

Higher-order energy-conserving gyrokinetic theory

Alexey Mishchenko¹ and Alain J. Brizard²

¹*Max-Planck-Institut für Plasmaphysik, EURATOM-Association, D-17491, Greifswald, Germany*

²*Department of Chemistry and Physics, Saint Michael's College, Colchester, VT 05439, USA*

A higher-order self-consistent energy-conserving gyrokinetic system of equations is derived. It is shown that additional terms appear in the quasineutrality condition. These terms are nonlinear in the electric field. The derivation includes higher-order terms in the gyrokinetic Hamiltonian (needed for the energy conservation) and employs a variational principle which automatically provides all the conservation laws through the Noether theorem. The equations derived here can be applied in certain transition layers such as the stellarator transport barriers caused by the transition between the electron and ion root regimes. The theory may also be of interest for the edge plasma, where the nonlinear terms in the quasineutrality equation could be relevant. The equations derived are simple enough and can readily be used in gyrokinetic codes.

I. INTRODUCTION

An initial motivation for this work was to derive the gyrokinetic system of equations in presence of narrow shear layers. As *shear layers* we understand zones of an abrupt change in a background radial electric field (on a spatial scale down to the ion gyroradius ρ_i) which is, however, not accompanied by a similarly rapid variation in the plasma pressure [1]. This in contrast to *pedestals*, where both the radial electric field and the plasma pressure change on the same scale of the order of poloidal ion gyroradius $\rho_{i(\text{pol})}$ (see Ref. [2]). The gyrokinetic theory in pedestals has been developed in Refs. [3, 4] employing the ordering $\rho_{i(\text{pol})} \gg \rho_i$ (here, ρ_i is the ion gyroradius) which is usually the case for a typical tokamak plasma. The $\mathbf{E} \times \mathbf{B}$ velocity $U_E = E_r/B$ has been assumed to be much larger than the ion magnetic drift velocity v_{di} (strong background electric field ordering; recall that $U_E \sim v_{di}$ in the conventional drift ordering). In this paper, we adopt the “intermediate” background electric field ordering assuming U_E to be much larger than the drift velocity but at the same time smaller than the ion thermal velocity $v_{di} \ll U_E \ll v_{thi}$. The spatial scale of the associated electric field can be comparable with the ion gyroradius. The ambient magnetic field, however, is only allowed to vary slowly on the system scale.

Shear layers can develop, for example, in stellarators. In non-axisymmetric magnetic geometry, the neoclassical fluxes become ambipolar only at a certain radial electric field E_r [5, 6]. The ambipolarity condition is known to have two stable solutions (the ion and the electron roots). Frequently, these two solutions of the ambipolarity condition coexist [7–9] in the same plasma at different radial positions. A shear layer is then formed at the transition between the ion and electron roots. The radial electric field can change rapidly inside the layer (on a scale down to the ion gyroradius). At the same time, the temperature and density profiles remain relatively smooth (in contrast to a tokamak pedestal). This happens because the radial electric field vanishes inside the shear layer which makes the neoclassical diffusion coefficients large there. As a result, the increased neoclassical diffusion smooths the temperature and density profiles. Another example is a shear layer on the magnetic separatrix (both in tokamaks and stellarators). It can be formed because the background electric field outside the separatrix is set by the edge boundary condition (scrape-off-layer physics) whereas inside it is defined by the momentum transport (in tokamaks) or ambipolarity condition (in stellarators).

Shear layers can strongly affect the plasma microturbulence. For example, in stellarators stabilization of the turbulence and reduction of the anomalous transport has been observed [9] in the transition region from the electron to the ion root. There is an interesting question whether bifurcations in turbulent transport can be initiated in non-axisymmetric geometry by the preconditioning of neoclassical ambipolarity [10]. As mentioned above, the available formulations of the gyrokinetic theory [3, 4, 11, 12] in presence of a large (beyond the drift ordering) background electric field have adopted the ordering $L_E \gg \rho_i$ with L_E the characteristic scale of the background electric field. In this paper we derive the gyrokinetic system of equations which can be used to describe the microturbulence in the presence of the shear layers with the width comparable to the ion gyroradius.

We will see that higher-order gyrokinetic corrections must be considered in order to include both the microturbulence (satisfying the usual gyrokinetic ordering) and the shear-layer field. Technically, the calculation presented here includes the derivation of the third-order correction to the gyrokinetic Hamiltonian. Recently, a number of publications [13, 14] argued that the higher-order gyrokinetic corrections are essential to address correctly the toroidal momentum transport in tokamak plasma. While still being controversial [15], this issue certainly deserves further study. In this paper, we derive a self-consistent energy-conserving gyrokinetic theory accurate to the third order. The higher-order corrections

may also be of interest for edge plasma where the turbulence can be rather strong so that the low-order perturbation theory may not be sufficient.

We perform our derivation using the Lie transform approach. There are several versions of the Lie transform available. For the first time, Lie series were introduced into the perturbation theory by Hori [16]. The first power series representation of the Lie transform was developed by Deprit (Deprit series, see Ref. [17]). Dewar gave the general representation for non-analytic transformations (such that the generating function need not to have a power representation, see Ref. [18]). Dragt and Finn have proven [19] that a large class of canonical transformations can be represented as a product of the Lie operators with the associated Lie algebra defined by the Poisson bracket structure. Based on this insight they suggested a formulation of the Lie transform alternative to the Deprit approach. This formulation does not easily fit into the general representation of Dewar [18]. However, it has sometimes practical benefits since there may be fewer terms in the perturbation series. The Deprit series formulation was used by Dubin [20] and Littlejohn [21] for an early derivations of the Hamiltonian gyrokinetic theory. In the following, the Dragt-Finn formulation was extended to the completely general non-canonical Lie transform [22] generated by an arbitrary vector field on a differentiable manifold (representing the extended phase space). Obviously, the canonical Lie transform is a subset of this general transformation where the vector field is determined by the external derivative of the generating function and the corresponding symplectic structure. The non-canonical Lie transform was used by Hahm [23, 24] to derive the gyrokinetic theory in general magnetic geometry. The next important step in the formulation of the gyrokinetic theory was taken by Brizard introducing the extended phase space in the gyrokinetic theory [25]. This facilitated the time-dependent calculations and allowed the theory to be cast into a variational form [25, 26] (using the principle of least action from the classical field theory). In this formulation, symmetries of the action automatically deliver the associated conserved quantities through the Noether theorem. A detailed description of this development can be found in the recent review [27].

In this paper, we derive higher-order corrections to the gyrokinetic Hamiltonian theory using both the Deprit series approach and the more conventional Dragt-Finn formulation. We find that these different approaches to the Lie transform give identical results. We derive the nonlinear corrections to the gyrokinetic polarization density using the variational approach. Similar terms appeared in the quasineutrality equation already in Dubin's derivation [20]. However, this early derivation was not energy-conserving to the sufficient order because the gyrokinetic Hamiltonian was derived only to the second order and no variational formalism was used (the pullback was applied to the distribution function in order to compute the polarization density). Recently, Dubin's derivation [20] has been reproduced by Parra and Catto [14] using the reduced-Vlasov approach. In Ref. [28], it has been shown that the reduced-Vlasov calculation can be reformulated in terms of the Lie transform.

The paper is organised as follows. In Sec. II, the derivation based on the Deprit series is presented. In Sec. III, the derivation employs the Dragt-Finn formulation. The gyrokinetic quasineutrality equation and the gyrokinetic equations of motion are derived in Sec. IV. Finally, in Sec. V, conclusions are made. Many details of the calculations are shown in the Appendices A-C.

II. DERIVATION USING DEPRIT SERIES

We split the electrostatic potential into the microturbulence potential ϕ and the mesoscale potential Φ (set by neoclassical theory etc). For the microturbulence, the usual gyrokinetic ordering [27] is assumed:

$$\epsilon = q\phi/T \sim k_{\parallel}/k_{\perp} \sim \rho_i/L_n \sim \omega/\omega_c \ll 1 \quad (1)$$

Here, k_{\perp} and k_{\parallel} are the perpendicular and parallel wave numbers characterising the microturbulence, ρ_i is the ion gyroradius, L_n is the spatial scale of the unperturbed plasma density, ω is the characteristic frequency of the field perturbation and $\omega_c = eB/m$ is the cyclotron frequency. For the mesoscale potential, we assume $e\Phi/T \sim \sqrt{\epsilon} \ll 1$; the characteristic spatial scale of Φ can be comparable to the gyroradius $L_E \sim \rho_i$. In this ordering, the mesoscale $\mathbf{E} \times \mathbf{B}$ velocity $U_E = |\nabla_r \Phi|/B \sim \sqrt{\epsilon} v_{thi}$ is much larger than the drift velocity $v_{di}/U_E \sim \sqrt{\epsilon}$ (but still much smaller than the thermal speed $U_E/v_{thi} \sim \sqrt{\epsilon}$). For such a field, the standard gyrokinetic theory [23, 24] cannot be used since the $\mathbf{E} \times \mathbf{B}$ velocity is comparable to the drift one there. It is also beyond the scope of theories [3, 4, 11, 12] where the ordering $L_E \gg \rho_i$ has been adopted. Note that the standard gyrokinetic theory [23, 24] is truncated at the second order in the gyrokinetic parameter ϵ . In what follows, we will derive the higher-order gyrokinetic theory for the mesoscale potential Φ using $\sqrt{\epsilon}$ as a small parameter [we will truncate our derivation at the fourth order in $\sqrt{\epsilon}$; note that $(\sqrt{\epsilon})^4 = \epsilon^2$ so that the microturbulence will be treated at the usual level of accuracy]. The ambient magnetic field is assumed to change on a long spatial scale L_B so that $\rho_i/L_B = \epsilon_B \ll 1$. Following Refs. [23, 24, 27] we carry our derivation only to the first order in ϵ_B . In practice, it means that the Poisson brackets (the symplectic structure) employed in this paper coincide with the usual guiding center Poisson brackets; the magnetic curvature and gradient drifts are given by the usual expressions (as in Ref. [23]). Extension of this derivation to the higher orders

in ϵ_B is possible [29] but leads to rather bulky expressions [14, 30] which we would like to avoid here since the focus of this paper is on the energy-conserving derivation of the nonlinear corrections to the polarization density.

Derivation of the gyrokinetic theories proceeds usually in two steps [27]. First, the initial particle phase space is transformed into the so-called guiding-center phase space which results from the perturbation theory eliminating the gyro-dependencies associated with the inhomogeneous ambient magnetic field (this transformation includes an expansion in the small parameter ϵ_B). The second step is the gyrokinetic Lie transform. This transformation includes an expansion in the small parameter $\sqrt{\epsilon}$. It will be considered in this paper (the guiding-center transformation to the order desired has already been done [27, 29]).

In the ordering chosen here, the guiding-center Hamiltonian function has the form:

$$h = \left(\frac{mv_{\parallel}^2}{2} + \mu B \right) + \sqrt{\epsilon} e\Phi + \epsilon e\phi \quad (2)$$

As mentioned, this ordering can be applied in two cases. First, to study the gyrokinetic microturbulence (described by ϕ) in presence of certain equilibrium structures such as the shear layers which can develop in the transition area between the ion and electron roots of the ambipolarity condition in stellarator plasmas. The electric field corresponding to these structures (described by Φ) has indeed other temporal scales comparing to that of the microturbulence ϕ . In the stellarator example, Φ is a neoclassical equilibrium electric field which can be found solving the equilibrium ambipolarity condition $\Gamma_i = \Gamma_e$ with $\Gamma_{i,e}$ being the radial particle fluxes of the ions and electrons, respectively. Another situation for which the ordering used in Eq. (2) could be of interest is, for example, the plasma edge. The edge turbulence can be rather strong so that the low-order perturbation theory may be not sufficient and higher-order corrections may be relevant. In this case, Φ describes the strong edge turbulence (ordered with $\sqrt{\epsilon}$) and ϕ (which would correspond to a weaker *core turbulence*) becomes obsolete and can be neglected everywhere in the equations.

Taking $\sqrt{\epsilon}$ as the small parameter for the perturbation theory to be constructed, one can split the Hamiltonian Eq. (2) into parts of different orders:

$$h_0 = \frac{mv_{\parallel}^2}{2} + \mu B, \quad h_1 = e\Phi, \quad h_2 = e\phi \quad (3)$$

where $\mu = mv_{\perp}^2/(2B)$ is the guiding-center magnetic moment, v_{\parallel} is the parallel velocity and B is the ambient magnetic field. The gyrokinetic Lie transform maps the guiding-center phase space into the gyrocenter phase space [27] so that the gyro-dependencies associated with the short-spatial-scale electric field are eliminated (here, the microturbulence field ϕ and the mesoscale field Φ are to be considered). The new gyrokinetic phase space coordinates can be represented as the old guiding-center coordinates plus small corrections associated with the electrostatic short-scale field [see Eqs. (49) and (50) for example]. Expressions for this transformation can be straightforwardly derived to the order desired once the corresponding gyrokinetic Hamiltonian and the Poisson brackets are known (as in Ref. [28]), although the explicit form of these expressions is not needed anywhere in the derivation of the theory (in contrast to the reduced-Vlasov approach which relies on the explicit expressions for the coordinate transformation). All the quantities have to be evaluated in terms of the transformed gyrokinetic phase-space coordinates [27] after the Lie transform has been applied. This is perhaps one of the biggest advantages of the Lie transform technique that it does not mix the old and new phase-space coordinates. One automatically operates in terms of the new variables which in addition are accurate to the order desired. For example, the notation μ in Eq. (2) corresponds to the guiding-center magnetic moment $\mu = mv_{\perp}^2/(2B)$ but automatically becomes the gyrokinetic magnetic moment Eq. (49) in the course of the transform. Similarly, the guiding-center gyrophase $\theta = \arctan(v_{\perp}/v_{\parallel})$ becomes the gyrokinetic gyrophase Eq. (50) etc. Thus, all the gyro-averages in the derivation below are automatically computed in terms of the gyrokinetic gyrophase accurate to the order of the perturbation theory applied, the derivatives are taken with respect to the gyrokinetic variables, the fields are computed at the gyrocenter positions and so on.

In this Section, we use the Deprit series version of the Lie transform formalism [17] to remove the gyro-dependencies to the desired order from the Hamiltonian function (the Dragt-Finn version [19] is employed in Sec. III). It has been shown in Refs. [25, 26] that an accuracy up to ϵ^2 is needed to recover the polarization density associated with the microturbulence (using the variational formalism). This corresponds to fourth order in terms of $\sqrt{\epsilon}$ which is the small

parameter used in this paper. In this case, Deprit series [17, 31] are given by the expansions:

$$h(\sqrt{\epsilon}) = \sum_{n=0}^{\infty} \epsilon^{n/2} h_n, \quad H(\sqrt{\epsilon}) = \sum_{n=0}^{\infty} \epsilon^{n/2} H_n, \quad \bar{H}_0 = H_0 \quad (4)$$

$$S(\sqrt{\epsilon}) = \sum_{n=0}^{\infty} \epsilon^{n/2} S_{n+1}, \quad L(\sqrt{\epsilon}) = \sum_{n=0}^{\infty} \epsilon^{n/2} L_{n+1}, \quad L_n = \{S_n, \cdot\} \quad (5)$$

$$T(\sqrt{\epsilon}) = \exp \left[- \int_0^{\sqrt{\epsilon}} L(\nu) d\nu \right] = \sum_{n=0}^{\infty} \epsilon^{n/2} T_n, \quad T_0 = 1 \quad (6)$$

Here, $T(\sqrt{\epsilon})$ is the Lie transform operator, $S(\sqrt{\epsilon})$ is the scalar function generating the transformation, $L(\sqrt{\epsilon})$ is the associated Lie derivative (defined through the Poisson brackets) and $H(\sqrt{\epsilon})$ is the transformed Hamiltonian. The Poisson brackets are given by the standard guiding-center expressions [27] since the symplectic structure is unperturbed (the ambient magnetic field is assumed to vary slowly).

The “master equation” for the Lie transform using the Deprit series has the following form [31]:

$$\hat{D}_0 S_n = n(H_n - h_n) - \sum_{m=1}^{n-1} (L_{n-m} H_m + m T_{n-m}^{-1} h_m) \quad (7)$$

Here, $\hat{D}_0 = \partial/\partial t + \{\cdot, h_0\}$ is the zero-order evolution operator and the inverse Lie transform operators can be found using the recurrence relation:

$$T_0^{-1} = 1, \quad T_n^{-1} = \frac{1}{n} \sum_{m=0}^{n-1} L_{n-m} T_m^{-1}, \quad n > 0 \quad (8)$$

As usual, the gyrokinetic Hamiltonian is required to be gyro-independent $H_k = \langle H_k \rangle$ (here, $\langle \cdot \rangle$ is the usual gyro-average defined, as mentioned above, in terms of the *gyrokinetic* gyrophase) and the generating functions are assumed to satisfy $\langle S_k \rangle = 0$. Satisfying these conditions gives the gyrokinetic Hamiltonian function to the desired order and the generating functions can be found solving Eq. (7):

$$H_1 = e \langle \Phi \rangle, \quad H_2 = e \left[\langle \phi \rangle + \frac{1}{2} \langle \{S_1, \tilde{\Phi}\} \rangle \right], \quad H_3 = \frac{e}{3!} \langle \{S_1, \{S_1, \Phi\}\} + \{S_2, \tilde{\Phi}\} + 4\{S_1, \tilde{\phi}\} \rangle \quad (9)$$

$$H_4 = \frac{e}{4!} \langle \{S_1, \{S_2, \Phi\}\} + 2\{S_2, \{S_1, \Phi\}\} + 2\{S_3, \tilde{\Phi}\} + 6\{S_2, \tilde{\phi}\} + 6\{S_1, \{S_1, \phi\}\} + \{S_1, \{S_1, \{S_1, \Phi\}\}\} \rangle \quad (10)$$

The equation for the first-order generating functions and its lowest-order solution can be written in the form:

$$\hat{D}_0 S_1 = \frac{\partial S_1}{\partial t} + \{S_1, h_0\} \approx \omega_c \frac{\partial S_1}{\partial \theta} = -e \tilde{\Phi}, \quad S_1 \approx -\frac{e \tilde{\Phi}^{(1)}}{\omega_c}, \quad \tilde{\Phi}^{(1)} = \int \tilde{\Phi} d\theta \quad (11)$$

Here, θ is the gyrokinetic (transformed) gyrophase [27], $\omega_c = eB/m$ is the cyclotron frequency (evaluated at the gyrocenter position) and $\tilde{\Phi} = \Phi - \langle \Phi \rangle$. Substituting the solution into Eq. (9), one obtains in the lowest order:

$$H_2 = e \langle \phi \rangle - \frac{e^2}{2B} \frac{\partial \langle \tilde{\Phi}^2 \rangle}{\partial \mu} \quad (12)$$

Here, μ is the gyrokinetic magnetic moment. The ambient magnetic field B is evaluated at the gyrocenter position.

In the second order of the Deprit perturbation theory, the equation for the generating function is as follows:

$$\hat{D}_0 S_2 \approx \omega_c \frac{\partial S_2}{\partial \theta} = -e \left(2e \tilde{\phi} + e \left[2\{S_1^{(0)}, \langle \Phi \rangle\} + \{S_1^{(0)}, \tilde{\Phi}\} - \langle \{S_1^{(0)}, \tilde{\Phi}\} \rangle \right] \right) \quad (13)$$

Solution of this equation is straightforward:

$$S_2 = -\frac{2e \tilde{\phi}^{(1)}}{\omega_c} - \frac{e}{\omega_c} \int_0^\theta d\theta' \left[2\{S_1, \langle \Phi \rangle\} + \{S_1, \tilde{\Phi}\} - \langle \{S_1, \tilde{\Phi}\} \rangle \right], \quad \tilde{\phi}^{(1)} = \int \tilde{\phi} d\theta \quad (14)$$

The resulting third-order Hamiltonian takes the form (see Appendix A for details):

$$H_3 = \frac{e^2}{B} \frac{\partial}{\partial \mu} \left[-\langle \tilde{\Phi} \tilde{\phi} \rangle + \frac{e}{2B} \langle \tilde{\Phi}^2 \rangle \frac{\partial \langle \Phi \rangle}{\partial \mu} + \frac{e}{6B} \frac{\partial}{\partial \mu} \langle \tilde{\Phi}^3 \rangle \right] \quad (15)$$

The lowest equation for the third-order generating function is

$$\begin{aligned} \omega_c \frac{\partial S_3^{(0)}}{\partial \theta} = & -e \left[\left\{ S_1^{(0)}, \langle \phi \rangle - \frac{e}{2B} \frac{\partial \langle \tilde{\Phi}^2 \rangle}{\partial \mu} \right\} + \{ S_2^{(0)}, \langle \Phi \rangle \} + \frac{1}{2} \{ S_2^{(0)}, \Phi \} + 2 \{ S_1^{(0)}, \phi \} + \right. \\ & \left. + \frac{1}{2} \{ S_1^{(0)}, \{ S_1^{(0)}, \Phi \} \} - \left\langle \frac{1}{2} \{ S_2^{(0)}, \Phi \} + 2 \{ S_1^{(0)}, \phi \} + \frac{1}{2} \{ S_1^{(0)}, \{ S_1^{(0)}, \Phi \} \} \right\rangle \right] \end{aligned} \quad (16)$$

In what follows, we will need the fourth-order gyrokinetic Hamiltonian only in order to compute the variational derivative with respect to the potential ϕ . Thus, it is sufficient to compute only such contributions into H_4 that explicitly contain ϕ . Substituting in Eq. (10) the Poisson brackets and neglecting the terms containing $\{R_j, R_k\}$ (the Poisson brackets of the gyrocenter coordinates), we finally obtain:

$$H_4 = -\frac{e^2}{2B} \frac{\partial \langle \tilde{\phi}^2 \rangle}{\partial \mu} + \frac{e^3}{2B^2} \frac{\partial}{\partial \mu} \left(\frac{\partial}{\partial \mu} \langle \tilde{\Phi}^2 \tilde{\phi} \rangle + 2 \langle \tilde{\phi} \tilde{\Phi} \rangle \frac{\partial \langle \Phi \rangle}{\partial \mu} + \langle \tilde{\Phi}^2 \rangle \frac{\partial \langle \phi \rangle}{\partial \mu} \right) + \mathcal{N}_4(\Phi) \quad (17)$$

Here the quantity $\mathcal{N}_4(\Phi)$ represents the contributions which contain solely Φ . We skip an explicit derivation of this quantity since it does not make any contribution into the quasineutrality equation. Note that the ϕ -dependent parts in H_2 , H_3 and H_4 can be derived more easily in the following way. First, one neglects ϕ and derives H_1 , H_2 and H_3 keeping only Φ . Then, replacing $\Phi \rightarrow \Phi + \sqrt{\epsilon} \phi$, one can derive the remaining ϕ -terms by carefully checking the order of each term. One can see that the result coincides with that derived above using the Lie transform technique. Further details of the derivation of the higher-order Hamiltonian functions are described in Appendix A.

III. DERIVATION USING DRAGT-FINN FORMULATION

The Dragt-Finn formulation of the Lie transform has been used in many gyrokinetic papers (summarised in Ref. [27]). Here, we show that the final physically-relevant results obtained with this approach are identical with the expressions found using the Deprit series although these two formulations cannot easily be brought into the same form (say, the general formulation of Dewar [18]). We apply an ordering identical to that used in Sec. II. The initial Hamiltonian function is given by Eq. (2). We use the Lie transform formalism [19, 27] to remove the gyro-dependencies to desired order out of the Hamiltonian function. It has been shown in Refs. [25, 26] that an accuracy up to ϵ^2 is needed to recover the polarization density associated with the microturbulence (this is the fourth order in terms of the small parameter $\sqrt{\epsilon}$ used here). The push-forward operator $\mathcal{T}_{\sqrt{\epsilon}}^{-1}$ that relates a function f on the guiding-center phase space to a new function $F \equiv \mathcal{T}_{\sqrt{\epsilon}}^{-1} f$ on the new gyrocenter phase space is defined as

$$\begin{aligned} \mathcal{T}_{\sqrt{\epsilon}}^{-1} & \equiv \cdots \exp(-\epsilon^2 \mathcal{L}_4) \exp(-\epsilon^{3/2} \mathcal{L}_3) \exp(-\epsilon \mathcal{L}_2) \exp(-\epsilon^{1/2} \mathcal{L}_1) \\ & = 1 - \epsilon^{1/2} \mathcal{L}_1 - \epsilon \left(\mathcal{L}_2 - \frac{1}{2} \mathcal{L}_1^2 \right) - \epsilon^{3/2} \left(\mathcal{L}_3 - \mathcal{L}_2 \mathcal{L}_1 + \frac{1}{6} \mathcal{L}_1^3 \right) \\ & \quad - \epsilon^2 \left(\mathcal{L}_4 - \mathcal{L}_3 \mathcal{L}_1 + \frac{1}{2} \mathcal{L}_2 \mathcal{L}_1^2 - \frac{1}{2} \mathcal{L}_2^2 - \frac{1}{24} \mathcal{L}_1^4 \right) + \cdots, \end{aligned} \quad (18)$$

where the action of the n th-order Lie derivative \mathcal{L}_n on an arbitrary scalar field G is defined as

$$\mathcal{L}_n G \equiv \{ \sigma_n, G \} = \frac{\partial \sigma_n}{\partial z^\alpha} J^{\alpha\beta} \frac{\partial G}{\partial z^\beta}, \quad (19)$$

where the scalar field σ_n generates the phase-space transformation at the n^{th} -order. It is clear that the functions σ_n generating the Lie transform Eq. (18) are different from the functions S_n generating the Lie transform Eq. (6) in the Deprit series approach. The Poisson brackets $\{ , \}$ are defined in the extended guiding-center phase space which includes the canonically-conjugated time (t) and energy (w) coordinates in addition to the usual phase-space coordinates. The extended phase space is needed in order to treat the time-dependent systems (see Refs. [25, 31] for a detailed description). The Hamiltonian function on the extended phase space must be replaced by the *extended*

Hamiltonian function $h^x = h - w$. Since in the Lie-transform perturbation theory only the lowest order is affected by this procedure [31], it is sufficient just to make a formal replacement $h_0 \rightarrow h_0^x = h_0 - w$ everywhere in our notations.

If we expand the “extended” guiding-center Hamiltonian $h^x \equiv h_0^x + \sqrt{\epsilon} h_1 + \epsilon h_2$ [see Eq. (3)], the new “extended” gyrokinetic Hamiltonian function becomes $H^x \equiv H_0^x + \epsilon^{1/2} H_1 + \epsilon H_2 + \epsilon^{3/2} H_3 + \epsilon^2 H_4 + \dots$, where

$$H_0^x = h_0^x, \quad (20)$$

$$H_1 = h_1 - \mathcal{L}_1 h_0^x, \quad (21)$$

$$H_2 = h_2 - \left(\mathcal{L}_2 - \frac{1}{2} \mathcal{L}_1^2 \right) h_0^x - \mathcal{L}_1 h_1, \quad (22)$$

$$H_3 = - \left(\mathcal{L}_3 - \mathcal{L}_2 \mathcal{L}_1 + \frac{1}{6} \mathcal{L}_1^3 \right) h_0^x - \left(\mathcal{L}_2 - \frac{1}{2} \mathcal{L}_1^2 \right) h_1 - \mathcal{L}_1 h_2, \quad (23)$$

$$H_4 = - \left(\mathcal{L}_4 - \mathcal{L}_3 \mathcal{L}_1 + \frac{1}{2} \mathcal{L}_2 \mathcal{L}_1^2 - \frac{1}{2} \mathcal{L}_2^2 - \frac{1}{24} \mathcal{L}_1^4 \right) h_0^x \\ - \left(\mathcal{L}_3 - \mathcal{L}_2 \mathcal{L}_1 + \frac{1}{6} \mathcal{L}_1^3 \right) h_1 - \left(\mathcal{L}_2 - \frac{1}{2} \mathcal{L}_1^2 \right) h_2, \quad (24)$$

where $\mathcal{L}_n \mathcal{L}_m G \equiv \{\sigma_n, \{\sigma_m, G\}\} \neq \mathcal{L}_m \mathcal{L}_n G$. All the quantities will be evaluated in terms of the new transformed gyrokinetic phase-space coordinates [27]. We now proceed with the perturbation analysis at each order from $n = 1$ to $n = 4$, where we use the definition that the generating scalar field σ_n is explicitly gyroangle-dependent. Hence, the Lie derivative of h_0^x generated by σ_n , defined as

$$\mathcal{L}_n h_0^x = \frac{d_0 \sigma_n}{dt} \equiv \frac{\partial \sigma_n}{\partial t} + \dot{\mathbf{X}} \cdot \nabla \sigma_n + \omega_c \frac{\partial \sigma_n}{\partial \theta}, \quad (25)$$

satisfies the property $\langle \mathcal{L}_n h_0^x \rangle \equiv 0$. Note that the energy coordinate w (entering the extended Hamiltonian function) does not make any contribution if appears with a nonlinear combination of the Lie derivatives, e.g. $\mathcal{L}_1^2 w = 0$, since the generating functions σ_n do not depend on w (per their definition [31]).

At first order, Eq. (21) yields the first-order Hamiltonian and the first-order equation for σ_1

$$H_1 \equiv e \langle \Phi \rangle, \quad \frac{d_0 \sigma_1}{dt} = e \tilde{\Phi}, \quad (26)$$

where we have expressed $\Phi \equiv \langle \Phi \rangle + \tilde{\Phi}$ in terms of its gyroangle-independent part $\langle \Phi \rangle$, which contributes to the first-order Hamiltonian (26), and its gyroangle-dependent part $\tilde{\Phi}$, which is used to define the first-order generating scalar field σ_1 . We formally introduce the solution for σ_1 as an integration along unperturbed Hamiltonian orbits

$$\sigma_1 \equiv \left(\frac{d_0}{dt} \right)^{-1} e \tilde{\Phi} \simeq \frac{e}{\omega_c} \tilde{\Phi}^{(1)}, \quad (27)$$

where the lowest-order solution is shown using the definitions $\tilde{\Phi}^{(n+1)} \equiv \int \tilde{\Phi}^{(n)} d\theta$, $\tilde{\Phi}^{(0)} \equiv \tilde{\Phi}$ and $\partial \tilde{\Phi}^{(n+1)} / \partial \theta \equiv \tilde{\Phi}^{(n)}$.

At second order, we can rearrange Eq. (22) as $H_2 = h_2 - \mathcal{L}_2 h_0^x - (e/2) \mathcal{L}_1 (\tilde{\Phi} + 2 \langle \Phi \rangle)$, so that the second-order Hamiltonian is expressed as

$$H_2 \equiv e \langle \phi \rangle + e \langle \Psi_2 \rangle, \quad (28)$$

where the second-order ponderomotive potential is defined as

$$\Psi_2 \equiv - \frac{1}{2} \left\{ \sigma_1, \tilde{\Phi} \right\} = - \frac{e}{2\omega_c} \left\{ \tilde{\Phi}^{(1)}, \tilde{\Phi} \right\}, \quad (29)$$

It is easy to show that Eq. (28) [together with the ponderomotive potential Eq. (29)] coincides with the second-order gyrokinetic Hamiltonian Eq. (12) derived using the Deprit series.

The second-order equation for σ_2 is

$$\frac{d_0 \sigma_2}{dt} = e \tilde{\phi} - \frac{e^2}{\omega_c} \left\{ \tilde{\Phi}^{(1)}, \langle \Phi \rangle \right\} + e \tilde{\Psi}_2. \quad (30)$$

The formal (lowest-order) solution of Eq. (30) is expressed as

$$\sigma_2 \simeq \frac{e}{\omega_c} \tilde{\phi}^{(1)} + \left(\frac{e}{\omega_c} \tilde{\Psi}_2^{(1)} - \frac{e^2}{\omega_c^2} \left\{ \tilde{\Phi}^{(2)}, \langle \Phi \rangle \right\} \right) \equiv \sigma_2[\phi] + \sigma_2[\Phi]. \quad (31)$$

At third order, we can rearrange Eq. (23) as

$$H_3 = -\mathcal{L}_3 h_0^x - e \mathcal{L}_2 \langle \Phi \rangle - e \mathcal{L}_1 \phi + \frac{e}{6} \mathcal{L}_1^2 \left(3 \langle \Phi \rangle + 2 \tilde{\Phi} \right), \quad (32)$$

so that the third-order Hamiltonian is $H_3 = e \langle \Psi_3 \rangle + e \langle \psi_3 \rangle$ where the third-order ponderomotive potentials are

$$\Psi_3 \equiv \frac{1}{2} \left\{ \sigma_1, \{ \sigma_1, \langle \Phi \rangle \} \right\} + \frac{1}{3} \left\{ \sigma_1, \{ \sigma_1, \tilde{\Phi} \} \right\}, \quad (33)$$

$$\psi_3 \equiv - \left\{ \sigma_1, \tilde{\phi} \right\}, \quad (34)$$

while the third-order equation for σ_3 is

$$\frac{d_0 \sigma_3}{dt} = -e \{ \sigma_2, \langle \Phi \rangle \} - e \{ \sigma_1, \langle \phi \rangle \} + e \tilde{\Psi}_3 + e \tilde{\psi}_3. \quad (35)$$

We will not need the solution for σ_3 in what follows. We note that the third-order ponderomotive potential Eq. (33) introduces nonlinear corrections to the polarization density associated with Φ , while the third-order ponderomotive potential Eq. (34) introduces coupling terms in the polarization densities associated with Φ and ϕ .

Substituting the first-order generating function Eq. (27) and evaluating the Poisson brackets, one can show that

$$H_3 = \left\langle -\frac{e^2}{\omega_c} \{ \tilde{\Phi}^{(1)}, \phi \} + \frac{e^3}{3\omega_c^2} \left\{ \tilde{\Phi}^{(1)}, \left\{ \tilde{\Phi}^{(1)}, \Phi + \frac{1}{2} \langle \Phi \rangle \right\} \right\} \right\rangle \quad (36)$$

One can use the identity $\langle \{ \tilde{\Phi}^{(1)}, A \} \rangle = (e/m) \partial \langle A \tilde{\Phi} \rangle / \partial \mu$ (here the $\{ \mu, \theta \}$ -part of the Poisson bracket is dominant) and the Leibniz rule for the Poisson brackets $a\{b, c\} + b\{a, c\} = \{ab, c\}$ to write the following:

$$H_3 = \frac{e^2}{B} \left(-\frac{\partial}{\partial \mu} \langle \tilde{\Phi} \tilde{\phi} \rangle + \frac{e}{2B} \frac{\partial}{\partial \mu} \left[\langle \tilde{\Phi}^2 \frac{\partial \langle \Phi \rangle}{\partial \mu} \rangle + \frac{1}{3} \frac{\partial}{\partial \mu} \langle \tilde{\Phi}^3 \rangle \right] \right) \quad (37)$$

This expression coincides exactly with Eq. (15) derived using the Deprit series approach to the Lie transform.

At fourth order, we can rearrange Eq. (24) as

$$\begin{aligned} H_4 = & -\mathcal{L}_4 h_0^x - e \mathcal{L}_3 \langle \Phi \rangle - \frac{1}{2} \mathcal{L}_2 \left[\left(e \phi + H_2 \right) - e \mathcal{L}_1 \left(\langle \Phi \rangle + \frac{1}{2} \tilde{\Phi} \right) \right] \\ & + \frac{e}{2} \mathcal{L}_1^2 \left[\phi - \frac{1}{12} \mathcal{L}_1 \left(4 \langle \Phi \rangle + 3 \tilde{\Phi} \right) \right], \end{aligned} \quad (38)$$

so that the fourth-order Hamiltonian is expressed as

$$H_4 \equiv e \langle \psi_4 \rangle + e \langle \Psi_4 \rangle, \quad (39)$$

where the fourth-order ponderomotive potential

$$\Psi_4 \equiv -\frac{1}{24} \mathcal{L}_1^3 \left(4 \langle \Phi \rangle + 3 \tilde{\Phi} \right) + \frac{1}{2} \left\{ \sigma_2[\Phi], \mathcal{L}_1 \left(\langle \Phi \rangle + \frac{1}{2} \tilde{\Phi} \right) \right\}$$

only involves Φ (and will thus be omitted here since it only contributes higher-order terms in the dynamics associated with Φ), while the fourth-order potential ψ_4 is defined as

$$\begin{aligned} \psi_4 \equiv & -\frac{1}{2} (\mathcal{L}_2 - \mathcal{L}_1^2) \phi + \frac{1}{2} \left\{ \sigma_2[\phi], \left[\mathcal{L}_1 \left(\langle \Phi \rangle + \frac{1}{2} \tilde{\Phi} \right) \right] \right\} \\ = & \frac{1}{2} \left\{ \sigma_1, \{ \sigma_1, \phi \} \right\} - \frac{1}{2} \left\{ \sigma_2, \phi \right\} + \frac{1}{2} \left\{ \sigma_2[\phi], \left\{ \sigma_1, \left(\langle \Phi \rangle + \frac{1}{2} \tilde{\Phi} \right) \right\} \right\}. \end{aligned} \quad (40)$$

The fourth-order potential ψ_4 introduces nonlinear corrections to the polarization density associated with ϕ . The solution for the fourth-order generating scalar field σ_4 is not needed and will be omitted.

Applying the same approach as described in Appendix A, one can derive the following relations:

$$\left\langle \{\sigma_1, \{\sigma_1, \phi\}\} \right\rangle = \frac{e^2}{B\omega_c} \frac{\partial}{\partial \mu} \left\langle \tilde{\Phi} \{\tilde{\Phi}^{(1)}, \tilde{\phi}\} \right\rangle + \frac{e^2}{B^2} \frac{\partial}{\partial \mu} \left[\left\langle \tilde{\Phi}^2 \right\rangle \frac{\partial \langle \phi \rangle}{\partial \mu} \right] \quad (41)$$

$$\left\{ \sigma_2, \phi \right\} = \frac{e}{B} \frac{\partial \langle \tilde{\phi}^2 \rangle}{\partial \mu} - \frac{e^2}{2B\omega_c} \frac{\partial}{\partial \mu} \left\langle \tilde{\phi} \{\tilde{\Phi}^{(1)}, \tilde{\Phi}\} \right\rangle - \frac{e^2}{B^2} \frac{\partial}{\partial \mu} \left[\left\langle \tilde{\phi} \tilde{\Phi} \right\rangle \frac{\partial \langle \Phi \rangle}{\partial \mu} \right] \quad (42)$$

$$\left\{ \sigma_2[\phi], \left\{ \sigma_1, \left(\langle \Phi \rangle + \frac{1}{2} \tilde{\Phi} \right) \right\} \right\} = \frac{e^2}{2B\omega_c} \frac{\partial}{\partial \mu} \left\langle \tilde{\phi} \{\tilde{\Phi}^{(1)}, \tilde{\Phi}\} \right\rangle + \frac{e^2}{B^2} \frac{\partial}{\partial \mu} \left[\left\langle \tilde{\phi} \tilde{\Phi} \right\rangle \frac{\partial \langle \Phi \rangle}{\partial \mu} \right] \quad (43)$$

Substituting Eqs. (41)-(42) into Eq. (40), using the identity $\left\langle \{\tilde{\Phi}^{(1)}, A\} \right\rangle = (e/m) \partial \langle A \tilde{\Phi} \rangle / \partial \mu$ and employing the Leibniz rule for the Poisson brackets, one obtains:

$$\langle \psi_4 \rangle = -\frac{e}{2B} \frac{\partial \langle \tilde{\phi}^2 \rangle}{\partial \mu} + \frac{e^2}{2B^2} \frac{\partial}{\partial \mu} \left(\frac{\partial}{\partial \mu} \left\langle \tilde{\Phi}^2 \tilde{\phi} \right\rangle + 2 \left\langle \tilde{\phi} \tilde{\Phi} \right\rangle \frac{\partial \langle \Phi \rangle}{\partial \mu} + \left\langle \tilde{\Phi}^2 \right\rangle \frac{\partial \langle \phi \rangle}{\partial \mu} \right) \quad (44)$$

This is in exact agreement with the fourth-order gyrokinetic Hamiltonian Eq. (17) derived using the Deprit series. Summarizing, we have shown that the higher-order gyrokinetic Hamiltonian derived using the conventional (Dragt-Finn) approach to the gyrokinetic theory coincides exactly with that derived in Sec. II using the Deprit series. It is clear that the quasineutrality equations resulting from the variational principle (see Sec. IV) will coincide, too.

IV. GYROKINETIC SYSTEM OF EQUATIONS

Having derived the gyrokinetic Hamiltonian function to the fourth order, we are now ready to write down the gyrokinetic field equations using the variational approach developed in Ref. [25, 26]. The gyrokinetic electrostatic action functional can be written in the extended phase space [25] as follows:

$$\mathcal{A}_{\text{gy}} = - \sum_{s=i,e,\dots} \int (H[Z, \Phi, \phi] - w) \mathcal{F}(Z) d^8 Z \quad (45)$$

Here, $Z = (\mathbf{R}, v_{\parallel}, \mu, \theta, t, w)$ are the extended phase-space coordinates, $\mathcal{F}(Z) = f(Z) \delta(H - w)$ is the distribution function in the extended phase space, $f(Z)$ is the usual gyrokinetic distribution function and w is the phase-space coordinate canonically conjugated to the time. The gyrokinetic field equation is then given by $\delta \mathcal{A}_{\text{gy}} / \delta \phi(\mathbf{x}) = 0$ (see Appendix B for details). Neglecting, as usual, the Debye screening, we obtain the gyrokinetic quasineutrality equation for the microturbulence potential ϕ as follows:

$$\sum_{s=i,e,\dots} e \int \left\{ f + \frac{e}{B} (\tilde{\Phi} + \tilde{\phi}) \frac{\partial f}{\partial \mu} + \frac{e^2}{2B^2} \left[\tilde{\Phi}^2 \frac{\partial^2 f}{\partial \mu^2} + \left(\frac{\partial \langle \Phi^2 \rangle}{\partial \mu} - 2 \Phi \frac{\partial \langle \Phi \rangle}{\partial \mu} \right) \frac{\partial f}{\partial \mu} \right] \right\} \delta(\mathbf{R} + \boldsymbol{\rho} - \mathbf{x}) d^6 Z = 0 \quad (46)$$

Here, \mathbf{R} is the gyrocenter position, $\boldsymbol{\rho} = \sqrt{2m\mu/(e^2 B)} (-\mathbf{e}_1 \cos \theta + \mathbf{e}_2 \sin \theta)$ is the gyroradius [which is computed in terms of the gyrokinetic variables μ and θ , see Eqs. (49) and (50)] and \mathbf{x} is the coordinate in the real space (an Euler spatial position). It is interesting that the expression identical to Eq. (46) can be derived by taking the variational derivative $\delta \mathcal{A}_{\text{gy}} / \delta \Phi(\mathbf{x}) = 0$ (a detailed derivation is presented in Appendix B). In this case, the third-order contribution to the gyrokinetic Hamiltonian is sufficient (clearly, it is sufficient only for the derivation of the quasineutrality equation).

One sees that the resulting equation contains the polarization density due to the background electrostatic potential, the polarization density caused by the microturbulence and an additional term which is quadratic in Φ (*nonlinear* polarization density). Note that this equation becomes a higher-order version of the usual gyrokinetic quasineutrality equation when the potential ϕ is neglected and the potential Φ is associated with the usual gyrokinetic self-consistent field. Such higher-order gyrokinetic quasineutrality equation was for the first time derived by Dubin [20] using the Lie transform approach (Deprit series) and recently reproduced by Parra and Catto [14] using the reduced-Vlasov approach. Their result coincides with Eq. (46) plus additional terms which can be shown to appear from the spatial $\{R_i, R_j\}$ -component of the Poisson tensor neglected here for simplicity (the associated terms are small). Unfortunately, the nonlinear corrections to the quasineutrality equation in both Refs. [20] and [14] destroy the energy conservation because the gyrokinetic Hamiltonian was not derived to the sufficient order (the derivation was truncated at the second order and the pullback of the distribution function was used in Ref. [20] in order to obtain the polarization

density). In contrast, the derivation in this paper is carried to the higher orders and relies on the variational principle so that all invariants (including the energy) are automatically conserved.

In the long-wavelength approximation (i.e. $k_\perp \rho \ll 1$ where ρ is the gyroradius), one can write the quasineutrality equation in the simple form (here \bar{n}_s is the usual gyrokinetic density; see Appendix C for a detailed derivation):

$$-\nabla_\perp \cdot \left[\frac{\bar{n}_i}{B\omega_c} \nabla_\perp (\phi + \Phi) \right] - \frac{3\bar{n}_i \rho_{thi}^2}{4B\omega_c} \nabla_\perp^2 [\nabla_\perp^2 (\phi + \Phi)] - \frac{1}{2} \nabla_\perp \cdot \left[\frac{\bar{n}_i}{B^2 \omega_c^2} [\nabla_\perp \Phi \nabla_\perp^2 \Phi - (\nabla_\perp \Phi \cdot \nabla_\perp) \nabla_\perp \Phi] \right] = \bar{n}_i - \bar{n}_e \quad (47)$$

Here, $\rho_{thi} = v_{thi}/\omega_c$ is the thermal gyroradius. The first term in Eq. (47) is the conventional polarization density. The second term is the well known finite-gyroradius correction which, for example, has been invoked to describe the radiative damping of the Toroidal Alfvén Eigenmodes [32]. In addition, a nonlinear contribution associated with the higher-order gyrokinetic Hamiltonian appears in the long-wavelength approximation (the third term) but expansion to the fourth order in $k_\perp \rho$ is needed. In the following, we will omit the second term for simplicity since the focus of this paper is on the nonlinear contribution to the gyrokinetic quasineutrality equation.

Note that Eq. (47) is different compared with the long-wavelength approximation of the quasineutrality equation derived by Parra and Catto (see Ref. [13]). One difference is related to the terms of the fourth order in $k_\perp \rho$ which is not surprising since Ref. [13] employed only the second-order accuracy in the gyroradius expansion. Another difference is in the nonlinear correction appearing in Ref. [13] already in the second order in $k_\perp \rho$ but absent here. It is interesting that one can derive this correction choosing an appropriate distribution function in the polarization density entering Eq. (47). The correction appears in this case from the subtle difference between the Maxwellian distribution function defined in terms of the *gyrokinetic* magnetic moment μ and the Maxwellian distribution function defined in terms of the *guiding-center* magnetic moment μ_0 (or, equivalently, the difference between the *gyrocenter* Maxwellian and the *guiding-center* Maxwellian). In the reduced-Vlasov approach Ref. [13], the zeroth-order distribution function is defined as a function of the *guiding-center* magnetic moment $\mu_0 = mv_\perp^2/(2B)$ (it is a *guiding-center* Maxwellian):

$$f_0(\mu_0) = n_0 \left(\frac{m}{2\pi T} \right)^{3/2} \exp \left(- \frac{mv_\perp^2}{2T} \right) \exp \left(- \frac{\mu_0 B}{T} \right) \quad (48)$$

The *guiding-center* magnetic moment μ_0 is related to the *gyrokinetic* magnetic moment μ as follows [28]:

$$\mu_0 = \mu - \frac{e}{B} \tilde{\Phi}(\mu, \theta) + \dots \quad (49)$$

Here, θ is the *gyrokinetic* gyrophase which can be expressed through the *guiding-center* gyrophase $\theta_0 = \arctan(v_\perp/v_\parallel)$:

$$\theta_0 = \theta + \frac{e}{B} \frac{\partial \tilde{\Phi}^{(1)}(\mu, \theta)}{\partial \mu} + \dots \quad (50)$$

Using Eq. (49), one can write the *guiding-center* Maxwellian Eq. (48) in the *gyrokinetic* variables:

$$f_0(\mu, \theta) = n_0 \left(\frac{m}{2\pi T} \right)^{3/2} \exp \left(- \frac{mv_\perp^2}{2T} \right) \exp \left(- \frac{\mu B}{T} \right) \exp \left(\frac{e\tilde{\Phi}}{T} \right) \approx f_M(\mu) \left[1 + \frac{e\tilde{\Phi}(\mu, \theta)}{T} \right] \quad (51)$$

Here, f_M is the *gyrokinetic* Maxwellian (defined in terms of the gyrokinetic moment μ). Note that Eq. (48) can also be derived as the Lie transform (push-forward) of the *guiding-center* Maxwellian distribution function using the explicit expression for the Lie transform operator through the Lie derivative and generating function: $f_0 = T f_M$. Clearly, the *guiding-center* Maxwellian depends on the *gyrokinetic* gyrophase, i.e. $f_0 = f_0(\mu, \theta)$. This choice is, in fact, not completely consistent with the “gyrokinetic philosophy” since the gyrokinetic distribution function must be gyro-independent (to all orders). It is perhaps one of the most important differences between the Lie transform and the recursive methods that the old and the new phase-space coordinates do not mix in the course of the Lie transform. Note that in the zeroth-order distribution function Eq. (48), the guiding-center definitions must have been used also for the parallel velocity and the guiding-center position, but the associated corrections are small (see Ref. [28]).

Proceeding with the derivation (as in Appendix C), it is now straightforward to obtain:

$$-\nabla_\perp \cdot \left[\frac{n_0}{B\omega_c} \nabla_\perp (\phi + \Phi) \right] + \frac{m_i n_0}{2T_i B^2} |\nabla_\perp \Phi|^2 = \bar{n}_i - \bar{n}_e \quad (52)$$

Note that n_0 here is the density which appears in the definition of the guiding-center Maxwellian distribution function Eq. (48). Only the terms $\sim (k_\perp \rho)^2$ are kept in the polarization density. One sees that the nonlinear correction in Eq. (52) coincides with that derived by Parra and Catto [13].

Having derived the long-wavelength approximation of the gyrokinetic quasineutrality equation, it is now instructive to consider the long-wavelength approximation for the gyrokinetic Hamiltonian (details in Appendix C). To the third order, it can be cast into the compact expression (here we omit ϕ and neglect for simplicity the term $\sim (k_\perp \rho)^4$ in the quadratic nonlinearity):

$$H = \frac{mv_\parallel^2}{2} + \mu B + e\langle\Phi\rangle - \frac{m}{2B^2} (\nabla_\perp \Phi)^2 + \frac{e}{2B^2\omega_c^2} (\nabla_\perp \Phi \cdot \nabla_\perp) \left[\frac{(\nabla_\perp \Phi)^2}{2} \right] \quad (53)$$

Introducing the perpendicular electric field $\mathbf{E}_\perp = -\nabla_\perp \Phi$, the $\mathbf{E} \times \mathbf{B}$ velocity $\mathbf{u}_E^2 = (\nabla_\perp \Phi)^2/B^2$ and the gyrokinetic polarization vector $\mathbf{P}_E = -\nabla_\perp \Phi/(B\omega_c)$, one can rewrite the Hamiltonian in a physically transparent form:

$$H = \frac{mv_\parallel^2}{2} + \mu B + e\langle\Phi\rangle - \frac{m\mathbf{u}_E^2}{2} - \frac{e}{2} (\mathbf{E}_\perp \cdot \nabla_\perp) \left(\frac{\mathbf{P}_E^2}{2} \right) \quad (54)$$

One sees that two terms appear in addition to the usual kinetic energy and the gyro-averaged energy of the electrostatic field. One term is well-known and corresponds to the kinetic energy of the flow associated with the $\mathbf{E} \times \mathbf{B}$ motion. Another term is new. This is the energy associated with the gyrokinetic polarization. All the terms in the gyrokinetic Hamiltonian are calculated at the gyrocenter position. Since the gyrokinetic Hamiltonian is known, it is straightforward to derive the equations of motion. For example, keeping the quadratic nonlinearity, one obtains the following equations:

$$\dot{\mathbf{R}} = v_\parallel \mathbf{b}^* + \frac{\mathbf{b} \times \mu \nabla B}{eB_\parallel^*} + \frac{\mathbf{E} \times \mathbf{b}}{B_\parallel^*} + \frac{e}{2m\omega_c^2} \frac{\nabla \mathbf{E}_\perp^2 \times \mathbf{b}}{B_\parallel^*} \quad (55)$$

$$m\dot{v}_\parallel = -\mu \mathbf{b}^* \cdot \nabla B + e \mathbf{b}^* \cdot \mathbf{E} + \frac{e^2}{2m\omega_c^2} \mathbf{b}^* \cdot \nabla \mathbf{E}_\perp^2 \quad (56)$$

One sees that the “ponderomotive” drift motion and the “ponderomotive” parallel acceleration (both associated with the second-order gyrokinetic Hamiltonian) appear in addition to the conventional terms (see also Ref. [33]). Note that extension of the equations of motion to higher orders is straightforward since the Hamiltonian function is known. It is interesting that these “ponderomotive” contributions to the equations of motion can be shown to vanish in Eqs. (55)-(56) if one includes the classical polarization into the definition of the gyrocenter $\mathbf{R} \rightarrow \mathbf{R} + \mathbf{E}_\perp/(\omega_c B)$. In addition, the polarization drift will explicitly enter the equations of the modified gyrocenter motion [34]. The higher-order gyrocenter orbits [resulting e.g. from Eq. (54)] must be used to advance the gyrokinetic distribution function according to the Vlasov equation. This will assure the energetic consistency of the gyrokinetic theory.

For the energetic consistency, it is important that one can obtain the long-wavelength approximation of the quasineutrality equation also using the long-wavelength approximation of the gyrokinetic Hamiltonian Eq. (54) in the variational principle (see Appendix C). One can easily check that the resulting expression for the nonlinear polarization density coincides with Eq. (47). Using the notations introduced above, one can write the quasineutrality equation in the compact form (here, for simplicity, we omit ϕ and the higher-order finite-gyroradius correction to the linear polarization density):

$$\nabla_\perp \cdot \left(\bar{n}_i \left\{ \mathbf{P}_E - \frac{1}{2} \left[\mathbf{P}_E (\nabla_\perp \cdot \mathbf{P}_E) - (\mathbf{P}_E \cdot \nabla_\perp) \mathbf{P}_E \right] \right\} \right) = \bar{n}_i - \bar{n}_e \quad (57)$$

The nonlinear contribution to the quasineutrality equation is of course small. However, it may be important for the toroidal momentum transport [13]. Also, this contribution enters the weak turbulence theory [20]. The nonlinearity in the quasineutrality equation described here must be added to the well known nonlinearity resulting from the self-consistent contribution ($\sim \delta n \sim \phi^2$) to the first-order polarization density $\delta n_{\text{pol}} = \int d^3Z \tilde{\phi} \partial \delta f / \partial \mu \delta(\mathbf{R} + \boldsymbol{\rho} - \mathbf{x})$.

V. CONCLUSIONS

In conclusion, we have derived a higher-order self-consistent energy-conserving gyrokinetic system of equations. The derivation has been carried using the Deprit series [17] and the conventional Lie transform approach to the gyrokinetic theory (initially proposed by Dragt and Finn [19] in the context of canonical maps). The resulting expressions are identical. We have shown that a nonlinear (in the electric field) term appears in the quasineutrality condition. A similar expression has been derived in Ref. [20] using the Lie transform approach and recently reproduced in Ref. [14] with the reduced-Vlasov calculation. However, these derivations were not energy-conserving since the gyrokinetic

Hamiltonian was not derived to the sufficient order. In contrast, the derivation presented in this paper includes the higher-order terms needed for the energy conservation and employs the variational principle which automatically provides all the conservation laws through the Noether theorem. The long-wavelength approximation of the gyrokinetic Hamiltonian and the quasineutrality equation has been derived. This formulation is more compact and accessible for the physical intuition than the general expressions accurate at all orders in the perpendicular wavenumber. The resulting equations are simple enough and can readily be used in gyrokinetic codes.

The equations derived here can be applied in certain transition layers such as the stellarator neoclassical turbulent transport barriers (the transition zone between the electron and ion roots of the ambipolarity condition in stellarator geometry) where the ambient neoclassical radial electric field may change on extremely short scale (few ion gyroradii) while the plasma density remains smooth. The theory may also be of interest for the edge plasma where the nonlinear terms in the quasineutrality equation may matter. The nonlinearity in the quasineutrality equation described here should be considered in addition to the well known nonlinearity resulting from the self-consistent contribution ($\sim \delta n \sim \phi^2$) to the first-order polarization density (see an early discussion in [20]). Finally, the higher-order gyrokinetic terms derived in this paper may be relevant to the tokamak toroidal momentum transport problem [13].

It is clear that there are other higher-order terms which might also be of relevance especially when the tokamak toroidal momentum transport problem is considered. One example for this are the higher-order cross-terms which simultaneously account for the gyro-dependency due to the ambient non-homogeneous magnetic field and the self-consistent field [14, 30]. Another example are the higher-order magnetic drifts [29] ($\sim \epsilon_B^2$ where $\epsilon_B = \rho/L_B$ with ρ the gyroradius and L_B the characteristic length of the ambient magnetic field). Another natural extension of the work presented here is to include the magnetic component of the self-consistent field which should result in a general electromagnetic formulation of the higher-order gyrokinetic theory. In this respect, an approach similar to that discussed in this paper can perhaps be used to study the magnetic islands which can be ordered as a mesoscale electromagnetic perturbation of the ambient magnetic field.

ACKNOWLEDGEMENTS

We acknowledge fruitful and informative discussions with P. Catto, T. S. Hahm, P. Helander, F. Parra and B. Scott. An interest of J. Krommes to our work is appreciated. Part of this work has been done in the Isaac Newton Institute, Cambridge, in the course of program ‘‘Gyrokinetics in Laboratory and Astrophysical Plasmas’’. We acknowledge hospitality of the Institute and thank the organisers (Alex Schekochihin in particular) for their help during our stay.

APPENDIX A: DERIVATION OF THE HIGHER-ORDER GYROKINETIC HAMILTONIANS

In this appendix, we present some details of the derivation for the higher-order gyrokinetic Hamiltonian functions using the *Deprit series* approach [17] (see Sec. II). Here, the generating functions are given by the expressions:

$$S_1 = -\frac{e\tilde{\Phi}^{(1)}}{\omega_c}, \quad S_2 = -\frac{2e\tilde{\phi}^{(1)}}{\omega_c} - \frac{e}{\omega_c} \int_0^\theta d\theta' \left[2\{S_1, \langle \Phi \rangle\} + \{S_1, \tilde{\Phi}\} - \langle \{S_1, \tilde{\Phi}\} \rangle \right] \quad (A1)$$

$$\omega_c \frac{\partial S_3}{\partial \theta} = \frac{e^2}{\omega_c} \left[\langle \tilde{\Phi}^{(1)}, \langle \phi \rangle \rangle + 2\langle \tilde{\phi}^{(1)}, \langle \Phi \rangle \rangle + \langle \tilde{\phi}^{(1)}, \Phi \rangle + 2\langle \tilde{\Phi}^{(1)}, \phi \rangle - \langle \{ \tilde{\phi}^{(1)}, \tilde{\Phi} \} + 2\langle \tilde{\Phi}^{(1)}, \tilde{\phi} \rangle \rangle \right] + \mathcal{N}_4(\Phi) \quad (A2)$$

In Eq. (A2) and throughout the text, $\mathcal{N}_4(\Phi)$ are the terms which appear in the fourth-order Hamiltonian and do not contain ϕ ; this terms are not going to be important in the derivation of the quasineutrality equation (see Appendix B).

The third-order Hamiltonian can be written as follows:

$$H_3 = \frac{e}{3!} \langle \{S_1, \{S_1, \Phi\}\} + \{S_2, \tilde{\Phi}\} + 4\{S_1, \tilde{\phi}\} \rangle \quad (A3)$$

Substituting explicit expression for the generating functions Eq. (A1) and neglecting the terms proportional to the spatial Poisson brackets $\{R_i, R_j\}$, one can write:

$$\langle \{S_1, \{S_1, \Phi\}\} \rangle = \frac{e^2}{B^2} \frac{\partial}{\partial \mu} \left(\langle \tilde{\Phi}^2 \rangle \frac{\partial \langle \Phi \rangle}{\partial \mu} \right) + \frac{e^2}{B\omega_c} \frac{\partial}{\partial \mu} \langle \tilde{\Phi} \{ \tilde{\Phi}^{(1)}, \tilde{\Phi} \} \rangle \quad (A4)$$

$$\langle \{S_2, \tilde{\Phi}\} \rangle = -\frac{2e}{B} \frac{\partial}{\partial \mu} \langle \tilde{\Phi} \tilde{\phi} \rangle + \frac{2e^2}{B^2} \frac{\partial}{\partial \mu} \left(\langle \tilde{\Phi}^2 \rangle \frac{\partial \langle \Phi \rangle}{\partial \mu} \right) + \frac{e^2}{B\omega_c} \frac{\partial}{\partial \mu} \langle \tilde{\Phi} \{ \tilde{\Phi}^{(1)}, \tilde{\Phi} \} \rangle \quad (A5)$$

$$\langle \{S_1, \tilde{\phi}\} \rangle = -\frac{e}{B} \frac{\partial}{\partial \mu} \langle \tilde{\Phi} \tilde{\phi} \rangle \quad (A6)$$

Substituting these expressions into Eq. (A3) and using the relation $\tilde{\Phi}\{\tilde{\Phi}^{(1)}, \tilde{\Phi}\} = 1/2\{\tilde{\Phi}^{(1)}, \tilde{\Phi}^2\}$, one obtains the third-order gyrokinetic Hamiltonian:

$$H_3 = \frac{e^2}{B} \frac{\partial}{\partial \mu} \left[-\langle \tilde{\Phi} \tilde{\phi} \rangle + \frac{e}{2B} \langle \tilde{\Phi}^2 \rangle \frac{\partial \langle \Phi \rangle}{\partial \mu} + \frac{e}{6B} \frac{\partial}{\partial \mu} \langle \tilde{\Phi}^3 \rangle \right] \quad (\text{A7})$$

The fourth-order gyrokinetic Hamiltonian has the form:

$$H_4 = \frac{e}{4!} \left\langle \{S_1, \{S_2, \Phi\}\} + 2\{S_2, \{S_1, \Phi\}\} + 2\{S_3, \tilde{\Phi}\} + 6\{S_2, \tilde{\phi}\} + 6\{S_1, \{S_1, \phi\}\} + \{S_1, \{S_1, \{S_1, \Phi\}\}\} \right\rangle \quad (\text{A8})$$

Substituting the expressions for the generating functions Eqs. (A1)-(A2) into Eq. (A8), one can write the following:

$$\langle \{S_1, \{S_2, \Phi\}\} \rangle = \frac{2e^2}{B^2} \frac{\partial}{\partial \mu} \left(\frac{1}{2} \frac{\partial}{\partial \mu} \langle \tilde{\Phi}^2 \tilde{\phi} \rangle + \langle \tilde{\Phi} \tilde{\phi} \rangle \frac{\partial \langle \Phi \rangle}{\partial \mu} \right) + \mathcal{N}_4(\Phi) \quad (\text{A9})$$

$$\langle \{S_3, \tilde{\Phi}\} \rangle = \frac{e^2}{B^2} \frac{\partial}{\partial \mu} \left(3 \langle \tilde{\Phi}^2 \rangle \frac{\partial \langle \phi \rangle}{\partial \mu} + 3 \langle \tilde{\Phi} \tilde{\phi} \rangle \frac{\partial \langle \Phi \rangle}{\partial \mu} + \frac{1}{2} \frac{\partial}{\partial \mu} \langle \tilde{\Phi}^2 \tilde{\phi} \rangle \right) + \frac{2e^2}{B\omega_c} \frac{\partial}{\partial \mu} \langle \tilde{\Phi} \{ \tilde{\Phi}^{(1)}, \tilde{\phi} \} \rangle + \mathcal{N}_4(\Phi) \quad (\text{A10})$$

$$\langle \{S_2, \{S_1, \Phi\}\} \rangle = \frac{2e^2}{B\omega_c} \frac{\partial}{\partial \mu} \langle \tilde{\phi} \{ \tilde{\Phi}^{(1)}, \tilde{\Phi} \} \rangle + \frac{2e^2}{B^2} \frac{\partial}{\partial \mu} \left[\langle \tilde{\Phi} \tilde{\phi} \rangle \frac{\partial \langle \Phi \rangle}{\partial \mu} \right] + \mathcal{N}_4(\Phi) \quad (\text{A11})$$

Here, the relations $\langle \tilde{\Phi} \{ \tilde{\Phi}^{(1)}, \tilde{\Phi} \} \rangle = 1/2 \langle \{ \tilde{\Phi}^{(1)}, \tilde{\Phi}^2 \} \rangle$ and $\langle \{S_3, \tilde{\Phi}\} \rangle = (e/m) \partial \langle \tilde{\Phi} S_{3,\theta} \rangle / \partial \mu$ has been used (with $S_{3,\theta} = \partial S_3 / \partial \theta$). Adding Eqs. (A10) and (A11) together, applying the Leibniz rule $\{f, gh\} = h\{f, g\} + g\{f, h\}$ and using the relation $\langle \{ \tilde{\Phi}^{(1)}, \tilde{\phi} \tilde{\Phi} \} \rangle = (e/m) \partial \langle \tilde{\Phi}^2 \tilde{\phi} \rangle / \partial \mu$, one can obtain the following:

$$\langle \{S_3, \tilde{\Phi}\} \rangle + \langle \{S_2, \{S_1, \Phi\}\} \rangle = \frac{e^2}{B^2} \frac{\partial}{\partial \mu} \left[\frac{5}{2} \frac{\partial}{\partial \mu} \langle \tilde{\Phi}^2 \tilde{\phi} \rangle + 5 \langle \tilde{\phi} \tilde{\Phi} \rangle \frac{\partial \langle \Phi \rangle}{\partial \mu} + 3 \langle \tilde{\Phi}^2 \rangle \frac{\partial \langle \phi \rangle}{\partial \mu} \right] + \mathcal{N}_4(\Phi) \quad (\text{A12})$$

Similarly, employing $\langle \{S_2, \tilde{\phi}\} \rangle = (e/m) \partial \langle \tilde{\phi} S_{2,\theta} \rangle / \partial \mu$ (with $S_{2,\theta} = \partial S_2 / \partial \theta$) one can write:

$$\langle \{S_2, \tilde{\phi}\} \rangle = -\frac{2e}{B} \frac{\partial \langle \tilde{\phi}^2 \rangle}{\partial \mu} + \frac{e^2}{B\omega_c} \frac{\partial}{\partial \mu} \langle \tilde{\phi} \{ \tilde{\Phi}^{(1)}, \tilde{\Phi} \} \rangle + \frac{2e^2}{B^2} \frac{\partial}{\partial \mu} \langle \tilde{\phi} \tilde{\Phi} \rangle \frac{\partial \langle \Phi \rangle}{\partial \mu} \quad (\text{A13})$$

$$\langle \{S_1, \{S_1, \phi\}\} \rangle = \frac{e^2}{B\omega_c} \frac{\partial}{\partial \mu} \langle \tilde{\Phi} \{ \tilde{\Phi}^{(1)}, \tilde{\phi} \} \rangle + \frac{e^2}{B^2} \frac{\partial}{\partial \mu} \left[\langle \tilde{\Phi}^2 \rangle \frac{\partial \langle \phi \rangle}{\partial \mu} \right] \quad (\text{A14})$$

Adding these two expressions together and applying again the Leibniz rule, one obtains

$$\langle \{S_2, \tilde{\phi}\} \rangle + \langle \{S_1, \{S_1, \phi\}\} \rangle = \frac{e^2}{B} \frac{\partial}{\partial \mu} \left(\frac{\partial}{\partial \mu} \langle \tilde{\Phi}^2 \tilde{\phi} \rangle + \langle \tilde{\Phi}^2 \rangle \frac{\partial \langle \phi \rangle}{\partial \mu} + 2 \langle \tilde{\phi} \tilde{\Phi} \rangle \frac{\partial \langle \Phi \rangle}{\partial \mu} \right) - \frac{2e}{B} \frac{\partial \langle \tilde{\phi}^2 \rangle}{\partial \mu} \quad (\text{A15})$$

Finally, the expression for the fourth-order gyrokinetic Hamiltonian is obtained from Eqs. (A9), (A12) and (A15):

$$H_4 = -\frac{e^2}{2B} \frac{\partial \langle \tilde{\phi}^2 \rangle}{\partial \mu} + \frac{e^3}{2B^2} \frac{\partial}{\partial \mu} \left(\frac{\partial}{\partial \mu} \langle \tilde{\Phi}^2 \tilde{\phi} \rangle + 2 \langle \tilde{\phi} \tilde{\Phi} \rangle \frac{\partial \langle \Phi \rangle}{\partial \mu} + \langle \tilde{\Phi}^2 \rangle \frac{\partial \langle \phi \rangle}{\partial \mu} \right) + \mathcal{N}_4(\Phi) \quad (\text{A16})$$

APPENDIX B: DERIVATION OF THE NONLINEAR POLARIZATION DENSITY

In this appendix we show details of the variational calculation needed to derive the nonlinear contribution to the quasineutrality equation. Note that the *second-order* contribution to the polarization density can be derived either taking the variational derivative of the *third-order* Hamiltonian function with respect to Φ or taking the variational derivative of the *fourth-order* Hamiltonian with respect to ϕ . In what follows we show that these both calculations lead to the same result.

In the third-order, the calculation is straightforward. There are few variational derivatives needed [cf. Eq. (15)]:

$$\frac{\delta}{\delta \Phi(\mathbf{x})} \int d^6 Z f \frac{\partial}{\partial \mu} \langle \tilde{\Phi} \tilde{\phi} \rangle = - \int \frac{\partial f}{\partial \mu} \tilde{\phi} \delta(\mathbf{R} + \boldsymbol{\rho} - \mathbf{x}) d^6 Z \quad (\text{B1})$$

$$\frac{\delta}{\delta \Phi(\mathbf{x})} \int d^6 Z f \frac{\partial}{\partial \mu} \left[\langle \tilde{\Phi}^2 \rangle \frac{\partial \langle \Phi \rangle}{\partial \mu} \right] = \int \left(-2 \frac{\partial f}{\partial \mu} \tilde{\Phi} \frac{\partial \langle \Phi \rangle}{\partial \mu} + \frac{\partial}{\partial \mu} \left[\langle \tilde{\Phi}^2 \rangle \frac{\partial f}{\partial \mu} \right] \right) \delta(\mathbf{R} + \boldsymbol{\rho} - \mathbf{x}) d^6 Z \quad (\text{B2})$$

$$\frac{\delta}{\delta \Phi(\mathbf{x})} \int d^6 Z f \frac{\partial^2}{\partial \mu^2} \langle \tilde{\Phi}^3 \rangle = 3 \int \frac{\partial^2 f}{\partial \mu^2} (\tilde{\Phi}^2 - \langle \tilde{\Phi}^2 \rangle) \delta(\mathbf{R} + \boldsymbol{\rho} - \mathbf{x}) d^6 Z \quad (\text{B3})$$

Here, the well-known relation $\delta\Phi(\mathbf{R} + \boldsymbol{\rho})/\delta\Phi(\mathbf{x}) = \delta(\mathbf{R} + \boldsymbol{\rho} - \mathbf{x})$ has been used. Combining Eqs. (B1)-(B3) together, one immediately obtains the expression for the nonlinear polarization density:

$$\frac{\delta}{\delta\Phi(\mathbf{x})} \int d^6Z f H_3 = \int d^6Z \frac{e^3}{2B^2} \left[\tilde{\Phi}^2 \frac{\partial^2 f}{\partial \mu^2} + \left(\frac{\partial \langle \tilde{\Phi}^2 \rangle}{\partial \mu} - 2\tilde{\Phi} \frac{\partial \langle \Phi \rangle}{\partial \mu} \right) \frac{\partial f}{\partial \mu} \right] \delta(\mathbf{R} + \boldsymbol{\rho} - \mathbf{x}) \quad (\text{B4})$$

In the fourth order, there are more variational derivatives needed:

$$-\frac{\delta}{\delta\phi(\mathbf{x})} \int \frac{e^2}{B} f \frac{\partial}{\partial \mu} \langle \tilde{\Phi} \tilde{\phi} \rangle d^6Z = \int \frac{e^2}{B} \frac{\partial f}{\partial \mu} \tilde{\Phi} \delta(\mathbf{R} + \boldsymbol{\rho} - \mathbf{x}) d^6Z \quad (\text{B5})$$

$$-\frac{\delta}{\delta\phi(\mathbf{x})} \int \frac{e^2}{2B} f \frac{\partial \langle \tilde{\phi}^2 \rangle}{\partial \mu} d^6Z = \int \frac{e^2}{B} \frac{\partial f}{\partial \mu} \tilde{\phi} \delta(\mathbf{R} + \boldsymbol{\rho} - \mathbf{x}) d^6Z \quad (\text{B6})$$

$$\frac{\delta}{\delta\phi(\mathbf{x})} \int f \frac{\partial^2}{\partial \mu^2} \langle \tilde{\Phi}^2 \tilde{\phi} \rangle d^6Z = \int \frac{\partial^2 f}{\partial \mu^2} (\tilde{\Phi}^2 - \langle \tilde{\Phi}^2 \rangle) \delta(\mathbf{R} + \boldsymbol{\rho} - \mathbf{x}) d^6Z \quad (\text{B7})$$

$$\frac{\delta}{\delta\phi(\mathbf{x})} \int f \frac{\partial}{\partial \mu} \left(\langle \tilde{\phi} \tilde{\Phi} \rangle \frac{\partial \langle \Phi \rangle}{\partial \mu} \right) d^6Z = - \int \frac{\partial f}{\partial \mu} \tilde{\Phi} \frac{\partial \langle \Phi \rangle}{\partial \mu} \delta(\mathbf{R} + \boldsymbol{\rho} - \mathbf{x}) d^6Z \quad (\text{B8})$$

$$\frac{\delta}{\delta\phi(\mathbf{x})} \int f \frac{\partial}{\partial \mu} \left(\langle \tilde{\Phi}^2 \rangle \frac{\partial \langle \phi \rangle}{\partial \mu} \right) d^6Z = \int \frac{\partial}{\partial \mu} \left(\langle \tilde{\Phi}^2 \rangle \frac{\partial f}{\partial \mu} \right) \delta(\mathbf{R} + \boldsymbol{\rho} - \mathbf{x}) d^6Z \quad (\text{B9})$$

Combining Eqs. (B5)-(B9) according to Eq.(A16), one obtains the nonlinear part of the polarization density:

$$\frac{\delta}{\delta\phi(\mathbf{x})} \int d^6Z f H_4 = \int d^6Z \frac{e^3}{2B^2} \left[\tilde{\Phi}^2 \frac{\partial^2 f}{\partial \mu^2} + \left(\frac{\partial \langle \tilde{\Phi}^2 \rangle}{\partial \mu} - 2\tilde{\Phi} \frac{\partial \langle \Phi \rangle}{\partial \mu} \right) \frac{\partial f}{\partial \mu} \right] \delta(\mathbf{R} + \boldsymbol{\rho} - \mathbf{x}) \quad (\text{B10})$$

One can show that Eq. (B10) exactly coincides with Eq. (B4) using the following relation:

$$\frac{\partial \langle \tilde{\Phi}^2 \rangle}{\partial \mu} - 2\tilde{\Phi} \frac{\partial \langle \Phi \rangle}{\partial \mu} = \frac{\partial \langle \Phi^2 \rangle}{\partial \mu} - 2\Phi \frac{\partial \langle \Phi \rangle}{\partial \mu} \quad (\text{B11})$$

This relation can be derived employing $\tilde{\Phi} = \Phi - \langle \Phi \rangle$ and $\langle \tilde{\Phi}^2 \rangle = \Phi^2 - \langle \Phi \rangle^2$.

APPENDIX C: LONG-WAVELENGTH APPROXIMATION

This appendix describes details of the long-wavelength approximation of the nonlinear polarization density and the third-order gyrokinetic Hamiltonian function. The most straightforward approach is to derive the approximation for the Hamiltonian first and then obtain the approximation for the polarization density using the variational principle. Consider the third-order gyrokinetic Hamiltonian Eq. (15). We will be interested only in the cubic part of H_3 because this is the higher-order contribution which is directly related to the nonlinear polarization density:

$$H_3 = \frac{e^2}{B} \frac{\partial}{\partial \mu} \left[\frac{e}{2B} \langle \tilde{\Phi}_{gy}^2 \rangle \frac{\partial \langle \Phi_{gy} \rangle}{\partial \mu} + \frac{e}{6B} \frac{\partial}{\partial \mu} \langle \tilde{\Phi}_{gy}^3 \rangle \right] \quad (\text{C1})$$

Here, the notation Φ_{gy} is used in order to stress the dependence of the electrostatic potential on the gyro-phase. We expand the potential in the powers of the gyroradius:

$$\Phi_{gy} = \sum_n \frac{\rho^n L_\Phi^n}{n!}, \quad L_\Phi^n = (\hat{\zeta} \cdot \nabla)^n \Phi_{gc}, \quad \Phi_{gy} = \Phi(\mathbf{R} + \boldsymbol{\rho}), \quad \Phi_{gc} = \Phi(\mathbf{R}) \equiv \Phi \quad (\text{C2})$$

Here $\rho = \sqrt{2m\mu/(e^2B)}$, $\hat{\zeta} = -\mathbf{e}_1 \cos \theta + \mathbf{e}_2 \sin \theta$, the gyroradius $\boldsymbol{\rho} = \rho \hat{\zeta}$, the *gyrokinetic* magnetic moment μ given by Eq. (49) (the first order), the *gyrokinetic* gyrophase θ defined in Eq. (50) (also the first order) and a pair of unit vectors orthogonal to the ambient magnetic field ($\mathbf{e}_1, \mathbf{e}_2$). We will need the fourth order of accuracy in $k_\perp \rho$ for our derivation (the lower orders vanish because of the μ -derivatives).

Substituting Eq. (C2) into Eq. (C1) and truncating the series on the fourth order, one obtains:

$$\langle \tilde{\Phi}_{gy}^2 \rangle = \rho^2 \langle L_\Phi L_\Phi \rangle = \frac{\rho^2}{2} (\nabla_\perp \Phi)^2, \quad \frac{\partial \langle \Phi_{gy} \rangle}{\partial \mu} = \frac{1}{2} \frac{\partial \rho^2}{\partial \mu} \langle L_\Phi^2 \rangle = \frac{1}{2e\omega_c} \nabla_\perp^2 \Phi \quad (C3)$$

$$\tilde{\Phi}_{gy}^3 = \rho^3 L_\Phi L_\Phi L_\Phi + \frac{3\rho^4}{2} L_\Phi L_\Phi \Delta_\Phi, \quad \langle L_\Phi L_\Phi \Delta_\Phi \rangle = -\frac{1}{4} \left[(\mathbf{E}_\perp \cdot \nabla_\perp) \left(\frac{\mathbf{E}_\perp^2}{2} \right) - \frac{\mathbf{E}_\perp^2}{2} (\nabla_\perp \cdot \mathbf{E}_\perp) \right] \quad (C4)$$

$$\langle \tilde{\Phi}_{gy}^3 \rangle = -\frac{3\rho^4}{8} \left[(\mathbf{E}_\perp \cdot \nabla_\perp) \left(\frac{\mathbf{E}_\perp^2}{2} \right) - \frac{\mathbf{E}_\perp^2}{2} (\nabla_\perp \cdot \mathbf{E}_\perp) \right] \quad (C5)$$

Here, $\Delta_\Phi = L_\Phi^2 - \langle L_\Phi^2 \rangle$ and $\mathbf{E}_\perp = -\nabla_\perp \Phi$ the perpendicular electric field. The following relations have been used:

$$\langle \cos^2 \theta \rangle = \langle \sin^2 \theta \rangle = 1/2, \quad \langle \cos^4 \theta \rangle = \langle \sin^4 \theta \rangle = 3/8, \quad \langle \cos^2 \theta \sin^2 \theta \rangle = 1/8 \quad (C6)$$

The third-order Hamiltonian can be cast into the compact and transparent form:

$$H_3 = \frac{e}{2B^2\omega_c^2} (\nabla_\perp \Phi \cdot \nabla_\perp) \left[\frac{(\nabla_\perp \Phi)^2}{2} \right] = -\frac{e}{2} (\mathbf{E}_\perp \cdot \nabla_\perp) \left(\frac{\mathbf{P}_E^2}{2} \right) \quad (C7)$$

Here, the gyrokinetic polarization vector $\mathbf{P}_E = -\nabla_\perp \Phi / (B\omega_c)$ is introduced.

Finally, employing the variational principle to Eq. (C7), we can derive the long-wavelength approximation of the nonlinear polarization density. The corresponding variational derivative of the gyrokinetic action is

$$e n_{\text{pol}}^{(\text{nl})} = \frac{\delta A_3}{\delta \Phi(\mathbf{x})} = \frac{\delta}{\delta \Phi(\mathbf{x})} \int H_3 \bar{f} d^6 Z = \frac{e}{2} \nabla_\perp \cdot \left(\frac{\bar{n}}{B^2\omega_c^2} [\nabla_\perp \Phi \nabla_\perp^2 \Phi - (\nabla_\perp \Phi \cdot \nabla_\perp) \nabla_\perp \Phi] \right) \quad (C8)$$

One can show that the same expression can be derived directly expanding the quasineutrality equation (46) in powers of $k_\perp \rho$ (fourth order of accuracy is needed). The nonlinear polarization density can be expressed as follows:

$$\begin{aligned} n_{\text{pol}}^{(\text{nl})} &= \int \frac{e^2}{2B^2} \left[\tilde{\Phi}_{gy}^2 \frac{\partial^2 f}{\partial \mu^2} + \left(\frac{\partial \langle \tilde{\Phi}_{gy}^2 \rangle}{\partial \mu} - 2 \tilde{\Phi}_{gy} \frac{\partial \langle \Phi_{gy} \rangle}{\partial \mu} \right) \frac{\partial f}{\partial \mu} \right] \delta(\mathbf{R} + \boldsymbol{\rho} - \mathbf{x}) d^6 Z = \\ &= \int d^6 Z \frac{e^2}{2B^2} f \left[\frac{\partial^2}{\partial \mu^2} \langle \tilde{\Phi}_{gy}^2 \delta_{gy} \rangle + 2 \frac{\partial}{\partial \mu} \left(\langle \tilde{\Phi}_{gy} \tilde{\delta}_{gy} \rangle \frac{\partial \langle \Phi_{gy} \rangle}{\partial \mu} \right) - \frac{\partial}{\partial \mu} \left(\langle \delta_{gy} \rangle \frac{\partial \langle \tilde{\Phi}_{gy}^2 \rangle}{\partial \mu} \right) \right] \end{aligned} \quad (C9)$$

One can formally expand the delta function in powers of $k_\perp \rho$:

$$\delta_{gy} = \sum_n \frac{\rho^n L_\delta^n}{n!}, \quad L_\delta^n = (\hat{\zeta} \cdot \nabla)^n \delta_{gc}, \quad \delta_{gy} = \delta(\mathbf{R} + \boldsymbol{\rho} - \mathbf{x}), \quad \delta_{gc} = \delta(\mathbf{R} - \mathbf{x}) \equiv \delta \quad (C10)$$

Using Eqs. (C2) and (C10), one can obtain:

$$n_{\text{pol}}^{(\text{nl})} = \int d^6 Z \frac{e^2}{2B^2} f \left[4 \langle L_\Phi L_\Phi L_\delta^2 \rangle + 8 \langle L_\Phi L_\delta L_\Phi^2 \rangle - 2 \langle L_\Phi L_\Phi \rangle \langle L_\delta^2 \rangle - 4 \langle L_\Phi L_\delta \rangle \langle L_\Phi^2 \rangle \right] \quad (C11)$$

$$2 \langle L_\Phi L_\Phi \rangle \langle L_\delta^2 \rangle = \frac{1}{2} P^2 (\nabla_\perp \cdot \mathbf{D}), \quad 4 \langle L_\Phi L_\delta \rangle \langle L_\Phi^2 \rangle = (\mathbf{P} \cdot \mathbf{D}) (\nabla_\perp \cdot \mathbf{P}), \quad \mathbf{P} = \nabla_\perp \Phi, \quad \mathbf{D} = \nabla_\perp \delta \quad (C12)$$

$$4 \langle L_\Phi L_\Phi L_\delta^2 \rangle + 8 \langle L_\Phi L_\delta L_\Phi^2 \rangle = \frac{3}{2} \nabla_\perp \cdot (P^2 \mathbf{D}) - (\mathbf{P} \times \nabla_\perp) \cdot (\mathbf{P} \times \mathbf{D}), \quad P^2 = \mathbf{P} \cdot \mathbf{P} \quad (C13)$$

$$(\mathbf{P} \times \nabla_\perp) \cdot (\mathbf{P} \times \mathbf{D}) = \frac{1}{2} \nabla_\perp \cdot (P^2 \mathbf{D}) + \frac{P^2}{2} (\nabla_\perp \cdot \mathbf{D}) - (\mathbf{P} \cdot \mathbf{D}) (\nabla_\perp \cdot \mathbf{P}) - \mathbf{P} \cdot [(\mathbf{P} \cdot \nabla_\perp) \mathbf{D}] \quad (C14)$$

Finally, the nonlinear polarization density takes the form ($\nabla_i P_j - \nabla_j P_i = 0$ has been used here):

$$n_{\text{pol}}^{(\text{nl})} = \int d^6 Z \frac{f}{2B^2\omega_c^2} (2\mathbf{P} \cdot [(\mathbf{D} \cdot \nabla_\perp) \mathbf{P}] + \mathbf{P} \cdot [(\mathbf{P} \cdot \nabla_\perp) \mathbf{D}]) = \nabla_\perp \cdot \left(\frac{\bar{n}}{2B^2\omega_c^2} [\mathbf{P} (\nabla_\perp \cdot \mathbf{P}) - (\mathbf{P} \cdot \nabla_\perp) \mathbf{P}] \right) \quad (C15)$$

Substituting $\mathbf{P} = \nabla_\perp \Phi$ in this expression one easily obtains Eq. (C8).

- [2] K. H. Burrell, E. J. Doyle, P. Gohil, R. J. Groebner, J. Kim, R. J. La Haye, L. L. Lao, R. A. Moyer, T. H. Osborne, W. A. Peebles, C. L. Rettig, T. H. Rhodes, and D. M. Thomas, *Phys. Plasmas* **1**, 1536 (1994).
- [3] T. S. Hahm, *Phys. Plasmas* **3**, 4658 (1996).
- [4] T. S. Hahm, L. Wang, and J. Madsen, *Phys. Plasmas* **16**, 022305 (2009).
- [5] D. E. Hastings, W. A. Houlberg, and K. C. Shaing, *Nucl. Fusion* **25**, 445 (1985).
- [6] D. E. Hastings, *Phys. Fluids* **28**, 334 (1985).
- [7] H. Idei, K. Ida, H. Sanuki, H. Yamada, H. Iguchi, S. Kubo, R. Akiyama, H. Arimoto, M. Fujiwara, M. Hosokawa, K. Matsuoka, S. Morita, K. Nishimura, K. Ohkubo, S. Okamura, S. Sakakibara, C. Takahashi, Y. Takita, K. Tsumori, and I. Yamada, *Phys. Rev. Lett.* **71**, 2220 (1993).
- [8] H. Maassberg, C. D. Beidler, U. Gasparino, M. Rome, W7-AS Team, K. S. Dyabilin, N. B. Marushchenko, and S. Murakami, *Phys. Plasmas* **7**, 295 (2000).
- [9] U. Stroth, K. Itoh, S.-I. Itoh, H. Hartfuss, H. Laqua, the ERCH team, and the W7-AS team, *Phys. Rev. Lett.* **86**, 5910 (2001).
- [10] F. Wagner, M. Hirsch, H.-J. Hartfuss, H. P. Laqua, and H. Maassberg, *Plasma Phys. Controlled Fusion* **48**, A217 (2006).
- [11] A. J. Brizard, *Phys. Plasmas* **2**, 459 (1995).
- [12] H. Sugama and W. Horton, *Phys. Plasmas* **5**, 2560 (1998).
- [13] F. Parra and P. Catto, *Plasma Phys. Controlled Fusion* **50**, 065014 (2008).
- [14] F. Parra and P. Catto, *Plasma Phys. Controlled Fusion* **51**, 065002 (2009).
- [15] B. Scott and V. Smirnov, submitted to *Phys. Plasmas*.
- [16] J. I. Hori, *Pub. Astron. Soc. Japan* **18**, 287 (1966).
- [17] A. Deprit, *Cel. Mech.* **1**, 12 (1969).
- [18] R. L. Dewar, *J. Phys.* **A9**, 2043 (1976).
- [19] A. J. Dragt and J. M. Finn, *J. Math. Phys.* **17**, 2215 (1976).
- [20] D. Dubin, J. Krommes, C. Oberman, and W. W. Lee, *Phys. Fluids* **26**, 3524 (1983).
- [21] R. G. Littlejohn, *Phys. Fluids* **24**, 1730 (1981).
- [22] J. R. Cary and R. G. Littlejohn, *Annals of Physics* **151**, 1 (1983).
- [23] T. S. Hahm, *Phys. Fluids* **31**, 2670 (1988).
- [24] T. S. Hahm, W. W. Lee, and A. J. Brizard, *Phys. Fluids* **31**, 1940 (1988).
- [25] A. J. Brizard, *Phys. Plasmas* **7**, 4816 (2000).
- [26] H. Sugama, *Phys. Plasmas* **7**, 466 (2000).
- [27] A. J. Brizard and T. S. Hahm, *Reviews of Modern Physics* **79**, 421 (2007).
- [28] A. Brizard and A. Mishchenko, *J. Plasma Phys.* **75**, 675 (2009).
- [29] R. G. Littlejohn, *J. Plasma Physics* **29**, 111 (1983).
- [30] A. Dimits, *Phys. Plasmas* **17**, 055901 (2010).
- [31] J. R. Cary, *Phys. Rep.* **79**, 131 (1981).
- [32] J. Candy and M. N. Rosenbluth, *Phys. Plasmas* **1**, 356 (1994).
- [33] J. R. Cary and A. J. Brizard, *Reviews of Modern Physics* **81**, 693 (2009).
- [34] L. Wang and T. S. Hahm, *Phys. Plasmas* **17**, 082304 (2010).