathbf{b} + \frac{m_a U}{B}(\mathbf{u}_{ax})_\perp - e_a\left\{\frac{\partial\Psi_a}{\partial\mathbf{B}} - \frac{1}{F_a}\frac{\partial}{\partial x^l}\left(F_a\frac{\partial\Psi_a}{\partial(\partial_l\mathbf{B})}\right)\right\}\right]. \tag{122}$$

Now, one can confirm the validity of Noether's theorem again from Eq. (119) which takes the conservation form of the total canonical momentum in the direction specified by the constant vector $\mathbf{e}$ when the background magnetic field satisfies the symmetry condition $\mathbf{e}\cdot\nabla\mathbf{A} = 0$ and $\sum_a \int d^3v\, \mathscr{K}_a \mathbf{p}_a$ can be ignored. In Sec. VI, toroidal angular momentum conservation is derived in the case of the axisymmetric background field. It should be noted that no specific conditions to determine $\mathbf{A}$ are imposed from the variational principle in contrast to ϕ and $\widehat{\mathbf{A}}$ which are variationally determined. Thus, $\delta L_{GKF}/\delta\mathbf{A}$ given in Eq. (120) does not vanish generally.

The total canonical momentum balance equation in Eq. (119) can be rewritten as

$$\frac{\partial}{\partial t}\left(\sum_a \int d^3v\, F_a \mathbf{p}_a\right) - \sum_a \int d^3v\, \mathscr{K}_a \mathbf{p}_a + \nabla\cdot\boldsymbol{\Theta} = \frac{\delta L_{GKF}}{\delta\mathbf{A}}\times\mathbf{B} - \mathbf{A}\,\nabla\cdot\left(\frac{\delta L_{GKF}}{\delta\mathbf{A}}\right), \tag{123}$$

which is also deformed to

$$\frac{\partial}{\partial t}\left(\sum_a \int d^3v\, F_a m_a U\mathbf{b}\right) - \sum_a \int d^3v\, \mathscr{K}_a m_a U\mathbf{b} + \nabla\cdot\boldsymbol{\Theta} = \rho_c^{(gc)}\mathbf{E}_T + \frac{\delta L_{GKF}}{\delta\mathbf{A}}\times\mathbf{B}, \tag{124}$$

where $\rho_c^{(gc)} \equiv \sum_a e_a N_a^{(g)}$ and $\mathbf{E}_T \equiv -c^{-1}\partial\mathbf{A}/\partial t$. Equation (124) represents the total balance equation of the kinetic momentum instead of the canonical one. The effects of collisions (or external sources) and the total pressure tensor are shown on the left-hand side of Eq. (124) while the Lorentz forces due to the back ground inductive field and the background magnetic field appear on the right-hand side.

Finally, Eq. (124) can be deformed through vector calculus to

$$\frac{\partial}{\partial t}\left(\sum_a \int d^3v\, F_a m_a U\mathbf{b} + \frac{1}{4\pi c}(\mathbf{D}_L\times\mathbf{B})\right) + \nabla\cdot\left(\boldsymbol{\Theta} + \frac{\mathbf{D}_L\mathbf{E}_T + \mathbf{E}_T\mathbf{D}_L}{4\pi}\right) + \nabla\left(\frac{\mathbf{E}_T\cdot\mathbf{D}_L}{4\pi}\right) = \sum_a \int d^3v\, \mathscr{K}_a m_a U\mathbf{b} + \left(\frac{\delta L_{GKF}}{\delta\mathbf{A}}\right)_T\times\mathbf{B}, \tag{125}$$

which is written in more detail as

$$\begin{aligned}
&\frac{\partial}{\partial t}\left(\sum_a \int d^3v\, F_a m_a U\mathbf{b} + \frac{1}{4\pi c}(\mathbf{D}_L\times\mathbf{B})\right) \\
&+\nabla\cdot\left(\mathbf{P}_{\mathrm{CGL}} + \boldsymbol{\pi}_\wedge + \boldsymbol{\pi}_{\parallel\Psi} + \mathbf{P}_\Psi\right) + \nabla\left(\frac{|\mathbf{E}_L|^2}{8\pi} + \frac{\mathbf{E}_T\cdot\mathbf{D}_L}{4\pi}\right) \\
&-\nabla\cdot\left(\frac{\mathbf{E}_L\mathbf{E}_L + \mathbf{D}_L\mathbf{E}_T + \mathbf{E}_T\mathbf{D}_L}{4\pi}\right) \\
&+\nabla\left(\frac{|\mathbf{B}+\widehat{\mathbf{B}}|^2}{8\pi}\right) - \nabla\cdot\left(\frac{(\mathbf{B}+\widehat{\mathbf{B}})(\mathbf{B}+\widehat{\mathbf{B}})}{4\pi}\right) \\
&-\nabla\left(\frac{\nabla\lambda\cdot\widehat{\mathbf{A}}}{4\pi c}\right) + \nabla\cdot\left(\frac{(\nabla\lambda)\widehat{\mathbf{A}} + \widehat{\mathbf{A}}(\nabla\lambda)}{4\pi c}\right) \\
&= \sum_a \int d^3v\, \mathscr{K}_a m_a U\mathbf{b} + \left(\frac{(\mathbf{j}_\#)_T}{c} - \frac{\nabla\times(\mathbf{B}+\widehat{\mathbf{B}})}{4\pi}\right)\times\mathbf{B}.
\end{aligned} \tag{126}$$

In Eq. (126), $\mathbf{D}_L$ is the longitudinal part of the displacement vector defined by Eq. (C4) and $\mathbf{j}_\#$ is defined in Eq. (121). The change rate of the kinetic momentum density plus the electromagnetic momentum density $(\mathbf{D}_L\times\mathbf{B})/(4\pi c)$ is described by Eq. (126). The left-hand side of Eq. (126) shows all terms of momentum transport written as the divergence of pressure tensors due to particles' motion and Maxwell stresses including both average and fluctuating parts of the electromagnetic field. Except for the terms on the right-hand side, Eq. (126) takes the conservation form similar to that of the total momentum conservation equation of the Vlasov-Darwin model derived in Refs.[45]. Since $\mathbf{j}_\# = \sum_a e_a\boldsymbol{\Gamma}_{\#a}$ does not accurately represents the fluctuating part of the current density, $(\mathbf{j}_\#)_T/c - (4\pi)^{-1}\nabla\times(\mathbf{B}+\widehat{\mathbf{B}})$ cannot be neglected as far as turbulent electromagnetic fields exist. In Sec. VII, the self-consistency condition to determine the average field $\mathbf{B}$ is considered to make the ensemble average of Eq. (126) take the conservation form.

VI. MOMENTUM BALANCE IN TOROIDAL SYSTEMS

In this section, we investigate the momentum balance in toroidal systems based on the results obtained in Sec. V. The background magnetic field $\mathbf{B}$ with and without symmetry are considered and the momentum balance equations averaged over the ensemble and the flux surface are examined. Here, the background magnetic field $\mathbf{B}$ is assumed to satisfy the toroidal MHD equilibrium equation,

$$\frac{1}{4\pi}(\nabla\times\mathbf{B})\times\mathbf{B} = \nabla P_0, \tag{127}$$

where P_0 is a magnetic flux surface function representing equilibrium pressure.

A. Axisymmetric systems

The axisymmetric toroidal background magnetic field is represented by

$$\mathbf{B} = I\nabla\zeta + \nabla\zeta\times\nabla\chi, \tag{128}$$

where ζ and χ represents the toroidal angle and the poloidal flux (divided by 2π), respectively, and the covariant toroidal component I is a flux surface function which is independent of the toroidal and poloidal angles. Denoting the major radius by R and writing the contravariant basis vector $\mathbf{e}_\zeta$ in the toroidal direction as

$$\mathbf{e}_\zeta \equiv R^2\nabla\zeta \equiv \frac{\partial\mathbf{x}}{\partial\zeta}, \tag{129}$$

one obtains the following relation,

$$\nabla\mathbf{e}_\zeta = R^{-1}[(\nabla R)\mathbf{e}_\zeta - \mathbf{e}_\zeta(\nabla R)] = \mathbf{n}\times\mathbf{I} = \mathbf{I}\times\mathbf{n}, \tag{130}$$

where $\mathbf{n}\equiv R^{-1}(\mathbf{e}_\zeta\times\nabla R)$ is the unit vector parallel to the direction of the major axis and $\mathbf{I}$ is the unit tensor. It is shown from Eq. (130) that an arbitrary symmetric tensor $\mathbf{S}$ ($S^{ij}=S^{ji}$) satisfies

$$(\nabla\cdot\mathbf{S})\cdot\mathbf{e}_\zeta = \nabla\cdot(\mathbf{S}\cdot\mathbf{e}_\zeta). \tag{131}$$

It is also noted that $\nabla\cdot\mathbf{e}_\zeta = 0$ and

$$\mathbf{e}_\zeta\cdot\nabla S = \nabla\cdot(S\mathbf{e}_\zeta), \tag{132}$$

where S is an arbitrary scalar function.

In the axisymmetric background field $\mathbf{B}$, $\mathbf{A}$ can also be given by the axisymmetric field which satisfies

$$\mathbf{e}_\zeta\cdot\nabla\mathbf{A} = \frac{1}{R}\mathbf{A}\times\mathbf{n}. \tag{133}$$

Then, the inner product of $\mathbf{e}_\zeta$ and Eq. (114) is taken and Eqs. (130), (131), and (133) are used to derive

$$\frac{\partial}{\partial t}\left(\sum_a\int d^3v\,F_a\mathbf{p}_a\cdot\mathbf{e}_\zeta\right) + \nabla\cdot\left[\left(\mathbf{\Theta}+\frac{\delta L_{GKF}}{\delta\mathbf{A}}\mathbf{A}\right)\cdot\mathbf{e}_\zeta\right] = \sum_a\int d^3v\,\mathscr{K}_a\mathbf{p}_a\cdot\mathbf{e}_\zeta. \tag{134}$$

Except for the the right-hand side, Eq. (134) takes the conservation form of the canonical momentum conjugate to the toroidal angle as expected from Noether's theorem. It is found from the assumption given in Eq. (103) that $\sum_a\int d^3v\,\mathscr{K}_a\mathbf{p}_a\cdot\mathbf{e}_\zeta = \sum_a\int d^3v\,\mathscr{K}_a m_a U b_\zeta$ where $b_\zeta = \mathbf{b}\cdot\mathbf{e}_\zeta$. In the zero-gyroradius limit, when using a particle collision operator for $\mathscr{K}_a$, one finds that $\sum_a\int d^3v\,\mathscr{K}_a m_a U b_\zeta$ to vanish because the momentum of particles is conserved in collisions. Furthermore, it can be shown that when $\mathscr{K}_a$ is given by the collision operator which appropriately includes the finite gyroradius effect,[57–59] $\sum_a\int d^3v\,\mathscr{K}_a m_a U b_\zeta$ can be written as a divergence of the sum of classical momentum transport fluxes. Therefore, without external momentum sources, Eq. (134) keeps the conservation form even though the collision term is present. In addition to the case of axisymmetry, the canonical momentum conservation is confirmed in other cases of symmetry under continuous isometric transformations such as a translational symmetry and a helical (or screw) symmetry.[37]

Next, the toroidal component of the momentum balance equation in Eq. (126) is considered. The transverse part of $\mathbf{j}_\#$ on the right-hand side of Eq. (126) can be written in terms of a certain field $\mathbf{B}_\#$ as

$$(\mathbf{j}_\#)_T = \frac{c}{4\pi}\nabla\times\mathbf{B}_\# \tag{135}$$

which is combined with $\mathbf{B}\times\mathbf{e}_\zeta = \nabla\chi$ to derive

$$\big((\mathbf{j}_\#)_T\times\mathbf{B}\big)\cdot\mathbf{e}_\zeta = (\mathbf{j}_\#)_T\cdot\nabla\chi = \nabla\cdot\left(\frac{c}{4\pi}\mathbf{B}_\#\times\nabla\chi\right). \tag{136}$$

One also obtains

$$\begin{aligned}\Big(\big(\nabla\times(\mathbf{B}+\widehat{\mathbf{B}})\big)\times\mathbf{B}\Big)\cdot\mathbf{e}_\zeta &= \Big(\nabla\times(\mathbf{B}+\widehat{\mathbf{B}})\Big)\cdot\nabla\chi\\ &= \nabla\cdot\Big((\mathbf{B}+\widehat{\mathbf{B}})\times\nabla\chi\Big).\end{aligned} \tag{137}$$

Now, taking the inner product of Eq. (126) and $\mathbf{e}_\zeta$, it is found that the toroidal angular momentum balance equation can be written in the following form,

$$\begin{aligned}&\frac{\partial}{\partial t}\left(\sum_a\int d^3v\,F_a m_a U b_\zeta + \frac{1}{4\pi c}\mathbf{D}_L\cdot\nabla\chi\right)\\ &\quad+\nabla\cdot\left[\left(\mathbf{\Theta}+\frac{\mathbf{D}_L\mathbf{E}_T+\mathbf{E}_T\mathbf{D}_L}{4\pi}\right)\cdot\mathbf{e}_\zeta + \frac{\mathbf{E}_T\cdot\mathbf{D}_L}{4\pi}\mathbf{e}_\zeta\right]\\ &\quad+\nabla\cdot\left[\frac{1}{4\pi}(\mathbf{B}_\#-\mathbf{B}-\widehat{\mathbf{B}})\times\nabla\chi\right]\\ &\quad=\sum_a\int d^3v\,\mathscr{K}_a m_a U b_\zeta,\end{aligned} \tag{138}$$

where $(\mathbf{D}_L\times\mathbf{B})\cdot\mathbf{e}_\zeta = \mathbf{D}_L\cdot\nabla\chi$ is used. The expression of the divergence term on the left-hand side of Eq. (138) is straightforwardly derived from Eq. (126) using Eqs. (131), (132), (136), and (137). Without external momentum sources, Eq. (138) keeps the conservation form in the same way as Eq. (134).

It is shown above that the symmetry of the pressure tensor is essential in deriving the equation of the toroidal angular momentum conservation in an axisymmetric system. From this point of view, the derivation of the symmetric pressure tensor from the variational derivative of the Lagrangian with respect to the metric tensor is useful.

B. Non-axisymmetric systems

In non-axisymmetric toroidal systems, the background field $\mathbf{B}$ is expressed by

$$\mathbf{B} = \nabla\psi(s) \times \nabla\theta + \nabla\zeta \times \nabla\chi(s), \tag{139}$$

where θ and ζ are the poloidal and toroidal angles, respectively, and s is an arbitrarily chosen flux-surface label. Here, we assume the background $\mathbf{B}$ to be stationary, $\partial\mathbf{B}/\partial t = 0$, for simplicity, and use the Hamada coordinates[60] (s, θ, ζ), in which the Jacobian $\sqrt{g} \equiv [\nabla s \cdot (\nabla\theta \times \zeta)]^{-1}$, the poloidal field $B^\theta \equiv \mathbf{B}\cdot\nabla\theta$, and the toroidal field $B^\zeta \equiv \mathbf{B}\cdot\nabla\zeta$ are flux-surface functions. Then, the contravariant basis vector $\mathbf{e}_\zeta$ in the toroidal direction is written as

$$\mathbf{e}_\zeta \equiv \frac{\partial\mathbf{x}}{\partial\zeta} \equiv \sqrt{g}\nabla s \times \nabla\theta, \tag{140}$$

which satisfies $\nabla\cdot\mathbf{e}_\zeta = 0$ and $\mathbf{B}\times\mathbf{e}_\zeta = \nabla\chi$. Therefore, equations in the same form as Eqs. (132), (136), and (137) hold while Eq. (131) does not because Eq. (130) is not valid in the non-axisymmetric case. Now, taking the inner product of Eq. (126) and $\mathbf{e}_\zeta$ gives the toroidal angular momentum balance equation in the following form,

$$\begin{aligned}&\frac{\partial}{\partial t}\left(\sum_a \int d^3v F_a m_a U b_\zeta + \frac{1}{4\pi c}\mathbf{D}_L\cdot\nabla\chi\right)\\ &\quad + \nabla\cdot(\cdots) + \mathbf{e}_\zeta\cdot\big(\nabla\cdot(\mathbf{P}_{\rm CGL}+\cdots)\big)\\ &= 0,\end{aligned} \tag{141}$$

where the effects of external momentum sources are ignored. The conservation form is broken in Eq. (141) because of $\mathbf{e}_\zeta\cdot\big(\nabla\cdot(\mathbf{P}_{\rm CGL}+\cdots)\big)$ on the left-hand side of Eq. (141). Here, the CGL and other pressure tensors and anisotropic Maxwell stress terms are included in $\mathbf{e}_\zeta\cdot\big(\nabla\cdot(\mathbf{P}_{\rm CGL}+\cdots)\big)$ while, using Eq. (132), the tensors proportional to the unit tensor can be transferred to the inside of the divergence term $\nabla\cdot(\cdots)$ in Eq. (141). Using the flux-surface average defined by $\langle\cdots\rangle \equiv \oint d\theta\oint d\zeta\,\sqrt{g}\cdots/V'(s)$ with $V'(s) \equiv \oint d\theta\oint d\zeta\,\sqrt{g}$, it is found that the divergence term $\nabla\cdot(\cdots)$ in Eq. (141) is annihilated by the flux-surface average because, as seen from Eqs. (132), (136), (137), and (140), the inner products of the vectors in $(\cdots)$ and ∇s vanish and

$$\langle\nabla\cdot\mathbf{T}\rangle = \frac{1}{V'(s)}\frac{d}{ds}\big(V'(s)\langle\mathbf{T}\cdot\nabla s\rangle\big) \tag{142}$$

holds for any vector $\mathbf{T}$. Then, taking the flux-surface average, Eq. (141) is reduced to the more compact form,

$$\begin{aligned}&\frac{\partial}{\partial t}\left\langle\sum_a \int d^3v F_a m_a U b_\zeta + \frac{1}{4\pi c}\mathbf{D}_L\cdot\nabla\chi\right\rangle\\ &\quad + \big\langle\mathbf{e}_\zeta\cdot\big(\nabla\cdot(\mathbf{P}_{\rm CGL}+\cdots)\big)\big\rangle\\ &= 0.\end{aligned} \tag{143}$$

It is recalled that, in neoclassical theory for non-axisymmetric systems,[63] the lowest-order toroidal viscosity is given by $\big\langle\mathbf{e}_\zeta\cdot(\nabla\cdot\mathbf{P}_{\rm CGL})\big\rangle$. It is shown in Sec. VII that, when using the ensemble average and the gyroradius expansion of Eq. (143) in general non-axisymmetric toroidal systems, this neoclassical toroidal viscosity becomes a dominant term.

C. Quasi-axisymmetric systems

In this subsection, quasi-axisymmetric toroidal systems[35] are considered using Eq. (143) with the Hamada coordinates (s, θ, ζ) to represent the equilibrium magnetic field $\mathbf{B}$. The quasi-axisymmetry is characterized by $B = |\mathbf{B}|$ being independent of the toroidal angle, $\partial B/\partial\zeta = 0$, which is equivalent to $\partial B/\partial\zeta_B = 0$ in the Boozer coordinates[64] (s, θ_B, ζ_B) as proved in Ref.[65]. In the quasi-axisymmetric equilibrium field $\mathbf{B}$, the CGL-type pressure tensor $\mathbf{P}_{\rm CGL} \equiv P_\parallel\mathbf{bb} + P_\perp(\mathbf{I}-\mathbf{bb})$ is shown to satisfy

$$\big\langle\mathbf{e}_\zeta\cdot(\nabla\cdot\mathbf{P}_{\rm CGL})\big\rangle = \left\langle(P_\perp - P_\parallel)\frac{\partial\ln B}{\partial\zeta}\right\rangle = 0, \tag{144}$$

which implies that the dominant neoclassical toroidal viscosity which exists in general non-axisymmetric systems vanishes as in the axisymmetric systems. Thus, one of the factors preventing conservation of the toroidal angular momentum in Eq. (143) disappears even though perfect conservation is not allowed. Because of Eq. (144), the magnitude of the dominant terms in Eq. (143) becomes of higher order. Then, as mentioned in Sec. II, basic gyrokinetic equations including higher-order terms, which are not considered in the present study, are required to accurately describe the flux-surface-averaged momentum balance along the symmetry direction in stellarator-symmetric quasisymmetric stellarators as well as up-down symmetric tokamaks.[51–53]

VII. ENSEMBLE-AVERAGED MOMENTUM BALANCE

In this section, the momentum balance equation in Eq. (126) is ensemble-averaged, by which all terms in the equation are smoothed to make their space-time scales of variations much larger than those of fluctuations. The ensemble average is used as the basic method of statistical mechanics to obtain the macroscopic mean values of physical valuables and it can also be considered to equal the local space–time average, the definition of which is described in detail in Ref.[66]. The fluctuations are assumed to have wavelengths of the order of gyroradii in the directions perpendicular to the background magnetic field, and they are treated by the WKB representation in Sec. VIII.

Since the background magnetic field $\mathbf{B}$ is considered to include no fluctuations, one can write $\mathbf{B} = \langle\mathbf{B}\rangle_{\rm ens}$, where $\langle\cdots\rangle_{\rm ens}$ represents the ensemble average. It should also be noted here that no equation to determine the background magnetic field $\mathbf{B}$ is given from the variational principle while $\mathbf{B}$ is allowed to change with time in the present model. If the variational condition $\delta L_{GKF}/\delta\mathbf{A} = 0$ was employed, $\mathbf{B}$ would include fluctuation components as seen from Eq. (120). Then, $\langle(\delta L_{GKF}/\delta\mathbf{A})_T\rangle_{\rm ens} = 0$ is assumed here instead of $\delta L_{GKF}/\delta\mathbf{A} = 0$ as the condition for determining $\mathbf{B}$. Using Eq. (120), $\langle(\delta L_{GKF}/\delta\mathbf{A})_T\rangle_{\rm ens} = 0$ is written as

$$\nabla\times\mathbf{B} = \frac{4\pi}{c}\langle(\mathbf{j}_\#)_T\rangle_{\rm ens}, \tag{145}$$

which seems to be appropriate because the equilibrium part of $\mathbf{j}_\# \equiv c\delta L_{GK}/\delta\mathbf{A}$ [see Eq. (121)] is equal to that of $\mathbf{j} \equiv c\delta L_{GK}/\delta\widehat{\mathbf{A}}$ [see Eq. (D1)] to the lowest order in the gyroradius expansion and it represents the current density consistent with the MHD equilibrium. In the neoclassical transport theory[49], the self-consistency condition for the background field $\mathbf{B}$ evolving in the transport time scale is given from the MHD equilibrium equation. In passing, the condition for $\mathbf{B}$ can be derived as the variational equation in the drift kinetic model[44,61] and the variational derivation of the time-evolving axisymmetric background field is considered in the gyrokinetic model[21,62].

With the help of Eq. (121), Eq. (145) can be rewritten as

$$\nabla \times \langle \mathbf{H}_\# \rangle_{\rm ens} = \frac{4\pi}{c} \langle (\mathbf{j}^{(gc)})_T \rangle_{\rm ens}, \tag{146}$$

Here, $\mathbf{j}^{(gc)}$ is the gyrocenter current given by Eq. (C12) and $\mathbf{H}_\#$ is defined using $\mathbf{M}_\#$ in Eq. (122) as

$$\mathbf{H}_\# \equiv \mathbf{B} + \widehat{\mathbf{B}} - 4\pi \mathbf{M}_\#, \tag{147}$$

from which one has

$$\langle \mathbf{H}_\# \rangle_{\rm ens} = \mathbf{B} - 4\pi \langle \mathbf{M}_\# \rangle_{\rm ens}. \tag{148}$$

Using Eq. (145), the ensemble average of Eq. (126) is written as

$$\begin{aligned} &\frac{\partial}{\partial t} \left\langle \sum_a \int d^3v F_a m_a U \mathbf{b} + \frac{1}{4\pi c} (\mathbf{D}_L \times \mathbf{B}) \right\rangle_{\rm ens} \\ &+ \nabla \cdot \langle \mathbf{P}_{\rm CGL} + \boldsymbol{\pi}_\wedge + \boldsymbol{\pi}_{\parallel\Psi} + \mathbf{P}_\Psi \rangle_{\rm ens} + \nabla \left\langle \frac{|\mathbf{E}_L|^2}{8\pi} + \frac{\mathbf{E}_T \cdot \mathbf{D}_L}{4\pi} \right\rangle_{\rm ens} \\ &- \nabla \cdot \left\langle \frac{\mathbf{E}_L \mathbf{E}_L + \mathbf{D}_L \mathbf{E}_T + \mathbf{E}_T \mathbf{D}_L}{4\pi} \right\rangle_{\rm ens} \\ &+ \nabla \left\langle \frac{|\mathbf{B} + \widehat{\mathbf{B}}|^2}{8\pi} \right\rangle_{\rm ens} - \nabla \cdot \left\langle \frac{(\mathbf{B} + \widehat{\mathbf{B}})(\mathbf{B} + \widehat{\mathbf{B}})}{4\pi} \right\rangle_{\rm ens} \\ &- \nabla \left\langle \frac{\nabla\lambda \cdot \widehat{\mathbf{A}}}{4\pi c} \right\rangle_{\rm ens} + \nabla \cdot \left\langle \frac{(\nabla\lambda)\widehat{\mathbf{A}} + \widehat{\mathbf{A}}(\nabla\lambda)}{4\pi c} \right\rangle_{\rm ens} \\ &= \sum_a \left\langle \int d^3v \mathscr{K}_a m_a U \mathbf{b} \right\rangle_{\rm ens}. \end{aligned} \tag{149}$$

The ensemble-averaged momentum balance equation, Eq. (149), takes the conservation form when no external sources of momentum exist and the right-hand side is written as a divergence of the tensor representing classical momentum transport. It is emphasized here that the ensemble-averaged momentum conservation described above is satisfied even in non-axisymmetric systems when the background field $\mathbf{B}$ is determined by the equilibrium condition given in Eq. (145). It is interesting to compare Eq. (149) with the momentum conservation law shown by Eqs. (31)–(33) in Ref.[45] for the Vlasov-Poisson-Ampère (or Vlasov-Darwin) system in which collisional effects are ignored and the magnetic field is not divided into background and turbulent parts. One can see that kinetic and electromagnetic momenta, kinetic and electromagnetic pressure tensors, and longitudinal and transverse electric fields in the momentum conservation equation of the Vlasov-Darwin system appear in Eq. (149) in a similar manner and that Eq. (149) additionally includes polarization, magnetization, and other higher-order terms due to finite-gyroradius effects and electromagnetic microturbulence. The similarities and differences described above are regarded as natural results because the electromagnetic gyrokinetic systems are derived from the Vlasov-Darwin system through ordering assumptions regarding gyroradius scales and fluctuation amplitudes.

Hereafter in this section, Eq. (149) is expanded using the ordering parameter given by the normalized gyroradius $\delta = \rho/L$ which is the ratio of the gyroradius ρ to the equilibrium scale length L. The zeroth-order part F_{a0} of the distribution function is assumed to be given by the local Maxwellian as

$$\begin{aligned} F_{a0} = \langle F_{a0} \rangle_{\rm ens} = F_{aM} \equiv D_{a0} f_{aM} \\ \equiv N_{a0} D_{a0} \left(\frac{m_a}{2\pi T_{a0}} \right)^{3/2} \exp\left[-\frac{1}{T_{a0}} \left(\frac{1}{2} m_a U^2 + \mu B \right) \right], \end{aligned} \tag{150}$$

where N_{a0} and T_{a0} are the background density and temperature of the particle species a, respectively, and $D_{a0} \equiv B/m_a$ is the zeroth-order part of $D_a \equiv B^*_{a\parallel}/m_a$. Here, the transport time scale ordering is used for the ensemble-averaged variables in Eq. (149) which means that $\partial\langle\cdots\rangle_{\rm ens}/\partial t \sim \delta^2 (v_T/L) \langle\cdots\rangle_{\rm ens}$ where v_T is thermal velocity. The CGL pressure tensor defined in Eq. (93) is expanded in $\delta = \rho/L$ as

$$\mathbf{P}_{{\rm CGL}a} = P_{a0}\mathbf{I} + (\mathbf{P}_{{\rm CGL}a})_1 + \mathscr{O}(\delta^2), \tag{151}$$

where the zeroth-order part is isotropic and expressed by the scalar pressure, $P_{a0} \equiv N_{a0} T_{a0}$. Then, it is found that the zeroth-order part of the ensemble-averaged momentum conservation equation, Eq. (149), is given by

$$\nabla \left(P_0 + \frac{B^2}{8\pi} \right) - \frac{1}{4\pi} \nabla \cdot (\mathbf{BB}) = 0 \tag{152}$$

where the equilibrium pressure is defined by $P_0 \equiv \sum_a P_{a0} \equiv \sum_a N_{a0} T_{a0}$. Equation (152) is easily confirmed to be equivalent to Eq. (127) representing the MHD equilibrium.

The first-order part of Eq. (149) comes only from the CGL pressure tensor because the other pressure tensors and the turbulent Maxwell stress tensors are of the second order. Thus, one obtains

$$\nabla \cdot \langle (\mathbf{P}_{\rm CGL})_1 \rangle_{\rm ens} = 0. \tag{153}$$

The turbulent part $\widehat{F}_a$ of the distribution function F_a has no contribution to $(\mathbf{P}_{\rm CGL})_1$ and to Eq. (153) because $\langle \widehat{F}_a \rangle_{\rm ens} = 0$. In neoclassical transport theory,[48–50] the parallel component of $\nabla \cdot \langle (\mathbf{P}_{\rm CGL})_1 \rangle_{\rm ens}$ in Eq. (153) automatically vanishes because of a quasineutrality condition and the momentum conservation property of the collision term. Also, the flux-surface average of the toroidal component of Eq. (153),

$$\langle \langle \mathbf{e}_\zeta \cdot (\nabla \cdot (\mathbf{P}_{\rm CGL})_1) \rangle \rangle = 0, \tag{154}$$

automatically holds in axisymmetric and quasi-axisymmetric toroidal systems as described in Sec. VI. Here, $\langle\langle\cdots\rangle\rangle$ denotes the average over the flux surface and the ensemble. In general non-axisymmetric systems such as stellarator and heliotron plasmas, Eq. (154) is not automatically satisfied but it imposes an ambipolarity condition on neoclassical particle fluxes, from which the background radial electric field can be determined.[63]

The effects of electromagnetic microturbulence with perpendicular wavelengths on the gyroradius scale appear on the second order in Eq. (149). In Sec. VIII, the turbulence contributions to the momentum transport are investigated in detail using the WKB representation for turbulent fluctuations.

VIII. WKB REPRESENTATION

The WKB (or ballooning) representation[10] is useful for treating turbulent fluctuations which have small wavelengths of the order of the gyroradius ρ in the directions perpendicular to the background magnetic field. The WKB representation for the fluctuation part $\widehat{Q}$ of an arbitrary function $Q(\mathbf{x},t)$ takes the form,

$$\widehat{Q}(\mathbf{x},t) = \sum_{\mathbf{k}_\perp} \widehat{Q}_{\mathbf{k}_\perp}(\mathbf{x},t)\exp[iS_{\mathbf{k}_\perp}(\mathbf{x},t)], \tag{155}$$

where $\widehat{Q}_{\mathbf{k}_\perp}(\mathbf{x},t)$ has the equilibrium gradient scale length L while the eikonal $S_{\mathbf{k}_\perp}(\mathbf{x},t)$ represents a rapid variation with the wave number vector $\mathbf{k}_\perp \equiv \nabla S_{\mathbf{k}_\perp} (\sim \rho^{-1})$ which satisfies $\mathbf{k}_\perp \cdot \mathbf{b} = 0$.

In the gyroradius expansion using $\delta = \rho/L \ll 1$, the zeroth-order gyrocenter distribution function is assumed to be given by the local Maxwellian, $F_{a0} = F_{aM} = D_{a0} f_{aM}$, as shown in Eq. (150). The fluctuation part of F_a appears in the first order and it is given by the WKB representation as $\widehat{F}_{a1} = \sum_{\mathbf{k}_\perp} \widehat{F}_{a1\mathbf{k}_\perp} \exp[iS_{\mathbf{k}_\perp}(\mathbf{X},t)]$, where $\mathbf{X}$ is the gyrocenter position vector. As shown in Ref.[31], the $\mathbf{k}_\perp$-component $\widehat{F}_{a1\mathbf{k}_\perp}$ is written as $\widehat{F}_{a1\mathbf{k}_\perp} = D_{a0}\widehat{f}_{a1\mathbf{k}_\perp}$, where $\widehat{f}_{a1\mathbf{k}_\perp}$ is given by

$$\hat{f}_{a1\mathbf{k}_\perp} = -f_{aM}\frac{e_a}{T_a}\langle\widehat{\psi}_a\rangle_{\xi\mathbf{k}_\perp} + \widehat{h}_{a\mathbf{k}_\perp}. \tag{156}$$

Here, $\widehat{h}_{a\mathbf{k}_\perp}$ is the nonadiabatic part of the turbulent distribution function and the gyrophase-averaged potential $\langle\widehat{\psi}_a\rangle_{\xi\mathbf{k}_\perp}$ is defined by

$$\langle\widehat{\psi}_a\rangle_{\xi\mathbf{k}_\perp} = J_0\left(\frac{k_\perp v_\perp}{\Omega_a}\right)\left(\widehat{\phi}_{\mathbf{k}_\perp} - \frac{U}{c}\widehat{A}_{\parallel\mathbf{k}_\perp}\right) + J_1\left(\frac{k_\perp v_\perp}{\Omega_a}\right)\frac{v_\perp}{c}\frac{\widehat{B}_{\parallel\mathbf{k}_\perp}}{k_\perp}, \tag{157}$$

where J_0 and J_1 are the Bessel functions. In addition, another kind of gyrophase-averaged potential is defined by

$$\begin{aligned}\langle\widehat{\chi}_a\rangle_{\xi\mathbf{k}_\perp} &= -\frac{k_\perp v_\perp}{\Omega_a}J_1\left(\frac{k_\perp v_\perp}{\Omega_a}\right)\left(\widehat{\phi}_{\mathbf{k}_\perp} - \frac{U}{c}\widehat{A}_{\parallel\mathbf{k}_\perp}\right)\\ &\quad + \left[\frac{k_\perp v_\perp}{\Omega_a}J_0\left(\frac{k_\perp v_\perp}{\Omega_a}\right) - J_1\left(\frac{k_\perp v_\perp}{\Omega_a}\right)\right]\frac{v_\perp}{c}\frac{\widehat{B}_{\parallel\mathbf{k}_\perp}}{k_\perp}\end{aligned} \tag{158}$$

which is used later to express the turbulent pressure tensor in Eq. (166).

It is now recalled that the ensemble-averaged quantities are smooth spatial functions with the gradient scale length L. For arbitrary real-valued turbulent fluctuations $\hat{Q}$ and $\hat{Q}'$, $\langle\hat{Q}^*_{\mathbf{k}_\perp}\hat{Q}'_{\mathbf{k}'_\perp}\rangle_{\rm ens} = 0$ for $\mathbf{k}_\perp \neq \mathbf{k}'_\perp$ and $\langle\hat{Q}\hat{Q}'\rangle_{\rm ens} = \sum_{\mathbf{k}_\perp}\langle\hat{Q}^*_{\mathbf{k}_\perp}\hat{Q}'_{\mathbf{k}_\perp}\rangle_{\rm ens}$ hold. In the ensemble-averaged momentum balance equation given by Eq. (149), the effects of the electromagnetic turbulence on the momentum transport enter $\langle\boldsymbol{\pi}_{\parallel\Psi}\rangle_{\rm ens}$ $\langle\boldsymbol{\pi}_\wedge\rangle_{\rm ens}$, and $\langle\mathbf{P}_\Psi\rangle_{\rm ens}$ through the correlation between the turbulent distribution function and the turbulent potential. Using Eqs. (94), (95), and (117), and neglecting terms of higher orders in $\delta = \rho/L$, one finds that

$$\begin{aligned}\langle\boldsymbol{\pi}_{\parallel\Psi}\rangle_{\rm ens} &= -\mathbf{bb}\sum_a\left(\frac{n_{a0}e_a^2}{m_a c^2}\langle(\widehat{A}_\parallel)^2\rangle_{\rm ens}\right.\\ &\qquad\left. + \frac{e_a}{c}\int d^3v\, U\langle\widehat{h}_a\langle\widehat{A}_\parallel\rangle_\vartheta\rangle_{\rm ens}\right)\\ &= -\mathbf{bb}\left(\frac{\omega_p^2}{4\pi c^2}\langle(\widehat{A}_\parallel)^2\rangle_{\rm ens} + \frac{1}{c}\langle\widehat{j}_\parallel\widehat{A}_\parallel\rangle_{\rm ens}\right)\end{aligned} \tag{159}$$

and

$$\begin{aligned}\langle\boldsymbol{\pi}_\wedge\rangle_{\rm ens} &= \sum_a\int d^3v\,\langle F_a\rangle_{\rm ens} m_a U\left[b^i\langle(u_{ax})^j_\perp\rangle_{\rm ens}\right.\\ &\qquad\left. + \langle(u_{ax})^i_\perp\rangle_{\rm ens}b^j\right] + \langle\boldsymbol{\pi}^{\rm turb}_\wedge\rangle_{\rm ens},\end{aligned} \tag{160}$$

where the turbulent part $\langle\boldsymbol{\pi}^{\rm turb}_\wedge\rangle_{\rm ens}$ is given by

$$\begin{aligned}\langle\boldsymbol{\pi}^{\rm turb}_\wedge\rangle_{\rm ens} &= \sum_a\int d^3v\,\widehat{F}_a m_a U[b^i(\widehat{u}_{ax})^j_\perp + (\widehat{u}_{ax})^i_\perp b^j]\\ &= \frac{c}{B}\sum_{\mathbf{k}_\perp}[\mathbf{b}(\mathbf{k}_\perp\times\mathbf{b}) + (\mathbf{k}_\perp\times\mathbf{b})\mathbf{b}]\\ &\quad\times\sum_a\int d^3v\, m_a U{\rm Im}[\langle\widehat{h}^*_{a\mathbf{k}_\perp}\langle\widehat{\psi}_a\rangle_{\vartheta\mathbf{k}_\perp}\rangle_{\rm ens}].\end{aligned} \tag{161}$$

Using Eqs. (96) and (117), $\langle\mathbf{P}_\Psi\rangle_{\rm ens}$ is written as

$$\langle\mathbf{P}_\Psi\rangle_{\rm ens} = \langle\mathbf{P}_{\langle\phi\rangle_{\rm ens}}\rangle_{\rm ens} + \langle\mathbf{P}^{\rm turb}_\Psi\rangle_{\rm ens} \tag{162}$$

where the effects of the ensemble-averaged (or background) electric field $\langle\mathbf{E}_L\rangle_{\rm ens} \equiv -\nabla\langle\phi\rangle_{\rm ens}$ and the turbulent electromagnetic field are included in

$$\begin{aligned}\langle\mathbf{P}_{\langle\phi\rangle_{\rm ens}}\rangle_{\rm ens} &= \frac{n_{a0}m_a c^2}{B^2}(\mathbf{b}\times\nabla\langle\phi\rangle_{\rm ens})(\mathbf{b}\times\nabla\langle\phi\rangle_{\rm ens}) - \frac{m_a c^2}{2e_a B^2}\\ &\quad\times\Big[\nabla(n_{a0}T_{a0})\nabla\langle\phi\rangle_{\rm ens} + \nabla\langle\phi\rangle_{\rm ens}\nabla(n_{a0}T_{a0})\\ &\qquad - (\mathbf{I}-\mathbf{bb})\nabla(n_{a0}T_{a0})\cdot\nabla\langle\phi\rangle_{\rm ens}\Big] + n_{a0}m_{a0}\frac{c^2 T_{a0}}{2e_a B^2}\\ &\quad\times\Big[(\mathbf{I}-3\mathbf{bb})\mathbf{bb}:(\nabla\nabla\langle\phi\rangle_{\rm ens}) + \mathbf{b}(\mathbf{b}\cdot\nabla\nabla\langle\phi\rangle_{\rm ens})\\ &\qquad + (\mathbf{b}\cdot\nabla\nabla\langle\phi\rangle_{\rm ens})\mathbf{b} + \nabla\cdot(\mathbf{bb})\nabla\langle\phi\rangle_{\rm ens}\\ &\qquad + \nabla\langle\phi\rangle_{\rm ens}\nabla\cdot(\mathbf{bb}) - \nabla\langle\phi\rangle_{\rm ens}\cdot\nabla(\mathbf{bb})\\ &\qquad + 2\nabla_\perp\ln B\,\nabla\langle\phi\rangle_{\rm ens} + 2\nabla\langle\phi\rangle_{\rm ens}\,\nabla_\perp\ln B\\ &\qquad - 2(\nabla\langle\phi\rangle_{\rm ens}\cdot\nabla\ln B)(\mathbf{I}-\mathbf{bb})\Big]\end{aligned} \tag{163}$$

and

$$\langle \mathbf{P}_\Psi^{\rm turb} \rangle_{\rm ens} = \langle \mathbf{P}_\Psi^{\rm ad} \rangle_{\rm ens} + \langle \mathbf{P}_\Psi^{\rm nad} \rangle_{\rm ens}, \tag{164}$$

respectively, where

$$\begin{aligned}\langle \mathbf{P}_\Psi^{\rm ad} \rangle_{\rm ens} = \big(\mathbf{I} - \mathbf{bb}\big) \sum_a \sum_{\mathbf{k}_\perp} \frac{n_{a0} e_a^2}{2T_a} \Bigg[\left\langle |\widehat{\phi}_{\mathbf{k}_\perp}|^2 + \frac{T_a}{m_a c^2} |\widehat{A}_{\parallel \mathbf{k}_\perp}|^2 \right\rangle_{\rm ens} \\ \times \left\{1 - \Gamma_0(b_{a\mathbf{k}_\perp})\right\} + \frac{T_a}{m_a c^2} \left\langle \frac{|\widehat{B}_{\parallel \mathbf{k}_\perp}|^2}{k_\perp^2} \right\rangle_{\rm ens} \big[1 - 2b_{a\mathbf{k}_\perp} \\ \times \left\{\Gamma_0(b_{a\mathbf{k}_\perp}) - \Gamma_1(b_{a\mathbf{k}_\perp})\right\}\big] - \frac{v_{Ta}}{c} \,\mathrm{Re} \left\langle \frac{\widehat{\phi}^*_{\mathbf{k}_\perp} \widehat{B}_{\parallel \mathbf{k}_\perp}}{k_\perp} \right\rangle_{\rm ens} \\ \times (2b_{a\mathbf{k}_\perp})^{1/2} \left\{\Gamma_0(b_{a\mathbf{k}_\perp}) - \Gamma_1(b_{a\mathbf{k}_\perp})\right\} \Bigg] \end{aligned} \tag{165}$$

and

$$\begin{aligned}\langle \mathbf{P}_\Psi^{\rm nad} \rangle_{\rm ens} \;=\; & -\sum_a \sum_{\mathbf{k}_\perp} \left[\frac{1}{2} (\mathbf{I} - \mathbf{bb}) - \frac{\mathbf{k}_\perp \mathbf{k}_\perp}{k_\perp^2} \right] \\ & \times e_a \int d^3 v\, \mathrm{Re}[\langle \widehat{h}^*_{a\mathbf{k}_\perp} \langle \widehat{\chi}_a \rangle_{\vartheta \mathbf{k}_\perp} \rangle_{\rm ens}] - \frac{1}{c} \langle \widehat{\mathbf{j}}\, \widehat{\mathbf{A}} \rangle_{\rm ens} \\ & - \frac{1}{2c} \langle \widehat{\mathbf{j}}_\perp \cdot \widehat{\mathbf{A}}_\perp \rangle_{\rm ens} (\mathbf{I} - \mathbf{bb}) \end{aligned} \tag{166}$$

are derived from the adiabatic and nonadiabatic parts of the distribution function in Eq. (156), respectively. On the right-hand side of Eq. (165), $b_{a\mathbf{k}_\perp} \equiv k_\perp^2 T_a/(m_a \Omega_a^2)$ is used, and the functions Γ_0 and Γ_1 are defined by $\Gamma_0(b) \equiv I_0(b) \exp(-b)$ and $\Gamma_1(b) \equiv I_1(b) \exp(-b)$, respectively, where I_0 and I_1 are the modified Bessel functions.

One finds that the pressure tensor terms consisting of only **bb** and **I** parts cannot produce transport of toroidal and poloidal momenta across flux surfaces in toroidal magnetically confined systems. In the axisymmetric toroidal system, in which the magnetic field is given by Eq. (128), one can use $(\mathbf{k}_\perp \times \mathbf{b}) \cdot \nabla \chi = -BR^2 \nabla \zeta \cdot \mathbf{k}_\perp$ and Eqs. (161), (164), (165), and (166) to obtain the radial transport of the toroidal angular momentum due to the interaction of the nonadiabatic distribution function and the turbulent electromagnetic potential as

$$\begin{aligned}&\nabla \chi \cdot \langle \pi_\wedge^{\rm turb} + \mathbf{P}_\Psi^{\rm turb} \rangle_{\rm ens} \cdot R^2 \nabla \zeta \\ &= \sum_a \sum_{\mathbf{k}_\perp} \int d^3 v \left[-\frac{cI}{B} m_a U \,\mathrm{Im}[\langle \widehat{h}^*_{a\mathbf{k}_\perp} \langle \widehat{\psi}_a \rangle_{\vartheta \mathbf{k}_\perp} \rangle_{\rm ens}] (\mathbf{k}_\perp \cdot R^2 \nabla \zeta) \right. \\ &\quad \left. + e_a \mathrm{Re}[\langle \widehat{h}^*_{a\mathbf{k}_\perp} \langle \widehat{\chi}_a \rangle_{\vartheta \mathbf{k}_\perp} \rangle_{\rm ens}] \frac{(\mathbf{k}_\perp \cdot \nabla \chi)(\mathbf{k}_\perp \cdot R^2 \nabla \zeta)}{k_\perp^2} \right]. \end{aligned} \tag{167}$$

The flux-surface average of Eq. (167) agrees with the low-flow ordering limit of the result given in Eq. (53) of Ref.[67] where the turbulent radial transport of the toroidal angular momentum double-averaged over the ensemble and the flux surface is presented for the case of high-flow ordering.[21,66,67] The radial turbulent transport of the toroidal angular momentum in Eq. (167) is not the flux-surface average but a spatially-local expression. It is shown in the case of the low-flow ordering[51,68] that, in the axisymmetric configuration with up-down symmetry, the flux-surface average of Eq. (167) vanishes even though the local value of Eq. (167) does not. It can also be shown from Eq. (163) that the flux-surface average of $\nabla \chi \cdot \langle \mathbf{P}_{\langle \phi \rangle_{\rm ens}} \rangle_{\rm ens} \cdot R^2 \nabla \zeta$ vanishes as well.

Next, let us consider the Maxwell stress terms in the ensemble-averaged momentum balance equation in Eq. (149). The Maxwell stress due to the electric field is dominantly given by $\mathbf{E}_L$ because the magnitude of $\mathbf{E}_T \equiv -c^{-1} \partial \mathbf{A}/\partial t$ is smaller than that of $\mathbf{E}_L$ by a factor of δ. On the left-hand side of Eq. (149), the turbulent magnetic pressure tensor also appears as

$$\frac{1}{8\pi} \left\langle |\widehat{\mathbf{B}}|^2 \right\rangle_{\rm ens} \mathbf{I} - \frac{1}{4\pi} \left\langle \widehat{\mathbf{B}} \widehat{\mathbf{B}} \right\rangle_{\rm ens}, \tag{168}$$

which has the opposite sign to that of the Maxwell stress tensor due to turbulent magnetic fields.

It is also noted that the terms proportional to $\nabla \lambda$ and $\widehat{\mathbf{A}}$ in Eq. (149) are negligible compared with the other magnetic Maxwell stress terms because $c^{-1} |\nabla \lambda| |\widehat{\mathbf{A}}| / |\widehat{\mathbf{B}}|^2 \sim |\mathbf{j}_L|/(ck_\perp |\widehat{\mathbf{B}}|) \sim (\partial \mathbf{E}_L/\partial t)/(ck_\perp |\widehat{\mathbf{B}}|) \sim (v_T/L)(cT/eB)(e\widehat{\phi}/T)/(c^2 |\widehat{\mathbf{B}}|/B) \sim (\rho/L)(v_T^2/c^2) \ll 1$, where Eq. (37), $\nabla \cdot \mathbf{j}_L = -\partial \rho_c/\partial t = -(4\pi)^{-1} \partial (\nabla \cdot \mathbf{E}_L)/\partial t$, $\partial/\partial t \sim v_T/L$, and $e\widehat{\phi}/T \sim |\widehat{\mathbf{B}}|/B \sim \rho/L$ are used.

One can also find that the nonadiabatic distribution function produces the turbulent current $\widehat{\mathbf{j}}$ which correlates with the turbulent vector potential $\widehat{\mathbf{A}}$ and produces the pressure tensor given by

$$-\frac{1}{c} \langle \widehat{j}_\parallel \widehat{A}_\parallel \rangle_{\rm ens} \mathbf{bb} - \frac{1}{c} \langle \widehat{\mathbf{j}}\, \widehat{\mathbf{A}} \rangle_{\rm ens} - \frac{1}{2c} \langle \widehat{\mathbf{j}}_\perp \cdot \widehat{\mathbf{A}}_\perp \rangle_{\rm ens} (\mathbf{I} - \mathbf{bb}), \tag{169}$$

which are included in Eqs. (159) and (166). In the wavenumber representation, the turbulent magnetic field is given by

$$\widehat{\mathbf{B}}_{\mathbf{k}_\perp} = i\mathbf{k}_\perp \times \widehat{\mathbf{A}}_{\mathbf{k}_\perp} = \widehat{B}_{\parallel \mathbf{k}_\perp} \mathbf{b} + \widehat{\mathbf{B}}_{\parallel \mathbf{k}_\perp}, \tag{170}$$

where $\widehat{B}_{\parallel \mathbf{k}_\perp} = i\mathbf{b} \cdot (\mathbf{k}_\perp \times \widehat{\mathbf{A}}_{\mathbf{k}_\perp})$ and $\widehat{\mathbf{B}}_{\perp \mathbf{k}_\perp} = i(\mathbf{k}_\perp \times \mathbf{b}) \widehat{A}_{\parallel \mathbf{k}_\perp}$. The turbulent vector potential is written by

$$\widehat{\mathbf{A}}_{\mathbf{k}_\perp} = \widehat{A}_{\parallel \mathbf{k}_\perp} \mathbf{b} + \widehat{\mathbf{A}}_{\perp \mathbf{k}_\perp}, \tag{171}$$

where $\widehat{A}_{\parallel \mathbf{k}_\perp} = \mathbf{b} \cdot \widehat{\mathbf{A}}_{\mathbf{k}_\perp}$ and $\widehat{\mathbf{A}}_{\perp \mathbf{k}_\perp} = k_\perp^{-2} (\mathbf{b} \times \mathbf{k}_\perp)[(\mathbf{b} \times \mathbf{k}_\perp) \cdot \widehat{\mathbf{A}}_{\mathbf{k}_\perp}]$. Here, the Coulomb gauge condition $\mathbf{k}_\perp \cdot \widehat{\mathbf{A}}_{\mathbf{k}_\perp} = 0$ is used. Then, one has

$$|\widehat{\mathbf{B}}_{\mathbf{k}_\perp}|^2 = |\widehat{B}_{\parallel \mathbf{k}_\perp}|^2 + |\widehat{\mathbf{B}}_{\perp \mathbf{k}_\perp}|^2 = k_\perp^2 \left(|\widehat{\mathbf{A}}_{\perp \mathbf{k}_\perp}|^2 + |\widehat{A}_{\parallel \mathbf{k}_\perp}|^2 \right), \tag{172}$$

and Ampère's law is given by

$$\widehat{\mathbf{j}}_{\mathbf{k}_\perp} = \frac{c}{4\pi} k_\perp^2 \widehat{\mathbf{A}}_{\mathbf{k}_\perp}. \tag{173}$$

Using Eqs. (170)–(173), one obtains

$$\begin{aligned}-\frac{1}{4\pi} \langle \widehat{\mathbf{B}} \widehat{\mathbf{B}} \rangle_{\rm ens} - \frac{1}{c} \langle \widehat{\mathbf{j}} \widehat{\mathbf{A}} \rangle_{\rm ens} &= \frac{1}{4\pi} \sum_{\mathbf{k}_\perp} \langle |\widehat{\mathbf{A}}_{\mathbf{k}_\perp}|^2 \rangle_{\rm ens} \left(\mathbf{k}_\perp \mathbf{k}_\perp - k_\perp^2 \mathbf{I} \right) \\ &= \frac{1}{4\pi} \sum_{\mathbf{k}_\perp} \langle |\widehat{\mathbf{B}}_{\mathbf{k}_\perp}|^2 \rangle_{\rm ens} \left(\frac{\mathbf{k}_\perp \mathbf{k}_\perp}{k_\perp^2} - \mathbf{I} \right), \end{aligned} \tag{174}$$

and the summation of Eqs. (168) and (169) are written as

$$\begin{aligned}
&\text{Eq. (168)} + \text{Eq. (169)} \\
&= \frac{1}{8\pi}\sum_{\mathbf{k}_\perp}\left[-\langle|\widehat{\mathbf{B}}_{\mathbf{k}_\perp}|^2 + 2|\widehat{\mathbf{B}}_{\perp\mathbf{k}_\perp}|^2\rangle_{\text{ens}}\mathbf{bb} - \langle|\widehat{\mathbf{B}}_{\perp\mathbf{k}_\perp}|^2\rangle_{\text{ens}}\frac{\mathbf{k}_\perp\mathbf{k}_\perp}{k_\perp^2} \right. \\
&\qquad \left. - \langle|\widehat{\mathbf{B}}_{\mathbf{k}_\perp}|^2 + |\widehat{\mathbf{B}}_{\|\mathbf{k}_\perp}|^2\rangle_{\text{ens}}\frac{(\mathbf{b}\times\mathbf{k}_\perp)(\mathbf{b}\times\mathbf{k}_\perp)}{k_\perp^2}\right], \qquad (175)
\end{aligned}$$

which is given in terms of only the turbulent magnetic field.

Again, in considering the axisymmetric toroidal system, the radial transport of the toroidal angular momentum due to the turbulent magnetic field is represented by a component of Eq. (175) obtained from a double-dot product with a dyad $(\nabla\chi)(R^2\nabla\zeta)$, which is equivalent to that of Eq. (174) and written as

$$\begin{aligned}
&\nabla\chi\cdot\left[-\frac{1}{4\pi}\langle\widehat{\mathbf{B}}\widehat{\mathbf{B}}\rangle_{\text{ens}} - \frac{1}{c}\langle\widehat{\mathbf{j}}\widehat{\mathbf{A}}\rangle_{\text{ens}}\right]\cdot R^2\nabla\zeta \\
&= \frac{1}{4\pi}\sum_{\mathbf{k}_\perp}\langle|\widehat{\mathbf{A}}_{\mathbf{k}_\perp}|^2\rangle_{\text{ens}}(\mathbf{k}_\perp\cdot\nabla\chi)(\mathbf{k}_\perp\cdot R^2\nabla\zeta) \\
&= \frac{1}{4\pi}\sum_{\mathbf{k}_\perp}\langle|\widehat{\mathbf{B}}_{\mathbf{k}_\perp}|^2\rangle_{\text{ens}}\frac{(\mathbf{k}_\perp\cdot\nabla\chi)(\mathbf{k}_\perp\cdot R^2\nabla\zeta)}{k_\perp^2}. \qquad (176)
\end{aligned}$$

Taking the flux-surface average of Eq. (176) yields the same expression of the toroidal angular momentum transport across the flux surface caused by the turbulent magnetic field as derived in Refs.[67] and[66]. However, in the same manner as in the case of Eq. (167), the flux-surface average of Eq. (176) is shown to vanish in the axisymmetric configuration with up-down symmetry under the low-flow ordering[51,68].

It is also found from Eqs. (159) and (165) that the adiabatic part of the perturbed distribution function in Eq. (156) produces pressure tensor terms in $\langle\boldsymbol{\pi}_{\|\Psi}\rangle_{\text{ens}}$ and $\langle\mathbf{P}_\Psi^{\text{ad}}\rangle_{\text{ens}}$ which are given in terms of turbulent electrostatic and vector potentials although these terms consist of only the **bb** and **I** parts so that they cannot produce toroidal and poloidal momentum transport across flux surfaces in toroidal plasmas.

IX. CONCLUSIONS

In this paper, the Eulerian (or Euler-Poincaré) variational formulation is presented to obtain the governing equations of the electromagnetic turbulent gyrokinetic system, for which the local momentum balance equation is derived from the invariance of the Lagrangian of the system under an arbitrary spatial coordinate transformation. In addition, the effects of collisions and external sources are taken into account in the momentum balance equation.

Using the gyrokinetic Lagrangian which retains proper electromagnetic potential terms and taking the variational derivatives of the Lagrangian with respect to the electrostatic and vector potentials of the perturbed magnetic field, one can obtain the gyrokinetic Poisson equation and Ampère's law where the effects of the polarization and magnetization due to finite gyroradii and electromagnetic microturbulence are correctly included. Especially, the derived gyrokinetic Ampère's law can accurately express the current density from the microscopic gyroradius scale to the macroscopic equilibrium scale so that it is useful for long-time and global gyrokinetic turbulence simulations of high beta plasmas.

The local momentum balance equation obtained in the present work contains the symmetric pressure tensor which is derived from the variational derivative of the Lagrangian with respect to the metric tensor. It is shown that the pressure tensor obtained for the whole system consisting of all particles and fields involves the gyrokinetic and field parts; the neoclassical and turbulent momentum transport processes are described by the former part while the Maxwell stress is by the latter.

One can confirm from the momentum balance equation that, when the background magnetic field has a symmetry such as a translational one and an axisymmetry, the canonical momentum conjugate to the coordinate in the symmetry direction is conserved as predicted by Noether's theorem. The symmetry of the pressure tensor is found to be an important property for derivation of the momentum conservation in the symmetric background field. When the background field is assumed to satisfy the appropriate condition representing the macroscopic Ampère's law, the ensemble-averaged total momentum balance equation is found to take the conservation form even in the asymmetric background field. Thus, this condition can be conveniently applied to long-time gyrokinetic simulations in which the change in the background field occurs with the relaxation of high-beta plasmas. It is also shown that, in the toroidal systems with the quasi-axisymmetric background field, the toroidal angular momentum is not rigorously conserved although the flux-surface-averaged neoclassical toroidal viscosity, which is a dominant component for breaking the toroidal momentum conservation in general non-axisymmetric systems, vanishes.

The WKB representation is employed to derive detailed expressions of the ensemble-averaged pressure tensor due to the electromagnetic microturbulence, which provide a means for evaluating the local turbulent momentum transport by the local flux-tube gyrokinetic simulation. The radial transport fluxes of the toroidal angular momentum caused by the nonadiabatic distribution function and the turbulent electromagnetic fields in the axisymmetric system are represented as a non-diagonal component of the pressure tensor, which are shown to agree with the results from the previous works based on the classical gyrokinetic formulation. The local pressure tensor represented by a symmetric 3×3 matrix contains further information on momentum transport which is useful for more detailed analyses of transport processes by gyrokinetic simulations.

ACKNOWLEDGMENTS

This work is supported in part by the JSPS Grants-in-Aid for Scientific Research Grant No. 19H01879 and in part by the NIFS Collaborative Research Program NIFS23KIPT009.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts of interest to disclose.

Author Contributions

Hideo Sugama: Conceptualization (lead); Formal analysis (lead); Funding acquisition (lead); Writing – original draft (lead).

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Appendix A: DECOMPOSITION OF THE POTENTIAL FIELD

The potential field defined in Eq. (3) is decomposed here as

$$\begin{aligned}\Psi_a(\mathbf{Z},t) \;\equiv\; & \phi(\mathbf{X},t)+\Psi_{E1a}(\mathbf{X},\mu,t)+\Psi_{\widehat{A}1a}(\mathbf{X},U,\mu,t) \\ & +\Psi_{E2a}(\mathbf{X},\mu,t)+\Psi_{E\widehat{A}a}(\mathbf{X},U,\mu,t) \\ & +\Psi_{\widehat{A}2a}(\mathbf{X},U,\mu,t),\end{aligned} \tag{A1}$$

where

$$\Psi_{E1a}(\mathbf{X},\mu,t)\equiv\langle\phi(\mathbf{X}+\boldsymbol{\rho}_a,t)\rangle_\vartheta-\phi(\mathbf{X},t), \tag{A2}$$

$$\Psi_{\widehat{A}1a}(\mathbf{X},U,\mu,t)\equiv-\frac{1}{c}\left[\langle\mathbf{v}\cdot\widehat{\mathbf{A}}\rangle_\vartheta+\mathbf{v}_{Ba}\cdot\langle\widehat{\mathbf{A}}\rangle_\vartheta\right] \tag{A3}$$

$$\Psi_{E2a}(\mathbf{X},\mu,t)\equiv-\frac{e_a}{2B}\frac{\partial}{\partial\mu}\langle(\widetilde{\phi})^2\rangle_\vartheta, \tag{A4}$$

$$\Psi_{E\widehat{A}a}(\mathbf{X},U,\mu,t)\equiv\frac{e_a}{cB}\frac{\partial}{\partial\mu}\langle\widetilde{\phi}(\mathbf{v}\cdot\widehat{\mathbf{A}})\rangle_\vartheta, \tag{A5}$$

and

$$\begin{aligned}&\Psi_{\widehat{A}2a}(\mathbf{X},U,\mu,t)\\ &\equiv\frac{e_a}{2m_ac^2}\langle|\widehat{\mathbf{A}}|^2\rangle_\vartheta-\frac{e_a}{2c^2B}\frac{\partial}{\partial\mu}\langle[(\mathbf{v}\cdot\widehat{\mathbf{A}})-\langle\mathbf{v}\cdot\widehat{\mathbf{A}}\rangle_\vartheta]^2\rangle_\vartheta.\end{aligned} \tag{A6}$$

It should be noted that, in this Appendix, the Cartesian spatial coordinates and the conventional dyadic notation representing vectors and tensors in terms of boldface letters are used. Then, the electrostatic potential $\phi(\mathbf{X}+\boldsymbol{\rho}_a)$ and the perturbed vector potential $\widehat{\mathbf{A}}(\mathbf{X}+\boldsymbol{\rho}_a)$ are Taylor expanded about the gyrocenter position $\mathbf{X}$ as

$$\begin{bmatrix}\phi(\mathbf{X}+\boldsymbol{\rho}_a)\\ \widehat{\mathbf{A}}(\mathbf{X}+\boldsymbol{\rho}_a)\end{bmatrix}=\sum_{n=0}^{\infty}\frac{1}{n!}\rho_a^{j_1}\cdots\rho_a^{j_n}\partial_{j_1\cdots j_n}\begin{bmatrix}\phi(\mathbf{X})\\ \widehat{\mathbf{A}}(\mathbf{X})\end{bmatrix}, \tag{A7}$$

where X^j and ρ^j $(j=1,2,3)$ and the Cartesian spatial coordinates of the gyrocenter position vector $\mathbf{X}$ and the gyroradius vector $\boldsymbol{\rho}$, respectively, and the partial derivatives are represented using the simplified notation, $\partial_{j_1\cdots j_n}\equiv\partial^n/\partial X^{j_1}\cdots\partial X^{j_n}$. Here, for simplicity, we omit the t-dependence of ϕ and employ the summation convention that the same symbol used for a pair of upper and lower indices indicates a summation over the range $\{1,2,3\}$. Therefore, summation notations $\sum_{j_1=1}^3\cdots\sum_{j_n=1}^3$ are dropped in Eq. (A7).

Using Eqs. (A2) and (A7), one can write the part of the potential function which linearly depends on the electric field and its derivatives as

$$\begin{aligned}\Psi_{E1a} &= \sum_{n=1}^{\infty}\frac{\alpha_a^{j_1\cdots j_n}}{n!}\partial_{j_1\cdots j_n}\phi(\mathbf{X})\\ &= -\sum_{n=1}^{\infty}\frac{\alpha_a^{j_1\cdots j_n}}{n!}\partial_{j_1\cdots j_{n-1}}(E_L)_{j_n}(\mathbf{X}),\end{aligned} \tag{A8}$$

where the gyrophase average of a product of n gyroradius vector components is denoted by

$$\alpha_a^{j_1\cdots j_n}\equiv\langle\rho_a^{j_1}\cdots\rho_a^{j_n}\rangle_\vartheta. \tag{A9}$$

We see that $\alpha_a^{j_1\cdots j_n}$ is symmetric with respect to arbitrary permutations of the indices $j_1,\cdots,j_n$. We also find that $\alpha_a^{j_1\cdots j_n}=0$ for odd n and

$$\alpha_a^{j_1\cdots j_{2l}}=\frac{1}{(2l)!}\sum_{\sigma\in\mathfrak{S}_{2l}}\eta_a^{j_{\sigma(1)}\cdots j_{\sigma(2l)}}, \tag{A10}$$

where $\mathfrak{S}_{2l}$ is the symmetric group of permutations of the set $\{1,2,\cdots,2l\}$ and $\eta_a^{j_1\cdots j_{2l}}$ is defined by

$$\eta_a^{j_1\cdots j_{2l}}=\frac{(2l)!}{(l!)^2}\left(\frac{\rho_a}{2}\right)^{2l}h^{j_1j_2}h^{j_3j_4}\cdots h^{j_{2l-1}j_{2l}}, \tag{A11}$$

with $\rho_a\equiv(c/e_a)\sqrt{2m_a\mu/B}$ and $h^{ij}\equiv\delta^{ij}-b^ib^j$. Here, b^i is the ith component of $\mathbf{b}\equiv\mathbf{B}/B$ and δ^{ij} represents the Kronecker delta; $\delta^{ij}=1$ (for $i=j$), 0 (for $i\neq j$).

Next, using Eqs. (A3), the part which linearly depends on the vector potential and its derivatives is written as

$$\begin{aligned}\Psi_{\widehat{A}1a} &= -\frac{1}{c}\left[(U\mathbf{b}+\mathbf{v}_{Ba})\cdot\langle\widehat{\mathbf{A}}(\mathbf{X}+\boldsymbol{\rho}_a)\rangle_\vartheta+\langle\mathbf{v}_\perp\cdot\widehat{\mathbf{A}}(\mathbf{X}+\boldsymbol{\rho}_a)\rangle_\vartheta\right]\\ &= -\frac{1}{c}\left[(U\mathbf{b}+\mathbf{v}_{Ba})\cdot\langle\widehat{\mathbf{A}}(\mathbf{X}+\boldsymbol{\rho}_a)\rangle_\vartheta\right.\\ &\qquad\left.+\Omega_a\langle(\boldsymbol{\rho}_a\times\mathbf{b})\cdot\widehat{\mathbf{A}}(\mathbf{X}+\boldsymbol{\rho}_a)\rangle_\vartheta\right]\\ &= -\frac{1}{c}\sum_{n=0}^{\infty}\frac{1}{n!}\left[\alpha_a^{j_1\cdots j_n}\,(Ub^i+v_{Ba}^i)\right.\\ &\qquad\left.+\Omega_a\,\varepsilon_{klm}\,\alpha_a^{j_1\cdots j_nk}\,b^l\,\delta^{im}\right]\partial_{j_1\cdots j_n}\widehat{A}_i(\mathbf{X}),\end{aligned} \tag{A12}$$

Now, using Eqs. (A7), one has

$$\langle(\widetilde{\phi})^2\rangle_\vartheta=\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{\beta^{i_1\cdots i_m;j_1\cdots j_n}}{m!\,n!}\left(\partial_{i_1\cdots i_m}\phi(\mathbf{X})\right)\left(\partial_{j_1\cdots j_n}\phi(\mathbf{X})\right). \tag{A13}$$

Here, $\beta^{i_1\cdots i_m;j_1\cdots j_n}$ is defined by

$$\beta^{i_1\cdots i_m;j_1\cdots j_n} \equiv \alpha_a^{i_1\cdots i_m j_1\cdots j_n} - \alpha_a^{i_1\cdots i_m}\alpha_a^{j_1\cdots j_n}, \qquad \text{(A14)}$$

which satisfies $\beta^{i_1\cdots i_m;j_1\cdots j_n} = 0$ for odd $(m+n)$. Substituting Eq. (A13) into Eq. (A4), the part which is quadratically dependent on $\{\partial_{i_1\cdots i_{m-1}}(E_L)_{i_m}\}$ is derived as

$$\Psi_{E2a}(\mathbf{X},\mu,t) = \frac{1}{2}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} C_{E2a}^{i_1\cdots i_m;j_1\cdots j_n}\partial_{i_1\cdots i_{m-1}}(E_L)_{i_m} \times \partial_{j_1\cdots j_{n-1}}(E_L)_{j_n}, \qquad \text{(A15)}$$

where $C_{E2a}^{i_1\cdots i_m;j_1\cdots j_n}$ is given by

$$C_{E2a}^{i_1\cdots i_m;j_1\cdots j_n} \equiv -\frac{e_a}{2\mu B}\frac{(m+n)\beta^{i_1\cdots i_m;j_1\cdots j_n}}{m!\,n!}. \qquad \text{(A16)}$$

It is also found from Eqs. (A5) and (A6) that the remaining potential functions $\Psi_{E\widehat{A}a}$ and $\Psi_{\widehat{A}2a}$ take bilinear and quadratic forms which are given by

$$\Psi_{E\widehat{A}a}(\mathbf{X},\mu,t) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} C_{E\widehat{A}a}^{i_1\cdots i_m;j_1\cdots j_n}\partial_{i_1\cdots i_{m-1}}(E_L)_{i_m} \times \partial_{j_1\cdots j_{n-1}}\widehat{A}_{j_n}, \qquad \text{(A17)}$$

and

$$\Psi_{\widehat{A}2a}(\mathbf{X},\mu,t) = \frac{1}{2}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} C_{\widehat{A}2a}^{i_1\cdots i_m;j_1\cdots j_n}\partial_{i_1\cdots i_{m-1}}\widehat{A}_{i_m} \times \partial_{j_1\cdots j_{n-1}}\widehat{A}_{j_n}, \qquad \text{(A18)}$$

respectively. One can use Eq. (A9) to derive the coefficients $C_{E\widehat{A}a}^{i_1\cdots i_m;j_1\cdots j_n}$ and $C_{\widehat{A}2a}^{i_1\cdots i_m;j_1\cdots j_n}$ in Eqs. (A17) and (A18) as functions of $\alpha_a^{j_1\cdots j_n}$. The expressions given in Eqs. (A8), (A12), (A15), (A17), and (A18) are valid in the Cartesian spatial coordinates although they can be easily transformed into those in general spatial coordinates as shown in Sec. III.

Appendix B: THE ELECTROMAGNETIC INTERACTION PART OF THE GYROKINETIC LAGRANGIAN DENSITY

The gyrokinetic Lagrangian given by Eq. (18) in Sec.III is written as $L_{GK} \equiv \int d^3X\, \mathscr{L}_{GK}$ where the gyrokinetic Lagrangian density $\mathscr{L}_{GK}$ is defined as a function of $(\mathbf{X},t)$ by $\mathscr{L}_{GK} \equiv \sum_a \int d^3v\, F_a L_{GYa}$. The part of $\mathscr{L}_{GK}$ including the potential field Ψ_a is represented by

$$\mathscr{L}_\Psi \equiv \sum_a \mathscr{L}_{\Psi a} = -\sum_a \int d^3v\, F_a e_a \Psi_a$$
$$= -\rho_c^{(g)}(\mathbf{X},t)\phi(\mathbf{X},t) + \mathscr{L}_{E1} + \mathscr{L}_{\widehat{A}1} + \mathscr{L}_{E2} + \mathscr{L}_{E\widehat{A}} + \mathscr{L}_{\widehat{A}2}, \qquad \text{(B1)}$$

where Eq. (A1) is used. Equation (B1) describes the electromagnetic interaction of charged particles and is used to derive gyrokinetic expressions for polarization and magnetization as shown in Appendices C and D. In Eq. (B1), the gyrocenter charge density $\rho_c^{(g)}(\mathbf{X},t)$ is given by $\rho_c^{(g)}(\mathbf{X},t) \equiv \sum_a e_a N_a^{(g)}(\mathbf{X},t) \equiv \sum_a e_a \int d^3v\, F_a(\mathbf{X},U,\mu,t)$ and the other components of the Lagrangian density are defined by

$$\left[\mathscr{L}_{E1}, \mathscr{L}_{\widehat{A}1}, \mathscr{L}_{E2}, \mathscr{L}_{E\widehat{A}}, \mathscr{L}_{\widehat{A}2}\right] \equiv \sum_a \left[\mathscr{L}_{E1a}, \mathscr{L}_{\widehat{A}1a}, \mathscr{L}_{E2a}, \mathscr{L}_{E\widehat{A}a}, \mathscr{L}_{\widehat{A}2a}\right]$$
$$\equiv -\sum_a \int d^3v\, F_a e_a \left[\Psi_{E1a}, \Psi_{\widehat{A}1a}, \Psi_{E2a}\Psi_{E\widehat{A}a}, \Psi_{\widehat{A}2a}\right]. \qquad \text{(B2)}$$

Here, using Eqs. (A8) and (B2), $\mathscr{L}_{E1a}$ can be represented in the linear form of $\mathbf{E}_L$ and its spatial derivatives,

$$\mathscr{L}_{E1a} = \sum_{k=1}^{\infty} Q_{0a}^{j_1\cdots j_{2k}}\partial_{j_1\cdots j_{2k-1}}(E_L)_{j_{2k}}, \qquad \text{(B3)}$$

where $Q_{0a}^{j_1\cdots j_{2k}}$ represents the multipole moment of the electric charge distribution[54] of species a induced by finite gyroradius,

$$Q_{0a}^{j_1\cdots j_{2k}} \equiv e_a \int d^3v\, F_a \frac{\alpha_a^{j_1\cdots j_{2k}}}{(2k)!}. \qquad \text{(B4)}$$

Substituting Eq. (A12) into Eq. (B2) yields the linear form of $\widehat{\mathbf{A}}$ and its spatial derivatives,

$$\mathscr{L}_{\widehat{A}1a} = \sum_{n=1}^{\infty} R_{0a}^{j_1\cdots j_n}\partial_{j_1\cdots j_{n-1}}\widehat{A}_{j_n}, \qquad \text{(B5)}$$

where

$$R_{0a}^{j_1\cdots j_n i} \equiv \frac{e_a}{c}\int d^3v\, F_a \frac{1}{n!}\left[\alpha_a^{j_1\cdots j_n}(Ub^i + v_{Ba}^i) + \Omega_a\,\varepsilon_{klm}\,\alpha_a^{j_1\cdots j_n k}\, b^l \delta^{im}\right]. \qquad \text{(B6)}$$

Especially, in the cases of $n=0$ and 1, Eq. (B6) is written as

$$R_{0a}^{i} \equiv \frac{e_a}{c}\int d^3v\, F_a(Ub^i + v_{Ba}^i) = \left[\frac{e_a}{c}N_a^{(g)}V_{ag\parallel}\mathbf{b} + \frac{\mathbf{b}}{B}\times\left(P_{\parallel a}\mathbf{b}\cdot\nabla\mathbf{b} + P_{\perp a}\nabla\ln B\right)\right]^i, \qquad \text{(B7)}$$

and

$$R_{0a}^{ji} \equiv \frac{e_a}{c}\int d^3v\, F_a\,\Omega_a\,\varepsilon_{klm}\,\alpha_a^{jk}\, b^l\delta^{im} = \frac{P_{\perp a}}{B}(\mathbf{I}\times\mathbf{b})^{ji}, \qquad \text{(B8)}$$

respectively, where $[P_{\parallel a},\, P_{\perp a}] \equiv \int d^3v\, F_a\, [m_a U^2,\, \mu B]$. From Eqs. (A15), (A17), (A18), and (B2), one obtains the quadratic forms,

$$\mathscr{L}_{E2a} = \frac{1}{2}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\chi_{Ea}^{i_1\cdots i_m;j_1\cdots j_n}\partial_{i_1\cdots i_{m-1}}(E_L)_{i_m} \times \partial_{j_1\cdots j_{n-1}}(E_L)_{j_n}, \qquad \text{(B9)}$$

$$\mathscr{L}_{E\widehat{A}a} = \sum_{n=1}^{\infty} Q_{\widehat{A}a}^{j_1\cdots j_n}\partial_{j_1\cdots j_{n-1}}(E_L)_{j_n} = \sum_{n=1}^{\infty} R_{Ea}^{k_1\cdots k_n}\partial_{k_1\cdots k_{n-1}}\widehat{A}_{k_n} = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\chi_{E\widehat{A}a}^{j_1\cdots j_m;k_1,\cdots k_n}\partial_{j_1\cdots j_{m-1}}(E_L)_{j_m} \times \partial_{k_1\cdots k_{n-1}}\widehat{A}_{k_n}, \qquad \text{(B10)}$$

and

$$\mathscr{L}_{\widehat{A}2a} = \frac{1}{2}\sum_{n=1}^{\infty} R_{\widehat{A}a}^{j_1\cdots j_n}\partial_{j_1\cdots j_{n-1}}\widehat{A}_{j_n}$$
$$= \frac{1}{2}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\chi_{\widehat{A}a}^{j_1\cdots j_m;k_1,\cdots k_n}(\partial_{j_1\cdots j_{m-1}}\widehat{A}_{j_m})(\partial_{k_1\cdots k_{n-1}}\widehat{A}_{k_n}). \quad \text{(B11)}$$

Here, $\chi_{Ea}^{i_1\cdots i_m;j_1\cdots j_n}$, $\chi_{E\widehat{A}a}^{i_1\cdots i_m;j_1\cdots j_n}$, and $\chi_{\widehat{A}a}^{i_1\cdots i_m;j_1\cdots j_n}$ are regarded as the generalized electromagnetic susceptibilities which are defined by

$$\left[\chi_{Ea}^{i_1\cdots i_m;j_1\cdots j_n}, \chi_{E\widehat{A}a}^{i_1\cdots i_m;j_1\cdots j_n}, \chi_{\widehat{A}a}^{i_1\cdots i_m;j_1\cdots j_n}\right]$$
$$\equiv -e_a\int d^3v\, F_a\left[C_{E2a}^{i_1\cdots i_m;j_1\cdots j_n}, C_{E\widehat{A}a}^{i_1\cdots i_m;j_1\cdots j_n}, C_{\widehat{A}2a}^{i_1\cdots i_m;j_1\cdots j_n}\right]. \quad \text{(B12)}$$

Appendix C: CHARGE DENSITY

In this Appendix, it is shown in detail that the charge density of particles consists of the charge density of gyrocenters and other components including multipole moments which appear due to finite gyroradii of charged particles and turbulent electromagnetic fields. The charge density ρ_c which appears in Poisson's equation, Eq. (33), is given by the variational derivative of the gyrokinetic Lagrangian L_{GK} with respect to the electrostatic potential ϕ as shown in Eq. (34) where the delta function $\delta^3(\mathbf{X}+\rho_a-\mathbf{x})$ is used. Here, to obtain another expression of ρ_c, the variational derivative is written as

$$\rho_c = -\frac{\delta L_{GK}}{\delta\phi} = -\sum_{n=0}^{\infty}(-1)^n\partial_{j_1\cdots j_n}\left(\frac{\partial\mathscr{L}_\Psi}{\partial(\partial_{j_1\cdots j_n}\phi)}\right)$$
$$= -\frac{\partial\mathscr{L}_\Psi}{\partial\phi} + \sum_{n=1}^{\infty}(-1)^n\partial_{j_1\cdots j_n}\frac{\partial\mathscr{L}_\Psi}{\partial(\partial_{j_1\cdots j_{n-1}}(E_L)_{j_n})}$$
$$= \rho_c^{(gc)} - \nabla\cdot\mathbf{P}_G, \quad \text{(C1)}$$

where $\rho_c^{(gc)}$ and $\mathbf{P}_G$ are the gyrocenter charge density and the polarization density vector, defined by

$$\rho_c^{(gc)} \equiv -\frac{\partial\mathscr{L}_\Psi}{\partial\phi} \equiv \sum_a e_a\int d^3v\, F_a \equiv \sum_a e_a N_a^{(g)}, \quad \text{(C2)}$$

and

$$\mathbf{P}_G \equiv \frac{\delta L_{GK}}{\delta\mathbf{E}_L} \equiv \sum_{n=0}^{\infty}(-1)^n\partial_{j_1\cdots j_n}\frac{\partial\mathscr{L}_\Psi}{\partial(\partial_{j_1\cdots j_n}\mathbf{E}_L)}, \quad \text{(C3)}$$

respectively. Then, the electric displacement field $\mathbf{D}$ is given by

$$\mathbf{D} \equiv \mathbf{E} + 4\pi\mathbf{P}_G, \quad \text{(C4)}$$

in terms of which the gyrokinetic Poisson equation is written as

$$\nabla\cdot\mathbf{D} = 4\pi\rho_c^{(gc)}. \quad \text{(C5)}$$

From Eq. (C3), the ith component of the gyrokinetic polarization density vector is written as

$$P_G^i = \sum_{n=0}^{\infty}(-1)^n\partial_{i_1\cdots i_n}Q^{ii_1\cdots i_n}, \quad \text{(C6)}$$

where the multipole moments $Q^{ii_1\cdots i_n}$ $(n = 0,1,2,\cdots)$ are given using Eqs. (B1), (B2), (B3), (B9), and (B10) as

$$Q^{ii_1\cdots i_n} = \frac{\partial\mathscr{L}_\Psi}{\partial(\partial_{i_1\cdots i_n}(E_L)_i)} = \sum_a\left(Q_{0a}^{ii_1\cdots i_n} + Q_{Ea}^{ii_1\cdots i_n} + Q_{\widehat{A}a}^{ii_1\cdots i_n}\right). \quad \text{(C7)}$$

Here, $Q_{0a}^{ii_1\cdots i_n}$ is defined in Eq. (B4). The other multipole moments $Q_{Ea}^{ii_1\cdots i_m}$ and $Q_{\widehat{A}a}^{ii_1\cdots i_n}$ of the electric charge distribution of species a are written in terms of in the linear forms with respect to $(E_L)_i$, $\widehat{A}_i$, and their spatial derivatives as

$$Q_{Ea}^{ii_1\cdots i_m} \equiv \sum_{n=1}^{\infty}\chi_{Ea}^{ii_1\cdots i_m;j_1\cdots j_n}\partial_{j_1\cdots j_{n-1}}(E_L)_{j_n}, \quad \text{(C8)}$$

and

$$Q_{\widehat{A}a}^{ii_1\cdots i_n} = \sum_{m=1}^{\infty}\chi_{E\widehat{A}a}^{ii_1\cdots i_n;k_1\cdots k_m}\partial_{k_1\cdots k_{m-1}}\widehat{A}_{k_m}, \quad \text{(C9)}$$

respectively, where $\chi_{Ea}^{ii_1\cdots i_n;jj_1\cdots j_m}$ and $\chi_{E\widehat{A}a}^{ii_1\cdots i_n;k_1\cdots k_m}$ are defined in Eq. (B12). It is found from Eqs. (A14), (A16), (B4), and (B12) that $Q_{0a}^{i_1\cdots i_m}$ and $Q_{Ea}^{i_1\cdots i_m}$ are both symmetric with respect to arbitrary permutations of the indices $i_1,\cdots,i_m$ because $\alpha_a^{i_1\cdots i_m}$ defined in Eq. (A9) has the same symmetry.

When retaining only the $n=0$ term in Eq. (C9) and using the lowest order distribution function F_{a0} given by the local Maxwellian, the polarization density vector is approximated by

$$P_G^i \simeq \sum_a Q_{Ea}^i \simeq \sum_n \frac{n_{a0}m_a c^2}{B^2}\mathbf{E}_L = \frac{c^2}{4\pi v_A^2}\mathbf{E}_L, \quad \text{(C10)}$$

where $n_{a0} \equiv \int d^3v\, F_{a0}$ and $v_A \equiv B^2/(4\pi\sum_a n_{a0}m_a)$ represent the equilibrium density and the Alfvén velocity, respectively. Equation (C10) presents a well-known expression of polarization. It should be noted that Eq. (C1) with Eq. (C3) including all multipole moments gives the charge density which is equivalent to that presented in Eq. (34). Then, as shown in Ref.[31], the gyrokinetic Poisson equation given by the classical gyrokinetic theory[1–3] based on the WKB formalism can be derived as well from the turbulent part of Eq. (33) using the charge density given by Eq. (34) [or Eq. (C1) with Eqs. (C2) and (C3)].

It is remarked here that the gyrocenter charge density $\rho_c^{(gc)}$ defined in Eq. (C2) satisfies

$$\frac{\partial\rho_c^{(gc)}}{\partial t} + \nabla\cdot\mathbf{j}^{(gc)} = 0, \quad \text{(C11)}$$

where the gyrocenter current density $\mathbf{j}^{(gc)}$ is given by

$$\mathbf{j}^{(gc)} \equiv \sum_a e_a\mathbf{\Gamma}_a^{(gc)} \equiv \sum_a e_a\int d^3v\, F_a\mathbf{u}_{ax}. \quad \text{(C12)}$$

Equation (C11) is derived by taking the velocity-space integral and the species summation of Eq. (102) with the help of Eq. (103). Combining Eqs. (C5) and (C11), one obtains

$$\nabla \cdot \left(\frac{\partial \mathbf{D}}{\partial t} + 4\pi \, \mathbf{j}^{(gc)} \right) = 0, \tag{C13}$$

from which one can write

$$4\pi \, \mathbf{j}_L^{(gc)} = -\frac{\partial \mathbf{D}_L}{\partial t} = \frac{\partial}{\partial t} \left(\nabla \phi_D \right). \tag{C14}$$

In Eq. (C14), the subscript L denotes the longitudinal (or irrotational) vector part, and the potential ϕ_D for $\mathbf{D}_L$ is defined such that $\mathbf{D}_L = -\nabla \phi_D$.

Appendix D: CURRENT DENSITY

In a similar way to Appendix C, this Appendix shows how the current density of particles is expressed by the sum of the current density of gyrocenters and other components induced by finite gyroradii of charged particles and turbulent electromagnetic fields. The gyrokinetic Ampère's law is presented in Eq. (38) which contains the transverse part $\mathbf{j}_T$ of the current density $\mathbf{j}$ given by the variational derivative of the gyrokinetic Lagrangian L_{GK} with respect to the perturbed vector potential $\widehat{\mathbf{A}}$ as shown in Eq. (36). Here, another expression of $\mathbf{j}$ is obtained by writing the variational derivative as

$$\begin{aligned} \frac{1}{c}\mathbf{j} &= \frac{\delta L_{GK}}{\delta \widehat{\mathbf{A}}} = \sum_{n=0}^{\infty} (-1)^n \partial_{j_1 \cdots j_n} \left(\frac{\partial \mathscr{L}_\Psi}{\partial (\partial_{j_1 \cdots j_n} \widehat{\mathbf{A}})} \right) \\ &= \sum_a e_a \int d^3v \, F_a \sum_{n=0}^{\infty} (-1)^{n+1} \partial_{j_1 \cdots j_n} \left(\frac{\partial \Psi_a}{\partial (\partial_{j_1 \cdots j_n} \widehat{\mathbf{A}})} \right) \\ &= \sum_a \frac{e_a}{c} \mathbf{\Gamma}_a. \end{aligned} \tag{D1}$$

The particle flux $\mathbf{\Gamma}_a$ of species a in Eq. (D1) is written as

$$\mathbf{\Gamma}_a = \sum_{n=0}^{\infty} \mathbf{\Gamma}_a^{(n)}, \tag{D2}$$

where $\mathbf{\Gamma}_a^{(n)}$ is defined by

$$\mathbf{\Gamma}_a^{(n)} = c \int d^3v \, F_a (-1)^{n+1} \partial_{j_1 \cdots j_n} \left(\frac{\partial \Psi_a}{\partial (\partial_{j_1 \cdots j_n} \widehat{\mathbf{A}})} \right). \tag{D3}$$

The zeroth-order flux $\mathbf{\Gamma}_a^{(0)}$ is written as

$$\begin{aligned} \mathbf{\Gamma}_a^{(0)} &\equiv -c \int d^3v \, F_a \left(\frac{\partial \Psi_a}{\partial \widehat{\mathbf{A}}} \right) \\ &= \int d^6Z \, \delta^3(\mathbf{X} - \mathbf{x}) \bigg[F_a(\mathbf{Z}, t) \left(\mathbf{v} - \frac{e_a}{m_a c} \widehat{\mathbf{A}} + \mathbf{v}_{Ba} \right) \\ &\quad + \frac{e_a \widetilde{\psi}_a}{B} \frac{\partial F_a}{\partial \mu} \mathbf{v} \bigg] \\ &= \int d^3v \, F_a \left[\left(U - \frac{e_a}{m_a c} \widehat{A}_\parallel \right) \mathbf{b} + \mathbf{v}_{Ba} + \frac{c}{B} \mathbf{b} \times \nabla \langle \psi_a \rangle_\vartheta \right] \\ &\quad + \mathscr{O}(\delta^2), \end{aligned} \tag{D4}$$

which is equivalent to $\mathbf{\Gamma}_a^{(gc)} \equiv \int d^3v \, F_a \mathbf{u}_{ax}$ to the lowest order in δ.

Using Eqs. (B1), (B5), (B10), and (B11), one can represent the derivatives $\partial \mathscr{L}_\Psi / \partial (\partial_{j_1 \cdots j_n} \widehat{A}_k)$ by

$$\begin{aligned} R^{j_1 \cdots j_n k} &\equiv \frac{\partial \mathscr{L}_\Psi}{\partial (\partial_{j_1 \cdots j_n} \widehat{A}_k)} = -\sum_a e_a \int d^3v \, F_a \frac{\partial \Psi_a}{\partial (\partial_{j_1 \cdots j_n} \widehat{A}_k)} \\ &= \sum_a \left(R_{0a}^{j_1 \cdots j_n k} + R_{Ea}^{j_1 \cdots j_n k} + R_{\widehat{A}a}^{j_1 \cdots j_n k} \right), \end{aligned} \tag{D5}$$

where

$$R_{Ea}^{j_1 \cdots j_n k} = \sum_{m=0}^{\infty} \chi_{E\widehat{A}a}^{i_1 \cdots i_m i; j_1 \cdots j_n k} \partial_{i_1 \cdots i_m} (E_L)_i, \tag{D6}$$

$$R_{\widehat{A}a}^{j_1 \cdots j_n k} = \sum_{m=0}^{\infty} \chi_{\widehat{A}a}^{j_1 \cdots j_n k; l_1 \cdots l_m l} \partial_{l_1 \cdots l_m} \widehat{A}_l, \tag{D7}$$

and $R_{0a}^{j_1 \cdots j_n k}$ is defined by Eq. (B6). The coefficients $\chi_{E\widehat{A}a}^{i_1 \cdots i_m i; j_1 \cdots j_n k}$ and $\chi_{\widehat{A}a}^{j_1 \cdots j_n k; l_1 \cdots l_m l}$ are given by Eq. (B12).

Now, it is found from Eqs. (D1)–(D4) that the lth component of $\mathbf{j}$ can be expressed as

$$j^l = (j^{(0)})^l + c \, \partial_k N^{kl}, \tag{D8}$$

where

$$\mathbf{j}^{(0)} \equiv c \frac{\partial \mathscr{L}_\Psi}{\partial \widehat{\mathbf{A}}} = -c \sum_a e_a \int d^3v \, F_a \frac{\partial \Psi_a}{\partial \widehat{\mathbf{A}}} = \sum_a e_a \mathbf{\Gamma}_a^{(0)}, \tag{D9}$$

is regarded as the current of gyrocenters and

$$N^{kl} \equiv \sum_{n=0}^{\infty} (-1)^{n+1} \partial_{j_1 \cdots j_n} R^{j_1 \cdots j_n kl}. \tag{D10}$$

Here, it should be recalled that $\mathbf{j}^{(0)} \equiv \sum_a e_a \mathbf{\Gamma}_a^{(0)}$ defined above equals $\mathbf{j}^{(gc)} \equiv \sum_a e_a \mathbf{\Gamma}_a^{(gc)} \equiv \sum_a e_a \int d^3v \, F_a \mathbf{u}_{ax}$ to the lowest order in δ although the equality does not rigorously holds. When assuming $|\boldsymbol{\rho_a} \cdot \nabla| < 1$ and retaining only the lowest order of N^{kl} in the expansion with respect to $|\boldsymbol{\rho_a} \cdot \nabla|$, one just has the nonturbulent contribution to N^{kl} as

$$N^{kl} \simeq -R^{kl} \simeq -\sum_a R_{0a}^{kl}. \tag{D11}$$

Then, using Eqs. (B8) and (D11) leads to

$$c \partial_k N^{kl} \simeq -\sum_a \left[\nabla \times \left(\frac{c P_{\perp a}}{B} \mathbf{b} \right) \right]^l, \tag{D12}$$

where $P_{\perp a}$ is defined after Eq. (B8). Equation (D12) is a well-known expression of a magnetization current [see Ref.[69]]. Then, from using Eqs. (D4), (D8), (D9), and (D12) with the lowest order distribution function F_{a0} given by the local Maxwellian, the perpendicular component of the equilibrium current can be derived as

$$\begin{aligned} \mathbf{j}_\perp &= \sum_a e_a \int d^3v \, F_{a0} \left(\mathbf{v}_{Ba} + \frac{c}{B} \mathbf{b} \times \nabla \langle \phi \rangle_{\text{ens}} \right) \\ &\quad - \left[\nabla \times \left(\frac{c P_0}{B} \mathbf{b} \right) \right]_\perp \\ &= \frac{c}{B} \mathbf{b} \times \nabla P_0, \end{aligned} \tag{D13}$$

where $P_0 \equiv \sum_a P_{a0} \equiv \sum_a \int d^3v\, F_{a0}\mu B$ denotes the equilibrium pressure and $\sum_a e_a \int d^3v\, F_{a0} = 0$ is used. Equation (D13) presents the magnetization law[69] and one can see that, as pointed out in Ref.[31], the diamagnetic current consistent with the MHD equilibrium $c^{-1}\mathbf{j}\times\mathbf{B} = \nabla P_0$ is correctly derived from the variational formulation with the $\mathbf{v}_{Ba}$ term retained in the potential part of the Hamiltonian given by Eq. (2) with Eq. (3). It is also shown in Ref.[31] that the gyrokinetic Ampère's law given by the classical gyrokinetic theory[1–3] based on the WKB formalism can be derived from the turbulent part of Eq. (38) using Eq. (36) which is equivalent to Eq. (D1) with all the gyroradius expansion terms retained.

Here, $\mathbf{j}^{(0)}$ defined in Eq. (D9) is divided into the longitudinal and transverse parts as $\mathbf{j}^{(0)} = \mathbf{j}_L^{(0)} + \mathbf{j}_T^{(0)}$, where the longitudinal part can be written in terms of the scalar function $\lambda^{(0)}$ as $\mathbf{j}_L^{(0)} = \frac{1}{4\pi}\nabla\lambda^{(0)}$. Similarly, N^{kl} in Eq. (D10) is represented by the sum of two parts,

$$N^{kl} = N_L^{kl} + N_T^{kl}, \tag{D14}$$

where N_L^{kl} and N_T^{kl} are defined such that $\varepsilon_{mnl}\partial^n N_L^{kl} = 0$ and and $\partial_l N_T^{kl} = 0$ are satisfied. Then, there exist V^k and W_n^k in terms of which N_L^{kl} and N_T^{kl} are given by

$$N_L^{kl} = \partial^l V^k, \quad N_T^{kl} = \varepsilon^{lmn}\partial_m W_n^k. \tag{D15}$$

Using Eq. (37) and $\mathbf{j}_L^{(0)} \equiv \frac{1}{4\pi}\nabla\lambda^{(0)}$, one obtains

$$\frac{\lambda}{4\pi} = \frac{\lambda^{(0)}}{4\pi} + c\,\nabla\cdot\mathbf{V}, \tag{D16}$$

and

$$\mathbf{j}_T = \mathbf{j}_T^{(0)} + c\nabla\times\mathbf{M}_W, \tag{D17}$$

where $\mathbf{V}$ is the vector with the components V^k $(k = 1,2,3)$, and the jth component of the vector $\mathbf{M}_W$ is defined by

$$(M_W)_j = \partial_k W_j^k. \tag{D18}$$

It is found from Eq. (D17) that the magnetization field can be represented by $\mathbf{M}_W$ up to a gradient of an arbitrary scalar function. Using $\mathbf{M}_W$, the magnetic intensity field $\mathbf{H}$ is defined by

$$\mathbf{H} = \mathbf{B} + \widehat{\mathbf{B}} - 4\pi\mathbf{M}_W \tag{D19}$$

which is distinguished from $\mathbf{H}_\#$ given in Eq. (147). Then, using Eq. (D19), the gyrokinetic Ampère's law, Eq. (38) can be rewritten as

$$\nabla\times\mathbf{H} = \frac{4\pi}{c}\mathbf{j}_T^{(0)}. \tag{D20}$$

Appendix E: ENERGY BALANCE IN ELECTROMAGNETIC GYROKINETIC TURBULENCE

This Appendix presents energy balance equations in electromagnetic gyrokinetic turbulence. Here, the Cartesian coordinate system is used and three-dimensional vectors are written in terms of boldface letters. The energy of a single charged particle of species a is denoted by $\mathscr{E}_a$ which is equal to the gyrocenter Hamiltonian H_{GYa} in Eq. (2) and written as

$$\mathscr{E}_a \equiv \frac{1}{2}m_a U^2 + \mu B + e_a\Psi_a \equiv H_{GYa} = \frac{\partial L_{GYa}}{\partial \mathbf{u}_{aZ}}\cdot\mathbf{u}_{aZ} - L_{GYa}. \tag{E1}$$

It can be shown from Eqs. (28), (32), and (E1) that the total derivative of H_{GYa} is written as

$$\begin{aligned}\dot{\mathscr{E}}_a &\equiv \left(\frac{d}{dt}\right)_a H_{GYa} \equiv \left(\frac{\partial}{\partial t} + \mathbf{u}_{aZ}\cdot\frac{\partial}{\partial \mathbf{Z}}\right) H_{GYa} \\ &= -\left(\frac{\partial L_{GYa}}{\partial t}\right)_u = e\frac{\partial \Psi_a}{\partial t} + \mu\frac{\partial B}{\partial t} - \frac{e_a}{c}\mathbf{u}_{ax}\cdot\frac{\partial \mathbf{A}_a^*}{\partial t},\end{aligned} \tag{E2}$$

where $(\partial L_{GYa}/\partial t)_u$ denotes the time derivatives of L_{GYa} with $\mathbf{u}_{aZ}$ kept fixed in L_{GYa}. Multiplying Eq. (E2) with F_a and taking its velocity-space integral, the local energy balance equation for the system of the single particle species is obtained as

$$\begin{aligned}&\frac{\partial}{\partial t}\left(\int d^3v\, F_a\mathscr{E}_a\right) + \nabla\cdot\left(\int d^3v\, F_a\mathscr{E}_a\mathbf{u}_{ax}\right) \\ &\quad = \int d^3v\, \left(F_a\dot{\mathscr{E}}_a + \mathscr{K}_a\mathscr{E}_a\right),\end{aligned} \tag{E3}$$

where the gyrokinetic Boltzmann equation shown in Eq. (102) is used and $\mathbf{u}_{ax} = (d/dt)_a\mathbf{X}$ represents the gyrocenter velocity defined at the right-hand side of Eq. (11).

Next, the energy balance in the whole system including particles of all species and the turbulent electromagnetic fields is considered. From Eq. (40), one finds

$$\begin{aligned}&\sum_a\int d^3v\, F_a\frac{\partial L_{GYa}}{\partial t} + \frac{\partial \mathscr{L}_F}{\partial t} \\ &\quad = \sum_J\left\{\frac{\partial \mathscr{L}_{GKF}}{\partial(\partial_J\phi)}\frac{\partial(\partial_J\phi)}{\partial t} + \frac{\partial \mathscr{L}_{GKF}}{\partial(\partial_J\mathbf{A})}\cdot\frac{\partial(\partial_J\mathbf{A})}{\partial t}\right. \\ &\qquad \left. + \frac{\partial \mathscr{L}_{GKF}}{\partial(\partial_J\widehat{\mathbf{A}})}\cdot\frac{\partial(\partial_J\widehat{\mathbf{A}})}{\partial t}\right\} + \frac{\partial \mathscr{L}_{GKF}}{\partial\lambda}\frac{\partial\lambda}{\partial t}.\end{aligned} \tag{E4}$$

Then, taking the species summation of Eq. (E3) and using Eqs. (E1), (E2), and (E4), they yield

$$\begin{aligned}&\frac{\partial}{\partial t}\left(\sum_a\int d^3v\, F_a\, H_{GYa} - \mathscr{L}_F\right) \\ &\quad + \nabla\cdot\left(\sum_a\int d^3v\, F_a\, H_{GYa}\,\mathbf{u}_{ax}\right) + \sum_J\left\{\frac{\partial \mathscr{L}_{GKF}}{\partial(\partial_J\phi)}\frac{\partial(\partial_J\phi)}{\partial t}\right. \\ &\quad \left. + \frac{\partial \mathscr{L}_{GKF}}{\partial(\partial_J\mathbf{A})}\cdot\frac{\partial(\partial_J\mathbf{A})}{\partial t} + \frac{\partial \mathscr{L}_{GKF}}{\partial(\partial_J\widehat{\mathbf{A}})}\cdot\frac{\partial(\partial_J\widehat{\mathbf{A}})}{\partial t}\right\} + \frac{\partial \mathscr{L}_{GKF}}{\partial\lambda}\frac{\partial\lambda}{\partial t} \\ &\quad = \sum_a\int d^3v\, \mathscr{K}_a\, H_{GYa}.\end{aligned} \tag{E5}$$

Now, the variational equations, $\delta L_{GK}/\delta\phi = 0$, $\delta L_{GK}/\delta\widehat{\mathbf{A}} = 0$, and $\delta L_{GK}/\delta\lambda = 0$, which are equivalent to the gyrokinetic Poisson equation, Ampère's law, and the Coulomb gauge condition, are used to rewrite Eq. (E5) as

$$\frac{\partial}{\partial t}\left(\sum_a \int d^3v\, F_a\, H_{GYa} - \mathscr{L}_F\right) + \nabla\cdot\left(\sum_a \int d^3v\, F_a\, H_{GYa}\,\mathbf{u}_{ax}\right) + \sum_{n=1}^{\infty}\sum_{k=1}^{n}(-1)^{k-1} \times \partial_{j_k}\left\{\partial_{j_1\cdots j_{k-1}}\left(\frac{\partial \mathscr{L}_{GKF}}{\partial(\partial_{j_1\cdots j_n}\phi)}\right)\frac{\partial(\partial_{j_{k+1}\cdots j_n}\phi)}{\partial t} + \partial_{j_1\cdots j_{k-1}}\left(\frac{\partial \mathscr{L}_{GKF}}{\partial(\partial_{j_1\cdots j_n}\widehat{A}_l)}\right)\frac{\partial(\partial_{j_{k+1}\cdots j_n}\widehat{A}_l)}{\partial t}\right\} + \nabla\cdot\left[\frac{c}{4\pi}\mathbf{E}_T\times(\mathbf{B}+\widehat{\mathbf{B}}-4\pi\mathbf{M}_\#)\right] + \partial_i\left[\Xi^i_j\frac{\partial B^j}{\partial t}\right] = \left[\mathbf{j}_\# - \frac{c}{4\pi}\nabla\times(\mathbf{B}+\widehat{\mathbf{B}})\right]\cdot\mathbf{E}_T + \sum_a\int d^3v\,\mathscr{K}_a\,H_{GYa}, \quad \text{(E6)}$$

where Ξ^i_j on the left-hand side is defined by

$$\Xi^i_j = \frac{\partial \mathscr{L}_{GK}}{\partial(\partial_i B^j)} = \sum_a \int d^3v\, F_a \frac{\partial L_{GYa}}{\partial(\partial_i B^j)}. \quad \text{(E7)}$$

Equation (E6) can be further deformed to obtain the local energy balance equation of the whole system as

$$\frac{\partial}{\partial t}\left[\sum_a\int d^3v F_a\left(\frac{1}{2}m_a\left|\mathbf{v}-\frac{e_a}{m_a c}\widehat{\mathbf{A}}\right|^2 - \frac{e_a}{c}\mathbf{v}_{Ba}\cdot\widehat{\mathbf{A}} - \frac{e_a^2}{2c^2B}\frac{\partial}{\partial\mu}\left[\widetilde{(\mathbf{v}\cdot\widehat{\mathbf{A}})^2}\right]\right) + \frac{|\mathbf{E}_L|^2}{8\pi} + \frac{\mathbf{E}_L\cdot\mathbf{P}_D}{2} + \frac{1}{2}\sum_{n=1}^{\infty}(\partial_{j_1\cdots j_n}(E_L)_i)Q_E^{ij_1\cdots j_n} + \frac{\mathbf{E}_T\cdot\mathbf{D}_L}{4\pi} + \frac{|\mathbf{B}+\widehat{\mathbf{B}}|^2}{8\pi}\right] + \nabla\cdot\left[\sum_a\int d^3v F_a\left\{\frac{1}{2}m_aU^2 + \mu B + e_a(\Psi_a - \phi(x))\right\}\mathbf{u}_{ax} + \frac{c}{4\pi}\mathbf{E}\times\mathbf{H}_\# + \frac{1}{4\pi}\phi_D\frac{\partial \mathbf{E}_T}{\partial t} + \phi\frac{\partial(\mathbf{P}_G)_T}{\partial t}\right] + \partial_i\left[\Xi^i_j\frac{\partial B^j}{\partial t}\right] + \partial_i\left[\sum_{n=1}^{\infty}\sum_{k=0}^{n-1}(-1)^{n-k}\left\{(\partial_{j_{n-k}\cdots j_{n-1}}(E_L)_{j_n}) \times \partial_{j_1\cdots j_{n-k-1}}\left(\frac{\partial Q^{ij_1\cdots j_n}}{\partial t}\right) + c(\partial_{j_{n-k}\cdots j_n}(\widehat{E}_T)_l) \times \partial_{j_1\cdots j_{n-k-1}}R^{ij_1\cdots j_n l}\right\}\right] + \nabla\cdot\left[\frac{c}{4\pi}\widehat{\mathbf{E}}_T\times(\mathbf{B}+\widehat{\mathbf{B}}) - \frac{\lambda}{4\pi}\widehat{\mathbf{E}}_T\right] + \partial_i\left[cN^{ij}(\widehat{E}_T)_j\right] = \left[(\mathbf{j}_\#)_T - \frac{c}{4\pi}\nabla\times(\mathbf{B}+\widehat{\mathbf{B}})\right]\cdot(\mathbf{E}_T+\mathbf{E}_L) + \sum_a\int d^3v\,\mathscr{K}_a\left\{\frac{1}{2}m_aU^2 + \mu B + e_a(\Psi_a - \phi(x))\right\}, \quad \text{(E8)}$$

where $\mathbf{H}_\#$ is defined in Eq. (147). The rate of change in the sum of the kinetic and electromagnetic energy densities is described by Eq. (E8). There appear the effects of polarization and magnetization including all multipole moments. The energy flux on the left-hand side of Eq. (E8) contains the kinetic energy flow due to the gyrocenter motion, the Poynting vector, and the extra contributions due to the electromagnetic microturbulence. The last terms on the left-hand side of Eq. (E8) can be deformed into

$$\nabla\cdot\left[\frac{c}{4\pi}\widehat{\mathbf{E}}_T\times(\mathbf{B}+\widehat{\mathbf{B}}) - \frac{\lambda}{4\pi}\widehat{\mathbf{E}}_T\right] + \partial_i\left[cN^{ij}(\widehat{E}_T)_j\right] = \nabla\cdot\left[\frac{c}{4\pi}\widehat{\mathbf{E}}_T\times\mathbf{H} - \frac{\lambda^{(0)}}{4\pi}\widehat{\mathbf{E}}_T + c\mathbf{V}\cdot\nabla\widehat{\mathbf{E}}_T\right] + \partial_i\left[c\varepsilon^{ijk}W^l_j\partial_l(\widehat{E}_T)_k + W^i_j\frac{\partial\widehat{B}^j}{\partial t}\right] \quad \text{(E9)}$$

where $\mathbf{H}$ is given in Eq. (D19). It is seen from Eqs. (E8) and (E9) that the energy balance equation two types of the Poynting vector, $(c/4\pi)\mathbf{E}\times\mathbf{H}_\#$ and $(c/4\pi)\widehat{\mathbf{E}}_T\times\mathbf{H}$, where $\mathbf{E}\equiv\mathbf{E}_L+\mathbf{E}_T$, $\mathbf{E}_L = -\nabla\phi$, $\mathbf{E}_T = -c^{-1}\partial\mathbf{A}/\partial t$ and $\widehat{\mathbf{E}}_T = -c^{-1}\partial\widehat{\mathbf{A}}/\partial t$. On the right-hand side of Eqs. (E6) and (E8), the effects of collisions and/or external sources are represented by terms including $\mathscr{K}_a$. It is noted that these terms can be written as the divergence of classical energy transport flux when $\mathscr{K}_a$ is given by the collision operator including the finite gyroradius effect. In addition, terms including $\mathbf{E}_L\equiv-\nabla\phi$ on the right-hand side of Eq. (E8) are given in the divergence form as

$$\left[(\mathbf{j}_\#)_T - \frac{c}{4\pi}\nabla\times(\mathbf{B}+\widehat{\mathbf{B}})\right]\cdot\mathbf{E}_L = \nabla\cdot\left[\frac{c}{4\pi}\nabla\phi\times(\mathbf{B}_\# - \mathbf{B} - \widehat{\mathbf{B}})\right]. \quad \text{(E10)}$$

Therefore, in the stationary background magnetic field where $\mathbf{E}_T\equiv -c^{-1}\partial\mathbf{A}/\partial t = 0$ and with no external energy sources, Eqs. (E6) and (E8) take the conservation form. Furthermore, even when $\mathbf{E}_T\equiv -c^{-1}\partial\mathbf{A}/\partial t \neq 0$, the ensemble average of Eq. (E6) and (E8) can take the form of total energy conservation on macroscopic spatiotemporal scales in the background field determined by the condition in Eq. (145). It can also be confirmed from comparison with Eq. (22) in Ref.[45] that the kinetic and electromagnetic energies, the kinetic energy flux, the Poynting vector, and the longitudinal and transverse electric fields in the the energy conservation equation of the Vlasov-Darwin system are retained in Eq. (E8). There, additional terms due to finite-gyroradius effects and electromagnetic microturbulence are included as well.

[1] T. M. Antonsen, Jr. and B. Lane, Phys. Fluids **23**, 1205 (1980).
[2] P.J. Catto, W.M. Tang and D.E. Baldwin, Plasma Phys. **23**, 639 (1981).
[3] E. A. Frieman and L. Chen, Phys. Fluids **25**, 502 (1982).
[4] A. J. Brizard and T. S. Hahm, Rev. Mod. Phys. **79**, 421 (2007).
[5] H. Sugama, Phys. Plasmas **7**, 466 (2000).
[6] A. A. Schekochihin, S. C. Cowley, W. Dorland, G. W. Hammett, G. G. Howes, E. Quataert, and T. Tatsuno, Astrophys. J., Suppl. **182**, 310 (2009).
[7] J. A. Krommes, Ann. Rev. Fluid Mech. **44**, 175 (2012).
[8] W. Horton, *Turbulent Transport in Magnetized Plasmas*, 2nd edition (World Scientific, Singapore, 2018), Chap. 12.
[9] H. Sugama, M. Okamoto, W. Horton, and M. Wakatani, Phys. Plasmas **3**, 2379 (1996).
[10] R. D. Hazeltine and J. D. Meiss, *Plasma Confinement* (Addison-Wesley, Redwood City, California, 1992), Chap. 7.10.
[11] A. M. Dimits, G. Bateman, M. A. Beer, B. I. Cohen, W. Dorland, G. W. Hammett, C. Kim, J. E. Kinsey, M. Kotschenreuther, A. H. Kritz, L. L. Lao, J. Mandrekas, W. M. Nevins, S. E. Parker, A. J. Redd, D. E. Shumaker, R. Sydora, and J. Weiland, Phys. Plasmas **7**, 969 (2000).
[12] F. Jenko, W. Dorland, M. Kotschenreuther, and B. N. Rogers, Phys. Plasmas **7**, 1904 (2000).

[13] J. Candy and R. E. Waltz, J. Comp. Phys. **186**, 545 (2003).
[14] T.-H. Watanabe and H. Sugama, Nucl. Fusion **46**, 24 (2006).
[15] A. G. Peeters, Y. Camenen, F. J. Casson, W. A. Hornsby, A. P. Snodin, D. Strintzi, G. Szepesi, Comp. Phys. Comm. **180**, 2650 (2009).
[16] R. G. Littlejohn, J. Math. Phys. **23** , 742 (1982).
[17] B. Scott and J. Smirnov, Phys. Plasmas **17**, 112302 (2010).
[18] A. J. Brizard and N. Tronko, Phys. Plasmas **18**, 082307 (2011).
[19] F. I. Parra and I. Calvo, Plasma Phys. Control. Fusion **53**, 045001, (2011).
[20] H. Qin, J. W. Burby, and R. C. Davidson, Phys. Rev. E **90**, 043102 (2014).
[21] H. Sugama, M. Nunami, M. Nakata, and T.-H. Watanabe Phys. Plasmas **24**, 020701 (2017); H. Sugama, Rev. Mod. Plasma Phys. 1:9 (2017).
[22] P. Fan, H. Qin, and J. Xiao, Plasma Sci. Technol. **23**, 105103 (2021).
[23] Z. Lin, T. S. Hahm, W. W. Lee, W. M. Tang, and R. B. White, Science **281**, 1835 (1998).
[24] Y. Idomura, Phys. Plasmas **24**, 080701 (2017).
[25] S. Ku, C.S. Chang, and P. H. Diamond, Nucl. Fusion **49**, 115021 (2009).
[26] W. X. Wang, T. S. Hahm, S. Ethier, G. Rewoldt, W. W. Lee, W. M. Tang, S. M. Kaye, and P. H. Diamond, Phys. Rev. Lett. **102**, 035005 (2009).
[27] V. Grandgirard, J. Abiteboul, J. Bigot, T. Cartier-Michaud, N. Crouseilles, G. Dif-Pradalier, Ch. Ehrlacher, D. Esteve, X. Garbet, Ph. Ghendrih, G. Latu, M. Mehrenberger, C. Norscini, Ch. Passeron, F. Rozar, Y. Sarazin, E. Sonnendrücker, A. Strugarek, and D. Zarzoso, Comput. Phys. Commun. **207**, 35 (2016).
[28] A. Bottino and E. Sonnendrücker, J. Plasma Phys. **81**, 435810501 (2015).
[29] J. A. Heikkinen, S. J. Janhunen, T. P. Kiviniemi, and F.Ogando, J. Comput. Phys. **227**, 5582 (2008).
[30] N. R. Mandell, A. Hakim, G. W. Hammett, M. Francisquez, J. Plasma Phys. **86**, 905860109 (2020).
[31] H. Sugama, S. Matsuoka, and M. Nunami, Phys. Plasmas **29**, 052509 (2022).
[32] K. H. Burrell, Phys. Plasmas **27**, 060501 (2020).
[33] A.H. Boozer, Phys. Fluids **26**, 496 (1983).
[34] J. Nührenberg and R. Zille, Phys. Lett. **129**, 113 (1988).
[35] D. A. Spong, Phys. Plasmas **12**, 056114 (2005).
[36] P. Helander, Rep. Prog. Phys. **77**, 087001 (2014).
[37] H. Sugama, S. Matsuoka, M. Nunami, and S. Satake, Phys. Plasmas **28**, 022312 (2021).
[38] W. A. Newcomb, Nucl. Fusion Suppl. Part 2, 451 (1962).
[39] A. J. Brizard and C. Tronci, Phys. Plasmas **23**, 062107 (2016).
[40] E. Hirvijoki, J. W. Burby, D. Pfefferlé, and A. J. Brizard, J. Phys. A: Math. Theor. **53**, 235204 (2020).
[41] H. Cendra, D. D. Holm, M. J. W. Hoyle, and J. E. Marsden, J. Math. Phys. **39**, 3138 (1998).
[42] J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems*, Texts in Applied Mathematics (Springer, 1999), Chap. 13.5.
[43] J. Squire, H. Qin, W.M. Tang, and C. Chandre, Phys. Plasmas **20**, 022501 (2013).
[44] H. Sugama, M. Nunami, S. Satake, and T.-H. Watanabe, Phys. Plasmas **25**, 102506 (2018).
[45] H. Sugama, T.-H. Watanabe, and M. Nunami, Phys. Plasmas **20**, 024503 (2013).
[46] A. J. Brizard, J. Plasma Phys. **87**, 905870307 (2021).
[47] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (Pergamon, New York, 1975), Chap. 11.
[48] F. L. Hinton and R. D. Hazeltine, Rev. Mod. Phys. **48**, 239 (1976).
[49] S. P. Hirshman and D. J. Sigmar, Nucl. Fusion **21**, 1079 (1981).
[50] P. Helander and D. J. Sigmar, *Collisional Transport in Magnetized Plasmas* (Cambridge University Press, Cambridge, 2002), Chap. 8.
[51] H. Sugama, T.-H. Watanabe, M. Nunami, and S. Nishimura, Plasma Phys. Control. Fusion **53**, 024004 (2011).
[52] F. I. Parra, M. Barnes, I. Calvo, and P. J. Catto, Phys. Plasmas **19**, 056116 (2012).
[53] I. Calvo and F. I. Parra, Plasma Phys. Control. Fusion **57**, 075006 (2015).
[54] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1998), Chap. 6.
[55] H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics*, 3rd ed. (Addison-Wesley, San Francisco, 2002), p. 593.
[56] J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems*, Texts in Applied Mathematics (Springer, 1999), Chap. 4.3.
[57] H. Sugama, T.-H. Watanabe, and M. Nunami, Phys. Plasmas **22**, 082306 (2015).
[58] A. J. Brizard, Phys. Plasmas **11**, 4429 (2004).
[59] J. W. Burby, A. J. Brizard, and H. Qin, Phys. Plasmas **22**, 100707 (2015).
[60] S. Hamada, Nucl. Fusion **2**, 23 (1962).
[61] A. J. Brizard, Phys. Plasmas **30**, 102106 (2023).
[62] H. Sugama, T.-H. Watanabe, and M. Nunami, Phys. Plasmas **21**, 012515 (2014).
[63] M. Wakatani, *Stellarator and Heliotron Devices* (Oxford University Press, New York, 1998), Chap. 7.
[64] A. H. Boozer, Phys. Fluids **26**, 1288 (1983).
[65] H. Sugama and S. Nishimura, Phys. Plasmas **9**, 4637 (2002).
[66] I. G. Abel, G. G. Plunk, E. Wang, M. Barnes, S. C. Cowley, W. Dorland, and A. A. Schekochihin, Rep. Prog. Phys. **76**, 116201 (2013).
[67] H. Sugama and W. Horton, Phys. Plasmas **5**, 2560 (1998).
[68] F. I. Parra, M. Barnes, and A. G. Peeters, Phys. Plasmas **18**, 062501 (2011).
[69] R. D. Hazeltine and J. D. Meiss, *Plasma Confinement* (Addison-Wesley, Redwood City, California, 1992), Chap. 4.5.