



HAL
open science

Second order nonlinear gyrokinetic theory: From the particle to the gyrocenter

Natalia Tronko, Cristel Chandre

► **To cite this version:**

Natalia Tronko, Cristel Chandre. Second order nonlinear gyrokinetic theory: From the particle to the gyrocenter. 2017. hal-01587560v1

HAL Id: hal-01587560

<https://hal.science/hal-01587560v1>

Preprint submitted on 14 Sep 2017 (v1), last revised 23 Apr 2018 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Second order nonlinear gyrokinetic theory : From the particle to the gyrocenter

N. Tronko^{1,2} and C. Chandre³

¹*Max-Planck Institute for Plasma Physics, 85748, Garching, Germany*

²*TU Munich, Mathematics Center, 85747, Garching, Germany*

³*Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France*

The standard two-step perturbative approach of the gyrokinetic reduction usually consists in a *guiding-center* reduction and a subsequent *gyrocenter* reduction. At each step, a Lie transform procedure is applied aiming to eliminate the fast gyroangle dependencies from the one-form of the particle motion. Here we present an alternative two-step procedure: Using a translation in velocity, we embed the Lie transform performed on the symplectic part of the gyrocenter reduction in the guiding-center one. Using a canonical Lie transform, we then eliminate the gyroangle dependence from the Hamiltonian. As a result we present a systematic and self-contained derivation of the second order fully electromagnetic gyrokinetic equations for the code ordering, assuming that the perturbations related to the curvature of the background magnetic field are of higher order with respect to the relative amplitudes of the fluctuating fields.

I. INTRODUCTION

A strongly magnetized plasma forms a complex multi-scaled system in space and time. Building reduced models allows the identification of the main physical mechanisms in different regimes and configurations.

After more than three decades of active development, gyrokinetic theory is nowadays one of the important theoretical frameworks for the investigation of strongly magnetized plasmas. The main idea behind the gyrokinetic dynamical reduction consists in a systematic elimination of the fastest scales of motion, which leads to a drastic reduction of computational time.

From a theoretical viewpoint, the gyrokinetic reduction provides access to accurate predictions on long-scale processes such as transport which is one of the main issues for fusion plasma confinement [1]. In astrophysical plasmas, the gyrokinetic theory is also of interest [2]: Recently, gyrokinetic simulations have been used to access small scale spectra, in order to fill the gaps when magnetohydrodynamics approximations fail.

The gyrokinetic dynamical reduction exploits the fact that the particle dynamics is decomposed into a fast rotation around the magnetic field lines and a slow drift motion. The temporal scale of the gyromotion is set by the cyclotron frequency $\Omega = eB/mc$, where e and m are, respectively, the charge and mass of the particles, B is the magnetic field amplitude and c the speed of light. The gyromotion is described by a fast gyroangle variable θ to which corresponds to a canonically conjugate slowly varying magnetic moment μ , which is, at the lowest order,

$$\mu = \frac{mv_{\perp}^2}{2B},$$

where v_{\perp} is the perpendicular velocity of the charged particle with respect to the magnetic field lines. The magnetic moment is a measure of the area enclosed by the motion of a particle rotating around a magnetic field line. From this geometrical picture comes the idea of using μ as an action variable canonically conjugated to the fast gyromotion around the magnetic field lines. In the case of a constant and uniform background magnetic field, μ is an exact dynamical invariant.

The sources for the violation of the conservation of the magnetic moment come from two different origins: first, spatial variations of the background quantities such as magnetic field non-uniformities and curvature and, second, the presence of electromagnetic fluctuations generated by the plasma. The goal of the gyromagnetic reduction is to average the fast dynamics taking into account these two sources of perturbations, and restore a conserved quantity, a modified magnetic moment.

There exist several ways to derive the gyrokinetic theory: direct ones and systematic ones. A direct approach [3] consists in performing directly an asymptotic expansion on the Vlasov equation, making the control of orderings and consistency rather cumbersome. Systematic derivations of gyrokinetic reductions use the Hamiltonian/Lagrangian framework [4, 5]. The main advantage of systematic approaches is the consistency of the reduced particle model. Moreover, the coupling between the reduced particle dynamics with electromagnetic fields within the field-particle Lagrangian provides a framework for the consistent derivation of the models. The implementation of the Hamiltonian formalism for the particle dynamical reduction towards the guiding-center motion began with the works of Littlejohn in the Hamiltonian framework [6, 7] and in the Lagrangian one [8]. Subsequent developments of the gyrokinetic field theory, coupling the reduced particle dynamics with electromagnetic fields were done in Refs. [9, 10]. Nowadays this

framework is widely used for the derivation of consistently reduced models suitable for the numerical implementation [5, 11, 12].

More specifically, the goal of the systematic gyrokinetic dynamical reduction consists in building a new set of phase-space variables, such that the θ -dependence is completely decoupled from the other variables, and that the magnetic moment μ has a trivial dynamics, i.e., $\dot{\mu} = 0$. Therefore, the reduced particle dynamics is described in the 5-dimensional phase space with variables $(\mathbf{X}, v_{\parallel}, \mu)$ where \mathbf{X} represents the reduced particle position (the gyrocenter), v_{\parallel} is a variable which mirrors the parallel velocity of the particle and μ is the conserved magnetic moment. This change of coordinates is constructed via a perturbative series of phase-space transformations. The main advantage of this approach is that these transformations are invertible at each step of the perturbative procedure, allowing to recover information on the particle dynamics from the averaged one.

The phase-space Lagrangian is the starting point of the gyrokinetic derivation. It is given by

$$L_0(\mathbf{x}, \mathbf{v}; t) = \left(\frac{e}{\epsilon c} \mathbf{A}(\mathbf{x}) + m\mathbf{v} \right) \cdot \dot{\mathbf{x}} - H(\mathbf{x}, \mathbf{v}, t), \quad (1)$$

where \mathbf{x} and \mathbf{v} are the position and velocity of the charged particle. The term in $\dot{\mathbf{x}}$ represents the symplectic part and the second term the Hamiltonian part.

Electromagnetic perturbations are introduced via the phase-space Lagrangian perturbation:

$$L_1(\mathbf{x}, \mathbf{v}; t) = \epsilon_{\delta} \left(\frac{e}{c} \mathbf{A}_1(\mathbf{x}) \cdot \dot{\mathbf{x}} - e\phi_1(\mathbf{x}) \right), \quad (2)$$

where we notice that the perturbation associated with the magnetic potential only affects the symplectic part of the phase-space Lagrangian and the electrostatic potential, its Hamiltonian part.

A fundamental aspect of the gyrokinetic theory is the assumption on the ordering of the various small parameters present in the system. Three small parameters are identified: ϵ_B , ϵ_{δ} and ϵ_{\parallel} . First, the background magnetic field is supposed to vary slowly in space and time. A small parameter of pure geometrical meaning is associated with this hypothesis: $\epsilon_B = \rho_{\text{th}}/L_B$, where ρ_{th} is the thermal Larmor radius and $L_B = \|\nabla B/B\|$ defines the spatial scale on which the magnetic field exhibits significant variations. Second, the gyrokinetic theory assumes that the amplitude of the fluctuating electromagnetic fields is small, i.e., $\epsilon_{\delta} = (k_{\perp} \rho_{\text{th}}) e\phi_1/T_i$, where ϕ_1 is the amplitude of the fluctuating electrostatic potential and T_i the ion temperature. Similarly to that we assume that $\phi_1 \sim A_{1\parallel}/v_{\parallel}$ is of order ϵ_{δ} . Finally, experimental observations report that the most dangerous instabilities occur in the plane perpendicular to the background magnetic field for strongly magnetized plasmas. We take this into account by assuming that $\epsilon_{\parallel} = E_{1\parallel}/E_{1\perp} \sim k_{\parallel}/k_{\perp}$ is also of order ϵ_{δ} .

The only two main assumptions on the small parameters which sustain the derivation of the gyrokinetic model we propose are $\epsilon_B \sim \epsilon_{\delta}^2$ and $\epsilon_{\parallel} \sim \epsilon_{\delta}$. In particular, we notice that we do not make any assumption on ϵ which is the ratio of the gyroradius to scale length, indicating how strong the magnetic field is.

In the literature (see, e.g., Ref. [5] for a review), the standard gyrokinetic dynamical reduction is organized in two consecutive steps: the *guiding-center* reduction, where the new set of phase-space variables is constructed to evidence the conservation of the magnetic moment μ with respect to the background magnetic field nonlinearities, and the subsequent *gyrocenter* reduction which builds a modified set of variables to restore the magnetic moment conservation broken by the introduction of the electromagnetic fluctuations. This gyrokinetic dynamical reduction splits the difficulties in two steps with respect to the small parameter of the system: ϵ_B for the guiding-center step and ϵ_{δ} for the gyrocenter step. Each of these steps consists in eliminating the gyrophase dependence from the symplectic and the Hamiltonian part, simultaneously.

In this work, we propose an alternative derivation of the gyrokinetic equations by combining the guiding-center and the gyrocenter reductions in a single step, treating the symplectic and the Hamiltonian part of the Lagrangian in two consecutive steps. This is achieved by an apt shift of the velocity, followed by a modified guiding-center reduction to move the θ -dependence of the symplectic part of the Lagrangian to the Hamiltonian part. Subsequently, the θ -dependence of the Hamiltonian is removed by some canonical Lie transforms. This approach follows the one developed in Ref. [12] in the particular case where the electromagnetic potential does not have any perpendicular component.

This paper is organised as follows: In Sec. II we introduce the velocity shift which allows the application of the guiding-center theory following Ref. [8]. In Sec. III we recall the general procedure of canonical Lie transforms and we apply this perturbative procedure to derive the reduced Hamiltonian dynamics at the second order in ϵ_{δ} , explicitly providing the new set of variables at order ϵ_{δ} . Finally, in Sec. IV, we present the gyrocenter characteristics, from which the gyrokinetic Vlasov equation is reconstructed, and we briefly remind the variational method to derive the reduced Maxwell equations. For completeness, we add an appendix containing a slightly revisited guiding-center derivation, with the full details of the derivation, which is suited for our choice or ordering. As a result, the derivation of the gyrokinetic model is self-consistent, and does not rely on previous knowledge of the guiding-center theory.

II. GYROCENTER AS A MODIFIED GUIDING-CENTER

The perturbed one-form associated with the perturbed Lagrangian (1)-(2) is given by

$$\gamma_{\text{pert}} = \left(\frac{e}{\epsilon c} \mathbf{A}(\mathbf{x}) + \epsilon_\delta \frac{e}{c} \mathbf{A}_1(\mathbf{x}) + m\mathbf{v} \right) \cdot d\mathbf{x} - Hdt,$$

where

$$H = \frac{1}{2}m\mathbf{v}^2 + \epsilon_\delta e\phi_1(\mathbf{x}). \quad (3)$$

The part in dt is referred to as the Hamiltonian part of the one-form, whereas the part in $d\mathbf{x}$ is its symplectic part. We translate the particle velocity \mathbf{v} :

$$\bar{\mathbf{v}} = \mathbf{v} + \epsilon_\delta \frac{e}{mc} \mathbf{A}_1(\mathbf{x}).$$

This velocity shift allows us to apply directly and readily Littlejohn's guiding-center theory [8] without performing the reduction calculations twice (first, the guiding-center reduction without the perturbation \mathbf{A}_1 and ϕ_1 , and then the gyrocenter reduction) in the symplectic part of the one-form. Therefore we can now apply the guiding-center results on γ_{pert} in the variables $(\mathbf{x}, \bar{\mathbf{v}})$. We decompose $\bar{\mathbf{v}}$ into

$$\bar{\mathbf{v}} = \bar{v}_\parallel \hat{\mathbf{b}}(\mathbf{x}) + \bar{v}_\perp \hat{\perp}(\bar{\theta}, \mathbf{x}),$$

where $\hat{\mathbf{b}} = \mathbf{B}/B$. The above-relation defines the two vectors $\hat{\perp}$ and $\hat{\rho}$ of the orthonormal basis $(\hat{\perp}, \hat{\rho}, \hat{\mathbf{b}})$. Given an orthonormal basis $(\hat{\mathbf{b}}_1(\mathbf{x}), \hat{\mathbf{b}}_2(\mathbf{x}), \hat{\mathbf{b}}(\mathbf{x}))$, the following vectors

$$\begin{aligned} \hat{\perp}(\bar{\theta}, \mathbf{x}) &= -\hat{\mathbf{b}}_1(\mathbf{x}) \sin \bar{\theta} - \hat{\mathbf{b}}_2(\mathbf{x}) \cos \bar{\theta}, \\ \hat{\rho}(\bar{\theta}, \mathbf{x}) &= \hat{\mathbf{b}}_1(\mathbf{x}) \cos \bar{\theta} - \hat{\mathbf{b}}_2(\mathbf{x}) \sin \bar{\theta}, \end{aligned}$$

define the angle $\bar{\theta}$. The crucial step in the guiding-center theory is a shift of the particle position, i.e., $\mathbf{x} = \bar{\mathbf{X}} + \boldsymbol{\rho}$ with

$$\boldsymbol{\rho} = \epsilon \frac{m\bar{v}_\perp c}{eB(\bar{\mathbf{X}})} \hat{\rho}(\bar{\theta}, \bar{\mathbf{X}}) + \bar{\boldsymbol{\rho}}_1 + \mathcal{O}(\epsilon_B^2),$$

where the explicit expression of $\bar{\boldsymbol{\rho}}_1$ which is of order ϵ_B is given in Appendix (for more details on the guiding-center theory, we refer to Refs. [8, 13, 14]). Here the expression for $\bar{\boldsymbol{\rho}}_1$ is not explicitly needed since we only provide the change of coordinates at order ϵ_δ (see Sec. III). After the modified guiding-center reduction, the one-form becomes

$$\bar{\gamma}_{\text{pert}} = \left(\frac{e}{\epsilon c} \mathbf{A}(\bar{\mathbf{X}}) + m\bar{v}_\parallel \hat{\mathbf{b}}(\bar{\mathbf{X}}) - \epsilon \frac{mc}{e} \mu \mathbf{R}^* \right) \cdot d\bar{\mathbf{X}} + \epsilon \frac{mc}{e} \bar{\mu} d\bar{\theta} - Hdt + \mathcal{O}(\epsilon_B^2),$$

where $\bar{\mu} = m\bar{v}_\perp^2/(2B(\bar{\mathbf{X}}))$, and $\mathbf{R}^* = \nabla \hat{\mathbf{b}}_1 \cdot \hat{\mathbf{b}}_2 + (\hat{\mathbf{b}} \cdot \nabla \times \hat{\mathbf{b}}) \hat{\mathbf{b}}/2$. We notice that the symplectic part of $\bar{\gamma}_{\text{pert}}$ has no explicit dependence on θ , which was the objective of the guiding-center reduction. The Poisson bracket associated with the symplectic part of $\bar{\gamma}_{\text{pert}}$ is given by

$$\{F, G\}_{\text{gc}} = \frac{e}{\epsilon mc} \left(\frac{\partial F}{\partial \theta} \frac{\partial G}{\partial \bar{\mu}} - \frac{\partial F}{\partial \bar{\mu}} \frac{\partial G}{\partial \theta} \right) + \frac{\mathbf{B}^*}{mB_\parallel^*} \cdot \left(\nabla F \frac{\partial G}{\partial \bar{v}_\parallel} - \frac{\partial F}{\partial \bar{v}_\parallel} \nabla G \right) - \epsilon \frac{c\hat{\mathbf{b}}}{eB_\parallel^*} \cdot (\nabla F \times \nabla G), \quad (4)$$

for observables F and G , functions of $(\bar{\mathbf{X}}, \bar{\theta}, \bar{\mu}, \bar{v}_\parallel)$, and where

$$\mathbf{B}^* = \mathbf{B} + \epsilon \frac{mc}{e} \bar{v}_\parallel \nabla \times \hat{\mathbf{b}} - \epsilon^2 \frac{mc^2}{e^2} \bar{\mu} \nabla \times \mathbf{R}^*,$$

and $B_\parallel^* = \hat{\mathbf{b}} \cdot \mathbf{B}^*$. After the translation in velocity, Hamiltonian (3) becomes

$$H = \frac{1}{2}m\bar{v}_\parallel^2 + \bar{\mu}B(\bar{\mathbf{X}}) + \epsilon_\delta e\psi_1(\bar{\mathbf{X}}, \bar{\theta}, \bar{\mu}, \bar{v}_\parallel) + \epsilon_\delta^2 \frac{e^2}{2mc^2} \|\mathbf{A}_1(\bar{\mathbf{X}} + \boldsymbol{\rho})\|^2, \quad (5)$$

where there is an explicit θ -dependence at order ϵ_δ and ϵ_δ^2 through the potentials ψ_1 and \mathbf{A}_1 . The modified potential ψ_1 is given by

$$\psi_1(\bar{\mathbf{X}}, \bar{\theta}, \bar{\mu}, \bar{v}_\parallel) = \phi_1(\bar{\mathbf{X}} + \boldsymbol{\rho}) - \frac{\bar{v}_\parallel}{c} \hat{\mathbf{b}}(\bar{\mathbf{X}} + \boldsymbol{\rho}) \cdot \mathbf{A}_1(\bar{\mathbf{X}} + \boldsymbol{\rho}) - \sqrt{\frac{2\bar{\mu}B(\bar{\mathbf{X}})}{mc^2}} \hat{\perp}(\bar{\theta}, \bar{\mathbf{X}} + \boldsymbol{\rho}) \cdot \mathbf{A}_1(\bar{\mathbf{X}} + \boldsymbol{\rho}).$$

It is important to note that all the fluctuating part has been removed from the symplectic part of the one-form and moved to the Hamiltonian part. In this way, the averaging over the fast variable $\bar{\theta}$ has to be performed only on the Hamiltonian and not on the symplectic part of the one-form. By using canonical transformations, the symplectic part of the one-form is not affected (up to an exact one-form).

As a final remark, we notice that the shift in the position contains the Larmor radius from the guiding center and also the displacement generated by the perturbation field \mathbf{A}_1 , and explicitly depends on \mathbf{A}_1 . At the leading order we have:

$$\mathbf{x} = \bar{\mathbf{X}} + \epsilon \frac{mc}{eB} \hat{\mathbf{b}} \times \bar{\mathbf{v}} = \bar{\mathbf{X}} + \epsilon \frac{mc}{eB} \hat{\mathbf{b}} \times \mathbf{v} + \epsilon \epsilon_\delta \frac{1}{B} \hat{\mathbf{b}} \times \mathbf{A}_1 + \mathcal{O}(\epsilon_\delta^2). \quad (6)$$

Furthermore, the averaging procedure performed in the Hamiltonian will modify the position $\bar{\mathbf{X}}$. We will come back to the expressions of the new coordinates after performing the averaging procedure.

III. AVERAGING PROCEDURE OF THE HAMILTONIAN

A. Canonical Lie transforms

In order to perform the averaging with respect to the fast variable $\bar{\theta}$ in the Hamiltonian, we use canonical Lie transforms. These transforms are near-identity canonical changes of coordinates which do not modify the expression of the symplectic part of the one-form (up to an exact one form), or equivalently, do not change the expression of the Poisson bracket. A canonical Lie transform only affects the Hamiltonian, and with an apt choice of generating function eliminates the unwanted part of the Hamiltonian, in our case, its fast-varying part. For more details on Lie transforms, we refer to Ref. [15]. The invertible change of coordinates from the old variables \mathbf{z} to the new ones \mathbf{Z} is defined as

$$\mathbf{Z} = e^{\mathcal{L}_S} \mathbf{z},$$

where S is the scalar generating function of the transformation, and the operator \mathcal{L}_S is defined as $\mathcal{L}_S = \{S, \cdot\}_{\text{gc}}$. The bracket $\{\cdot, \cdot\}_{\text{gc}}$ is given by Eq. (4). This change of coordinates transforms any observable $F(\mathbf{z})$ into $\bar{F}(\mathbf{Z})$ as

$$\bar{F}(\mathbf{Z}) = e^{-\mathcal{L}_S} F(\mathbf{Z}) = F - \{S, F\}_{\text{gc}} + \frac{1}{2} \{S, \{S, F\}_{\text{gc}}\}_{\text{gc}} + \mathcal{O}(S^3), \quad (7)$$

which is obtained from the scalar invariance $\bar{F}(\mathbf{Z}) = F(\mathbf{z})$ (and the fact that the Poisson bracket satisfies the Leibniz rule) and where the right-hand-side of Eq. (7) is evaluated at \mathbf{Z} .

We recall that the guiding-center Poisson bracket (4) is decomposed into

$$\{F, G\}_{\text{gc}} = \epsilon^{-1} \{F, G\}_{-1} + \{F, G\}_0 + \epsilon \{F, G\}_1,$$

and the Hamiltonian $H = H_0 + \epsilon_\delta H_1 + \epsilon_\delta^2 H_2$ [see Eq. (5)]. In order to remove the $\bar{\theta}$ -dependence from H , we consider a generating function of the type $S = \epsilon \epsilon_\delta S_1 + \epsilon \epsilon_\delta^2 S_2$. The purpose of S_1 is to eliminate the fluctuating part of the Hamiltonian H at order ϵ_δ (i.e., present in H_1) and S_2 eliminates the fluctuating terms at order ϵ_δ^2 . In order to illustrate the method, we first consider the order ϵ_δ . We decompose H_1 in a fluctuating and an averaged part: $H_1 = \widetilde{H}_1 + \langle H_1 \rangle$, where $\langle H_1 \rangle = (2\pi)^{-1} \int_0^{2\pi} d\bar{\theta} H_1$. At order ϵ_δ , it leads to

$$\bar{H} = H_0 + \epsilon_\delta \left(\langle H_1 \rangle + \widetilde{H}_1 - \{S_1, H_0\}_{-1} - \epsilon \{S_1, H_0\}_0 - \epsilon^2 \{S_1, H_0\}_1 \right) + \mathcal{O}(\epsilon_\delta^2).$$

By inspecting the various terms in the above-equation, we notice that $\{S_1, H_0\}_1$ is of order ϵ_B since it involves ∇B . This term is thus neglected even at the next order since $\epsilon_B \sim \epsilon_\delta^2$. The term $\{S_1, H_0\}_0$ involves a term proportional to $\mathbf{B}^* \cdot \nabla S_1$. Up to order ϵ_B , this term is the parallel gradient of the generating function. Since the generating function is a function of the fluctuating fields, this term will be of order of the parallel gradients of the fluctuating fields ϕ_1 and \mathbf{A}_1 . We assumed that these fluctuations are of order ϵ_δ , so the term $\{S_1, H_0\}_0$ is moved to order ϵ_δ^2 . Therefore the resulting equation which determines the generating function S_1 is

$$\{S_1, H_0\}_{-1} = \frac{eB}{mc} \frac{\partial S_1}{\partial \bar{\theta}} = \widetilde{H}_1. \quad (8)$$

The same line of reasoning leads to the expansion of the new Hamiltonian at the second order where we have neglected all terms involving $\{S_i, H_0\}_1$ and the terms $\{S_i, H_0\}_0$ and $\{S_i, H_1\}_0$ (for $i = 1, 2$) present at order ϵ_δ^2 :

$$\begin{aligned} \bar{H} = H_0 + \epsilon_\delta \langle H_1 \rangle - \epsilon \epsilon_\delta \{S_1, H_0\}_0 + \epsilon_\delta^2 \left(\langle H_2 \rangle + \widetilde{H}_2 - \{S_2, H_0\}_{-1} - \{S_1, H_1\}_{-1} - \epsilon^2 \{S_1, H_1\}_1 \right. \\ \left. + \frac{1}{2} \{S_1, \{S_1, H_0\}_{-1}\}_{-1} + \frac{\epsilon^2}{2} \{S_1, \{S_1, H_0\}_{-1}\}_1 \right) + \mathcal{O}(\epsilon_\delta^3). \end{aligned}$$

The equation which determines S_2 is

$$\{S_2, H_0\}_{-1} = \widetilde{H}_2 - \{S_1, \langle H_1 \rangle\}_{-1} - \frac{1}{2} \overline{\{S_1, \widetilde{H}_1\}_{-1}} - \epsilon \epsilon_\delta^{-1} \{S_1, H_0\}_0 - \epsilon^2 \{S_1, \langle H_1 \rangle\}_1 - \frac{\epsilon^2}{2} \overline{\{S_1, \widetilde{H}_1\}_1},$$

where we have used Eq. (8). Consequently, the new Hamiltonian becomes

$$\bar{H} = H_0 + \epsilon_\delta \langle H_1 \rangle + \epsilon_\delta^2 \left(\langle H_2 \rangle - \frac{1}{2} \langle \{S_1, \widetilde{H}_1\}_{-1} \rangle - \frac{\epsilon^2}{2} \langle \{S_1, \widetilde{H}_1\}_1 \rangle \right) + \mathcal{O}(\epsilon_\delta^3).$$

B. Application to Hamiltonian (5)

We rewrite the Poisson bracket $\{S_1, \widetilde{H}_1\}_{-1}$ as

$$\{S_1, \widetilde{H}_1\}_{-1} = \frac{e}{mc} \frac{\partial}{\partial \bar{\mu}} \left(\widetilde{H}_1 \frac{\partial S_1}{\partial \bar{\theta}} \right) - \frac{e}{mc} \frac{\partial}{\partial \bar{\theta}} \left(\widetilde{H}_1 \frac{\partial S_1}{\partial \bar{\mu}} \right).$$

From the expression of S_1 , we conclude that the averaged Hamiltonian \bar{H} obtained from $H = H_0 + \epsilon_\delta H_1 + \epsilon_\delta^2 H_2$ is

$$\bar{H} = H_0 + \epsilon_\delta \langle H_1 \rangle + \epsilon_\delta^2 \left(\langle H_2 \rangle - \frac{1}{2B} \frac{\partial}{\partial \bar{\mu}} \langle \widetilde{H}_1^2 \rangle + \epsilon^2 \frac{c}{2eB} \hat{\mathbf{b}} \cdot \langle \nabla S_1 \times \nabla \widetilde{H}_1 \rangle \right) + \mathcal{O}(\epsilon_\delta^3),$$

where we have used Eq. (8) and where we notice the presence of two additional second order terms, compared to the naïve average of the Hamiltonian. Next, we apply this result to Hamiltonian (5) where

$$\begin{aligned} H_0 &= \bar{\mu} B(\bar{\mathbf{X}}) + \frac{1}{2} m v_\parallel^2, \\ H_1 &= e \psi_1(\bar{\mathbf{X}}, \bar{\theta}, \bar{\mu}, v_\parallel), \\ H_2 &= \frac{e^2}{2mc^2} \|\mathbf{A}_1(\bar{\mathbf{X}} + \boldsymbol{\rho})\|^2. \end{aligned}$$

The new coordinates are denoted $(\mathbf{X}_{\text{gy}}, \theta_{\text{gy}}, \mu_{\text{gy}}, v_{\parallel \text{gy}})$. The reduced Hamiltonian is then

$$\begin{aligned} H_{\text{gy}} &= H_{\text{gc}} + \epsilon_\delta e \langle \psi_1 \rangle \\ &+ \epsilon_\delta^2 \left(\frac{e^2}{2mc^2} \langle \|\mathbf{A}_1(\mathbf{X}_{\text{gy}} + \boldsymbol{\rho})\|^2 \rangle - \frac{e^2}{2B(\mathbf{X}_{\text{gy}})} \frac{\partial}{\partial \mu_{\text{gy}}} \langle \widetilde{\psi}_1^2 \rangle - \epsilon^2 \frac{mc^2}{2B(\mathbf{X}_{\text{gy}})^2} \hat{\mathbf{b}}(\mathbf{X}_{\text{gy}}) \cdot \left\langle \nabla \widetilde{\psi}_1 \times \int d\theta_{\text{gy}} \nabla \widetilde{\psi}_1 \right\rangle \right), \end{aligned}$$

where ψ_1 is evaluated at $(\mathbf{X}_{\text{gy}}, \theta_{\text{gy}}, \mu_{\text{gy}}, v_{\parallel \text{gy}})$ and where

$$H_{\text{gc}} = \mu_{\text{gy}} B(\mathbf{X}_{\text{gy}}) + \frac{1}{2} m v_{\parallel \text{gy}}^2.$$

The averaging has been performed using the generating function S_1 given by

$$S_1(\mathbf{X}_{\text{gy}}, \theta_{\text{gy}}, \mu_{\text{gy}}, v_{\parallel \text{gy}}) = \frac{mc}{B} \int d\theta_{\text{gy}} \widetilde{\psi}_1(\mathbf{X}_{\text{gy}}, \theta_{\text{gy}}, \mu_{\text{gy}}, v_{\parallel \text{gy}}),$$

which is essential in order to determine, at the leading order, the change of coordinates which has realized the reduction.

We notice that the fields ϕ_1 and \mathbf{A}_1 are evaluated at $\mathbf{X}_{\text{gy}} + \boldsymbol{\rho}$ which is the position of the particle, which in turn depends on the fields ϕ_1 and \mathbf{A}_1 through $\boldsymbol{\rho}$ [see Eq. (6)]. We want to evaluate these fields at the position $\mathbf{X}_{\text{gy}} + \boldsymbol{\rho}_0$, where $\boldsymbol{\rho}_0$ is the Larmor radius given by

$$\boldsymbol{\rho}_0 = \epsilon \frac{mc}{eB} \hat{\mathbf{b}} \times \mathbf{v},$$

which no longer depends on ϕ_1 and \mathbf{A}_1 . We expand the term $\epsilon_\delta e \langle \psi_1 \rangle$ in H_{gy} at order ϵ_δ^2 [see Eq. (6)], using expansions of the form

$$\epsilon_\delta e \phi_1(\mathbf{X}_{\text{gy}} + \boldsymbol{\rho}) = \epsilon_\delta e \phi_1(\mathbf{X}_{\text{gy}} + \boldsymbol{\rho}_0) + \epsilon \epsilon_\delta^2 \frac{1}{B} \hat{\mathbf{b}} \times \mathbf{A}_1(\mathbf{X}_{\text{gy}} + \boldsymbol{\rho}_0) \cdot \nabla \phi_1(\mathbf{X}_{\text{gy}} + \boldsymbol{\rho}_0) + \mathcal{O}(\epsilon_\delta^3),$$

where the background-related fields (e.g., B and $\hat{\mathbf{b}}$) are now evaluated at \mathbf{X}_{gy} since their spatial dependences provide contributions to order ϵ_δ^3 in the Hamiltonian. The shift due to the fluctuating fields induces an additional ϵ_δ^2 term in the Hamiltonian H_{gy} which becomes

$$\begin{aligned} H_{\text{gy}} = & H_{\text{gc}} + \epsilon_\delta e \langle \Psi_1 \rangle + \epsilon_\delta^2 e^2 \left(\frac{1}{2mc^2} \langle \|\mathbf{A}_1\|^2 \rangle - \frac{1}{2B} \frac{\partial}{\partial \mu_{\text{gy}}} (\langle \Psi_1^2 \rangle - \langle \Psi_1 \rangle^2) \right) \\ & + \epsilon_\delta^2 \left(\epsilon \frac{e}{B} \langle \hat{\mathbf{b}} \times \mathbf{A}_1 \cdot \nabla \Psi_1 \rangle - \epsilon^2 \frac{mc^2}{2B^2} \hat{\mathbf{b}} \cdot \left\langle \nabla \widetilde{\Psi}_1 \times \int d\theta_{\text{gy}} \nabla \widetilde{\Psi}_1 \right\rangle \right), \end{aligned} \quad (9)$$

where the fields ϕ_1 and \mathbf{A}_1 are now evaluated at $\mathbf{X}_{\text{gy}} + \boldsymbol{\rho}_0$. The magnetic field is evaluated at \mathbf{X}_{gy} , and the vector $\hat{\perp}$ at $(\theta_{\text{gy}}, \mathbf{X}_{\text{gy}})$. Here the modified potential is given by

$$\Psi_1(\mathbf{X}_{\text{gy}}, \theta_{\text{gy}}, \mu_{\text{gy}}, v_{\parallel \text{gy}}) = \phi_1(\mathbf{X}_{\text{gy}} + \boldsymbol{\rho}_0) - \frac{v_{\parallel \text{gy}}}{c} \hat{\mathbf{b}}(\mathbf{X}_{\text{gy}}) \cdot \mathbf{A}_1(\mathbf{X}_{\text{gy}} + \boldsymbol{\rho}_0) - \sqrt{\frac{2\mu_{\text{gy}} B(\mathbf{X}_{\text{gy}})}{mc^2}} \hat{\perp}(\theta_{\text{gy}}, \mathbf{X}_{\text{gy}}) \cdot \mathbf{A}_1(\mathbf{X}_{\text{gy}} + \boldsymbol{\rho}_0),$$

which is a slightly different function than ψ_1 since the fields are now evaluated at $\mathbf{X}_{\text{gy}} + \boldsymbol{\rho}_0$ and not at the position of the particle. The last two terms in Eq. (9) are often discarded using an assumption on the small parameter ϵ .

C. Changes of coordinates

Next, we look at the expression of the change of coordinates which links the particle dynamics with the gyrocenter dynamics. We recall that, at the leading order, this change is a result of two steps: a translation of the velocity by the perturbation fields and of the position by a modified Larmor radius, and an averaging performed at the Hamiltonian level using a canonical Lie transform. Below, we provide the explicit expressions at order ϵ_δ .

Given our choice of generating function and the ordering of the bracket, the old coordinates \mathbf{z} as functions of the new ones \mathbf{Z} are written as

$$\mathbf{z} = \mathbf{Z} - \epsilon \epsilon_\delta \{S_1, \mathbf{Z}\}_{\text{gc}} + \mathcal{O}(\epsilon_\delta^2).$$

We remind that there was a first step (a modified guiding-center step) which mapped (\mathbf{x}, \mathbf{v}) into $(\bar{\mathbf{X}}, \bar{\theta}, \bar{\mu}, \bar{v}_{\parallel})$:

$$\begin{aligned} \mathbf{x} = & \bar{\mathbf{X}} + \epsilon \frac{mc}{eB(\bar{\mathbf{X}})} \sqrt{\frac{2\bar{\mu}B(\bar{\mathbf{X}})}{m}} \hat{\boldsymbol{\rho}}(\bar{\theta}, \bar{\mathbf{X}}), \\ \mathbf{v} = & \bar{v}_{\parallel} \hat{\mathbf{b}}(\bar{\mathbf{X}}) + \sqrt{\frac{2\bar{\mu}B(\bar{\mathbf{X}})}{m}} \hat{\perp}(\bar{\theta}, \bar{\mathbf{X}}) - \epsilon_\delta \frac{e}{mc} \mathbf{A}_1(\bar{\mathbf{X}} + \boldsymbol{\rho}). \end{aligned} \quad (10)$$

We recall that the expressions of $\bar{\theta}$ and $\bar{\mu}$ depend on \mathbf{A}_1 . In order to identify this dependence at the leading order, we use the identity

$$\sqrt{\frac{2\bar{\mu}B(\bar{\mathbf{X}})}{m}} \hat{\boldsymbol{\rho}} = \hat{\mathbf{b}} \times \mathbf{v} + \epsilon_\delta \frac{e}{mc} \hat{\mathbf{b}} \times \mathbf{A}_1(\bar{\mathbf{X}} + \boldsymbol{\rho}),$$

which is obtained by the cross product of Eq. (10) with $\hat{\mathbf{b}}$. We define $\boldsymbol{\rho}_0$ as

$$\boldsymbol{\rho}_0 = \epsilon \frac{mc}{eB} \hat{\mathbf{b}} \times \mathbf{v}.$$

Therefore, up to order ϵ_δ^2 terms,

$$\begin{aligned} \mathbf{x} = & \bar{\mathbf{X}} + \boldsymbol{\rho}_0 + \epsilon \epsilon_\delta \frac{1}{B} \hat{\mathbf{b}} \times \mathbf{A}_1(\bar{\mathbf{X}} + \boldsymbol{\rho}_0), \\ \mathbf{v} = & \bar{v}_{\parallel} \hat{\mathbf{b}} - \frac{eB}{\epsilon mc} \hat{\mathbf{b}} \times \boldsymbol{\rho}_0 - \epsilon_\delta \frac{e}{mc} \hat{\mathbf{b}} A_{1\parallel}(\bar{\mathbf{X}} + \boldsymbol{\rho}_0), \end{aligned}$$

where $A_{1\parallel} = \hat{\mathbf{b}} \cdot \mathbf{A}_1$.

The second step performed the averaging of the Hamiltonian using a canonical Lie transform. It mapped $(\bar{\mathbf{X}}, \bar{\theta}, \bar{\mu}, \bar{v}_{\parallel})$ into $(\mathbf{X}_{\text{gy}}, \theta_{\text{gy}}, \mu_{\text{gy}}, v_{\parallel\text{gy}})$. Up to order $\mathcal{O}(\epsilon_\delta^2)$, the expressions for this second change of coordinates are

$$\begin{aligned}\mathbf{X}_{\text{gy}} &= \bar{\mathbf{X}} + \epsilon\epsilon_\delta \{S_1, \bar{\mathbf{X}}\}_0 + \epsilon^2\epsilon_\delta \{S_1, \bar{\mathbf{X}}\}_1 = \bar{\mathbf{X}} - \epsilon\epsilon_\delta \left(\frac{1}{m} \hat{\mathbf{b}} \frac{\partial S_1}{\partial \bar{v}_{\parallel}} + \epsilon \frac{c}{eB} \hat{\mathbf{b}} \times \nabla S_1 \right), \\ \theta_{\text{gy}} &= \bar{\theta} + \epsilon_\delta \{S_1, \bar{\theta}\}_{-1} = \bar{\theta} - \epsilon_\delta \frac{e}{mc} \frac{\partial S_1}{\partial \bar{\mu}}, \\ \mu_{\text{gy}} &= \bar{\mu} + \epsilon_\delta \{S_1, \bar{\mu}\}_{-1} = \bar{\mu} + \epsilon_\delta \frac{e}{B} \tilde{\psi}_1, \\ v_{\parallel\text{gy}} &= \bar{v}_{\parallel}.\end{aligned}$$

By combining the two changes of coordinates, we obtain

$$\begin{aligned}\mathbf{x} &= \mathbf{X}_{\text{gy}} + \boldsymbol{\rho}_0 + \epsilon\epsilon_\delta \left(\frac{1}{m} \hat{\mathbf{b}} \frac{\partial S_1}{\partial \bar{v}_{\parallel\text{gy}}} + \epsilon \frac{c}{eB} \hat{\mathbf{b}} \times \nabla S_1 + \frac{1}{B} \hat{\mathbf{b}} \times \mathbf{A}_1 \right), \\ \mathbf{v} &= v_{\parallel\text{gy}} \hat{\mathbf{b}} - \frac{eB}{\epsilon mc} \hat{\mathbf{b}} \times \boldsymbol{\rho}_0 - \epsilon_\delta \frac{e}{mc} \hat{\mathbf{b}} A_{1\parallel},\end{aligned}$$

where the functions $\hat{\mathbf{b}}$, B , $\hat{\boldsymbol{\rho}}$ and $\hat{\perp}$ are taken at \mathbf{X}_{gy} , and the functions $\tilde{\psi}_1$, S_1 and its derivatives, at $(\mathbf{X}_{\text{gy}}, \theta_{\text{gy}}, \mu_{\text{gy}}, v_{\parallel\text{gy}})$. The fields \mathbf{A}_1 and ϕ_1 are evaluated at $\mathbf{X}_{\text{gy}} + \boldsymbol{\rho}_0$. The terms in $\boldsymbol{\rho}_0$ in the above equations are slightly misleading since they are not fully expressed as functions of the new variables $(\mathbf{X}_{\text{gy}}, \theta_{\text{gy}}, \mu_{\text{gy}}, v_{\parallel\text{gy}})$. For consistency, we introduce the variables $(\theta, \mu, v_{\parallel})$ associated with the velocity \mathbf{v} :

$$\mathbf{v} = v_{\parallel} \hat{\mathbf{b}}(\mathbf{x}) + \sqrt{\frac{2\mu B(\mathbf{x})}{m}} \hat{\perp}(\theta, \mathbf{x}).$$

The expressions of the variables $(\theta, \mu, v_{\parallel})$ are given by

$$\begin{aligned}\theta &= \theta_{\text{gy}} + \epsilon_\delta \frac{e}{mc} \left(\frac{\partial S_1}{\partial \mu_{\text{gy}}} + \sqrt{\frac{m}{2\mu_{\text{gy}} B}} \hat{\boldsymbol{\rho}} \cdot \mathbf{A}_1 \right), \\ \mu &= \mu_{\text{gy}} - \epsilon_\delta \frac{e}{B} \left(\tilde{\Psi}_1 + \sqrt{\frac{2\mu_{\text{gy}} B}{m}} \hat{\perp} \cdot \mathbf{A}_1 \right), \\ v_{\parallel} &= v_{\parallel\text{gy}} - \epsilon_\delta \frac{e}{mc} A_{1\parallel}.\end{aligned}$$

We recover the expressions of Refs. [5, 16].

IV. GYROKINETIC VLASOV-MAXWELL EQUATIONS

The electromagnetic gyrokinetic equations consist in a Vlasov equation which is defined from its characteristics and a set of Maxwell equation, a Poisson and an Ampère equation. They are obtained from Hamiltonian H_{gy} given by Eq. (9).

A. Gyrokinetic particle characteristics and the gyrokinetic Vlasov equation

The particle characteristics are given by:

$$\begin{aligned}\dot{\mathbf{X}}_{\text{gy}} &= \{\mathbf{X}_{\text{gy}}, H_{\text{gy}}\}_{\text{gc}} = \epsilon \frac{c\hat{\mathbf{b}}}{eB_{\parallel}^*} \times \nabla H_{\text{gy}} + \frac{\mathbf{B}^*}{mB_{\parallel}^*} \frac{\partial H_{\text{gy}}}{\partial v_{\parallel\text{gy}}}, \\ \dot{v}_{\parallel\text{gy}} &= \{v_{\parallel\text{gy}}, H_{\text{gy}}\}_{\text{gc}} = -\frac{\mathbf{B}^*}{mB_{\parallel}^*} \cdot \nabla H_{\text{gy}},\end{aligned}$$

where H_{gy} is given by Eq. (9). The fully nonlinear gyrokinetic Vlasov equation for the Vlasov distribution $F(\mathbf{X}_{\text{gy}}, v_{\parallel\text{gy}}, \mu_{\text{gy}}, t)$ is reconstructed from the gyrocenter particle characteristics by stating that it is constant along

the characteristics:

$$0 = \frac{dF}{dt} = \frac{\partial F}{\partial t} + \{\mathbf{X}_{\text{gy}}, H_{\text{gy}}\}_{\text{gc}} \cdot \nabla F + \{v_{\parallel \text{gy}}, H_{\text{gy}}\}_{\text{gc}} \frac{\partial F}{\partial v_{\parallel \text{gy}}},$$

where we notice that the term in $\partial F/\partial \mu_{\text{gy}}$ is absent since $\{\mu_{\text{gy}}, H_{\text{gy}}\}_{\text{gc}} = 0$. The Vlasov equation can be rewritten with the help of the Poisson bracket (4):

$$\frac{\partial F}{\partial t} = -\{F, H_{\text{gy}}\}_{\text{gc}}. \quad (11)$$

B. Gyrokinetic Vlasov-Maxwell equations

The expression for the action functional in the quasineutrality approximation is given by:

$$\begin{aligned} \mathcal{A} = \int_{t_0}^{t_1} dt \mathcal{L} = \int dt \left[\int d\Omega \left(\left(\frac{e}{\epsilon c} \mathbf{A} + m v_{\parallel \text{gy}} \hat{\mathbf{b}} \right) \cdot \dot{\mathbf{X}}_{\text{gy}} + \epsilon \frac{mc}{e} \mu_{\text{gy}} \dot{\theta}_{\text{gy}} - H_{\text{gy}} \right) F(\mathbf{X}_{\text{gy}}, v_{\parallel \text{gy}}, \mu_{\text{gy}}, t) \right. \\ \left. - \epsilon_\delta^2 \int dV \frac{|\nabla \times \mathbf{A}_1|^2}{8\pi} \right], \end{aligned}$$

where $d\Omega = dV dW$ with $dV = d^3 \mathbf{X}_{\text{gy}}$ and $dW = B_{\parallel}^* dv_{\parallel \text{gy}} d\mu_{\text{gy}} d\theta_{\text{gy}}$. The gyrokinetic quasineutrality equation is obtained in the weak form using the functional derivatives of the action:

$$\begin{aligned} 0 = \frac{\delta \mathcal{L}}{\delta \hat{\phi}_1} \circ \hat{\phi}_1 = \epsilon_\delta \int d\Omega F \left(-e \langle \hat{\phi}_1 \rangle + \epsilon_\delta \frac{e^2}{B} \frac{\partial}{\partial \mu_{\text{gy}}} \left(\langle \Psi_1 \hat{\phi}_1 \rangle - \langle \Psi_1 \rangle \langle \hat{\phi}_1 \rangle \right) \right. \\ \left. - \epsilon \epsilon_\delta \frac{e}{B} \langle \hat{\mathbf{b}} \times \mathbf{A}_1 \cdot \nabla \hat{\phi}_1 \rangle + \epsilon^2 \epsilon_\delta \frac{mc^2}{2B^2} \hat{\mathbf{b}} \cdot \left\langle \nabla \hat{\phi}_1 \times \int d\theta_{\text{gy}} \nabla \tilde{\Psi}_1 + \nabla \tilde{\Psi}_1 \times \int d\theta_{\text{gy}} \nabla \hat{\phi}_1 \right\rangle \right), \quad (12) \end{aligned}$$

where $\hat{\phi}_1$ is a test function, evaluated at $\mathbf{X}_{\text{gy}} + \boldsymbol{\rho}_0$.

The parallel component of the gyrokinetic Ampère equation is given by

$$\begin{aligned} 0 = \frac{\partial \mathcal{L}}{\partial A_{1\parallel}} \circ \hat{A}_{1\parallel} = -\frac{\epsilon_\delta^2}{4\pi} \int dV \nabla \hat{A}_{1\parallel} \cdot [\hat{\mathbf{b}} \times (\nabla \times \mathbf{A}_1)] - \epsilon_\delta^2 \frac{e^2}{mc^2} \int d\Omega F \langle A_{1\parallel} \hat{A}_{1\parallel} \rangle \\ + e \epsilon_\delta \int d\Omega F \frac{v_{\parallel \text{gy}}}{c} \left(\langle \hat{A}_{1\parallel} \rangle - \epsilon_\delta \frac{e}{B} \frac{\partial}{\partial \mu_{\text{gy}}} \left(\langle \Psi_1 \hat{A}_{1\parallel} \rangle - \langle \Psi_1 \rangle \langle \hat{A}_{1\parallel} \rangle \right) + \epsilon \epsilon_\delta \frac{1}{B} \langle \hat{\mathbf{b}} \times \mathbf{A}_1 \cdot \nabla \hat{A}_{1\parallel} \rangle \right) \\ - \epsilon^2 \epsilon_\delta^2 \int d\Omega F \frac{mc^2}{2B^2} \frac{v_{\parallel \text{gy}}}{c} \hat{\mathbf{b}} \cdot \left\langle \nabla \hat{A}_{1\parallel} \times \int d\theta_{\text{gy}} \nabla \tilde{\Psi}_1 + \nabla \tilde{\Psi}_1 \times \int d\theta_{\text{gy}} \nabla \hat{A}_{1\parallel} \right\rangle, \quad (13) \end{aligned}$$

where $\hat{A}_{1\parallel}$ is a test function, evaluated at $\mathbf{X}_{\text{gy}} + \boldsymbol{\rho}_0$. The perpendicular component of the gyrokinetic Ampère equation is given by:

$$\begin{aligned} 0 = \frac{\delta \mathcal{L}}{\delta \mathbf{A}_{1\perp}} \circ \hat{\mathbf{A}}_{1\perp} = -\frac{\epsilon_\delta^2}{4\pi} \int dV (\nabla \times \mathbf{A}_1) \cdot (\nabla \times \hat{\mathbf{A}}_{1\perp}) - \epsilon_\delta^2 \frac{e^2}{mc^2} \int d\Omega F \langle \mathbf{A}_1 \cdot \hat{\mathbf{A}}_{1\perp} \rangle \\ + e \epsilon_\delta \int d\Omega F \left[\sqrt{\frac{2\mu_{\text{gy}} B}{mc^2}} \langle \hat{\perp} \cdot \hat{\mathbf{A}}_{1\perp} \rangle - \epsilon_\delta \frac{e}{B} \frac{\partial}{\partial \mu_{\text{gy}}} \left(\sqrt{\frac{2\mu_{\text{gy}} B}{mc^2}} (\langle \Psi_1 \hat{\perp} \cdot \hat{\mathbf{A}}_{1\perp} \rangle - \langle \Psi_1 \rangle \langle \hat{\perp} \cdot \hat{\mathbf{A}}_{1\perp} \rangle) \right) \right] \\ + \epsilon \epsilon_\delta^2 \int d\Omega F \frac{e}{B} \left(-\langle \hat{\mathbf{b}} \times \hat{\mathbf{A}}_{1\perp} \cdot \nabla \Psi_1 \rangle + \sqrt{\frac{2\mu_{\text{gy}} B}{mc^2}} \langle \hat{\mathbf{b}} \times \mathbf{A}_1 \cdot \nabla (\hat{\perp} \cdot \hat{\mathbf{A}}_{1\perp}) \rangle \right) \\ - \epsilon^2 \epsilon_\delta^2 \int d\Omega F \frac{mc^2}{2B^2} \sqrt{\frac{2\mu_{\text{gy}} B}{mc^2}} \hat{\mathbf{b}} \cdot \left\langle \nabla (\hat{\perp} \cdot \hat{\mathbf{A}}_{1\perp}) \times \int d\theta_{\text{gy}} \nabla \tilde{\Psi}_1 + \nabla \tilde{\Psi}_1 \times \int d\theta_{\text{gy}} \nabla (\hat{\perp} \cdot \hat{\mathbf{A}}_{1\perp}) \right\rangle, \quad (14) \end{aligned}$$

where $\hat{\mathbf{A}}_{1\perp}$ is a test function, evaluated at $\mathbf{X}_{\text{gy}} + \boldsymbol{\rho}_0$.

Equation (11) for the Vlasov equation and Eqs. (12), (13) and (14) for the Maxwell equations constitute the second order gyrokinetic Vlasov Maxwell equations associated with the second order Hamiltonian for the gyrocenters given by Eq. (9), consistent with the ordering $\epsilon_B \sim \epsilon_\delta^2$.

ACKNOWLEDGMENTS

The authors would like to thank A. J. Brizard, B. D. Scott and E. Sonnendrücker for helpful discussions. This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training Programme No. 2014- 2018 under Grant Agreement No. 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

APPENDIX: LITTLEJOHN'S GUIDING-CENTER THEORY

In this Appendix, we revisit Littlejohn's guiding-center theory [8]. We begin with the one-form for the motion of the particle:

$$\gamma = \left(\frac{e}{\epsilon c} \mathbf{A}(\mathbf{x}) + m\mathbf{v} \right) \cdot d\mathbf{x} - H dt.$$

The revisit comes from the fact that we do not consider ϵ as a small parameter as it was done in Refs. [6–8]. Here the small parameter is ϵ_B which relates to the spatial variation of the external magnetic field. The parameter ϵ is kept in our expressions for coherence with the existing literature in order to better compare our expressions with the ones obtained by Littlejohn in Refs. [7, 8].

The idea of the guiding center is to find a change of coordinates which removes the fluctuating part from the one-form. The guiding-center transformation is based on two different changes of coordinates, a far-from-identity transformation which consists in a shift of the position by the Larmor radius, and a near-identity transformation which eliminates the fluctuating terms of the one-form at order ϵ_B .

The first step is a translation by the Larmor radius, inspired from the situation where the magnetic field is constant and uniform:

$$\mathbf{x} = \bar{\mathbf{x}} + \boldsymbol{\rho}_0.$$

We translate the velocity, by defining

$$\mathbf{w} = \mathbf{v} + \frac{e}{\epsilon mc} \left[\mathbf{A}(\bar{\mathbf{x}} + \boldsymbol{\rho}_0) - \mathbf{A}(\bar{\mathbf{x}}) - (\boldsymbol{\rho}_0 \cdot \nabla) \mathbf{A}(\bar{\mathbf{x}}) - \frac{1}{2} (\boldsymbol{\rho}_0 \boldsymbol{\rho}_0 : \nabla \nabla) \mathbf{A}(\bar{\mathbf{x}}) \right],$$

where the last term written in indices is $-(1/2)\rho_i\rho_j\partial^2\mathbf{A}/\partial\bar{x}_i\partial\bar{x}_j$. We notice that the quantity with which the velocity has been translated is of order ϵ_B^2 (since it involves third derivatives of the vector potential, i.e., second derivatives of the magnetic field). We decompose the velocity \mathbf{w} in the following way:

$$\mathbf{w} = w_{\parallel} \hat{\mathbf{b}}(\mathbf{x}) + w_{\perp} \hat{\perp}(\theta, \mathbf{x}),$$

where the orthonormal basis $(\hat{\mathbf{b}}, \hat{\perp}, \hat{\boldsymbol{\rho}})$ is defined in Sec. II. We decompose the one-form as

$$\gamma = \gamma_0 + \gamma_1,$$

where

$$\gamma_0 = \left(\frac{e}{\epsilon c} \mathbf{A}(\bar{\mathbf{x}}) + mw_{\parallel} \hat{\mathbf{b}}(\bar{\mathbf{x}}) \right) \cdot d\bar{\mathbf{x}} - H dt,$$

and

$$\begin{aligned} \gamma_1 = & \left(\frac{e}{\epsilon c} (\boldsymbol{\rho}_0 \cdot \nabla) \mathbf{A} + \frac{e}{2\epsilon c} (\boldsymbol{\rho}_0 \boldsymbol{\rho}_0 : \nabla \nabla) \mathbf{A} + mw_{\perp} \hat{\perp}(\theta, \bar{\mathbf{x}} + \boldsymbol{\rho}_0) + mw_{\parallel} [\hat{\mathbf{b}}(\bar{\mathbf{x}} + \boldsymbol{\rho}_0) - \hat{\mathbf{b}}(\bar{\mathbf{x}})] \right) \cdot d\bar{\mathbf{x}} \\ & + \left(\frac{e}{\epsilon c} \mathbf{A} + \frac{e}{\epsilon c} (\boldsymbol{\rho}_0 \cdot \nabla) \mathbf{A} + \frac{e}{2\epsilon c} (\boldsymbol{\rho}_0 \boldsymbol{\rho}_0 : \nabla \nabla) \mathbf{A} + mw_{\perp} \hat{\perp}(\theta, \bar{\mathbf{x}} + \boldsymbol{\rho}_0) + mw_{\parallel} \hat{\mathbf{b}}(\bar{\mathbf{x}} + \boldsymbol{\rho}_0) \right) \cdot d\boldsymbol{\rho}_0. \end{aligned}$$

We use a gauge-invariance to simplify the one form: γ can be replaced by $\gamma + d\sigma$ where σ is any scalar function of $(\bar{\mathbf{x}}, w_{\parallel}, \mathbf{w}_{\perp})$. Looking at the shape of γ_1 , especially the terms in $d\boldsymbol{\rho}_0$, some terms are removed by considering

$$\sigma_1 = -\frac{e}{\epsilon c} \mathbf{A} \cdot \boldsymbol{\rho}_0 - \frac{e}{2\epsilon c} (\boldsymbol{\rho}_0 \cdot \nabla) \mathbf{A} \cdot \boldsymbol{\rho}_0 - \frac{e}{6\epsilon c} (\boldsymbol{\rho}_0 \boldsymbol{\rho}_0 : \nabla \nabla) \mathbf{A} \cdot \boldsymbol{\rho}_0. \quad (15)$$

In what follows and otherwise specified, the dependence of the functions on the variables is omitted when these functions are unambiguously expressed in the current set of variables. The one-form becomes

$$\begin{aligned} \gamma_1 + d\sigma_1 = & \left(\frac{eB}{\epsilon c} \hat{\mathbf{b}} \times \boldsymbol{\rho}_0 + \frac{e}{2\epsilon c} (\boldsymbol{\rho}_0 \cdot \nabla)(B\hat{\mathbf{b}}) \times \boldsymbol{\rho}_0 + mw_{\perp} \hat{\perp} + mw_{\perp} (\boldsymbol{\rho}_0 \cdot \nabla) \hat{\perp} + mw_{\parallel} (\boldsymbol{\rho}_0 \cdot \nabla) \hat{\mathbf{b}} \right) \cdot d\bar{\mathbf{x}} \\ & + \left(\frac{eB}{2\epsilon c} \hat{\mathbf{b}} \times \boldsymbol{\rho}_0 + \frac{e}{3\epsilon c} (\boldsymbol{\rho}_0 \cdot \nabla)(B\hat{\mathbf{b}}) \times \boldsymbol{\rho}_0 + mw_{\perp} \hat{\perp} + mw_{\perp} (\boldsymbol{\rho}_0 \cdot \nabla) \hat{\perp} + mw_{\parallel} \hat{\mathbf{b}} + mw_{\parallel} (\boldsymbol{\rho}_0 \cdot \nabla) \hat{\mathbf{b}} \right) \cdot d\boldsymbol{\rho}_0, \end{aligned}$$

where we have neglected the contributions of order ϵ_B^2 . In the above expression we have used the identity

$$(\boldsymbol{\rho}_0 \cdot \nabla) \mathbf{A} - \nabla \mathbf{A} \cdot \boldsymbol{\rho}_0 = B\hat{\mathbf{b}} \times \boldsymbol{\rho}_0.$$

The condition for $\boldsymbol{\rho}_0$ is

$$\frac{eB}{\epsilon c} \hat{\mathbf{b}} \times \boldsymbol{\rho}_0 + mw_{\perp} \hat{\perp} = 0,$$

i.e.,

$$\boldsymbol{\rho}_0 = \epsilon \frac{mw_{\perp} c}{eB} \hat{\boldsymbol{\rho}}.$$

The one-form becomes

$$\begin{aligned} \gamma_1 + d\sigma_1 = & \epsilon \frac{m^2 c}{eB} \left(\frac{w_{\perp}^2}{2B} (\hat{\boldsymbol{\rho}} \cdot \nabla)(B\hat{\mathbf{b}}) \times \hat{\boldsymbol{\rho}} + w_{\perp}^2 (\hat{\boldsymbol{\rho}} \cdot \nabla) \hat{\perp} + w_{\parallel} w_{\perp} (\hat{\boldsymbol{\rho}} \cdot \nabla) \hat{\mathbf{b}} \right) \cdot d\bar{\mathbf{x}} \\ & + \epsilon \frac{m^2 c}{eB} \left[w_{\parallel} \hat{\mathbf{b}} + \frac{1}{2} w_{\perp} \hat{\perp} + \epsilon \frac{mc}{eB} \left(\frac{w_{\perp}^2}{3B} (\hat{\boldsymbol{\rho}} \cdot \nabla)(B\hat{\mathbf{b}}) \times \hat{\boldsymbol{\rho}} + w_{\perp}^2 (\hat{\boldsymbol{\rho}} \cdot \nabla) \hat{\perp} + w_{\parallel} w_{\perp} (\hat{\boldsymbol{\rho}} \cdot \nabla) \hat{\mathbf{b}} \right) \right] \cdot d(w_{\perp} \hat{\boldsymbol{\rho}}). \end{aligned}$$

The leading order term (in the small parameter ϵ_B) is

$$\epsilon \frac{m^2 c}{eB} \left(w_{\parallel} \hat{\mathbf{b}} + \frac{1}{2} w_{\perp} \hat{\perp} \right) \cdot d(w_{\perp} \hat{\boldsymbol{\rho}}) = \epsilon \frac{m^2 c}{eB} \left[\frac{w_{\perp}^2}{2} d\theta - \left(\frac{w_{\perp}^2}{2} \nabla \hat{\perp} \cdot \hat{\boldsymbol{\rho}} - w_{\perp} w_{\parallel} \nabla \hat{\boldsymbol{\rho}} \cdot \hat{\mathbf{b}} \right) \cdot d\bar{\mathbf{x}} \right],$$

since $\hat{\perp} = \partial \hat{\boldsymbol{\rho}} / \partial \theta$ and where we have used $\nabla \hat{\boldsymbol{\rho}} \cdot \hat{\perp} = -\nabla \hat{\perp} \cdot \hat{\boldsymbol{\rho}}$. We notice that the terms in $d\bar{\mathbf{x}}$ are of order ϵ_B . We rewrite the one-form γ as

$$\gamma + d\sigma_1 = \tilde{\gamma}_0 + \tilde{\gamma}_1,$$

where

$$\tilde{\gamma}_0 = \left(\frac{e}{\epsilon c} \mathbf{A}(\bar{\mathbf{x}}) + mw_{\parallel} \hat{\mathbf{b}}(\bar{\mathbf{x}}) \right) \cdot d\bar{\mathbf{x}} + \epsilon \frac{m^2 w_{\perp}^2 c}{2eB} d\theta - H dt,$$

is the leading order and

$$\begin{aligned} \tilde{\gamma}_1 = & \epsilon \frac{m^2 c}{eB} \left(\frac{w_{\perp}^2}{2B} (\hat{\boldsymbol{\rho}} \cdot \nabla)(B\hat{\mathbf{b}}) \times \hat{\boldsymbol{\rho}} + w_{\perp}^2 (\hat{\boldsymbol{\rho}} \cdot \nabla) \hat{\perp} - \frac{w_{\perp}^2}{2} \nabla \hat{\perp} \cdot \hat{\boldsymbol{\rho}} + w_{\parallel} w_{\perp} (\nabla \times \hat{\mathbf{b}}) \times \hat{\boldsymbol{\rho}} \right) \cdot d\bar{\mathbf{x}} \\ & + \epsilon^2 \frac{m^3 c^2}{e^2 B^2} \left(\frac{w_{\perp}^2}{3B} (\hat{\boldsymbol{\rho}} \cdot \nabla)(B\hat{\mathbf{b}}) \times \hat{\boldsymbol{\rho}} + w_{\perp}^2 (\hat{\boldsymbol{\rho}} \cdot \nabla) \hat{\perp} + w_{\parallel} w_{\perp} (\hat{\boldsymbol{\rho}} \cdot \nabla) \hat{\mathbf{b}} \right) \cdot (\hat{\boldsymbol{\rho}} dw_{\perp} + w_{\perp} \hat{\perp} d\theta), \end{aligned} \quad (16)$$

is of order ϵ_B . Here we have used the identity $(\hat{\boldsymbol{\rho}} \cdot \nabla) \hat{\mathbf{b}} + \nabla \hat{\boldsymbol{\rho}} \cdot \hat{\mathbf{b}} = (\nabla \times \hat{\mathbf{b}}) \times \hat{\boldsymbol{\rho}}$. We notice that the order ϵ_B of the one-form, i.e., $\tilde{\gamma}_1$, contains fluctuating terms in θ . These terms are eliminated by a near-identity transformation which we consider at order one in ϵ_B :

$$\bar{\mathbf{x}} = \mathbf{X} + \boldsymbol{\xi}(\mathbf{X}, W_{\parallel}, W_{\perp}, \Theta), \quad (17a)$$

$$w_{\parallel} = W_{\parallel} + \mathcal{W}_{\parallel}(\mathbf{X}, W_{\parallel}, W_{\perp}, \Theta), \quad (17b)$$

$$w_{\perp} = W_{\perp} + \mathcal{W}_{\perp}(\mathbf{X}, W_{\parallel}, W_{\perp}, \Theta), \quad (17c)$$

$$\theta = \Theta + \mathcal{T}(\mathbf{X}, W_{\parallel}, W_{\perp}, \Theta), \quad (17d)$$

where the unknown functions $\boldsymbol{\xi}$, \mathcal{W}_{\parallel} , \mathcal{W}_{\perp} and \mathcal{T} are of order ϵ_B .

Since $\tilde{\gamma}_1$ is of order ϵ_B , its expression in the new coordinates $(\mathbf{X}, W_{\parallel}, W_{\perp}, \Theta)$ is exactly the same as in the old coordinates $(\bar{\mathbf{x}}, w_{\parallel}, w_{\perp}, \theta)$ up to ϵ_B^2 terms. It remains to expand $\tilde{\gamma}_0$ at the first order in ϵ_B . The expansion leads to

$$\tilde{\gamma}_0 = \left(\frac{e}{\epsilon c} \mathbf{A}(\mathbf{X}) + mW_{\parallel} \hat{\mathbf{b}}(\mathbf{X}) \right) \cdot d\mathbf{X} + \epsilon \frac{m^2 W_{\perp}^2 c}{2eB} d\Theta - H dt + \tilde{\gamma}_2,$$

where

$$\begin{aligned} \tilde{\gamma}_2 &= \left(\frac{e}{\epsilon c} (\boldsymbol{\xi} \cdot \nabla) \mathbf{A} + m\mathcal{W}_{\parallel} \hat{\mathbf{b}} \right) \cdot d\mathbf{X} + \left(\frac{e}{\epsilon c} \mathbf{A} + mW_{\parallel} \hat{\mathbf{b}} \right) \cdot d\boldsymbol{\xi} \\ &\quad + \epsilon \frac{m^2 c}{eB} W_{\perp} \mathcal{W}_{\perp} d\Theta + \epsilon \frac{m^2 W_{\perp}^2 c}{2eB} d\mathcal{T}. \end{aligned}$$

We notice that we have used the approximation $B(\mathbf{X} + \boldsymbol{\xi}) \approx B(\mathbf{X})$ since the gradients of B are of order ϵ_B and $\boldsymbol{\xi}$ is also of order ϵ_B , so the difference $B(\mathbf{X} + \boldsymbol{\xi}) - B(\mathbf{X})$ is of order ϵ_B^2 . The same approximation holds for $\hat{\mathbf{b}}$. Using a gauge transformation similar to Eq. (15), we simplify $\tilde{\gamma}_2$ into

$$\begin{aligned} \tilde{\gamma}_2 + d\sigma_2 &= \left(\frac{eB}{\epsilon c} \hat{\mathbf{b}} \times \boldsymbol{\xi} + m\mathcal{W}_{\parallel} \hat{\mathbf{b}} \right) \cdot d\mathbf{X} + mW_{\parallel} \hat{\mathbf{b}} \cdot d\boldsymbol{\xi} \\ &\quad + \epsilon \frac{m^2 c}{eB} W_{\perp} \mathcal{W}_{\perp} d\Theta + \epsilon \frac{m^2 W_{\perp}^2 c}{2eB} d\mathcal{T}, \end{aligned} \quad (18)$$

with

$$\sigma_2 = -\frac{e}{\epsilon c} \mathbf{A} \cdot \boldsymbol{\xi}.$$

We look at the one-form $\tilde{\gamma}_1 + \tilde{\gamma}_2 + d\sigma_2$ and determine the unknown functions $\boldsymbol{\xi}$, \mathcal{W}_{\parallel} , \mathcal{W}_{\perp} and \mathcal{T} such that this one-form no longer possesses Θ -dependent terms. We notice that the spatial derivatives of \mathcal{T} and $\boldsymbol{\xi}$ are of order ϵ_B ; therefore, the terms in $d\boldsymbol{\xi}$ and $d\mathcal{T}$ only contain terms in dW_{\parallel} , dW_{\perp} and $d\Theta$. For the same reason, the terms in $d\rho_0$ in $\tilde{\gamma}_1$ only involve terms in dW_{\perp} and $d\Theta$.

Combining Eqs. (16) and (18), the one-form $\tilde{\gamma}_1 + \tilde{\gamma}_2 + d\sigma_2$ can be written as

$$\tilde{\gamma}_1 + \tilde{\gamma}_2 + d\sigma_2 = \boldsymbol{\Gamma}_{\mathbf{X}} \cdot d\mathbf{X} + \Gamma_{\Theta} d\Theta + \Gamma_{\parallel} dW_{\parallel} + \Gamma_{\perp} dW_{\perp}.$$

The terms in $d\mathbf{X}$ are

$$\boldsymbol{\Gamma}_{\mathbf{X}} = \frac{eB}{\epsilon c} \hat{\mathbf{b}} \times \boldsymbol{\xi} + m\mathcal{W}_{\parallel} \hat{\mathbf{b}} + \epsilon \frac{m^2 c}{eB} \left[\frac{W_{\perp}^2}{2B} (\hat{\boldsymbol{\rho}} \cdot \nabla) (B\hat{\mathbf{b}}) \times \hat{\boldsymbol{\rho}} + W_{\perp}^2 (\hat{\boldsymbol{\rho}} \cdot \nabla) \hat{\perp} - \frac{W_{\perp}^2}{2} \nabla \hat{\perp} \cdot \hat{\boldsymbol{\rho}} + W_{\parallel} W_{\perp} (\nabla \times \hat{\mathbf{b}}) \times \hat{\boldsymbol{\rho}} \right].$$

The terms in $d\Theta$ are

$$\Gamma_{\Theta} = mW_{\parallel} \hat{\mathbf{b}} \cdot \frac{\partial \boldsymbol{\xi}}{\partial \Theta} + \epsilon \frac{m^2 W_{\perp}^2 c}{2eB} \frac{\partial \mathcal{T}}{\partial \Theta} + \epsilon \frac{m^2 c}{eB} W_{\perp} \mathcal{W}_{\perp} + \epsilon^2 \frac{m^3 c^2}{e^2 B^2} \left[-\frac{W_{\perp}^3}{3B} \hat{\boldsymbol{\rho}} \cdot \nabla B + W_{\parallel} W_{\perp}^2 (\hat{\boldsymbol{\rho}} \cdot \nabla) \hat{\mathbf{b}} \cdot \hat{\perp} \right],$$

where we have used the identities $\nabla \hat{\mathbf{b}} \cdot \hat{\mathbf{b}} = \nabla \hat{\perp} \cdot \hat{\perp} = 0$ since $\hat{\mathbf{b}}$ and $\hat{\perp}$ are of unit norm. The terms in dW_{\parallel} are

$$\Gamma_{\parallel} = mW_{\parallel} \hat{\mathbf{b}} \cdot \frac{\partial \boldsymbol{\xi}}{\partial W_{\parallel}} + \epsilon \frac{m^2 W_{\perp}^2 c}{2eB} \frac{\partial \mathcal{T}}{\partial W_{\parallel}},$$

and the terms in dW_{\perp} are

$$\Gamma_{\perp} = mW_{\parallel} \hat{\mathbf{b}} \cdot \frac{\partial \boldsymbol{\xi}}{\partial W_{\perp}} + \epsilon \frac{m^2 W_{\perp}^2 c}{2eB} \frac{\partial \mathcal{T}}{\partial W_{\perp}} + \epsilon^2 \frac{m^3 c^2}{e^2 B^2} \left[W_{\perp}^2 (\hat{\boldsymbol{\rho}} \cdot \nabla) \hat{\perp} \cdot \hat{\boldsymbol{\rho}} + W_{\parallel} W_{\perp} (\hat{\boldsymbol{\rho}} \cdot \nabla) \hat{\mathbf{b}} \cdot \hat{\boldsymbol{\rho}} \right].$$

The terms in $\boldsymbol{\Gamma}_{\mathbf{X}}$ perpendicular to $\hat{\mathbf{b}}$ determine the perpendicular component of $\boldsymbol{\xi}$, whereas the terms parallel to $\hat{\mathbf{b}}$ determine \mathcal{W}_{\parallel} . The perpendicular component of $\boldsymbol{\xi}$ is obtained from the cross-product of $\boldsymbol{\Gamma}_{\mathbf{X}}$ with $\hat{\mathbf{b}}$, and is given by

$$\boldsymbol{\xi}^{\perp} = \epsilon^2 \frac{m^2 c^2}{e^2 B^2} \left(-\frac{W_{\perp}^2}{2B} (\hat{\boldsymbol{\rho}} \cdot \nabla B) \hat{\boldsymbol{\rho}} + W_{\perp}^2 \hat{\mathbf{b}} \times (\hat{\boldsymbol{\rho}} \cdot \nabla) \hat{\perp} - W_{\parallel} W_{\perp} (\hat{\mathbf{b}} \cdot \nabla \times \hat{\mathbf{b}}) \hat{\boldsymbol{\rho}} \right),$$

where we have used the fact that $\hat{\mathbf{b}} \times [(\hat{\boldsymbol{\rho}} \cdot \nabla)\hat{\mathbf{b}} \times \hat{\boldsymbol{\rho}}] = 0$ which comes from $\nabla \hat{\mathbf{b}} \cdot \hat{\mathbf{b}} = 0$. The scalar product between $\Gamma_{\mathbf{X}}$ and $\hat{\mathbf{b}}$ leads to

$$\mathcal{W}_{\parallel} = \epsilon \frac{mc}{eB} \left(\frac{W_{\perp}^2}{2} (\hat{\boldsymbol{\rho}} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\perp} - W_{\parallel} W_{\perp} (\nabla \times \hat{\mathbf{b}}) \cdot \hat{\perp} \right),$$

where we have used the identity $\nabla \hat{\perp} \cdot \hat{\mathbf{b}} = -\nabla \hat{\mathbf{b}} \cdot \hat{\perp}$. In $\Gamma_{\mathbf{X}}$, which this choice of functions, it remains

$$\Gamma_{\mathbf{X}} = -\epsilon \frac{mc}{e} \mu \mathbf{R},$$

where $\mathbf{R} = \nabla \hat{\perp} \cdot \hat{\boldsymbol{\rho}} = \nabla \hat{\mathbf{b}}_1 \cdot \hat{\mathbf{b}}_2$ which is independent of the gyroangle Θ .

The first two terms in Γ_{Θ} , Γ_{\parallel} and Γ_{\perp} calls for a gauge transformation with

$$\sigma_3 = -m W_{\parallel} \hat{\mathbf{b}} \cdot \boldsymbol{\xi} - \epsilon \frac{m^2 W_{\perp}^2 c}{2eB} \mathcal{T}$$

The terms in Γ_{Θ} , Γ_{\parallel} and Γ_{\perp} becomes

$$\begin{aligned} \tilde{\Gamma}_{\Theta} &= \epsilon \frac{m^2 c}{eB} W_{\perp} \mathcal{W}_{\perp} + \epsilon^2 \frac{m^3 c^2}{e^2 B^2} \left[-\frac{W_{\perp}^3}{3B} \hat{\boldsymbol{\rho}} \cdot \nabla B + W_{\parallel} W_{\perp}^2 (\hat{\boldsymbol{\rho}} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\perp} \right], \\ \tilde{\Gamma}_{\parallel} &= -m \hat{\mathbf{b}} \cdot \boldsymbol{\xi}, \\ \tilde{\Gamma}_{\perp} &= -\epsilon \frac{m^2 c}{eB} W_{\perp} \mathcal{T} + \epsilon^2 \frac{m^3 c^2}{e^2 B^2} \left[W_{\perp}^2 (\hat{\boldsymbol{\rho}} \cdot \nabla)\hat{\perp} \cdot \hat{\boldsymbol{\rho}} + W_{\parallel} W_{\perp} (\hat{\boldsymbol{\rho}} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\boldsymbol{\rho}} \right]. \end{aligned}$$

Using this gauge transformation, the equations which determine the unknown functions \mathcal{W}_{\perp} , \mathcal{W}_{\parallel} and the parallel component of $\boldsymbol{\xi}$ have been decoupled. We choose \mathcal{T} such that it only eliminates the fluctuating terms. Using the following expression for the fluctuating part of $(\hat{\boldsymbol{\rho}} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\boldsymbol{\rho}}$:

$$\overline{(\hat{\boldsymbol{\rho}} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\boldsymbol{\rho}}} = \frac{1}{2} \left((\hat{\boldsymbol{\rho}} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\boldsymbol{\rho}} - (\hat{\perp} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\perp} \right), \quad (19)$$

we obtain the following expression for \mathcal{T} :

$$\mathcal{T} = \epsilon \frac{mc}{eB} \left[W_{\perp} (\hat{\boldsymbol{\rho}} \cdot \nabla)\hat{\perp} \cdot \hat{\boldsymbol{\rho}} + \frac{W_{\parallel}}{2} \left((\hat{\boldsymbol{\rho}} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\boldsymbol{\rho}} - (\hat{\perp} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\perp} \right) \right].$$

In the one form it remains a term in dW_{\perp} :

$$\epsilon^2 \frac{m^3 c^2}{2e^2 B^2} W_{\parallel} W_{\perp} \left[(\hat{\boldsymbol{\rho}} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\boldsymbol{\rho}} + (\hat{\perp} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\perp} \right] dW_{\perp}.$$

Using a gauge transformation with

$$\sigma_4 = -\epsilon^2 \frac{m^3 c^2}{4e^2 B^2} W_{\parallel} W_{\perp}^2 \left[(\hat{\boldsymbol{\rho}} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\boldsymbol{\rho}} + (\hat{\perp} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\perp} \right],$$

the term in dW_{\perp} is eliminated and a new term appears in dW_{\parallel} , which then determines the parallel component of $\boldsymbol{\xi}$:

$$\hat{\mathbf{b}} \cdot \boldsymbol{\xi} = -\epsilon^2 \frac{mc^2}{4e^2 B^2} W_{\perp}^2 \left[(\hat{\boldsymbol{\rho}} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\boldsymbol{\rho}} + (\hat{\perp} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\perp} \right].$$

For \mathcal{W}_{\perp} , we choose

$$\mathcal{W}_{\perp} = \epsilon \frac{mc}{eB} \left[\frac{W_{\perp}^2}{3B} \hat{\boldsymbol{\rho}} \cdot \nabla B - W_{\parallel} W_{\perp} (\hat{\boldsymbol{\rho}} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\perp} \right],$$

in order to eliminate terms in $d\Theta$. Other choices of functions $\boldsymbol{\xi}$, \mathcal{W}_{\parallel} , \mathcal{W}_{\perp} and \mathcal{T} are possible and lead to equivalent theories, also when pushed to high orders (see Refs. [14, 17–19] for more details). Using the change of variables defined by Eq. (17) with the functions defined above, the Θ -dependence has been removed from the symplectic part of the one-form. However there is still some dependence in Θ in the Hamiltonian

Now, we proceed with the last step which is a canonical Lie transform of the Hamiltonian as in Sec. III. This transformation does not affect the symplectic part of the one-form and is only designed to eliminate the Θ -dependence in the Hamiltonian. With the change of coordinates (17), the Hamiltonian has been changed into

$$H = \frac{1}{2}mW_{\parallel}^2 + \frac{1}{2}mW_{\perp}^2 + m(W_{\parallel}\mathcal{W}_{\parallel} + W_{\perp}\mathcal{W}_{\perp}) + \mathcal{O}(\epsilon_B^2).$$

We consider a canonical Lie transform with the following generating function

$$\frac{\partial S}{\partial \Theta} + \epsilon \frac{m^2 c}{eB} (W_{\parallel} \widetilde{\mathcal{W}}_{\parallel} + W_{\perp} \widetilde{\mathcal{W}}_{\perp}) = 0,$$

where $\widetilde{\mathcal{W}}_{\parallel}$ and $\widetilde{\mathcal{W}}_{\perp}$ are the fluctuating part of \mathcal{W}_{\parallel} and \mathcal{W}_{\perp} . We notice that the gradient of the generating function, i.e., ∇S , is of order ϵ_B^2 . The way to determine the generating function S follows from the same principle as briefly explained in Sec. III. The Hamiltonian becomes

$$H = \frac{1}{2}mW_{\parallel}^2 + \frac{1}{2}mW_{\perp}^2 + m(W_{\parallel}\langle\mathcal{W}_{\parallel}\rangle + W_{\perp}\langle\mathcal{W}_{\perp}\rangle) = \frac{1}{2}mW_{\parallel}^2 + \mu B + \epsilon \frac{mc}{2e} W_{\parallel} \mu \hat{\mathbf{b}} \cdot (\nabla \times \hat{\mathbf{b}}),$$

where $\mu = mW_{\perp}^2/(2B)$. Here we have used the identity $\langle(\hat{\boldsymbol{\rho}} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\perp}\rangle = -(\hat{\mathbf{b}} \cdot (\nabla \times \hat{\mathbf{b}}))/2$. In order to complete the guiding-center derivation, we need to specify the full change of coordinates, which is a combination of the change given by Eqs. (17a)-(17d) and the canonical Lie transform, which is, up to order ϵ_B^2 terms:

$$\begin{aligned} \bar{\mathbf{x}} &= \mathbf{X} + \boldsymbol{\xi} + \{S, \mathbf{X}\}_{\text{gc}} = \mathbf{X} + \boldsymbol{\xi}^{\perp} + (\hat{\mathbf{b}} \cdot \boldsymbol{\xi})\hat{\mathbf{b}} - \frac{\hat{\mathbf{b}}}{m} \frac{\partial S}{\partial W_{\parallel}}, \\ w_{\parallel} &= W_{\parallel} + \mathcal{W}_{\parallel} + \{S, W_{\parallel}\}_{\text{gc}} = W_{\parallel} + \mathcal{W}_{\parallel}, \\ w_{\perp} &= W_{\perp} + \mathcal{W}_{\perp} + \{S, W_{\perp}\}_{\text{gc}} = W_{\perp} + \mathcal{W}_{\perp} + \frac{eB}{\epsilon m^2 W_{\perp} c} \frac{\partial S}{\partial \Theta}, \\ \theta &= \Theta + \mathcal{T} + \{S, \Theta\}_{\text{gc}} = \Theta + \mathcal{T} - \frac{eB}{\epsilon m^2 W_{\perp} c} \frac{\partial S}{\partial W_{\perp}}, \end{aligned}$$

where the Poisson bracket is obtained from the symplectic part of the one-form, and is given by

$$\{F, G\}_{\text{gc}} = \frac{e}{\epsilon mc} \left(\frac{\partial F}{\partial \Theta} \frac{\partial G}{\partial \mu} - \frac{\partial F}{\partial \mu} \frac{\partial G}{\partial \Theta} \right) + \frac{\mathbf{B}^*}{mB_{\parallel}^*} \cdot \left(\nabla F \frac{\partial G}{\partial W_{\parallel}} - \frac{\partial F}{\partial W_{\parallel}} \nabla G \right) - \epsilon \frac{c\hat{\mathbf{b}}}{eB_{\parallel}^*} \cdot (\nabla F \times \nabla G),$$

where

$$\mathbf{B}^* = \mathbf{B} + \epsilon \frac{mc}{e} W_{\parallel} \nabla \times \hat{\mathbf{b}} - \epsilon^2 \frac{mc^2}{e^2} \mu \nabla \times \mathbf{R},$$

and $B_{\parallel}^* = \mathbf{B}^* \cdot \hat{\mathbf{b}}$. In order to obtain explicit expressions for the change of coordinates, we specify the fluctuating part of the functions \mathcal{W}_{\parallel} and \mathcal{W}_{\perp} :

$$\begin{aligned} \widetilde{\mathcal{W}}_{\perp} &= \epsilon \frac{mc}{eB} \left(\frac{W_{\perp}^2}{3B} \hat{\boldsymbol{\rho}} \cdot \nabla B - \frac{W_{\parallel} W_{\perp}}{2} \left((\hat{\boldsymbol{\rho}} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\perp} + (\hat{\perp} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\boldsymbol{\rho}} \right) \right), \\ \widetilde{\mathcal{W}}_{\parallel} &= \epsilon \frac{mc}{eB} \left(\frac{W_{\perp}^2}{4} \left((\hat{\boldsymbol{\rho}} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\perp} + (\hat{\perp} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\boldsymbol{\rho}} \right) - W_{\parallel} W_{\perp} (\nabla \times \hat{\mathbf{b}}) \cdot \hat{\perp} \right), \end{aligned}$$

where we have used the fact that the fluctuating part of $(\hat{\boldsymbol{\rho}} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\perp}$ is given by

$$\overline{(\hat{\boldsymbol{\rho}} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\perp}} = \frac{1}{2} \left((\hat{\boldsymbol{\rho}} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\perp} + (\hat{\perp} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\boldsymbol{\rho}} \right).$$

The generating function S is chosen to be purely fluctuating:

$$S = \epsilon^2 \frac{m^3 c^2}{e^2 B^2} \left[\frac{W_{\perp}^3}{3B} \hat{\perp} \cdot \nabla B + \frac{W_{\parallel} W_{\perp}^2}{8} \left((\hat{\boldsymbol{\rho}} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\boldsymbol{\rho}} - (\hat{\perp} \cdot \nabla)\hat{\mathbf{b}} \cdot \hat{\perp} \right) + W_{\parallel}^2 W_{\perp} (\nabla \times \hat{\mathbf{b}}) \cdot \hat{\boldsymbol{\rho}} \right],$$

where we have used the fluctuating part of $(\hat{\rho} \cdot \nabla) \hat{\mathbf{b}} \cdot \hat{\rho}$ is given by Eq. (19). Therefore the change of coordinates from (\mathbf{x}, \mathbf{v}) to $(\mathbf{X}, W_{\parallel}, W_{\perp}, \Theta)$ is given by

$$\begin{aligned} \mathbf{x} = \mathbf{X} + \epsilon \frac{mW_{\perp}c}{eB} \hat{\rho} + \epsilon^2 \frac{m^2c^2}{e^2B^2} & \left(-\frac{W_{\perp}^2}{2B} \left((\hat{\rho} \cdot \nabla B) \hat{\rho} + 2(\hat{\perp} \cdot \nabla B) \hat{\perp} \right) - W_{\parallel} W_{\perp} \left((\hat{\mathbf{b}} \cdot \nabla \times \hat{\mathbf{b}}) \hat{\rho} + 2(\hat{\rho} \cdot \nabla \times \hat{\mathbf{b}}) \hat{\mathbf{b}} \right) \right. \\ & - \frac{W_{\perp}^2}{8} \left(3\hat{\rho} \cdot \nabla \hat{\mathbf{b}} \cdot \hat{\rho} + \hat{\perp} \cdot \nabla \hat{\mathbf{b}} \cdot \hat{\perp} \right) \hat{\mathbf{b}} + W_{\parallel}^2 \left((\hat{\perp} \cdot \nabla \times \hat{\mathbf{b}}) \hat{\rho} - (\hat{\rho} \cdot \nabla \times \hat{\mathbf{b}}) \hat{\perp} \right) \\ & \left. + \frac{W_{\parallel} W_{\perp}}{4} \left(-3\hat{\rho} \cdot \nabla \hat{\mathbf{b}} \cdot \hat{\perp} + \hat{\perp} \cdot \nabla \hat{\mathbf{b}} \cdot \hat{\rho} \right) \hat{\rho} + \frac{W_{\parallel} W_{\perp}}{4} \left(\hat{\rho} \cdot \nabla \hat{\mathbf{b}} \cdot \hat{\rho} - \hat{\perp} \cdot \nabla \hat{\mathbf{b}} \cdot \hat{\perp} \right) \hat{\perp} \right), \end{aligned} \quad (20a)$$

$$v_{\parallel} = W_{\parallel} + \epsilon \frac{mc}{eB} \left(\frac{W_{\perp}^2}{2} \hat{\rho} \cdot \nabla \hat{\mathbf{b}} \cdot \hat{\perp} - W_{\parallel} W_{\perp} \nabla \times \hat{\mathbf{b}} \cdot \hat{\perp} \right), \quad (20b)$$

$$v_{\perp} = W_{\perp} + \epsilon \frac{mc}{eB} \left(-\frac{3W_{\parallel} W_{\perp}}{4} \hat{\rho} \cdot \nabla \hat{\mathbf{b}} \cdot \hat{\perp} + \frac{W_{\parallel} W_{\perp}}{4} \hat{\perp} \cdot \nabla \hat{\mathbf{b}} \cdot \hat{\rho} + W_{\parallel}^2 \nabla \times \hat{\mathbf{b}} \cdot \hat{\perp} \right), \quad (20c)$$

$$\theta = \Theta + \epsilon \frac{mc}{eB} \left(-\frac{W_{\perp}}{B} \hat{\perp} \cdot \nabla B + W_{\perp} \hat{\rho} \cdot \nabla \hat{\perp} \cdot \hat{\rho} + \frac{W_{\parallel}}{4} \left(\hat{\rho} \cdot \nabla \hat{\mathbf{b}} \cdot \hat{\rho} - \hat{\perp} \cdot \nabla \hat{\mathbf{b}} \cdot \hat{\perp} \right) - \frac{W_{\parallel}^2}{W_{\perp}} \nabla \times \hat{\mathbf{b}} \cdot \hat{\rho} \right), \quad (20d)$$

where the right hand side is evaluated at $(\mathbf{X}, W_{\parallel}, W_{\perp}, \Theta)$ and where

$$\mathbf{v} = v_{\parallel} \hat{\mathbf{b}}(\mathbf{x}) + v_{\perp} \hat{\perp}(\theta, \mathbf{x}).$$

The inversion of the change of variables given by Eq. (20), i.e., providing $(\mathbf{X}, W_{\parallel}, W_{\perp}, \Theta)$ as functions of $(\mathbf{x}, v_{\parallel}, v_{\perp}, \theta)$, gives the exact same equations given in Ref. [8].

We note that the term $W_{\parallel} \mu \hat{\mathbf{b}} \cdot (\nabla \times \hat{\mathbf{b}})$ is usually moved to the symplectic part of the one-form by a translation in W_{\parallel} . In the new coordinates, it leads to the Hamiltonian

$$H_{\text{gc}} = \mu B(\mathbf{X}) + \frac{1}{2} m W_{\parallel}^2,$$

and the one-form

$$\gamma_{\text{gc}} = \left[\frac{e}{\epsilon c} \mathbf{A}(\mathbf{X}) + m W_{\parallel} \hat{\mathbf{b}}(\mathbf{X}) - \epsilon \frac{mc}{e} \mu \mathbf{R}^* \right] \cdot d\mathbf{X} + \epsilon \frac{mc}{e} \mu d\Theta - H_{\text{gc}} dt,$$

where $\mathbf{R}^* = \nabla \hat{\mathbf{b}}_1 \cdot \hat{\mathbf{b}}_2 + (\hat{\mathbf{b}} \cdot \nabla \times \hat{\mathbf{b}}) \hat{\mathbf{b}}/2$.

-
- [1] X. Garbet, Y. Idomura, L. Villard, and T. H. Watanabe. Gyrokinetic simulations of turbulent transport. *Nuclear Fusion*, 50:043002, 2010.
- [2] A. A. Schekochihin, S. C. Cowley, W. Dorland, G. W. Hammett, G. G. Howes, E. Quataert, and T. Tatsuno. Astrophysical gyrokinetics: kinetic and fluid turbulent cascades in magnetized weakly collisional plasmas. *Astro. J. Suppl.*, 182:310, 2009.
- [3] E. A. Frieman and L. Chen. Nonlinear gyrokinetic equations for low-frequency electromagnetic waves in general plasma equilibria. *Phys. Fluids*, 25:502, 1982.
- [4] T. S. Hahm. Nonlinear gyrokinetic equations for tokamak microturbulence. *Phys. Fluids*, 31:2670, 1988.
- [5] A. J. Brizard and T. S. Hahm. Foundations of nonlinear gyrokinetic theory. *Rev. Mod. Phys.*, 79:421, 2007.
- [6] R. G. Littlejohn. A guiding center Hamiltonian: A new approach. *J. Math. Phys.*, 20:2445, 1979.
- [7] R. G. Littlejohn. Hamiltonian formulation of guiding-center motion. *Phys. Fluids*, 29:1730, 1981.
- [8] R. G. Littlejohn. Variational principles of guiding centre motion. *J. Plasma Phys.*, 29:111, 1983.
- [9] H. Sugama. Gyrokinetic field theory. *Phys. Plasmas*, 7:466, 2000.
- [10] A. J. Brizard. New variational principle for the Vlasov-Maxwell equations. *Phys. Rev. Lett.*, 84:5768, 2000.
- [11] N. Tronko, A. Bottino, and E. Sonnendrücker. Second order gyrokinetic theory for Particle-In-Cell codes. *Phys. Plasmas*, 23:082505, 2016.
- [12] N. Tronko, A. Bottino, C. Chandre, and E. Sonnendrücker. Hierarchy of second order gyrokinetic Hamiltonian models for Particle-In-Cell codes. *Phys. Plasma Contr. Fusion*, 59:064008, 2017.
- [13] J. R. Cary and A. J. Brizard. Hamiltonian theory of guiding-center motion. *Rev. Mod. Phys.*, 81:693, 2009.
- [14] N. Tronko and A. J. Brizard. Lagrangian and Hamiltonian constraints for guiding-center Hamiltonian theories. *Phys. Plasmas*, 22:112507, 2015.
- [15] J. R. Cary. Lie transform perturbation theory for Hamiltonian systems. *Phys. Rep.*, 79:129, 1981.

- [16] A. Brizard. Nonlinear gyrokinetic Maxwell-Vlasov equations using magnetic co-ordinates. *J. Plasma Phys.*, 41:541, 1989.
- [17] J. W. Burby, J. Squire and H. Qin. Automation of the guiding center expansion. *Phys. Plasmas*, 20:072105, 2013.
- [18] F. I. Parra and I. Calvo. Phase-space Lagrangian derivation of electrostatic gyrokinetics in general geometry. *Plasma Phys. Controlled Fusion*, 53:045001, 2011.
- [19] F. I. Parra, I. Calvo, J. W. Burby, J. Squire and H. Qin. Equivalence of two independent calculations of the higher order guiding center Lagrangian. *Phys. Plasmas*, 21:104506, 2014.