Corrigendum: Phase-space Lagrangian derivation of electrostatic gyrokinetics in general geometry (2011 *Plasma Phys. Control. Fusion* 53 045001)

Felix I Parra$^{1,3}$ and Iván Calvo$^2$

$^1$Rudolf Peierls Centre for Theoretical Physics, University of Oxford, Oxford, OX1 3NP, UK

$^2$Laboratorio Nacional de Fusión, CIEMAT, 28040 Madrid, Spain

E-mail: f.parradiaz1@physics.ox.ac.uk and ivan.calvo@ciemat.es

Received 9 July 2014

Accepted for publication 9 July 2014

Published 22 July 2014

There is an error in equation (135) of [1]. The second term in the second line of equation (135) should be

\[
\left( \frac{\mu^2}{8} - \frac{u^2 \mu}{4B} \right) \nabla_{R \perp} \hat{b} : (\nabla_{R \perp} \hat{b})^T
\]

(1)

instead of

\[
\left( \frac{\mu^2}{8} - \frac{u^2 \mu}{4B} \right) \nabla_{R} \hat{b} : (\nabla_{R} \hat{b})^T.
\]

(2)

that is, only the perpendicular component of the gradient enters this term. Equation (135) is obtained by replacing equation (D.31) into (D.26). In equation (D.31), the second term,

\[
\frac{\mu^2}{4} \nabla_{R \perp} \hat{b} : (\nabla_{R \perp} \hat{b})^T,
\]

(3)

is the one that gives the term in (1).

Note that equations (D.26) and (D.31) are correct. The typographical error is restricted to equation (135), where the subindex $\perp$ was mistakenly dropped from the tensorial expression $\nabla_{R \perp} \hat{b} : (\nabla_{R \perp} \hat{b})^T$.

**References**

[1] Parra F I and Calvo I 2011 *Plasma Phys. Control. Fusion* 53 045001

$^3$ Also at: Culham Centre for Fusion Energy, Abingdon, OX14 3DB, UK.
Phase-space Lagrangian derivation of electrostatic gyrokinetics in general geometry

Felix I Parra\textsuperscript{1,2} and Iván Calvo\textsuperscript{2,3}

\textsuperscript{1} Rudolf Peierls Centre for Theoretical Physics, University of Oxford, Oxford, OX1 3NP, UK
\textsuperscript{2} Isaac Newton Institute for Mathematical Sciences, University of Cambridge, Cambridge, CB3 0EH, UK
\textsuperscript{3} Laboratorio Nacional de Fusión, Asociación EURATOM-CIEMAT, 28040 Madrid, Spain

E-mail: f.parradiaz1@physics.ox.ac.uk and ivan.calvo@ciemat.es

Received 1 September 2010, in final form 24 November 2010
Published 17 February 2011
Online at stacks.iop.org/PPCF/53/045001

Abstract

Gyrokinetic theory is based on an asymptotic expansion in the small parameter $\epsilon$, defined as the ratio of the gyroradius and the characteristic length of variation of the magnetic field. In this paper, this ordering is strictly implemented to compute the electrostatic gyrokinetic phase-space Lagrangian in general magnetic geometry to order $\epsilon^2$. In particular, a new expression for the complete second-order gyrokinetic Hamiltonian is provided, showing that in a rigorous treatment of gyrokinetic theory magnetic geometry and turbulence cannot be dealt with independently. The new phase-space gyrokinetic Lagrangian gives a Vlasov equation accurate to order $\epsilon^2$ and a Poisson equation accurate to order $\epsilon$. The final expressions are explicit and can be implemented into any simulation without further computations.

1. Introduction

Gyrokinetics \cite{1} has proven a very useful tool to study turbulence in the core of fusion devices, making kinetic simulations of turbulent fluctuations possible in reasonable computational times \cite{2-7}. Its main advantage is averaging over the gyrofrequency time scale without losing the effect of the finite size of the gyroradius that is of the order of the typical wavelength of the turbulence. To perform this average, it is necessary to assume certain orderings that in the electrostatic limit can be summarized as

\begin{equation}
B(r) \quad \text{with} \quad \nabla \sim \frac{1}{L} \\
\phi(r, t) \quad \text{with} \quad \nabla_\perp \sim \frac{1}{\rho}, \quad \hat{b} \cdot \nabla \sim \frac{1}{L}, \quad \frac{\partial}{\partial t} \sim \omega \\
\frac{\omega}{\Omega} \sim \frac{\rho}{L} \sim \frac{Ze \phi}{Mv^2} \sim \epsilon \ll 1.
\end{equation}
where $\varphi(r,t)$ is the electrostatic potential, $B(r)$ is the magnetic field, $\omega$ is the characteristic frequency of the turbulent fluctuations, $L$ is a characteristic macroscopic scale, $v_t$, $\rho = v_t/\Omega$ and $\Omega = ZeB/Mc$ are the thermal speed, the gyroradius and the gyrofrequency of the species of interest, $Ze$ and $M$ are the charge and the mass, and $e$ and $c$ are the magnitude of the electron charge and the speed of light. Since this paper is about electrostatic gyrokinetics, we have assumed that the magnetic field is stationary and its characteristic length of variation is of the order of the macroscopic length $L$. The ordering in (1) implies that the electrostatic potential fluctuates with some characteristic frequency $\omega$ and has a strong gradient perpendicular to the magnetic field, on the order of the inverse of the gyroradius, whereas its gradient parallel to the direction of the magnetic field, $\hat{b} = B/B$, is on the order of the inverse of the larger scale $L$. The frequency $\omega$ of the turbulence is usually much smaller than the gyrofrequency, making the gyrokinetic average over the gyromotion valid. We have employed the small parameter $\epsilon \sim \omega/\Omega \ll 1$ to make this explicit. In most fusion experiments, the ratio of the gyroradius and the macroscopic length is another small parameter that we also order as $\epsilon$. It is easy to see that in drift wave turbulence, for which the characteristic frequency is $\omega \sim v_t/L$, the quantities $\omega/\Omega$ and $\rho/L$ are indeed of the same order. More importantly, to obtain the typical gyrokinetic formalism, it is necessary to order the electrostatic potential as small compared with the characteristic energy of the particles. This assumption is necessary to prove that the gyromotion of the particles is circular to lowest order. The most common gyrokinetic ordering assumes that the parameter $Ze\varphi/Mv_t^2$ is comparable to $\epsilon$, as is done in (1). In this way, magnetic geometry effects such as the $\nabla B$ and curvature drifts, of order $\rho/L$, are allowed to be comparable to the turbulent $E \times B$ drift, of order $Ze\varphi/Mv_t^2$. Ordering these effects so that they are comparable is very important in, for example, the core of tokamaks, where the curvature of the magnetic field lines is believed to be the most important drive for the turbulence [8]. The ordering in (1) contains the simplest assumptions that are still interesting, but it can be extended to include components of the potential that have perpendicular gradients of the order of the inverse of the macroscopic length $L$ [9, 10]. The results that we present in this paper can be easily extended to some of these more general orderings, but we leave this for future work.

There are different techniques to obtain gyrokinetics (and for that matter, drift kinetics [11–13], of which gyrokinetics is a natural extension). On the one hand, it is possible to obtain the gyrokinetic equation by working iteratively on the Vlasov equation [14–17]. We will call these iterative methods. On the other hand, it is possible to use phase-space Lagrangian/Hamiltonian methods that solve order by order for the motion of the particle in a given electromagnetic field, uncoupling the gyromotion from the slower time scales [18–25]. Once the motion of the particle is known, the Vlasov equation is simply obtained by its characteristics. We will call these Lagrangian methods. Both procedures are asymptotic expansions in the parameter $\epsilon$, and give equivalent equations order by order, but the Lagrangian methods have the advantage of giving the equations in a form that exactly conserves some energy-like quantities. This property may be very important for the global, full $f$ simulations that are being developed [26–29]. To have an energy-like invariant and at the same time obtain equations of motion and equations for the electromagnetic fields that are the same to first order in $\epsilon$ as those obtained with the iterative procedure, it is necessary to carry the expansion in $\epsilon$ to higher order. For example, in a slab [23], it is necessary to obtain the Hamiltonian to second order in $\epsilon$. The second-order piece of the Hamiltonian, quadratic in the electrostatic potential $\varphi$, gives second-order corrections to the equations of motion and hence it is in principle negligible to first order. However, if the lowest order quasineutrality equation that contains a linear term in $\varphi$ is employed, the second-order correction to the Hamiltonian must be kept to obtain an energy-like invariant.
The complete calculation to order $\epsilon^2$ has not been done for a general static magnetic field in either formalism so far\(^4\). In the most common Lagrangian formulation \([25]\), the calculation is done in two steps: first, the turbulent electromagnetic fields are ignored and only the background magnetic field is considered, giving the drift kinetic equation; in the second step, the turbulent electromagnetic fields are added and the corresponding corrections are calculated. Consider the case in which the magnetic field does not vary in time, i.e. electrostatic gyrokinetics. In the first step, the equations are expanded in the small parameter $\epsilon \sim \rho/L$, whereas in the second step, they are expanded in $\epsilon_\psi \sim Z e \psi / M v_i^2$. The expansion in $\epsilon$ is only performed to first order because the next order results are very tedious to calculate. The expansion in $\epsilon_\psi$ is continued to second order because the pieces quadratic in $\psi$ are needed to have an energy-like invariant. In the expansion in $\epsilon$, the fact that there has been a previous expansion in $\epsilon$ is ignored, and as a result the terms of order $\epsilon \epsilon_\psi$ are never calculated. The missing terms of order $\epsilon^2$ and $\epsilon \epsilon_\psi$ are comparable to the terms of order $\epsilon_\psi^2$ according to the gyrokinetic ordering in (1), making this expansion consistent only when $\epsilon_\psi \gg \epsilon$. In addition, since the cross-terms that contain both the background magnetic field and the turbulent electrostatic potential, of order $\epsilon \epsilon_\psi$, are always neglected when the two-step method is presented, it is not obvious how to calculate them following that procedure. In this paper, we present the complete phase-space Lagrangian calculation with the standard gyrokinetic ordering (1), emphasizing the self-consistent calculation of the terms of order $\epsilon^2$ and $\epsilon \epsilon_\psi$. In the gyrokinetic equations that result from the new Lagrangian, the magnetic geometry effects and the fluctuating potential appear together in the second-order terms, showing that geometry and turbulence cannot be separated and dealt with independently. Our main result is the explicit expression for the second-order gyrokinetic Hamiltonian given in equations (132), (133), (134) and (135). It clearly exhibits the interplay between geometry and turbulence inherent to gyrokinetic theory, possessing terms of three types: terms quadratic in the electrostatic potential, terms that include both the electrostatic potential and the magnetic geometry and terms that are purely geometrical.

At this point, it is fair to wonder about the motivations beyond formal coherence to carry out the expansion consistently to second order in $\epsilon$. Keeping the second-order piece of the Hamiltonian that is quadratic in the electrostatic potential is necessary for the conservation of an energy-like invariant, as already noted above. When the other second-order terms computed in this paper are included, they have two effects: (i) the gyrokinetic Poisson’s equation is modified by the effect of the non-uniform magnetic field on the gyro-orbits and (ii) the equations of motion are modified to second order. Both of these effects are not conventionally kept in gyrokinetic formulations, but they may be crucial for conservation of momentum. Conservation of momentum in full $f$ gyrokinetic formulations has been the center of a recent controversy \([17, 30–34]\). By assuming a gyroBohm level of turbulent transport of momentum at long wavelengths, Catto and one of us, FIP, have argued that to recover with a full $f$ model the correct transport of toroidal angular momentum in a tokamak, it is necessary to have gyrokinetic Fokker–Planck and Poisson’s equations correct to third order in $\epsilon$ in the high flow ordering, for which the average velocity of the ions $V_i$ is of the order of the ion thermal speed $v_{ti}$, and correct to fourth order in the low flow ordering, for which $V_i \sim \epsilon v_{ti} \ll v_{ti}$. In the case of slab gyrokinetics, a consistent calculation of the transport of momentum in the low flow ordering requires the third order Hamiltonian \([33, 34]\). The requirements for a system with general geometry are still to be sorted out, and to do so it is necessary to study the new terms presented here and terms of even higher order.

In addition to the issues raised for full $f$ simulations, the formulation presented here will be very useful for $\delta f$ approaches to momentum transport in tokamaks in the low flow ordering.

\(^4\) In the particular case of a constant magnetic field the calculation to order $\epsilon^2$ was given in \([23]\).
Reference [35] presents a formulation of this problem in the electrostatic limit that requires the minimum number of modifications to existing $\delta f$ simulations. The most important conclusion in [35] is that the turbulent pieces of the distribution function and the electrostatic potential have to be calculated to an order higher in $\epsilon$ than usual because the contribution to momentum transport from the lowest order pieces vanishes due to symmetry arguments [36, 37]. These symmetry arguments do not hold if the higher order terms of the gyrokinetic equation are considered. To avoid calculating most of the next order corrections to the gyrokinetic Vlasov and Poisson’s equations, reference [35] has to resort to a subsidiary expansion based on the fact that in many tokamaks the poloidal component of the magnetic field is much smaller than the toroidal component. The new contribution to the Hamiltonian that we calculate here gives the self-consistent higher order contributions to the gyrokinetic Vlasov equation for the first time. Only a higher order gyrokinetic Poisson’s equation is then lacking to obtain a complete $\delta f$ formulation in the low flow ordering that does not require a small poloidal magnetic field; this will be the subject of a future publication.

The rest of this paper is organized as follows. In section 2 we write the non-dimensional phase-space Lagrangian of a particle in an electromagnetic field. The normalization shows explicitly the standard gyrokinetic ordering (1). In the first part of section 3 we review the phase-space Lagrangian approach to gyrokinetics to help the understanding of the calculation that follows. In the second half of this section we proceed to obtain the gyrokinetic Lagrangian to second order in our expansion parameter $\epsilon$. As mentioned above, this is our main result. In section 4 we obtain the Vlasov equation from this Lagrangian, and in section 5 we discuss the consequences of this formulation for Poisson’s equation. The new Vlasov and Poisson’s equations presented here are correct to second and first order in $\epsilon$, respectively. We should note that in the limit where the electrostatic potential has a scale of variation much larger than the gyroradius of the species of interest, our gyrokinetic equations provide the highest order guiding-center equations that we are aware of. In section 6 we borrow tools from classical field theory to obtain Poisson’s equation in a different way. We prove that there is an energy-like invariant and we discuss the stringent conditions on the equations to actually conserve it in a simulation. We finish with a discussion of our results and the future lines of research in section 7. The appendices contain the most cumbersome parts of the calculation as well as some material included for completeness. Finally, we would like to stress that in this paper we have given all our results in an explicit form that can be directly implemented in a computer code.

2. Normalized Lagrangian

The phase-space Lagrangian for the motion of a particle of mass $M$ and charge $Ze$ in an electromagnetic field is given by

$$L^X(r, v, \dot{r}, \dot{v}, t) = \left[ \frac{Ze}{c} A(r) + Mv \right] \cdot \frac{dr}{dt} - H^X(r, v, t),$$  \hspace{1cm} (2)

with the Hamiltonian

$$H^X(r, v, t) = \frac{1}{2}Mv^2 + Ze\phi(r, t).$$ \hspace{1cm} (3)

Here $A$ is the vector potential that is defined such that $\mathbf{B} = \nabla \times \mathbf{A}$. Note that the phase-space Lagrangian depends on the position of the particle $r$, its velocity $v$, the time derivatives of both the position and the velocity, $\dot{r} = dr/dt$ and $\dot{v} = dv/dt$, and the time $t$. For convenience, we will sometimes denote the phase-space coordinates $[r, v]$ as $[X^a]^0_{a=1} \equiv X = [r, v]$. We use the superscript $X$ in the Lagrangian (2) because it is a function of the phase-space coordinates $X$. 
The equations of motion are obtained by finding the stationary points of the action $\sigma^X[r(t), v(t)] = \int_{t_0}^{t_1} \mathcal{L}^X(r(t), v(t), \dot{r}(t), \dot{v}(t), t)$ with respect to variations of the functions $r(t)$ and $v(t)$ subject to the constraints $r(t = t_0) = r_0$, $v(t = t_0) = v_0$, $r(t = t_1) = r_1$ and $v(t = t_1) = v_1$. From this procedure we obtain six equations of motion, namely

$$\frac{d}{dt}(\nabla_r \mathcal{L}^X) = -\nabla_v \mathcal{L}^X \tag{4}$$

and

$$\frac{d}{dt}(\nabla_v \mathcal{L}^X) = -\nabla_r \mathcal{L}^X \tag{5}.$$  

This differs from the standard Lagrangian formalism where the Lagrangian function depends only on $r$, $\dot{r}$, and $t$. Actually, the phase-space Lagrangian formalism can be viewed as a variational formulation of Hamilton equations (see, for example, [38]). In plasma physics, it was first applied by Littlejohn to guiding-center dynamics in [21].

The Lagrangian (2) is non-dimensionalized using the characteristic thermal velocity of the species of interest $v_\text{th}$, the characteristic length $L^{-1} \sim |\nabla (\ln |A|)|$, and the characteristic time $L/v_\text{th}$. We assume that $\epsilon = \rho/L$ is a small parameter, with $\rho = v_\text{th}/\Omega$ and $\Omega = ZeB_0/Mc$ the gyrofrequency and the characteristic gyrofrequency of the species of interest, and $B_0 \sim |\nabla \times A|$ the characteristic magnitude of the magnetic field. We assume that the characteristic time and length scales in the electrostatic potential are the sound gyroradius $\rho/L$, and the sound characteristic time $L/v_\text{th}$, i.e., $\varphi(r/\rho, c_s t/L)$, where $\varphi$ has derivatives with respect to its arguments of order unity. Here, $c_s = \sqrt{T_\text{ei}/m_i}$ is the sound speed, $m_i$ and $\Omega_i = eB_0/m_i c$ are the mass and the gyrofrequency of the dominant ion species, usually singly charged, and $T_\text{e0}$ is the characteristic electron temperature. The assumption on the scales of the electrostatic potential can be easily relaxed to account for other time and spatial scales. Since the electrostatic potential $\varphi$ is a quantity that enters the equations of the different species, it is normalized using parameters that do not depend on the species, in particular the characteristic electron temperature $T_\text{e0}$, the magnitude of the electron charge $e$ and the mass of the dominant ion species $m_i$. This normalization will be useful in Poisson’s equation, where several species appear. The new, non-dimensionalized variables are

$$\tilde{r} = \frac{v_\text{th} t}{L}, \quad \tilde{r} = \frac{r}{L}, \quad \tilde{v} = \frac{v}{v_\text{th}}, \quad \tilde{A} = \frac{A}{B_0 L}, \quad \tilde{\varphi} = \frac{e\varphi}{\epsilon T_\text{e0}}, \quad \tilde{H} \tilde{X} = \frac{H X}{M v_\text{th}^2},$$

(6)

giving

$$\tilde{L} \tilde{X} (\tilde{r}, \tilde{v}, \dot{\tilde{r}}, \dot{\tilde{v}}, \tilde{t}) = \left[ \frac{1}{\epsilon} \tilde{A} (\tilde{r}) + \tilde{v} \right] \cdot \frac{d\tilde{r}}{dt} - \tilde{H} \tilde{X} (\tilde{r}, \tilde{v}, \tilde{t}), \tag{7}$$

with

$$\tilde{H} \tilde{X} (\tilde{r}, \tilde{v}, \tilde{t}) = \frac{1}{2} \tilde{v}^2 + \Lambda \epsilon \tilde{\varphi} (\tilde{r} / \epsilon t, \tilde{t}). \tag{8}$$

Here, $T_0 = M v_\text{th}^2$ is the characteristic temperature of the species of interest, $\epsilon_s = \rho_s/L$ is the ratio between the sound gyroradius and the characteristic scale length,

$$\lambda = \frac{\rho_s}{\rho} = Z \sqrt{\frac{T_\text{e0} m_i}{T_0 M}},$$

(9)

is the ratio between the sound gyroradius and the gyroradius of the species of interest,

$$\tau = \frac{v_\text{th}}{c_s} = \sqrt{\frac{T_\text{e0} m_i}{T_0 M}} \tag{10}$$
is the ratio between the thermal speed of the species of interest and the sound speed, and

$$\Lambda = \frac{Z T_{\vartheta_0}}{T_0} = Z^2 \left( \frac{T_{\vartheta_0}}{T_0} \right)^{3/2} \sqrt{\frac{m_i}{M}} \quad (11)$$

Even though the electrostatic potential is small, its perpendicular gradient is not. This assumption has been formally implemented by writing $\Lambda \epsilon \tilde{\psi}(\tilde{r}/\lambda \epsilon, \tilde{t}/\tau)$. Here and in what follows we assume $\Lambda \sim \lambda \sim \tau \sim 1$. This is the maximal ordering that contains in it several interesting regimes as subsidiary expansions in $\Lambda$, $\lambda$ and $\tau$. The form $\Lambda \epsilon \tilde{\psi}(\tilde{r}/\lambda \epsilon, \tilde{t}/\tau)$ is, however, somewhat deceiving because the gradients along the magnetic field lines must be small, that is, $\hat{b} \cdot \nabla \tilde{\psi} \sim 1 \ll 1/\lambda \epsilon$, with $\hat{b}(\tilde{r}) : = \hat{B}/\tilde{B}$ the unit vector parallel to the magnetic field. It is possible to formalize this condition by writing the functions in flux coordinates $s(\tilde{r})$, $\psi(\tilde{r})$ and $\alpha(\tilde{r})$ such that $\hat{b} = \partial \tilde{r}/\partial s$ and $\tilde{B} = \nabla \alpha \times \nabla \psi$. In these variables, the potential is given by

$$\tilde{\psi} = \psi(s(\tilde{r}), \psi(\tilde{r})/\lambda \epsilon, \alpha(\tilde{r})/\lambda \epsilon, \tilde{t}/\tau). \quad (12)$$

To simplify the notation, we will often use $\psi(\tilde{r}_s/\lambda \epsilon, \tilde{r}_\perp, \tilde{t}/\tau)$ instead of the most complete expression in (12). Where no confusion is possible, we will write $\psi(\tilde{r}, \tilde{t})$. In any case, we always assume

$$\hat{b}(\tilde{r}) \cdot \nabla \psi(\tilde{r}, \tilde{t}) \sim 1 \quad (13)$$

and

$$\nabla_{\tilde{r}_s} \psi(\tilde{r}, \tilde{t}) : = \hat{b}(\tilde{r}) \times (\nabla_\perp \psi(\tilde{r}, \tilde{t}) \times \hat{b}(\tilde{r})) \sim \frac{1}{\lambda \epsilon}. \quad (14)$$

Note that $\epsilon$ is species-dependent whereas $\epsilon_s = \lambda \epsilon$ is not. In sections 3 and 4 where we compute the gyrokinetic phase-space Lagrangian and the equations of motion of a single species, $\epsilon$ is the natural expansion parameter. However, in sections 5 and 6, devoted to the gyrokinetic Poisson’s equation, we need to consider several different species and $\epsilon_s$ is the appropriate, species-independent small parameter.

Finally, a notational remark is in order. In sections 3 and 4 we will be very careful to exhibit the dependence of our results on the mass, charge and temperature of the species through the parameters $\Lambda$, $\lambda$ and $\tau$. Although at some places this may seem unnecessary and awkward (and it would be if our objective were to treat always a single species), it is very convenient to write Poisson’s equation and the gyrokinetic phase-space Lagrangian for a mixture of species in sections 5 and 6.

From now on we will drop hats in the normalized expressions.

### 3. Phase-space Lagrangian perturbation theory

In this section we follow the general strategy of the applications of Hamiltonian and phase-space Lagrangian techniques to magnetized plasmas [18–25]. We search order by order in the small parameter $\epsilon$ for a change of phase-space variables such that only one of the variables has fast time dependence. The gyrophase $\theta$ is the fast variable that evolves in the gyrofrequency time scale. The rest of the phase-space variables (gyrocenter position $\mathbf{R}$, parallel velocity $u$ and magnetic moment $\mu$) evolve with the much slower characteristic time scale $L/v$. To achieve this, their time derivatives $d\mathbf{R}/dt$, $du/dt$ and $d\mu/dt$ will be made independent of the gyrophase to the order of interest. Had they depended on $\theta$, they would necessarily show rapid time fluctuations on top of the more physically interesting slow time evolution. To make the time derivatives $d\mathbf{R}/dt$, $du/dt$ and $d\mu/dt$ independent of the gyrophase $\theta$, we search for a Lagrangian that does not depend on $\theta$ (it will still depend on its time derivative $d\theta/dt$).
We first review briefly how to perform a change of variables in a phase-space Lagrangian in section 3.1. As we have already announced, the objective is the gyrokinetic Lagrangian to order \( \epsilon^2 \) in general magnetic geometry. Since the calculation is quite long and complicated, we have sketched the derivation in section 3.2 to offer the reader a global perspective of the formalism. This subsection also contains our own proof that the algorithm to find the gyrokinetic change of variables can be carried out to any order and that there exists an adiabatic invariant \( \mu \) to arbitrary order. Finally, in sections 3.3 and 3.4 we address the calculation of the gyrokinetic Lagrangian to second order in detail. Some of the algebra is relegated to appendix B and appendix D. The results to first order are compared with the iterative method in [17] in appendix C.

Before proceeding, we must mention that the phase-space Lagrangian (or Hamiltonian) approach to gyrokinetic theory has been geometrized [22, 25]. We have chosen not to use the language of differential geometry to make the paper accessible to a broader audience, without losing mathematical rigor. The reader familiar with the geometrical tools will realize that every step of our presentation can be translated into that language in an obvious way.

3.1. Transforming to new phase-space variables

Consider a transformation \( T \) that can be time dependent to a new set of gyrokinetic phase-space coordinates \( \{ Z_\alpha \}_{\alpha = 1}^6 \equiv Z \). We write \( X(Z, t) = (r(Z, t), v(Z, t)) = T(Z, t) \). The phase-space Lagrangian (7) can be easily written in the new set of variables by using the chain rule, giving

\[
L^Z(Z, \dot{Z}, t) = \sum_{\alpha = 1}^6 \Gamma_\alpha(Z, t) \frac{dZ_\alpha}{dt} - H^Z(Z, t),
\]

where

\[
\Gamma_\alpha(Z, t) = \left[ \frac{1}{\epsilon} A(r(Z, t)) + v(Z, t) \right] : \frac{\partial r(Z, t)}{\partial Z_\alpha}
\]

and

\[
H^Z(Z, t) = H^X(r(Z, t), v(Z, t), t) - \left[ \frac{1}{\epsilon} A(r(Z, t)) + v(Z, t) \right] : \frac{\partial r(Z, t)}{\partial t}.
\]

By finding the stationary points of the action \( \sigma^Z[Z(t)] = \int_{t_0}^{t_1} L^Z(Z(t), \dot{Z}(t), t) \, dt \) with respect to variations of \( Z(t) \) subject to the conditions \( Z(t = t_0) = Z_0 \) and \( Z(t = t_1) = Z_1 \), we obtain the new equations of motion

\[
\frac{d}{dt} \left( \frac{\partial L^Z}{\partial \dot{Z}_\alpha} \right) = \frac{\partial L^Z}{\partial Z_\alpha}, \quad \alpha = 1, 2, \ldots, 6.
\]

Note that the specific form in (15) implies that the equations of motion can be written as

\[
\sum_{\beta = 1}^6 L_{\alpha\beta} \frac{dZ_\beta}{dt} = \frac{\partial H^Z}{\partial Z_\alpha} + \frac{\partial \Gamma_\alpha}{\partial t}, \quad \alpha = 1, 2, \ldots, 6,
\]

with \( L_{\alpha\beta} \) the \( 6 \times 6 \) antisymmetric matrix

\[
L_{\alpha\beta} = \frac{\partial \Gamma_\beta}{\partial Z_\alpha} - \frac{\partial \Gamma_\alpha}{\partial Z_\beta}.
\]

Although our gyrokinetic change of variables has an explicit time dependence due to the contribution of the electrostatic potential, we will show that it is possible to choose the functions

5 Note that in part of the literature [25] \( T \) stands for the inverse of the transformation that we call \( T \).
such that \( \frac{\partial \Gamma_\alpha}{\partial t} \equiv 0 \). In our derivation we impose then that \( \frac{\partial \Gamma_\alpha}{\partial t} \equiv 0 \). Consequently, we drop the last term in (19) and write the equations of motion as

\[
\sum_{\beta=1}^{6} L_{\alpha\beta} \frac{dZ^\beta}{dt} = \frac{\partial H}{\partial Z^\alpha}, \quad \alpha = 1, 2, \ldots, 6. \tag{21}
\]

From expression (21) we define the Poisson bracket

\[
\{F, G\} = \sum_{\alpha,\beta=1}^{6} P_{\alpha\beta} \frac{\partial F}{\partial Z^\alpha} \frac{\partial G}{\partial Z^\beta}, \tag{22}
\]

with \( P_{\alpha\beta} = (L^{-1})_{\alpha\beta} \) the inverse of the antisymmetric matrix defined in (20). Then

\[
\frac{dZ^\alpha}{dt} = \{Z^\alpha, H\}, \quad \alpha = 1, 2, \ldots, 6. \tag{23}
\]

Noting that \( L_{\alpha\beta} \) satisfies (A.1) with \( n = 3 \), the proof in appendix A guarantees that (22) actually defines a Poisson bracket, i.e. that for any three functions \( F_1, F_2 \) and \( F_3 \), the bracket satisfies skew-symmetry

\[
\{F_1, F_2\} = -\{F_2, F_1\}, \tag{24}
\]

the Leibniz rule

\[
\{F_1, F_2 F_3\} = \{F_1, F_2\} F_3 + \{F_1, F_3\} F_2, \tag{25}
\]

and the Jacobi identity

\[
\{F_1, \{F_2, F_3\}\} + \{F_3, \{F_1, F_2\}\} + \{F_2, \{F_3, F_1\}\} = 0. \tag{26}
\]

In general, it is impossible to find a change of phase-space variables that makes the Lagrangian as written in (15) independent of gyrophase. However, the time derivatives \( \frac{dZ^\alpha}{dt} \) may be gyrophase independent even if the phase-space Lagrangian is not. This apparent discrepancy is easily solved considering that the equations of motion remain the same if instead of the Lagrangian (15) we employ

\[
\mathcal{L}(Z, \dot{Z}, t) = L^Z(Z, \dot{Z}, t) + \frac{dS}{dt}, \tag{27}
\]

where the function \( S(Z, t) \) depends on the phase-space variables \( Z \) and \( t \). Indeed, finding the stationary points of the action

\[
\bar{\sigma}[Z(t)] = \int_{t_0}^{t_1} \mathcal{L}(Z, \dot{Z}, t) \, dt = \int_{t_0}^{t_1} \mathcal{L}^Z(Z, \dot{Z}, t) \, dt + S(Z_1, t_1) - S(Z_0, t_0) \tag{28}
\]

with respect to variations of \( Z(t) \) subject to the conditions \( Z(t = t_0) = Z_0 \) and \( Z(t = t_1) = Z_1 \) gives the same equations of motion as finding the stationary points of the action \( \sigma^Z[Z(t)] \) because both actions differ only by terms that are held constant. Since the Lagrangian is not unique, we are not going to search for new phase-space variables \( Z \) such that the phase-space Lagrangian \( L^Z \) in (15) is gyrophase independent, but such that there exists a function \( S \) for which the phase-space Lagrangian \( \mathcal{L}^Z \) in (27) is gyrophase independent. This is equivalent to requiring that the time derivatives \( \frac{dZ^\alpha}{dt} \) be gyrophase independent. Thus, we are searching for both the change of variables \( Z \) and the function \( S \) such that the Lagrangian \( \mathcal{L} \) in (27) is gyrophase independent. Explicitly,

\[
\mathcal{L}(Z, \dot{Z}, t) = \sum_{\alpha=1}^{6} F_\alpha \frac{dZ^\alpha}{dt} - H(Z, t), \tag{29}
\]

\[6 \text{ In section 4 and appendix E we show that the matrix } L_{\alpha\beta} \text{ of our particular problem is indeed invertible.} \]
where
\[ \Gamma_{\alpha}(Z, t) = \left[ -\frac{1}{\epsilon} A(r(Z, t)) + v(Z, t) \right] \frac{\partial r(Z, t)}{\partial Z^\alpha} + \frac{\partial S(Z, t)}{\partial Z^\alpha} \] (30)
and
\[ H(Z, t) = H^X(r(Z, t), v(Z, t), t) - \left[ -\frac{1}{\epsilon} A(r(Z, t)) + v(Z, t) \right] \frac{\partial r(Z, t)}{\partial t} - \frac{\partial S(Z, t)}{\partial t} \] (31)
are gyrophase independent. In what follows, \( Z \) only refers to the gyrokinetic phase-space coordinates. Note that equations (19), (20), (21), (22) and (23) are valid for the new Lagrangian \( \overline{H} \). Simply replace \( \Gamma_{\alpha} \) by \( \Gamma_{\alpha} \) and \( H^Z \) by \( H \).

3.2. Obtaining the new gyrokinetic variables

As advanced in the introduction, the detailed computation to obtain the final form of the gyrokinetic Lagrangian to order \( \epsilon^2 \) is rather involved. This is why we devote this subsection to schematically show the steps leading to the determination of the change of variables and the function \( S \) order by order. We also give an easy proof that the algorithm can be carried out up to arbitrary order (although in practice the computations would become prohibitively difficult). The detailed calculation to second order is done in sections 3.3 and 3.4.

Our transformation to new phase-space coordinates will be denoted by \( T_{\epsilon} \),
\[ (r, v) = T_{\epsilon}(R, u, \mu, \theta, t) = T_{NP, \epsilon} T_{NP}(R, u, \mu, \theta, t), \] (32)
where \( R \) is the gyrocenter position, \( u \) is the gyrocenter parallel velocity, \( \mu \) is the magnetic moment and \( \theta \) is the gyrophase. For convenience, we have written the transformation \( T_{NP} \) as the composition of two other transformations that we call non-perturbative transformation \( T_{NP, \epsilon} \) and perturbative transformation \( T_{NP} \).

First, we perform a non-perturbative change of coordinates
\[ (r, v) = T_{NP, \epsilon}(Z_\epsilon) = T_{NP, \epsilon}(R_g, v_{||}, \mu_g, \theta_g), \] (33)
where \( R_g, v_{||}, \mu_g \) and \( \theta_g \) are lowest order approximations to the gyrocenter position, parallel velocity, magnetic moment and gyrophase. Their detailed definitions can be found in section 3.3. For simplicity, we will sometimes use the notation \( \{Z_\epsilon\}_\alpha^{\#}_{\text{local}} \equiv Z_g = \{R_g, v_{||}, \mu_g, \theta_g\} \). Physically, the non-perturbative change of variables in section 3.3 amounts to saying that to lowest order the gyromotion is circular. This is a consequence of the smallness of the gyroradius and the fact that the corrections due to the electrostatic potential are of next order. After the change of coordinates we add the total time derivative of a function \( S_{NP}(R_g, \mu_g, \theta_g) \) to the Lagrangian. The details of the calculation are given in section 3.3. The final result is
\[ L_{Z_g} = \left[ -\frac{1}{\epsilon} A(R_g) + v_{||} B(R_g) + \epsilon \Gamma^{(1)}_{R} + \epsilon^2 \Gamma^{(2)}_{R} + \cdots \right] \frac{dR_g}{dt} + \left[ -\mu_g + \epsilon \Gamma^{(1)}_{\mu} + \epsilon^2 \Gamma^{(2)}_{\mu} + \cdots \right] \frac{d\theta_g}{dt} = H^{(0)}(R_g, v_{||}, \mu_g) - \epsilon H^{(1)}, \] (34)
where
\[ H^{(0)}(R_g, v_{||}, \mu_g) = \frac{1}{2} v_{||}^2 + \mu_g B(R_g) \] (35)
and the rest of the terms are defined in section 3.3. Note that with the notation in (34) we have made explicit the fact that \( d\theta_g/dt \sim \epsilon^{-1} \), i.e. its time variation is of the order of the
gyrofrequency time scale and hence much faster than the evolution of the rest of the phase-space variables. The change of variables is non-perturbative and will give contributions to all orders in $\epsilon$. All the higher order terms $L_R^{(1)}(R, v_{||}, \mu_g, \theta_g)$, $L_R^{(2)}(R, v_{||}, \mu_g, \theta_g)$, ... $\Gamma^g$ \cite{19}, $\Gamma^g$ \cite{20}, $\Gamma^g$ \cite{21}, $H^{(1)}(R_{||}/\epsilon, R_{||}, \mu_g, \theta_g, t)$ depend on gyrophase. In the Lagrangian \cite{21}, the Hamiltonian has only the first order correction $H^{(1)}(R_{\perp}/\epsilon, R_{\perp}, \mu_g, \theta_g, t)$, with the higher order corrections being exactly zero. This correction $H^{(1)}(R_{\perp}/\epsilon, R_{\perp}, \mu_g, \theta_g, t)$ is the only term in the Lagrangian that has strong perpendicular gradients because it is the only contribution that depends on the electrostatic potential.

The gyrophase dependence in the Lagrangian \cite{21} must be eliminated with the definition of the gyrokinetic variables order by order. Since we only calculate the gyrokinetic variables order by order, we truncate the expansion in $\epsilon$ to the order of interest as

$$L^{(2)} = \left[ \frac{1}{\epsilon} A(R_g) + v_{||} \dot{R}_g + \sum_{i=1}^{n} \epsilon^{i+1} \Gamma^g \frac{dR_g}{dt} \right] \frac{dR_g}{dt}$$

(36)

Here, we have not written explicitly $d\theta_g/dt$ (or $d\epsilon/\epsilon$). Instead, we keep the terms that are multiplying $d\theta_g/dt$ to higher order in $\epsilon$. The notation $O(\epsilon^{n+1}, \epsilon^{n+2})$ will be extensively used in this paper and indicates that the terms of order $\epsilon^{n+1}$ that we have neglected are either proportional to $dR_g/dt$ or are in the Hamiltonian, and that the terms of order $\epsilon^{n+2}$ that we have neglected are proportional to $d\mu_g/dt$, $d\epsilon_g/dt$ and $d\theta_g/dt$ (the Lagrangian (36) does not contain terms proportional to $d\epsilon_g/dt$ or $d\mu_g/dt$). When we perform the expansion to obtain the gyrokinetic variables order by order, it will be apparent that this notation is convenient because we need to keep some terms to $O(\epsilon^n)$ and the rest to $O(\epsilon^{n+1})$.

Employing expression (36), we find the gyrokinetic variables by eliminating the gyrophase dependence order by order. For now, it is enough to simply write the transformation

$$(R_g, v_{||}, \mu_g, \theta_g) = T_{F, \epsilon}(Z, t) = T_{F, \epsilon}(R, u, \mu, \theta, t)$$

(37)

to the order of interest as

$$R_g = R + \sum_{i=1}^{n} \epsilon^{i+1} \tilde{R}_{i+1}, \quad v_{||} = u + \sum_{i=1}^{n} \epsilon^{i+1} \tilde{u}_i, \quad \mu_g = \mu + \sum_{i=1}^{n} \epsilon^{i+1} \tilde{\mu}_i, \quad \theta_g = \theta + \sum_{i=1}^{n} \epsilon^{i+1} \tilde{\theta}_i.$$  

(38)

In section 3.4 we show the connection of this expansion with Lie transforms. Sometimes we will use the abbreviated notation $[Z^{a}]_{a=1}^{6} \equiv Z = \{\tilde{R}, \tilde{u}, \tilde{\mu}, \tilde{\theta}\}$. Note that the variable $\tilde{R}$ is found to an order higher than the rest. The corrections $\tilde{R}_{i+1}$, $\tilde{u}_i$, $\tilde{\mu}_i$ and $\tilde{\theta}_i$ are obtained by imposing that the Lagrangian is gyrophase independent up to terms of order $O(\epsilon^n, \epsilon^{n+1})$, where the terms of order $\epsilon^n$ are in the Hamiltonian or are terms proportional to $dR/\epsilon dt$, and the terms of order $\epsilon^{n+1}$ are terms proportional to $d\mu/d\epsilon$, $d\epsilon/d\mu$ and $d\theta/d\epsilon$.

To demonstrate the procedure, we show schematically how to obtain the first corrections $\tilde{R}_2$, $\tilde{u}_1$, $\tilde{\mu}_1$ and $\tilde{\theta}_1$. We then argue that the same formalism can be extended to arbitrary order. The proof presented here demonstrates that the gyrokinetic variables can be consistently calculated order by order without running into problems. Recently, this has been a controversial issue \cite{39–41}. 

10
To calculate $\bar{R}_2$, $\bar{u}_1$, $\tilde{\mu}_1$ and $\tilde{\theta}_1$, we need to express the Lagrangian (36) in the new gyrokinetic variables to $O(\epsilon, \epsilon^3)$, giving

$$
\mathcal{L}^Z = \left[ \frac{1}{\epsilon} A(R) + u \dot{b}(R) + \epsilon (\bar{R}_2 \cdot \nabla R A(R) + \bar{u}_1 \dot{b}(R) + \Gamma^{(1)}_{\mu}(R, u, \mu, \theta)) \right] \cdot \frac{dR}{dt} + \left[ \epsilon^2 u \dot{b}(R) \right] \cdot \frac{d\dot{R}_2}{dt} + \left[ -\epsilon \mu + \epsilon^2 (-\tilde{\mu}_1 + \Gamma^{(1)}_{\theta}(R, u, \mu, \theta)) \right] \frac{d\theta}{dt} = \epsilon^2 \mu \frac{d\tilde{\theta}}{dt} - H^{(0)}(R, u, \mu) - \epsilon [u \bar{u}_1 + \tilde{\mu}_1 B(R) + H^{(1)}(R, u, \mu, \theta, t)] + O(\epsilon^2, \epsilon^3).
$$

(39)

Note that in the functions $\Gamma^{(1)}_{\mu}(R, \mu, \theta)$, $\Gamma^{(1)}_{\theta}(R, \mu, \theta)$, $H^{(0)}(R, \mu, \theta)$ and $H^{(1)}(R, \mu, \theta)$ the variables $R$, $\mu$, $\theta$ are replaced by the gyrokinetic variables $R$, $\mu$, and $\theta$. For example, from the definition of $H^{(0)}$ in (35), we find

$$
H^{(0)}(R, \mu, \theta) = \frac{1}{2} u^2 + \mu B(R).
$$

(40)

The correction $u \bar{u}_1 + \tilde{\mu}_1 B(R)$ to the Hamiltonian contains both the correction $H^{(1)}$ and the result of Taylor expanding $H^{(0)}(R, \mu, \theta) = H^{(0)}(R + \epsilon \dot{R}_2 + \cdots, u + \epsilon \bar{u}_1 + \cdots, \mu + \epsilon \tilde{\mu}_1 + \cdots)$ around $R$, $u$, and $\mu$, i.e.

$$
\bar{u}_1 \frac{dH^{(0)}}{du} + \tilde{\mu}_1 \frac{dH^{(0)}}{d\mu} + H^{(1)}(R, \mu, \theta, t) = u \bar{u}_1 + \tilde{\mu}_1 B(R) + H^{(1)}(R, \mu, \theta, t).
$$

(41)

In section 3.4 we show that the expansion around $R$ gives a term that is an order higher and can be ignored to this order.

As indicated in section 3.1, we can always add a time derivative to the Lagrangian (39). By doing so, we will get a form of the Lagrangian in which the derivatives of a function $S^{(2)}_{\mu}(R, u, \mu, \theta, t)$ enter. By imposing that the Lagrangian is of a specific form, we first obtain the corrections $\bar{R}_2, \bar{u}_1, \tilde{\mu}_1$, and $\tilde{\theta}_1$ as functions of $S^{(2)}_{\mu}$ and then we find $S^{(2)}_{\mu}$ by integrating a simple differential equation. We add to the Lagrangian (39) the total time derivative

$$
\frac{d}{dt} \left[ \epsilon^2 S^{(2)}_{\mu}(R, u, \mu, \theta, t) - \epsilon A(R) \cdot \dot{R}_2 - \epsilon^2 u \dot{b}(R) \cdot \dot{R}_2 + \epsilon^2 \mu \dot{\tilde{\theta}}_1 \right],
$$

(42)

where we have taken into account that $S^{(2)}_{\mu}$ depends on $R$ in two different ways: a fast dependence due to the potential that has wavelengths on the order of the gyroradius, and a slow dependence due to the background magnetic field. Note that adding the time derivative (42) to the Lagrangian (39) eliminates all the terms proportional to the time derivatives of the corrections $\bar{R}_2, \bar{u}_1, \tilde{\mu}_1$, and $\tilde{\theta}_1$, giving

$$
\mathcal{L} = \left[ \frac{1}{\epsilon} A(R) + u \dot{b}(R) + \epsilon \left( \bar{R}_2 \cdot \nabla R A(R) + \bar{u}_1 \dot{b}(R) + \Gamma^{(1)}_{\mu}(R, u, \mu, \theta) \right) \right] \cdot \frac{dR}{dt} + \left[ \epsilon^2 u \dot{b}(R) \right] \cdot \frac{d\dot{R}_2}{dt} + \left[ -\epsilon \mu + \epsilon^2 (-\tilde{\mu}_1 + \Gamma^{(1)}_{\theta}(R, u, \mu, \theta)) \right] \frac{d\theta}{dt} - \epsilon [u \bar{u}_1 + \tilde{\mu}_1 B(R) + H^{(1)}(R, u, \mu, \theta, t)] + O(\epsilon^2, \epsilon^3).
$$

(43)

$^8$ Observe that in principle, the term $\epsilon^2 A(R) \cdot d\dot{R}_1/dt$ should be included in $\mathcal{L}^Z$ (39). However, adding the time derivative of $-\epsilon^2 A(R) \cdot \dot{R}_1$ eliminates this order. The same can be said about (55) and (114), where the terms $\epsilon^2 A(R) \cdot d\dot{R}_2/dt$ and $\epsilon^2 A(R) \cdot d\dot{R}_3/dt$ could be added. It is very easy to show that the algorithm gives, of course, exactly the same results with either choice.
where we have used
\[ \hat{R}_2 \cdot (\nabla_R A - \nabla_R A) - \hat{R}_2 = (\nabla_R \times A) \times \hat{R}_2 = B \times \hat{R}_2, \] (44)

trivially deduced from the identity
\[ [\nabla_R A - (\nabla_R A)^T]_{ij} = \frac{\partial A_k}{\partial R_i} - \frac{\partial A_j}{\partial R_i} = \sum_{k=1}^3 \delta_{ijk}(\nabla_R \times A)_k. \] (45)

Here \( \delta_{ijk} \) is the Levi–Civita symbol and the superscript \( ^T \) stands for matrix transposition. Note that in equation (43) we have not made explicit the dependence of the functions on the gyrokinetic variables, but it is assumed that all the terms are functions of \( R, u, \mu \) and \( \theta \). By imposing that the Lagrangian (43) be equal to
\[ \mathcal{L} = \left( \frac{1}{\epsilon} + u \hat{b} + \epsilon \Gamma^{(1)}_R \right) \cdot \frac{dR}{dt} - \epsilon \mu \frac{d\theta}{dt} - H^{(0)} - \epsilon H^{(1)} + O(\epsilon^2, \epsilon^3), \] (46)

where \( \Gamma^{(1)}_R, H^{(0)} := H^{(0)}(R, u, \mu) \) and \( H^{(1)} \) are gyrophase independent, we obtain the equations
\[ \hat{R}_2 = \frac{\partial S_p^{(2)}}{\partial u} \hat{b}(R) + \frac{1}{B(R)} \hat{b}(R) \times \left[ \Gamma^{(1)}_R(R, u, \mu, \theta) - \Gamma^{(1)}_R + \nabla_{(R, /\epsilon)} S^{(2)}_P \right], \] (47)
\[ \tilde{u}_1 = \hat{b}(R) \cdot \left[ \Gamma^{(1)}_R(R, u, \mu, \theta) \right], \] (48)
\[ \tilde{\mu}_1 = \Gamma^{(1)}_\theta(R, u, \mu, \theta) + \frac{\partial S^{(2)}_P}{\partial \theta}, \] (49)
and
\[ \tilde{\theta}_1 = -\frac{\partial S^{(2)}_P}{\partial \mu}. \] (50)

The corrections \( \hat{R}_2, \tilde{u}_1, \tilde{\mu}_1 \) and \( \tilde{\theta}_1 \) can then be found if we obtain \( S^{(2)}_p \). To do so, we use that
\[ u \tilde{u}_1 + b(R) \tilde{\mu}_1 + H^{(1)}(R, /\epsilon, R, \mu, \theta, t) = \tilde{H}^{(1)}. \] (51)

First, we take the gyroaverage of this equation to obtain \( \tilde{H}^{(1)} \). Using the results in equations (48) and (49), the gyroaverage of equation (51) becomes
\[ \tilde{H}^{(1)} = u \hat{b}(R) \cdot \left[ \Gamma^{(1)}_R(R, u, \mu, \theta) \right] + B(R) \left[ \Gamma^{(1)}_\theta(R, u, \mu, \theta) \right] \]
\[ + \left\{ H^{(1)}(R, /\epsilon, R, \mu, \theta, t) \right\}, \] (52)

where \( \langle \ldots \rangle = (2\pi)^{-1} \int d\theta \langle \ldots \rangle \) is the gyroaverage holding \( R, u \) and \( \mu \) fixed. Note that we have been able to obtain \( \tilde{H}^{(1)} \) without explicitly finding \( S^{(2)}_p \). Solving for \( S^{(2)}_p \) is easy once \( \tilde{H}^{(1)} \) is known. Combining equations (51) and (52), we find
\[ \frac{\partial S^{(2)}_P}{\partial \theta} = \frac{u}{B(R)} \hat{b}(R) \cdot \left[ \Gamma^{(1)}_R(R, u, \mu, \theta) - \langle \Gamma^{(1)}_R(R, u, \mu, \theta) \rangle \right] \]
\[ -\langle \Gamma^{(1)}_\theta(R, u, \mu, \theta) - \langle \Gamma^{(1)}_\theta(R, u, \mu, \theta) \rangle \rangle \]
\[ - \frac{1}{B(R)} \left\{ H^{(1)}(R, /\epsilon, R, \mu, \theta, t) - \langle H^{(1)}(R, /\epsilon, R, \mu, \theta, t) \rangle \right\}. \] (53)

It is then straightforward to obtain \( S^{(2)}_p \) by integrating in the gyrophase \( \theta \). The specific calculation of the corrections \( \hat{R}_2, \tilde{u}_1, \tilde{\mu}_1, \tilde{\theta}_1 \) and \( \tilde{H}^{(1)} \) is done in section 3.4. Here we only want
to demonstrate the procedure. Note that the final solution depends on our choice of \( \Gamma_{R}^{(1)} \) and in general on our choice of the final expression for \( \mathcal{Z} \). However, once the choice is made, \( \mathcal{H}^{(1)} \) is completely determined.

Once \( \hat{R}_{n+1}, \hat{u}_{n}, \hat{\mu}_{n}, \hat{\theta}_{n} \) and \( S_{p}^{(2)} \) have been determined, we can prove by induction that the procedure gives the corrections to any order. Assume that the corrections to the phase-space Lagrangian to order \( O(\epsilon^{n}, \epsilon^{n+1}) \) have been obtained up to \( i = n - 1 \) in such a way that the Lagrangian is of the form

\[
\mathcal{L} = \left[ \frac{1}{\epsilon} A(R) + u \dot{b}(R) + \sum_{i=1}^{n-1} \epsilon^{i} \mathcal{T}_{R}^{(i)} \right] \cdot \frac{dR}{dt} - \epsilon \mu \frac{dH}{dt} - \mathcal{H}^{(0)} - \sum_{i=1}^{n-1} \epsilon^{i} \mathcal{H}^{(i)} + O(\epsilon^{n}, \epsilon^{n+1}).
\]  

(54)

Then it is possible to obtain the corrections \( \hat{R}_{n+1}, \hat{u}_{n}, \hat{\mu}_{n} \) and \( \hat{\theta}_{n} \), the function \( S_{p}^{(n+1)} \) and the phase-space Lagrangian to order \( O(\epsilon^{n}, \epsilon^{n+1}) \). To do so, the Lagrangian is written to \( O(\epsilon^{n}, \epsilon^{n+1}) \). In general, we cannot guess the exact form of the Lagrangian to such high order without doing the calculation order by order, but the terms that contain the corrections \( \hat{R}_{n+1}, \hat{u}_{n}, \hat{\mu}_{n} \) and \( \hat{\theta}_{n} \) are very easy to obtain. We find that

\[
\mathcal{L}^{Z} + \sum_{i=1}^{n-1} \frac{d}{dt} \left( \epsilon^{i+1} S_{p}^{(i+1)} - \epsilon^{i} A \cdot \hat{R}_{n+1} - \epsilon^{i} u \dot{b} \cdot \hat{R}_{n+1} + \epsilon^{i+1} \mu \hat{\theta}_{n} \right)
\]

\[
= \left[ \frac{1}{\epsilon} A(R) + u \dot{b}(R) + \sum_{i=1}^{n-1} \epsilon^{i} \mathcal{T}_{R}^{(i)} + \epsilon^{n} (\hat{R}_{n+1} \cdot \nabla_{R} A(R) + \hat{u}_{n} \dot{b}(R) + \hat{\Gamma}_{n+1}^{(n)}) \right] \cdot \frac{dR}{dt} 
\]

\[
+ [\epsilon^{n} A(R) + \epsilon^{n+1} u \dot{b}(R)] \cdot \frac{d\hat{R}_{n+1}}{dt} + \epsilon^{n+1} \hat{\Gamma}_{n}^{(n)} \frac{d\mu}{dt} + \epsilon^{n+1} \hat{\Gamma}_{n}^{(n)} \frac{d\mu}{dt} 
\]

\[
+ \left[ -\epsilon \mu + \epsilon^{n+1} (-\hat{\mu}_{n} + \hat{\Gamma}_{n}^{(n)}) \right] \frac{d\theta}{dt} - \epsilon^{n+1} \mu \frac{d\theta}{dt} - \mathcal{H}^{(0)} - \sum_{i=1}^{n-1} \epsilon^{i} \mathcal{H}^{(i)} 
\]

\[
- \epsilon^{n} [u \hat{\theta}_{n} + \hat{\mu}_{n} B(R) + \hat{H}^{(n)}] + O(\epsilon^{n+1}, \epsilon^{n+2}).
\]  

(55)

Here we have just separated the terms of order \( O(\epsilon^{n}, \epsilon^{n+1}) \) into those that depend on the corrections \( \hat{R}_{n+1}, \hat{u}_{n}, \hat{\mu}_{n} \) and \( \hat{\theta}_{n} \), and the rest that we have lumped into the terms \( \hat{\Gamma}_{n}^{(n)}, \hat{\Gamma}_{\mu}^{(n)}, \hat{\Gamma}_{\theta}^{(n)} \) and \( \hat{H}^{(n)} \).

The form of the Lagrangian (55) is very similar to the Lagrangian (39). We can then use the same procedure. We add the time derivative

\[
\frac{d}{dt} \left[ \epsilon^{n+1} S_{p}^{(n+1)} \right] = \left( \frac{d}{dt}/\epsilon, R, u, \mu, \theta, t \right) - \epsilon^{n} A(R) \cdot \hat{R}_{n+1} - \epsilon^{n+1} u \dot{b}(R) \cdot \hat{R}_{n+1} + \epsilon^{n+1} \mu \hat{\theta}_{n}
\]  

(56)

to cancel all the terms that are proportional to the time derivatives of the corrections \( \hat{R}_{n+1}, \hat{u}_{n}, \hat{\mu}_{n} \) and \( \hat{\theta}_{n} \). Then, by imposing that the Lagrangian (55) plus this time derivative be equal to

\[
\mathcal{Z} = \left( \frac{1}{\epsilon} A(R) + u \dot{b}(R) + \sum_{i=1}^{n} \epsilon^{i} \mathcal{T}_{R}^{(i)} \right) \cdot \frac{dR}{dt} - \epsilon \mu \frac{dH}{dt} - \mathcal{H}^{(0)} - \sum_{i=1}^{n} \epsilon^{i} \mathcal{H}^{(i)} + O(\epsilon^{n+1}, \epsilon^{n+2}),
\]  

(57)
we obtain the equations
\[
\hat{R}_{n+1} = \left( \hat{\Gamma}_{n}^{(n)} + \frac{\partial S_{p}^{(n+1)}}{\partial u} \right) \hat{b}(R) + \frac{1}{B(R)} \hat{b}(R) \times (\hat{\Gamma}_{R}^{(n)} - \hat{\Gamma}_{R}^{(n)} + \nabla (R_{\perp}/\epsilon) S_{p}^{(n+1)}),
\]
(58)
\[
\hat{u}_{n} = \hat{b}(R) \cdot (\tilde{\Gamma}_{R}^{(n)} - \hat{\Gamma}_{R}^{(n)}),
\]
(59)
\[
\tilde{\mu}_{n} = \tilde{\Gamma}_{\mu}^{(n)} + \frac{\partial S_{p}^{(n+1)}}{\partial \theta} 
\]
(60)
and
\[
\tilde{\theta}_{n} = - \tilde{\Gamma}_{\theta}^{(n)} - \frac{\partial S_{p}^{(n+1)}}{\partial \mu}.
\]
(61)

The $n$th correction to the Hamiltonian becomes
\[
H^{(n)} = u \hat{b}(R) \cdot (\tilde{\Gamma}_{R}^{(n)} - \langle \tilde{\Gamma}_{R}^{(n)} \rangle) + B(R) \langle \tilde{\Gamma}_{\theta}^{(n)} \rangle + \langle \tilde{H}^{(n)} \rangle,
\]
(62)
and the equation for $S_{p}^{(n+1)}$ is
\[
\frac{\partial S_{p}^{(n+1)}}{\partial \theta} = \frac{u}{B(R)} \hat{b}(R) \cdot (\tilde{\Gamma}_{R}^{(n)} - \langle \tilde{\Gamma}_{R}^{(n)} \rangle) - (\tilde{\Gamma}_{\theta}^{(n)} - \langle \tilde{\Gamma}_{\theta}^{(n)} \rangle) - \frac{1}{B(R)} (\tilde{H}^{(n)} - \langle \tilde{H}^{(n)} \rangle).
\]
(63)

In sections 3.3 and 3.4 we obtain the phase-space Lagrangian to $O(\epsilon^2, \epsilon^3)$. Specifically, in section 3.3 and appendix B we derive equation (34). In section 3.4 and appendix D we use the perturbation procedure explained here to go from equation (34) to the final result.

3.3. Non-perturbative change of variables

We perform a change of variables $(r, v) = T_{NP,\epsilon}(Z_{g}) = T_{NP,\epsilon}(R_{g}, v_{||g}, \mu_{g}, \theta_{g})$ defined by
\[
r = R_{g} + \epsilon \rho(R_{g}, \mu_{g}, \theta_{g})
\]
(64)
and
\[
v = v_{||g} \hat{b}(R_{g}) + \rho(R_{g}, \mu_{g}, \theta_{g}) \times B(R_{g}).
\]
(65)

with the gyroradius vector defined as
\[
\rho(R_{g}, \mu_{g}, \theta_{g}) = - \sqrt{\frac{2\mu_{g}}{B(R_{g})}} \left[ \sin \theta_{g} \hat{e}_{1}(R_{g}) - \cos \theta_{g} \hat{e}_{2}(R_{g}) \right].
\]
(66)

The unit vectors $\hat{e}_{1}(r)$ and $\hat{e}_{2}(r)$ are orthogonal to each other and to $\hat{b} = B/B$, and satisfy $\hat{e}_{1} \times \hat{e}_{2} = \hat{b}$ at every location $r$. Physically, $R_{g}$ is the guiding-center position, $v_{||g}$ the velocity parallel to the magnetic field at the guiding-center position, $\mu_{g}$ the lowest order magnetic moment, and $\theta_{g}$ the lowest order gyrophase. For a homogeneous static magnetic field and in the absence of electric field, the change of coordinates defined by (64), (65) and (66) exactly eliminates the gyrophase dependence. Note in passing that it is a well defined change of coordinates. These formulae explicitly give $(r, v)$ as a function of $Z_{g} = \{R_{g}, v_{||g}, \mu_{g}, \theta_{g}\}$ and it is easy to see that the transformation is invertible for small $\epsilon$: it is clearly invertible for $\epsilon = 0$ and the transformation is continuous in $\epsilon$.

We substitute the relations $X(Z_{g})$, given in (64) and (65), and
\[
\frac{dX^{u}}{dt}(Z_{g}, \dot{Z}_{g}) = \sum_{\beta=1}^{6} \frac{\partial X^{u}(Z_{g})}{\partial Z_{\beta}^{\gamma}} \dot{Z}_{\gamma}^{\beta}, \quad \alpha = 1, 2, \ldots, 6
\]
(67)
into the non-dimensionalized Lagrangian (7), $\mathcal{L}^\Sigma(X(X_\epsilon, X(Z_g), Z_g), t)$. The resulting Lagrangian that we denote as $\mathcal{L}^\Sigma(X(X_\epsilon, X(Z_g), Z_g), t)$ differs from the Lagrangian $\mathcal{L}^\Sigma$ in (34) and (85) by the time derivative of a function $\Sigma_{NP}$ and even though both Lagrangians give the same equations of motion, we have decided to stress the difference. The Lagrangian $\mathcal{L}^\Sigma(X(X_\epsilon, X(Z_g), Z_g), t)$ is

$$\mathcal{L}^\Sigma(X(X_\epsilon, X(Z_g), Z_g), t) = \left[ -1 \frac{A(R_g + \epsilon \rho)}{\lambda \epsilon} + v_{\| g} \hat{b}_g + \rho \times B_g \right] \cdot \frac{d}{dt} (R_g + \epsilon \rho)$$

with $H^{(1)}(R_g, v_{\| g}, \mu_g)$ defined in (35) and

$$H^{(1)}(R_g, \mu_g, \phi_g) = \Lambda \langle \phi(R_g, \lambda \epsilon, \mu_g) \rangle$$

where $\Theta(x)$ is the Heaviside step function, with $\Theta(x) = 1$ for $x > 0$ and $\Theta(x) = 0$ for $x < 0$. Any magnetic quantity with subindex $g$ is evaluated at $R_g$, e.g., $B_g := B(R_g)$. We write $\rho \equiv \rho(R_g, \mu_g, \phi_g)$ when no confusion is possible.

We have defined a new function $\phi(R_g, \mu_g, \phi_g, t)$ (note the difference in the font between $\phi$ and $\varphi$) given by

$$\phi(R_g, \mu_g, \phi_g, t) := \phi(R_g + \epsilon \rho(R_g, \mu_g, \phi_g), t).$$

Then $\langle \phi \rangle$ is the gyroaverage of $\phi$,

$$\langle \phi \rangle(R_g) = \frac{1}{2\pi} \int_{0}^{2\pi} d\theta_g \phi(R_g, \mu_g, \phi_g, t),$$

and $\tilde{\phi}$ the gyrophase-dependent piece,

$$\tilde{\phi}(R_g, \mu_g, \phi_g, t) = \phi(R_g, \mu_g, \phi_g, t) - \langle \phi \rangle(R_g, \mu_g, t).$$

We now prove that the notation in (69),

$$\phi(R_g, \mu_g, \phi_g, t) = \phi(R_g, \mu_g, \phi_g, t) - \langle \phi \rangle(R_g, \mu_g, t),$$

is appropriate. First, we show that $\mu_g$ is always divided by $\lambda^2$ and that the sign of $\lambda$ determines the phase of $\phi_g$ and later we demonstrate that if conditions (13) and (14) are satisfied, then

$$\tilde{b}(R_g) \cdot \nabla_R \phi(R_g, \mu_g, \phi_g, t) \sim 1$$

and

$$\nabla_{R_g} \phi(R_g, \mu_g, \phi_g, t) := \tilde{b}(R_g) \times (\nabla_{R_g} \tilde{\phi}(R_g, \mu_g, \phi_g, t) \times \tilde{b}(R_g)) \sim \frac{1}{\lambda \epsilon}$$

are also satisfied. To show that $\mu_g$ always appears divided by $\lambda^2$ and that we need to add $\pi$ to $\phi_g$ when $\lambda$ is negative, it is enough to realize that $\phi$ depends on $\mu_g$ and $\phi_g$ through the dependence of $\varphi$ on $r/\lambda \epsilon = R_g/\lambda \epsilon + \rho(R_g, \mu_g, \phi_g)/\lambda$ and that $\rho$ as defined in (66) only depends on $\mu_g$ through the multiplying term $\sqrt{\mu_g}$. It is then obvious that $r/\lambda \epsilon = R_g/\lambda \epsilon + \rho(R_g, \mu_g, \phi_g)/\lambda + \pi \Theta(-\lambda)$. To prove that (13) and (14) imply (74) and (75), we employ

$$\nabla_{R_g} \phi(R_g, \mu_g, \phi_g, t) = \nabla_{R_g} \phi(R_g + \epsilon \rho(R_g, \mu_g, \phi_g), t)$$

$$= \nabla_{\varphi} \varphi(r, t) + \epsilon \nabla_{R_g} \rho(R_g, \mu_g, \phi_g) \cdot \nabla_{\varphi} \varphi(r, t),$$

with $\epsilon \nabla_{R_g} \rho \cdot \nabla_{\varphi} \varphi \sim \epsilon \nabla_{R_g} \rho \cdot \nabla_{\varphi} \varphi \sim \lambda^{-1} \sim 1$. Then, using equations (13) and (14), it is easy to see that equations (74) and (75) are correct. Note that when we write $H^{(1)}(R_g, \mu_g, \phi_g, t)$ in (69) we are emphasizing the dependence on $\epsilon$ because the
asymptotic procedure is based on expanding in $\epsilon \ll 1$. The dependence on $\Lambda$, $\lambda$ and $\tau$ is only written explicitly in the function $\phi$.

We now show how to simplify (68). Employing

$$\nabla R_{e} \rho = -\frac{\nabla R_{e} B_{x}}{2 B_{x}} \rho - (\nabla R_{e} \hat{b}_{x} \cdot \rho) \hat{b}_{x} + \nabla R_{e} \hat{e}_{2g} \cdot \hat{e}_{1g} (\rho \times \hat{b}_{x}),$$

(77)

$$\frac{\partial \rho}{\partial \mu_{g}} = \frac{1}{2 \mu_{g}} \rho$$

(78)

and

$$\frac{\partial \rho}{\partial \theta_{g}} = -\rho \times \hat{b}_{x},$$

(79)

we write the Lagrangian in (68) as

$$L^{X}(X(Z_{g}), \dot{X}(Z_{g}, \dot{Z}_{g}), t) = \left[ \int_{t_{0}}^{t} A(R_{g} + \epsilon \rho) + v_{||g} \hat{b}_{g} + \rho \times B_{g} + \nabla R_{e} \rho \cdot A(R_{g} + \epsilon \rho) + \frac{1}{2 \mu_{g}} \frac{dR_{g}}{dt} + \frac{1}{2 \mu_{g}} \frac{\partial A(R_{g} + \epsilon \rho)}{\partial \mu_{g}} \frac{d\mu_{g}}{dt} + H(0) - \epsilon H^{(1)}, \right]$$

(80)

To obtain (77) we have used $\nabla R_{e} \hat{e}_{1g} = -(\nabla R_{e} \hat{b} \cdot \hat{e}_{1g}) \hat{b}_{g} - (\nabla R_{e} \hat{e}_{2g} \cdot \hat{e}_{1g}) \hat{e}_{2g}$ and $\nabla R_{e} \hat{e}_{2g} = -(\nabla R_{e} \hat{b} \cdot \hat{e}_{2g}) \hat{b}_{g} + (\nabla R_{e} \hat{e}_{2g} \cdot \hat{e}_{1g}) \hat{e}_{1g}$. To simplify the Lagrangian (80), we add the time derivative of

$$S_{NP}(R_{g}, \mu_{g}, \theta_{g}) = -\int_{t_{0}}^{t} \frac{\mu_{g}}{2 \mu_{g}} A(R_{g} + \epsilon \rho(R_{g}, \mu_{g}, \theta_{g})) \cdot \rho(R_{g}, \mu_{g}, \theta_{g}) \, dt.$$  

(81)

As a result we find

$$L^{Z_{g}} = \left[ \int_{t_{0}}^{t} \frac{1}{2} A(R_{g} + \epsilon \rho) + v_{||g} \hat{b}_{g} + \rho \times B_{g} + \nabla R_{e} \rho \cdot A(R_{g} + \epsilon \rho) + \nabla R_{e} S_{NP} \right. + \frac{1}{2 \mu_{g}} \frac{dR_{g}}{dt} + \frac{1}{2 \mu_{g}} \frac{\partial A(R_{g} + \epsilon \rho)}{\partial \mu_{g}} \frac{d\mu_{g}}{dt} + H(0) - \epsilon H^{(1)}, \right]$$

(82)

In appendix B we prove that

$$\nabla R_{e} S_{NP} = -\frac{1}{\epsilon} A(R_{g} + \epsilon \rho) + \frac{1}{\epsilon} A(R_{g} + \epsilon \rho) + \frac{1}{\epsilon} A(R_{g} + \epsilon \rho) + \frac{1}{\epsilon} A(R_{g} + \epsilon \rho) - \epsilon \mu_{g} \nabla R_{e} \hat{e}_{2g} \cdot \hat{e}_{1g}$$

$$- \int_{t_{0}}^{t} \frac{\mu_{g}}{2 \mu_{g}} \frac{d\mu_{g}'}{d\theta_{g}} \left[ \rho' \times [B(R_{g} + \epsilon \rho') - B_{g}] \right.$$

$$+ \epsilon (\rho' \times \hat{b}_{g} \cdot B(R_{g} + \epsilon \rho')) [\nabla R_{e} \hat{b}_{g} \cdot \rho']$$

$$+ \frac{2 \epsilon \mu_{g}'}{B_{g}} [\hat{b}_{g} \cdot B(R_{g} + \epsilon \rho') - B_{g}] [\nabla R_{e} \hat{e}_{2g} \cdot \hat{e}_{1g})]$$

(83)

and

$$\frac{\partial S_{NP}}{\partial \theta_{g}} = -\frac{\partial \rho}{\partial \theta_{g}} A(R_{g} + \epsilon \rho) + \frac{\partial \rho}{\partial \theta_{g}} + \frac{\epsilon}{B_{g}} \int_{t_{0}}^{t} \frac{\mu_{g}}{2 \mu_{g}} \frac{d\mu_{g}'}{d\theta_{g}} \left[ \hat{b}_{g} \cdot B(R_{g} + \epsilon \rho') - B_{g}, \right]$$

(84)
where we use the abbreviated notation \( \rho' \equiv \rho(R_g, \mu'_g, \theta_g) \). Substituting equations (83) and (84) into the Lagrangian (82) finally gives

\[
\mathcal{L}^{\mathcal{Z}_\epsilon} = \left( \frac{1}{\epsilon} A_f + v_{||} \mathbf{b}_g + \epsilon \Delta \Gamma_R \right) \cdot \frac{dR_g}{dt} + \left( -\epsilon \mu_g + \epsilon^2 \Delta \Gamma_\theta \right) \frac{d\theta_g}{dt} - H^{(0)} - \epsilon H^{(1)},
\]

with

\[
\Delta \Gamma_R = \mu_g \nabla_{R_g} \mathbf{e}_{2g} \cdot \hat{e}_{1g} - v_{||} \nabla_{R_g} \mathbf{b}_g \cdot \rho_g - \int_0^{\mu_g} \frac{d\mu_g'}{2\mu_g} \left\{ \frac{1}{\epsilon} \rho' \times [B(R_g + \epsilon \rho') - B_g] \\
+ [(\rho' \times \mathbf{b}_g) \cdot B(R_g + \epsilon \rho')] \nabla_{R_g} \mathbf{b}_g \cdot \rho_g \\
+ \frac{2\mu_g'}{B_g} [\mathbf{b}_g \cdot B(R_g + \epsilon \rho') - B_g] \nabla_{R_g} \mathbf{e}_{2g} \cdot \hat{e}_{1g} \right\}
\]

and

\[
\Delta \Gamma_\theta = \frac{1}{\epsilon B_g} \int_0^{\mu_g} \frac{d\mu_g'}{\mu_g} \left[ \mathbf{b}_g \cdot B(R_g + \epsilon \rho') - B_g \right].
\]

It is easy to write the Lagrangian (85) order by order. We use

\[
B(R_g + \epsilon \rho) = B_g + \epsilon \rho \cdot \nabla_{R_g} B_g + \frac{\epsilon^2}{2} \rho \rho : \nabla_{R_g} \nabla_{R_g} B_g + O(\epsilon^3),
\]

where our double-dot convention is \( ab \leftrightarrow \vec{M} = b \cdot \vec{M} \cdot a \), to obtain

\[
\mathcal{L}^{\mathcal{Z}_\epsilon} = \left( \frac{1}{\epsilon} A_f + v_{||} \mathbf{b}_g + \epsilon \Gamma_R^{(1)} + \epsilon^2 \Gamma_R^{(2)} \right) \cdot \frac{dR_g}{dt} + \left( -\epsilon \mu_g + \epsilon^2 \Gamma_\theta^{(1)} + \epsilon^3 \Gamma_\theta^{(2)} \right) \frac{d\theta_g}{dt} \\
- H^{(0)} - \epsilon H^{(1)} + O(\epsilon^3, \epsilon^4),
\]

where

\[
\Gamma_R^{(1)} = \mu_g \nabla_{R_g} \mathbf{e}_{2g} \cdot \hat{e}_{1g} - v_{||} \nabla_{R_g} \mathbf{b}_g \cdot \rho - \frac{1}{2} \left[ \rho \cdot \nabla_{R_g} \mathbf{b}_g \right] \rho \times \mathbf{b}_g \\
+ \frac{1}{2} \left[ \rho \cdot \nabla_{R_g} \mathbf{b}_g \right] \left( \rho \times \mathbf{b}_g \right) B_g,
\]

\[
\Gamma_R^{(2)} = \frac{1}{3} \rho \rho : \nabla_{R_g} \nabla_{R_g} B_g \times \rho - \frac{B_g}{3} \left[ \rho \cdot \nabla_{R_g} \mathbf{b}_g \right] \left( \rho \times \mathbf{b}_g \right) \nabla_{R_g} \mathbf{b}_g \cdot \rho \\
- \frac{2\mu_g}{3B_g} \left( \rho \cdot \nabla_{R_g} \mathbf{b}_g \right) \nabla_{R_g} \mathbf{e}_{2g} \cdot \hat{e}_{1g},
\]

\[
\Gamma_\theta^{(1)} = \frac{2\mu_g}{3B_g} \rho \cdot \nabla_{R_g} B_g
\]

and

\[
\Gamma_\theta^{(2)} = \frac{\mu_g}{4B_g} \rho \rho : \nabla_{R_g} \nabla_{R_g} B_g \cdot \mathbf{b}_g.
\]

### 3.4. Perturbative change of variables

In this subsection we find a new set of coordinates \( \{R, u, \mu, \theta\} \) that makes the Lagrangian (89) gyrophase independent. We employ the procedure described in section 3.2. The transformation \( \{R_g, v_{||}, \mu_g, \theta_g\} = T_{P,e} \{R, u, \mu, \theta, t\} \) is customarily written in the form of a Lie transform \([19, 25]\),

\[
T_{P,e} = T_1 T_2 T_3 \ldots ,
\]
where
\[ T_n = \exp \left[ \epsilon^{n+1} R_{n+1} \cdot \nabla R + \epsilon^n \left( \mu_n \frac{\partial}{\partial \mu} + \mu_n \frac{\partial}{\partial \mu} + \theta_n \frac{\partial}{\partial \theta} \right) \right]. \] (95)

Instead of this form, we use the form in (38) that we find more convenient. The connection between the two arrangements is trivial. To first order we find
\[ \hat{R}_1 = R_1, \quad \tilde{u}_1 = u_1, \quad \tilde{\mu}_1 = \mu_1, \quad \tilde{\theta}_1 = \theta_1. \] (96)

To second order, the relation is
\[ \hat{R}_2 = R_2, \quad \tilde{u}_2 = u_2, \quad \tilde{\mu}_2 = \mu_2, \quad \tilde{\theta}_2 = \theta_2. \] (97)

In section 3.4.1, the corrections \( R_2, u_1, \mu_1 \) and \( \theta_1 \) are calculated following the procedure in section 3.2, and the Lagrangian is obtained to \( O(\epsilon, \epsilon^2) \). In section 3.4.2 the Lagrangian is obtained to next order. It is possible to do so without explicitly obtaining \( R_3, \tilde{u}_2, \tilde{\mu}_2 \) and \( \tilde{\theta}_2 \).

### 3.4.1. Perturbative change of variables to first order

We obtain the first order gyrokinetic correction to the Hamiltonian, \( \tilde{H}^{(1)} \), by employing equation (52). We need to know that
\[ \langle \Gamma^{(1)}_R(R, u, \mu, \theta) \rangle = 0, \quad \langle H^{(1)} R, /\epsilon, R_1, \mu, \theta, t) \rangle = \Lambda(\phi)(R_1 /\epsilon, R_1, \mu /\lambda^2, t /\tau) \] and
\[ \langle \Gamma^{(1)}_R(R, u, \mu, \theta) \rangle = \mu \nabla_R \hat{e}_1 \cdot \hat{e}_1 + \frac{\mu}{2 \hat{b}} \times \nabla_B B - \frac{\mu}{2} \hat{b} \cdot \nabla_B \hat{b}, \] (101)

where we have used that
\[ \langle \rho \rho \rangle = \frac{\mu}{B} \tilde{F}^{(1)} - \hat{b} \hat{b}, \] (102)

with \( \tilde{F} \) the unit matrix. For the remainder of the section, whenever we do not write explicitly the arguments of the functions, it will be understood that they are evaluated at \( (R, u, \mu, \theta) \), i.e. \( \hat{b} \equiv b(R) \), \( \langle \phi \rangle \equiv \langle \phi \rangle(R_1 /\epsilon, R_1, \mu /\lambda^2, t /\tau) \), and so on. Substituting the values of \( \langle \Gamma^{(1)}_R(R, u, \mu, \theta) \rangle \), \( \langle \Gamma^{(1)}_\phi(R, u, \mu, \theta) \rangle \) and \( \langle H^{(1)}(R_1 /\epsilon, R_1, \mu, \theta, t) \rangle \) into equation (52), we find
\[ \tilde{H}^{(1)} = \Lambda(\phi) + \hat{a} \hat{b} \cdot \tilde{F}^{(1)} - u \mu \hat{b} \cdot \nabla_R \hat{e}_2 \cdot \hat{e}_1 + \frac{u \mu}{2} \hat{b} \cdot \nabla_R \times \hat{b}. \] (103)

Note that we have the freedom to choose \( \tilde{F}^{(1)} \) as we wish. Our choice will affect the corrections \( R_2 \) and \( u_1 \), and the final form of \( \tilde{H}^{(1)} \). To coincide with previous derivations in the literature [25], we choose
\[ \tilde{F}^{(1)}_R = \mu \nabla_R \hat{e}_2 \cdot \hat{e}_1 - \frac{\mu}{2} \hat{b} \hat{b} \cdot \nabla_R \times \hat{b}, \] (104)

giving
\[ \tilde{H}^{(1)} = \Lambda(\phi). \] (105)
In equation (104), we have chosen \( \Gamma_R^{(1)} = \mu \nabla_R \dot{e}_2 \dot{e}_1 + \cdots \) instead of \( \Gamma_R^{(1)} = \mu \dot{b} \cdot \nabla_R \dot{e}_2 \dot{e}_1 + \cdots \) to manifestly show that the equations of motion are independent of the choice of \( \dot{e}_1 \) and \( \dot{e}_2 \) [42].

The function \( S_p^{(2)} \) can be obtained by solving equation (53). Substituting equations (69), (90) and (92) into (53) gives

\[
\frac{\partial S_p^{(2)}}{\partial \theta} = -\frac{\mu^2}{B} \dot{b} \cdot \nabla_R \dot{b} \cdot \rho + \frac{u}{4} [\rho (\rho \times \dot{b}) + (\rho \times \dot{b}) \rho] : \nabla_R \dot{b} - \frac{2\mu}{3B} \rho \cdot \nabla_R B - \frac{\lambda \dot{\phi}}{B},
\]

where we have used that

\[
\rho \rho - \langle \rho \rho \rangle = \frac{1}{2} [\rho \rho - (\rho \times \dot{b}) (\rho \times \dot{b})].
\]

Integrating equation (106) in the gyrophase gives

\[
S_p^{(2)} = -\frac{u^2}{B} \dot{b} \cdot \nabla_R \dot{b} \cdot (\rho \times \dot{b}) - \frac{u}{8} [\rho \rho - (\rho \times \dot{b}) (\rho \times \dot{b})] : \nabla_R \dot{b}
- \frac{2\mu}{3B} (\rho \times \dot{b}) \cdot \nabla_R B - \frac{\lambda \dot{\phi}}{B},
\]

where the function \( \dot{\Phi} \) is the integral

\[
\dot{\Phi}(R_{\perp}/\lambda \epsilon, R_{||}, \mu/\lambda^2, \theta + \pi \Theta(-\lambda), t/\tau) = \int^{\theta}_0 d\theta' \dot{\phi}(R_{\perp}/\lambda \epsilon, R_{||}, \mu/\lambda^2, \theta' + \pi \Theta(-\lambda), t/\tau)
\]

such that \( \langle \dot{\Phi} \rangle = 0 \). Here we have used that \( \rho = \partial (\rho \times \dot{b})/\partial \theta \) and \( \rho (\rho \times \dot{b}) + (\rho \times \dot{b}) \rho = -((1/2) \partial [\rho \rho - (\rho \times \dot{b}) (\rho \times \dot{b})]/\partial \theta). \)

Using \( S_p^{(2)} \) in expressions (47), (48), (49) and (50), the first order corrections to the gyrokinetic variables become

\[
R_2 = -\frac{2\mu}{B} \dot{b} \cdot \nabla_R \dot{b} \cdot (\rho \times \dot{b}) - \frac{1}{8} \dot{b} [\rho \rho - (\rho \times \dot{b}) (\rho \times \dot{b})] : \nabla_R \dot{b} - \frac{u}{B} \dot{b} \times \nabla_R \dot{b} \cdot \rho
- \frac{1}{2B} \rho \rho \cdot \nabla_R B - \frac{\lambda}{\lambda B^2} \dot{b} \times \nabla_{(R_{\perp}/\lambda \epsilon)} \dot{\Phi},
\]

\[
u_1 = u \dot{b} \cdot \nabla_R \dot{b} \cdot \rho - \frac{B}{4} [\rho (\rho \times \dot{b}) + (\rho \times \dot{b}) \rho] : \nabla_R \dot{b},
\]

\[
\mu_1 = -\frac{u^2}{B} \dot{b} \cdot \nabla_R \dot{b} \cdot \rho + \frac{u}{4} [\rho (\rho \times \dot{b}) + (\rho \times \dot{b}) \rho] : \nabla_R \dot{b} - \frac{\lambda \dot{\phi}}{B}
\]

and

\[
\theta_1 = \frac{u^2}{2\mu B} \dot{b} \cdot \nabla_R \dot{b} \cdot (\rho \times \dot{b}) + \frac{u}{8\mu} [\rho \rho - (\rho \times \dot{b}) (\rho \times \dot{b})] : \nabla_R \dot{b}
+ \frac{1}{B} (\rho \times \dot{b}) \cdot \nabla_R B + \frac{\lambda}{\lambda^2 B} \partial \dot{\Phi}.
\]

In appendix C we show that this result is equivalent to the result obtained with the iterative method in [17].

3.4.2. Perturbative change of variables to second order. In this section we apply the change of variables (38) to the Lagrangian (89) to \( O(\epsilon^2, \epsilon^3) \). The idea is to write an expression similar to (55) with \( n = 2 \) so that we can use the technique demonstrated in section 3.2.
To $O(\epsilon^2, \epsilon^3)$, the Lagrangian (89) becomes

$$L^Z = \left[ \frac{1}{\epsilon} A + u \dot{b} + \epsilon \left( \Gamma^{(1)}_R + R_2 \cdot \nabla_R A + u_i \dot{b} \right) + \epsilon^2 \left( \Gamma^{(2)}_R + \dot{R}_3 \cdot \nabla_R A + \dddot{u}_2 \dot{b} \right) + u \dot{R}_2 \cdot \nabla_R \dot{b} + u_i \frac{\partial \Gamma^{(1)}_R}{\partial u} + \mu_1 \frac{\partial \Gamma^{(1)}_R}{\partial \mu} + \theta_1 \frac{\partial \Gamma^{(1)}_R}{\partial \theta} \right] \cdot \frac{dR}{dt} + \left[ \epsilon A + \epsilon^2 u \dot{b} + \epsilon^3 \left( \Gamma^{(1)}_R + R_2 \cdot \nabla_R A + u_i \dot{b} \right) \right] \cdot \frac{dR_2}{dt} + \left( \epsilon^2 A + \epsilon^3 u \dot{b} \right) \cdot \frac{d\dot{R}_3}{dt} + \left[ -\epsilon \mu + \epsilon^2 (-\mu_1 + \Gamma^{(1)}_{\mu}) + \epsilon^3 \left( -\mu_2 + \Gamma^{(2)}_{\mu} + \mu_1 \frac{\partial \Gamma^{(1)}_{\mu}}{\partial u} + \theta_1 \frac{\partial \Gamma^{(1)}_{\mu}}{\partial \theta} \right) \right] \frac{d\theta_1}{dt} - \epsilon^3 \mu \frac{d\dot{\theta}_1}{dt} - H^{(0)} - \epsilon (uu_1 + \mu_1 B + H^{(1)}) - \epsilon^2 \left( u \dddot{u}_2 + \frac{u_i^2}{2} + \mu_2 B + \mu R_2 \cdot \nabla_R B + R_2 \cdot \nabla_{(R,\epsilon)} H^{(1)} \right) + \mu_1 \frac{\partial H^{(1)}}{\partial \mu} + \theta_1 \frac{\partial H^{(1)}}{\partial \theta} \right] + O(\epsilon^3, \epsilon^4), (114)$$

where we have used that $\partial \Gamma^{(1)}_{\theta}/\partial u = 0$, that $\partial H^{(1)}/\partial u = 0$ and that $\Gamma^{(1)}_R$ and $\Gamma^{(1)}_{\theta}$ only depend slowly on $R$.

We have seen in section 3.2 that to first order we need to add to the Lagrangian (114) the time derivative

$$\frac{d}{dt}(\epsilon^2 S^{(2)}_p - \epsilon A \cdot R_2 - \epsilon^2 u \dot{b} \cdot R_2 + \epsilon^2 \mu \dot{\theta}_1), (115)$$

giving as a result the Lagrangian

$$L^Z + \frac{d}{dt}(\epsilon^2 S^{(2)}_p - \epsilon A \cdot R_2 - \epsilon^2 u \dot{b} \cdot R_2 + \epsilon^2 \mu \dot{\theta}_1) = \left[ \frac{1}{\epsilon} A + u \dot{b} + \epsilon \Gamma^{(1)}_R + \epsilon^2 \left( \dot{R}_3 \cdot \nabla_R A + \dddot{u}_2 \dot{b} + \Gamma^{(2)}_R \right) \right] \cdot \frac{dR}{dt} + \left( \epsilon^2 A + \epsilon^3 u \dot{b} \right) \cdot \frac{dR_2}{dt} + \left( \epsilon^2 A + \epsilon^3 u \dot{b} \right) \cdot \frac{d\dot{R}_3}{dt} + \left[ -\epsilon \mu + \epsilon^2 (-\mu_2 + \Gamma^{(2)}_{\mu} + \mu_1 \frac{\partial \Gamma^{(1)}_{\mu}}{\partial u} + \theta_1 \frac{\partial \Gamma^{(1)}_{\mu}}{\partial \theta} \right] \frac{d\theta_1}{dt} - \epsilon^3 \mu \frac{d\dot{\theta}_1}{dt} - H^{(0)} - \epsilon H^{(1)} - \epsilon^2 \left( u \dddot{u}_2 + \mu_2 B + \epsilon H^{(2)} \right) + O(\epsilon^3, \epsilon^4), (116)$$

with

$$\Gamma^{(2)}_R = \Gamma^{(2)}_p + u (\nabla_R \times \dot{b}) \times R_2 + u_i \frac{\partial \Gamma^{(1)}_R}{\partial u} + \mu_1 \frac{\partial \Gamma^{(1)}_R}{\partial \mu} + \theta_1 \frac{\partial \Gamma^{(1)}_R}{\partial \theta} + \nabla_R S^{(2)}_p$$

$$+ \nabla_{(R,\epsilon)} R_2 \cdot (\Gamma^{(1)}_R + R_2 \cdot \nabla_R A + u_i \dot{b}) + (-\mu_1 + \Gamma^{(1)}_{\epsilon}) \nabla_{(R,\epsilon)} \theta_1, \quad (117)$$

$$\Gamma^{(2)}_\mu = (\Gamma^{(1)}_R + R_2 \cdot \nabla_R A + u_i \dot{b}) \cdot \frac{\partial R_2}{\partial u} + \left( -\mu_1 + \Gamma^{(1)}_{\mu} \right) \frac{\partial \theta_1}{\partial u}, \quad (118)$$

$$\Gamma^{(2)}_\theta = (\Gamma^{(1)}_R + R_2 \cdot \nabla_R A + u_i \dot{b}) \cdot \frac{\partial \dot{R}_3}{\partial u} + \left( -\mu_1 + \Gamma^{(1)}_{\theta} \right) \frac{\partial \dot{\theta}_1}{\partial u}, \quad (119)$$

$$\Gamma^{(2)}_{\epsilon} = (\Gamma^{(1)}_p + R_2 \cdot \nabla_R A + u_i \dot{b}) \cdot \frac{\partial \dddot{u}_2}{\partial u} + \left( -\mu_1 + \Gamma^{(1)}_{\epsilon} \right) \frac{\partial \dddot{\theta}_1}{\partial u} \quad (120)$$
\[ \hat{H}^{(2)} = \frac{u^2}{2} + \mu R_2 \cdot \nabla_R B + R_2 \cdot \nabla_{(R_2,\omega)} H^{(1)} + \mu_1 \frac{\partial H^{(1)}}{\partial \mu} + \theta_1 \frac{\partial H^{(1)}}{\partial \theta} - \frac{\partial S_p^{(2)}}{\partial t}. \]  
\hspace{1cm} (121)

Note that in equation (117) we are using

\[ R_2 \cdot \nabla_R \hat{b} - \nabla_R \hat{b} \cdot R_2 = (\nabla_R \times \hat{b}) \times R_2, \]  
\hspace{1cm} (122)

and we have taken into account that \( S_p^{(2)} \) depends on \( R \) in two different ways. On the one hand, there is the dependence on \( R_1, \omega \), that was the only dependence that was taken into account in section 3.4.1. This dependence gives the strong gradient \( \nabla_{(R_1,\omega)} S_p^{(2)} = -\left( \Lambda / \lambda B \right) \nabla_{(R_1,\omega)} \Phi \).

On the other hand there is a slow dependence on \( R \) that gives the gradient

\[ \nabla_R S_p^{(2)} = \nabla_R \left[ -\frac{u^2}{B} \hat{b} \cdot \nabla_R \hat{b} \cdot (\rho \times \hat{b}) - \frac{u}{8} (\rho \rho - (\rho \times \hat{b})(\rho \times \hat{b})) : \nabla_R \hat{b} \right. \]
\[ \left. - \frac{2\mu}{3B} (\rho \times \hat{b}) \cdot \nabla_R B \right] + \frac{\Lambda \Phi}{B^2} \nabla_R B - \frac{\Lambda}{B} \hat{b} \cdot \nabla_R \Phi. \]  
\hspace{1cm} (123)

In section 3.2 we showed that by adding the total time derivative

\[ \frac{d}{dt}(\epsilon^3 S_p^{(3)}(R/\epsilon, R, u, \mu, \theta, t) - \epsilon^2 A \cdot \hat{R}_3 - \epsilon^3 u \hat{b} \cdot \hat{R}_3 + \epsilon^3 \mu \hat{b}_3) \]  
\hspace{1cm} (124)

to the Lagrangian (116) and making the result equal to

\[ \mathcal{L} = \left[ \frac{1}{\epsilon} A(R) + u \hat{b}(R) + \epsilon \Gamma_R^{(1)} \right] \frac{dR}{dt} - \epsilon \mu \frac{d\theta}{dt} - \mathcal{H}^{(1)} - \epsilon \mathcal{H}^{(2)} + O(\epsilon^3, \epsilon^4), \]  
\hspace{1cm} (125)

where we have explicitly set \( \Gamma_R^{(2)} = 0 \), we obtain the equations for the corrections

\[ \hat{R}_3 = \left( \dot{\Gamma}_u^{(2)} + \frac{\partial S_p^{(3)}}{\partial u} \right) \hat{b}(R) + \frac{1}{B(R)} \hat{b}(R) \times (\Gamma_R^{(2)} + \nabla_{(R_1/\omega)} S_p^{(3)}), \]  
\hspace{1cm} (126)

\[ \hat{u}_2 = -\dot{b}(R) \cdot \Gamma_R^{(2)}, \]  
\hspace{1cm} (127)

\[ \hat{\mu}_2 = \dot{\Gamma}_\theta^{(2)} + \frac{\partial S_p^{(3)}}{\partial \theta}, \]  
\hspace{1cm} (128)

and

\[ \hat{\theta}_2 = -\dot{\Gamma}_\mu^{(2)} - \frac{\partial S_p^{(3)}}{\partial \mu}. \]  
\hspace{1cm} (129)

The correction to the Hamiltonian is

\[ \mathcal{H}^{(2)} = -u \hat{b}(R) \cdot (\Gamma_R^{(2)}) + B(R)(\Gamma^2_\theta + \hat{\mathcal{H}}^{(2)}), \]  
\hspace{1cm} (130)

and the equation for \( S_p^{(3)} \) is

\[ \frac{\partial S_p^{(3)}}{\partial \theta} = \frac{u}{B(R)} \hat{b}(R) \cdot (\Gamma_R^{(2)} - (\Gamma_R^{(2)})) - (\Gamma_\theta^{(2)} - (\Gamma_\theta^{(2)})) - \frac{1}{B(R)} (\hat{\mathcal{H}}^{(2)} - (\hat{\mathcal{H}}^{(2)})). \]  
\hspace{1cm} (131)

In this paper we do not solve for \( S_p^{(3)} \) and hence we are not able to write explicitly the corrections \( \hat{R}_3, \hat{u}_2, \hat{\mu}_2, \text{ and } \hat{\theta}_2. \) We only obtain explicitly the correction \( \mathcal{H}^{(2)}. \) In appendix D we evaluate equation (130) in detail. The final result is

\[ \mathcal{H}^{(2)} = \Lambda^2 \Psi^{(2)}(R_1/\lambda, \epsilon, R, \mu/\lambda^2, t/\tau, \lambda) + \Lambda \Psi^{(2)}(R_1/\lambda, \epsilon, R, u, \mu/\lambda^2, t/\tau, \lambda) 
\]  
\[ + \Psi^{(2)}_B(R, u, \mu), \]  
\hspace{1cm} (132)
with

\[ \Psi^{(2)}_\phi = \frac{1}{2\lambda^2 B^2} \langle \nabla_{(R,\phi_{\lambda\epsilon})} \tilde{\Phi} \cdot (\tilde{b} \times \nabla_{(R,\phi_{\lambda\epsilon})} \tilde{\Phi}) \rangle - \frac{1}{2\lambda^2 B} \frac{\partial (\tilde{\Phi}^2)}{\partial (\mu/\lambda^2)} \]  

(133)

\[ \Psi^{(2)}_{\phi B} = -\frac{u}{\lambda B} \langle (\nabla_{(R,\phi_{\lambda\epsilon})} \tilde{\Phi} \times \tilde{b}) \cdot \nabla_R \tilde{b} \cdot \rho \rangle - \frac{\mu}{2\lambda^2 B^2} \nabla_R B \cdot \nabla_{(R,\phi_{\lambda\epsilon})} \langle \phi \rangle 
- \frac{1}{4\lambda B} \langle \nabla_{(R,\phi_{\lambda\epsilon})} \tilde{\Phi} \cdot [\rho \rho - (\rho \times \tilde{b})(\rho \times \tilde{b})] \cdot \nabla_R B \rangle - \frac{1}{B} \nabla_R B \cdot \langle \tilde{\Phi} \rho \rangle 
- \frac{u^2}{\lambda^2 B} \tilde{b} \cdot \nabla_R \tilde{b} \cdot \left\{ \frac{\partial \tilde{\Phi}}{\partial (\mu/\lambda^2)} \rho \right\} 
+ \frac{u}{4\lambda^2} \nabla_R \tilde{b} : \left\{ \frac{\partial \tilde{\Phi}}{\partial (\mu/\lambda^2)} [\rho (\rho \times \tilde{b}) + (\rho \times \tilde{b}) \rho] \right\} 
+ \frac{u}{4\mu} \nabla_R \tilde{b} : \langle \tilde{\Phi} [\rho (\rho \times \tilde{b}) + (\rho \times \tilde{b}) \rho] \rangle \]  

(134)

and

\[ \Psi^{(2)}_{B} = -\frac{3u^2 \mu}{2B^2} \tilde{b} \cdot \nabla_R \tilde{b} \cdot \nabla_R B + \frac{\mu^2}{4B} \langle (I - \tilde{b}\tilde{b}) : \nabla_R \nabla_R B \cdot \tilde{b} \rangle - \frac{3u^2 \mu}{4B^2} |\nabla_{R \perp} B|^2 
+ \frac{u^2 \mu}{2B} \nabla_R \tilde{b} : \nabla_R \tilde{b} + \left( \frac{\mu^2}{8B} - \frac{u^2 \mu}{4B} \right) \nabla_R \tilde{b} : (\nabla_R \tilde{b})^T - \left( \frac{3u^2 \mu}{8B} + \frac{\mu^2}{16} \right) (\nabla_R \cdot \tilde{b})^2 
+ \left( \frac{3u^2 \mu}{2B} - \frac{u^2}{2B^2} \right) |\tilde{b} \cdot \nabla_R \tilde{b}|^2 + \left( \frac{u^2 \mu}{8B} - \frac{\mu^2}{16} \right) (\tilde{b} \cdot \nabla_R \times \tilde{b})^2, \]  

(135)

where \( M^{**T} \) is the transpose of the matrix \( M^{**} \). The final phase-space Lagrangian is given then by (125). We can write it explicitly as

\[ \overline{\mathcal{L}} = \left[ \frac{1}{\varepsilon} \mathbf{A} (\mathbf{R}) + u \mathbf{b} (\mathbf{R}) - \varepsilon \mu \mathbf{K} (\mathbf{R}) \right] \cdot \frac{d\mathbf{R}}{dt} - \varepsilon \mu \frac{d\mathbf{b}}{dt} - \overline{\mathcal{H}} + O(\varepsilon^3, \varepsilon^4), \]

(136)

where

\[ \overline{\mathcal{H}} = \frac{1}{2} u^2 + \mu B (\mathbf{R}) + \lambda \varepsilon \langle \phi \rangle (\mathbf{R}_\perp / \lambda \epsilon, \mathbf{R}_\parallel / \lambda \epsilon, \mu / \lambda^2, t/\tau, \lambda) + \lambda^2 \varepsilon^2 \psi^{(2)}_\phi (\mathbf{R}_\perp / \lambda \epsilon, \mathbf{R}, \mu / \lambda^2, t/\tau, \lambda) 
+ \lambda \varepsilon^2 \psi^{(2)}_{\phi B} (\mathbf{R}_\perp / \lambda \epsilon, \mathbf{R}, u, \mu, \mu / \lambda^2, t/\tau, \lambda) + \varepsilon^2 \psi^{(2)}_{B} (\mathbf{R}, u, \mu) \]  

(137)

and

\[ \mathbf{K} (\mathbf{R}) = \frac{1}{2} b (\mathbf{R}) b (\mathbf{R}) \cdot \nabla_{\mathbf{R}} \times \tilde{b} (\mathbf{R}) - \nabla_{\mathbf{R}} \dot{e}_2 (\mathbf{R}) \cdot \dot{e}_1 (\mathbf{R}). \]

(138)

In previous work [24, 25], only the contribution \( \Psi^{(2)}_\phi \) was kept because the terms that contained the function \( \phi \) were assumed to be larger. With the more natural ordering (1), we find the new contributions \( \Psi^{(2)}_{\phi B} \) and \( \Psi^{(2)}_B \), demonstrating that magnetic geometry and electrostatic potential appear together and cannot be separated. In sections 4, 5 and 6 we show that \( \Psi^{(2)}_\phi \) and \( \Psi^{(2)}_B \) modify both the equations of motion and Poisson’s equation.

We end this section pointing out that we chose the final form of the Lagrangian (136), with \( \mathbf{R}^{(2)}_R = 0 \), to have the same Poisson brackets as previous authors [25]. There are other possible choices, e.g., making the second-order correction of the Hamiltonian independent of the parallel velocity, condition that can be achieved by defining the appropriate \( \mathbf{R}^{(2)}_R \).
4. Gyrokinetic equations of motion and Vlasov equation

The equations of motion are given by (23). To obtain them explicitly we need to find the Poisson bracket (22) that corresponds to the Lagrangian (136). Employing appendix E we find

\[ \{ F, G \} = \frac{1}{\epsilon} \left( \frac{\partial F}{\partial \mu} \frac{\partial G}{\partial \theta} - \frac{\partial F}{\partial \theta} \frac{\partial G}{\partial \mu} \right) + \frac{B^*}{B^*} \left( \nabla R^* \frac{\partial G}{\partial \theta} - \frac{\partial F}{\partial \theta} \nabla R^* G \right) + \frac{\epsilon}{B^*} \nabla R^* F \cdot (\hat{b} \times \nabla_R G), \]

(139)

where

\[ B^*(R, u, \mu) := B(R) + \epsilon u \nabla_R \times \hat{b}(R) - \epsilon^2 \mu \nabla_R \times K(R), \]

(140)

\[ B^*_{\parallel}(R, u, \mu) := B^*(R, u, \mu) \cdot \hat{b}(R) \]

(141)

and

\[ \nabla_R := \nabla_R - K(R) \frac{\partial}{\partial \theta}. \]

(142)

Employing the Poisson bracket in (139) and the Hamiltonian in (137), we find

\[ \dot{R} \equiv \frac{dR}{dt} = \left( u + \lambda \epsilon \frac{\partial \Psi_{\phi B}^{(2)}}{\partial \mu} + \epsilon^2 \frac{\partial \Psi_{\phi B}^{(2)}}{\partial u} \right) \frac{B^*}{B^*} + \frac{1}{B^*} \hat{b} \times \left( \epsilon \mu \nabla R B + \frac{\Lambda \epsilon}{\lambda} \nabla (R_{\parallel}/\lambda) \phi \right) \]

\[ + \frac{\Lambda^2 \epsilon^2}{\lambda} \nabla (R_{\parallel}/\lambda) \Psi_{\phi B}^{(2)} + \frac{\Lambda \epsilon^2}{\lambda} \nabla (R_{\parallel}/\lambda) \Psi_{\phi B}^{(2)} + \Lambda \epsilon^3 \nabla R \Psi_{\phi B}^{(2)} + \epsilon^3 \nabla R \Psi_{\phi B}^{(2)} \],

(143)

\[ \dot{u} \equiv \frac{du}{dt} = -\frac{\mu}{B^*} B^* \cdot \nabla R B - \Lambda \epsilon \hat{b} \cdot \nabla R \phi - \Lambda^2 \epsilon^2 \hat{b} \cdot \nabla R \Psi_{\phi B}^{(2)} - \Lambda \epsilon^2 \hat{b} \cdot \nabla R \Psi_{\phi B}^{(2)} \]

\[ - \epsilon^2 \hat{b} \cdot \nabla R \Psi_{\phi B}^{(2)} - \frac{1}{B^*} \left[ u \hat{b} \times \hat{b} \nabla R \right] - \epsilon \mu (\nabla R \times K)_{\parallel}, \]

(144)

\[ \dot{\mu} \equiv \frac{d\mu}{dt} = 0 \]

(145)

and

\[ \dot{\theta} \equiv \frac{d\theta}{dt} = -\frac{1}{\epsilon} \left( B - \frac{\Lambda}{\lambda^2} \frac{\partial \phi}{\partial (\mu/\lambda^2)} - \frac{\Lambda^2 \epsilon}{\lambda^2} \frac{\partial \Psi_{\phi B}^{(2)}}{\partial \mu} - \frac{\Lambda \epsilon}{\lambda} \frac{\partial \Psi_{\phi B}^{(2)}}{\partial \mu} - \epsilon \frac{\partial \Psi_{\phi B}^{(2)}}{\partial \mu} \right) \]

\[ - \frac{B^* \cdot K}{B^*} \left( u + \lambda \epsilon \frac{\partial \Psi_{\phi B}^{(2)}}{\partial u} + \epsilon^2 \frac{\partial \Psi_{\phi B}^{(2)}}{\partial u} \right) - \frac{1}{B^*} (K \times \hat{b}) \left( \epsilon \mu \nabla R B \right) \]

\[ + \frac{\Lambda \epsilon}{\lambda} \nabla (R_{\parallel}/\lambda) \phi \left( \phi \frac{\partial \Psi_{\phi B}^{(2)}}{\partial \mu} + \frac{\Lambda \epsilon^2}{\lambda} \nabla (R_{\parallel}/\lambda) \Psi_{\phi B}^{(2)} + \frac{\Lambda \epsilon^2}{\lambda} \nabla (R_{\parallel}/\lambda) \Psi_{\phi B}^{(2)} + \Lambda \epsilon^3 \nabla R \Psi_{\phi B}^{(2)} \right) \]

\[ + \epsilon^3 \nabla R \Psi_{\phi B}^{(2)} \],

(146)
the presence of collisions it is possible to estimate the size of eliminating the gyrofrequency time scales, a crucial step to obtain the gyrokinetic equation. In

Note that we have emphasized the fact that the dependence of the functions $F$, $\Psi^{(2)}_1$, and $\Psi^{(2)}_{\phi B}$ on $R$, $u$, and $\mu$ (recall equation (137)) can be fast or slow. For this reason we distinguish between derivatives with respect to the argument $R$, and between derivatives with respect to the argument $u$, and derivatives with respect to the argument $\mu/\lambda^2$.

The new correction to the Hamiltonian $\Psi^{(2)}_{\phi B}$ gives a contribution of order $\epsilon^2$ to the perpendicular and parallel motion of the gyrocenter, comparable to the contribution from $\Psi^{(2)}$, the term that is usually kept. The correction to the Hamiltonian $\Psi^{(2)}_{B}$ gives a negligible contribution to the perpendicular drift, but is needed for the parallel motion. Thus, both corrections must be kept to obtain the equations of motion to order $\epsilon^2$.

It is worth mentioning that equation (143) contains the Baños drift [43] in the definition of $u$. To make it clear, instead of choosing the first order Lagrangian as in (46) with $\Gamma_{R}$ given in (104), we can choose it to have

$$ \Gamma_R^{(1)} = \mu \nabla_R \hat{e}_2 \cdot \hat{e}_1 + \frac{\mu}{2} \hat{b} \cdot \nabla_R \times \hat{b} = \Gamma_R^{(1)} + \mu \hat{b} \cdot \nabla_R \times \hat{b}. \quad (147) $$

This choice gives a different parallel velocity $u' = u - \epsilon \mu \hat{b} \cdot \nabla_R \times \hat{b}$ and a different first order Hamiltonian $\Pi^{(1)} = \Pi^{(1)} + u \mu \hat{b} \cdot \nabla_R \times \hat{b}$. With this new choice, the equation for $R$ to first order is

$$ \dot{R} = (u' + \epsilon \mu \hat{b} \cdot \nabla_R \times \hat{b}) \hat{b} + \frac{\epsilon \mu}{B} \hat{b} \times \nabla_R B + \frac{B}{2} \hat{b} \times (\hat{b} \cdot \nabla_R \hat{b}) $$

$$ - \frac{\Lambda e}{\lambda B} \nabla_R (\phi \phi_R, \phi_R) \times \hat{b} + O(\epsilon^2). \quad (148) $$

Note that the Baños drift has been made explicit. From here on, we work only with our equations of motion (143), (144), (145) and (146) that are equivalent to the equations obtained with this alternative choice that makes the Baños drift manifest.

The gyrokinetic Vlasov equation is readily written for the phase-space distribution $F(R, u, \mu, \theta, t)$ in gyrokinetic coordinates, giving

$$ \frac{\partial F}{\partial t} + \dot{R} \cdot \nabla_R F + \frac{\partial F}{\partial u} + \phi \frac{\partial F}{\partial \phi} = 0, \quad (149) $$

or employing the Poisson bracket,

$$ \frac{\partial F}{\partial t} + \{F, \Pi\} = 0. \quad (150) $$

In the absence of collisions and making use of the fact that $\dot{R}$, $\dot{u}$ and $\dot{\theta}$ are independent of gyrophase, the gyrophase independent piece of the distribution function $\langle F \rangle$ and the gyrophase-dependent piece $\tilde{F} = F - \langle F \rangle$ are determined by two decoupled equations [23], namely,

$$ \frac{\partial \langle F \rangle}{\partial t} + \dot{R} \cdot \nabla_R \langle F \rangle + \dot{u} \frac{\partial \langle F \rangle}{\partial u} = 0 $$

$$ \frac{\partial \tilde{F}}{\partial t} + \dot{R} \cdot \nabla_R \tilde{F} + \dot{u} \frac{\partial \tilde{F}}{\partial u} + \phi \frac{\partial \tilde{F}}{\partial \phi} = 0. \quad (151) $$

Since $\dot{\theta} \sim 1/\epsilon$ $\gg \partial/\partial t$, $\dot{R} \cdot \nabla_R \langle u(\partial/\partial u)$, it is usually assumed that any dependence on the gyrophase disappears in a very short time and $\tilde{F}$ can be neglected. As a result, we assume that $F = \langle F \rangle$. By zeroing the gyrophase-dependent piece of the distribution function we are eliminating the gyrofrequency time scales, a crucial step to obtain the gyrokinetic equation. In the presence of collisions it is possible to estimate the size of $\tilde{F}$ and argue that it is small [17,44].
It is possible to write the gyrokinetic equation in conservative form. Indeed, the equations of motion obtained from a phase-space Lagrangian conserve phase-space volume. Therefore, the determinant of the Jacobian matrix of the gyrokinetic transformation, \( J_{\text{c}} \), found in appendix F to be 
\[
\text{det}(J_{\text{c}}) = B^* || \frac{\partial \dot{R}}{\partial \theta} + \frac{\partial \dot{u}}{\partial u} (B^* || \dot{u}) = 0.
\] (153)

For completeness, we prove this equation in appendix G. Since \( \dot{\theta} \) and \( B^* || \) do not depend on \( \theta \), this equation reduces to
\[
\nabla_R \cdot (B^* || \dot{R}) + \frac{\partial}{\partial u} (B^* || \dot{u}) = 0.
\] (154)

Using this expression and equation (151), and taking into account that our choice of Lagrangian (136) implies that \( \partial B^* || / \partial t = 0 \), we find the Vlasov equation in conservative form, i.e.
\[
\partial_t (B^* || \langle F \rangle) + \nabla_R \cdot (B^* || \dot{R} \langle F \rangle) + \frac{\partial}{\partial u} (B^* || \dot{u} \langle F \rangle) = 0.
\] (155)

5. Gyrokinetic Poisson’s equation

In sections 3 and 4 we have obtained the gyrokinetic equations of motion and the gyrokinetic Vlasov equation keeping the electrostatic potential as an unspecified function. The system of equations of electrostatic gyrokinetic theory is closed by coupling the gyrokinetic Vlasov equation to Poisson’s equation, which is the subject of study of this section.

First, in section 5.1 we present the normalization that we employ for Poisson’s equation. Since several species enter in Poisson’s equation, the normalization given in section 2 must be modified. With this new normalization, we obtain the gyrokinetic Poisson’s equation in section 5.2 by simply changing from the coordinates \( r \) and \( v \) to their gyrokinetic counterparts \( R, u, \mu \) and \( \theta \). In section 6 we provide another way of obtaining Poisson’s equation that is based on field theory.

5.1. Normalized Lagrangian in a system with several species

Poisson’s equation in Gaussian units reads
\[
\nabla^2 \varphi(r, t) = -4\pi \left[ e \int d^3 v_i f_i(r, v_i, t) + \sum_I Z_I e \int d^3 v_I f_I(r, v_I, t) - e \int d^3 v_e f_e(r, v_e, t) \right],
\] (156)

where \( f_i(r, v_i, t) \) is the particle distribution of the dominant ions, \( Z_I e \) and \( f_i(r, v_I, t) \) are the charge and the distribution function of the impurity \( I \), and \( f_e(r, v_e, t) \) is the distribution function of the electrons. Using the species-independent normalization
\[
L = \frac{v_i L}{c}, \quad r = \frac{r}{L}, \quad A = \frac{A}{B_0 L}, \quad \varphi = \frac{e \varphi}{\epsilon_0 T_e}
\] (157)

for time, space, vector potential and electrostatic potential, and the species-dependent normalization
\[
v_i = \frac{v_i}{v_i}, \quad v_I = \frac{v_I}{v_i}, \quad v_e = \frac{v_e}{v_e}, \quad f_i = \frac{v_i^3}{n_e} f_i, \quad f_I = \frac{v_I^3}{n_e} f_I, \quad f_e = \frac{v_e^3}{n_e} f_e
\] (158)
The gyrokinetic variables obtained in the previous sections and results of sections 2 and 3, the constants \( L_{MI} \) are the characteristic temperatures of the ion species and the impurity species \( I \) for the velocities and the distribution functions, we get

\[
\frac{-\epsilon_i^2 \lambda_{De}^2}{L^2} \nabla^2 \phi(\mathbf{r}, t) = \int d^3 v_i f_i(\mathbf{r}, \mathbf{v}_i, \mathbf{l}) + \sum_{I} Z_I \int d^3 v_I f_I(\mathbf{r}, \mathbf{v}_I, \mathbf{l}) - \int d^3 v_e f_e(\mathbf{r}, \mathbf{v}_e, \mathbf{l}).
\]

(159)

Here, \( n_{e0} \) is a characteristic equilibrium value for the electron density and

\[
\lambda_{De} = \frac{T_{e0}}{4\pi e^2 n_{e0}}
\]

(160)

is the electron Debye length. Recall that \( \epsilon_i = \rho_i / L \). This new normalization is somewhat different from the normalization in section 2. We have that \( \hat{r} = \mathbf{r}, \hat{A} = A_\phi = \phi \) and \( \hat{v}_p = v_p \) (the subindex \( p \) can take the value i, e, and any of the values of \( I \)), but the normalization for the time is different, giving \( I = \tau_L \). The gyrokinetic variables obtained in the previous sections can be employed here without further changes because both the normalization used in section 2 and the normalization in equations (157) and (158) give the same normalization for ions, impurities and electrons. In the results of sections 2 and 3, the constants \( \epsilon_i, \lambda, \tau \) and \( \epsilon \) depend on the species, and it will be useful to use subscripts for them as well. Their values for ions, the impurity species \( I \) and electrons are given in table 1. Note the sign in the definition of \( \epsilon_i \) and \( \lambda_e \). In table 1, \( T_0 \) and \( T_{10} \) are the characteristic temperatures of the ion species and the impurity species \( I \), respectively, and \( M_I \) is the mass of the impurity species \( I \).

| Species | Impurity \( I \) | Electrons |
|---------|-----------------|----------|
| \( \epsilon_i = \epsilon_i / \lambda_i \) | \( \epsilon_I = \epsilon_I / \lambda_I \) | \( -\epsilon_e = -\epsilon_e / \lambda_e \ll \epsilon_e \) |
| \( \lambda_i = \sqrt{T_{i0}/T_0} \) | \( \lambda_I = Z_I \sqrt{T_{i0} m_i / T_{10} M_I} \) | \( -\lambda_e = \sqrt{m_e / m_e} \gg 1 \) |
| \( \tau_i = \sqrt{T_{i0}/\tau_0} \) | \( \tau_I = \sqrt{T_{i0} m_i / \tau_0 M_I} \) | \( \tau_e = \sqrt{m_e / m_e} \gg 1 \) |
| \( \Lambda_i = (T_{i0}/T_0)^{1/2} \) | \( \Lambda_I = Z_I^2 (T_{i0}/T_{10})^{1/2} \sqrt{m_i / M_I} \) | \( \Lambda_e = \sqrt{m_e / m_e} \gg 1 \) |

Table 1. Values of the species-dependent parameters \( \epsilon, \Lambda, \lambda, \tau \) for ions, the impurity species \( I \) and electrons.

The function \( \phi \) in (70) is now species-dependent and given by

\[
\phi(\hat{R}_{p1}, \lambda_p \epsilon_p, \hat{R}_{p||}, \mu_p / \lambda^2_p, \theta_{pI}, \tau_{pI}) = \phi(\hat{R}_{p1}, \epsilon_p, \hat{R}_{p||}, \mu_p / \lambda^2_p, \theta_{pI}, \tau_{pI}).
\]

(161)

Later on we will use the more compact notation

\[
\phi_p(R, \mu_p, \theta_p, \tau_p) = \phi_p(R_{p1}, \epsilon_p, \hat{R}_{p||}, \mu_p / \lambda^2_p, \theta_{pI}, \tau_{pI}).
\]

(162)

To write the total Lagrangian of the system, it is necessary to normalize all the particle Lagrangians and Hamiltonians by the same quantity. We choose that the normalized Lagrangian and Hamiltonian be \( \mathcal{L} = \mathcal{L} / T_0 e^{\phi} \) and \( \mathcal{H} = \mathcal{H} / T_0 e^{\phi} \). The gyrokinetic Lagrangian (136), dependent on the species, becomes in the new normalization

\[
\mathcal{L}_p = \frac{T_{p0}}{T_0 \tau_p} \left[ \frac{\lambda_p}{\epsilon_p} \hat{A}(R_p) + \mu_p \hat{b}(R_p) - \frac{\epsilon_p}{\lambda_p} \mu_p K(R_p) \right] + \frac{dR_p}{dt} = \frac{T_{p0} \epsilon_p}{T_0 \lambda_p \tau_p} \mu_p \frac{d\theta_p}{dt} - \frac{\theta_p}{\tau_p},
\]

(163)
where
\[ H_p = \frac{T_p \rho_0}{T_0} \left[ \frac{1}{2} \nabla^2 u_p \right]^2 + \frac{T_p \rho_0}{T_0} e^2 \frac{\psi^{(2)}_p}{\lambda_p} (R_p, u_p, \mu_p) \]
\[ + Z_p \epsilon (\phi(R_p, u_p, \mu_p, \lambda_p)) \]
\[ + Z_p \lambda_p^2 \epsilon^2 \frac{\psi^{(2)}_p}{\lambda_p} (R_p, u_p, \mu_p, \lambda_p) \]
\[ + Z_p \lambda_p^2 \epsilon^2 \frac{\psi^{(2)}_p}{\lambda_p} (R_p, u_p, \mu_p, \lambda_p) \]
\[ (164) \]

and
\[ K(R_p) = \frac{1}{2} \hat{b}(R_p) \hat{b}(R_p) \cdot \nabla_{R_p} \times \hat{b}(R_p) - \nabla_{R_p} \hat{e}_z(R_p) \cdot \hat{e}_z(R_p). \]
\[ (165) \]

Here, obviously, \( Z_1 = 1 \) and \( Z_2 = -1 \). Analogously to (162) we define
\[ \psi^{(2)}_p(R_p, \mu_p, \lambda_p) = \psi^{(2)}_p(R_p, u_p, \mu_p, \lambda_p, \lambda_p) \]
\[ (166) \]

and
\[ \psi^{(2)}_p R_p, u_p, \mu_p, \lambda_p) = \psi^{(2)}_p(R_p, u_p, \mu_p, \lambda_p, \lambda_p). \]
\[ (167) \]

The equations of motion are those obtained in (143), (144), (145) and (146) multiplied by \( \tau_p \). As a result, the Vlasov equation for each species is
\[ \frac{\partial F_p}{\partial t} + \tau_p \hat{R} \cdot \nabla_{R_p} F_p + \tau_p \hat{u} \frac{\partial F_p}{\partial u_p} = 0. \]
\[ (168) \]

In what follows we work in non-dimensional variables but do not underline them.

### 5.2. Gyrokinetic Poisson’s equation via the gyrokinetic change of coordinates

Our objective is to write Poisson’s equation (159) in terms of \( F_p(R_p, u_p, \mu_p, t) \), i.e. the solution of equation (168). Since9 \( (r, v_p) = \tilde{T}_{n,p}(R_p, u_p, \mu_p, \theta_p, t) \),
\[ F_p(R_p, u_p, \mu_p, t) = f_p(T_{n,p}(R_p, u_p, \mu_p, \theta_p), t). \]
\[ (169) \]

Using the obvious identity
\[ \int d^3 v_p f_p(r, v_p, t) = \int d^3 r' \int d^3 v_p f_p(r', v_p, t) \delta(r' - r) \]
\[ (170) \]

and the change of variables formula we can write Poisson’s equation as
\[ - \frac{\epsilon^2}{L^2} \nabla^2 \phi(r, t) = \int d^3 R_i d u_i d \mu_i d \theta_i | \det(J_{\tilde{T}_{n,p}}(R_i, u_i, \mu_i, \theta_i, t))| \]
\[ + \sum_i Z_i \int d^3 R_i d u_i d \mu_i d \theta_i | \det(J_{\tilde{T}_{n,p}}(R_i, u_i, \mu_i, \theta_i, t))| \]
\[ - \int d^3 R_e d u_e d \mu_e d \theta_e | \det(J_{\tilde{T}_{n,e}}(R_e, u_e, \mu_e, \theta_e, t))| \]
\[ (171) \]

where \( J_{\tilde{T}_{n,p}}(R_p, u_p, \mu_p) \) is the Jacobian matrix of the transformation \( T_{n,p} \), calculated in appendix F, and \( \pi^p(r, v_p) := r \) is the projection onto the spatial part of the coordinates.

9 We stress that the transformation \( T_{n} \) depends on the species through the values of \( \epsilon_p, \Lambda_p \) and \( \lambda_p \) and denote it by \( T_{n,p} \).
(r, v_p). From an abstract viewpoint this is, perhaps, the simplest way of writing (159) in terms of \( F_p(R_p, u_p, \mu_p, \theta_p, t) \). Observe that (171) is an exact relation. However, in practical terms, we have only computed the explicit expression of \( T_{p, p} \) up to a certain order in \( \epsilon_p \), or equivalently, in the species-independent parameter \( \epsilon_s \). Making use of the results of this paper we can give an explicit expression for the change of variables up to order \( \epsilon_s^3 \), namely

\[
\pi '(T_{p, p}(R_p, u_p, \mu_p, \theta_p, t)) = R_p + \frac{\epsilon_s}{\lambda_p} \rho(R_p, \mu_p, \theta_p)
\]

\[
+ \frac{\epsilon_s^2}{\lambda_p^2} \left[ R_{p, 2} + \mu_{p, 1} \frac{\partial \rho}{\partial \mu_p} + \theta_{p, 1} \frac{\partial \rho}{\partial \theta_p} \right] + O(\epsilon_s^3),
\]

which allows us to write the identity

\[
\delta(\pi '(T_{p, p}(R_p, u_p, \mu_p, \theta_p, t)) - r) = \delta \left( R_p + \frac{\epsilon_s}{\lambda_p} \rho(R_p, \mu_p, \theta_p) - r \right)
\]

\[
+ \frac{\epsilon_s^2}{\lambda_p^2} \left( R_{p, 2} + \mu_{p, 1} \frac{\partial \rho}{\partial \mu_p} + \theta_{p, 1} \frac{\partial \rho}{\partial \theta_p} \right) \cdot \nabla \delta \left( R_p + \frac{\epsilon_s}{\lambda_p} \rho(R_p, \mu_p, \theta_p) - r \right)
\]

\[
+ O(\epsilon_s^3),
\]

where \( \nabla \delta \) denotes the gradient of the Dirac delta function with respect to its natural arguments and the subscript \( p = i, I, e \) in the corrections \( R_{p, 2}, \mu_{p, 2}, \theta_{p, 2} \) indicates that these corrections depend on the species through \( \lambda_p, \lambda_s \) and \( \phi_p(R_p, \mu_p, \theta_p, t) \) in (162). Substituting (173) into (171) finally gives

\[
- \frac{\epsilon_s \lambda_p^2}{L^2} \nabla^2 \psi(r, t) = \sum_p Z_p \int d^3 R_p \, du_p \, d\mu_p \, d\theta_p \, B_{||p}^* \, F_p \left[ \delta \left( R_p + \frac{\epsilon_s}{\lambda_p} \rho(R_p, \mu_p, \theta_p) - r \right) \right]
\]

\[
+ \frac{\epsilon_s^2}{\lambda_p^2} \left( R_{p, 2} + \mu_{p, 1} \frac{\partial \rho}{\partial \mu_p} + \theta_{p, 1} \frac{\partial \rho}{\partial \theta_p} \right) \cdot \nabla \delta \left( R_p + \frac{\epsilon_s}{\lambda_p} \rho(R_p, \mu_p, \theta_p) - r \right)
\]

\[
+ \cdots .
\]

Expression (174) may seem accurate to order \( \epsilon_s^3 \), i.e., that on the right side of (174) we are dropping only terms which are of order \( \epsilon_s^3 \) or higher. However, it is easy to see that this is not true. Using

\[
\nabla \delta \left( R_p + \frac{\epsilon_s}{\lambda_p} \rho - r \right) = \left( \mathbf{T} + \frac{\epsilon_s}{\lambda_p} \nabla R_p, \rho \right)^{-1} \cdot \nabla R_p \delta \left( R_p + \frac{\epsilon_s}{\lambda_p} \rho - r \right)
\]

\[
= \nabla R_p \delta \left( R_p + \frac{\epsilon_s}{\lambda_p} \rho - r \right) + O(\epsilon_s),
\]

\[
\frac{\epsilon_s}{\lambda_p} \frac{\partial \rho}{\partial \mu_p} \cdot \nabla \delta \left( R_p + \frac{\epsilon_s}{\lambda_p} \rho - r \right) = \frac{\partial}{\partial \mu_p} \left[ \delta \left( R_p + \frac{\epsilon_s}{\lambda_p} \rho - r \right) \right]
\]

and

\[
\frac{\epsilon_s}{\lambda_p} \frac{\partial \rho}{\partial \theta_p} \cdot \nabla \delta \left( R_p + \frac{\epsilon_s}{\lambda_p} \rho - r \right) = \frac{\partial}{\partial \theta_p} \left[ \delta \left( R_p + \frac{\epsilon_s}{\lambda_p} \rho - r \right) \right],
\]
and integrating by parts so that the delta function does not appear differentiated, we find that (174) becomes

\[
- \frac{\epsilon_s^2 \lambda_p}{L^2} \nabla^2 \varphi(r, t) = \sum_p Z_p \int d^3 R_p \, du_p \, d\mu_p \, d\theta_p \, \delta \left( R_p + \frac{\epsilon_s}{\lambda_p} \rho(R_p, \mu_p, \theta_p) - r \right) \left\{ B^*_{||,p} F_p - \frac{\epsilon_s^3}{\lambda_p^3} \frac{1}{\det(I + (\epsilon_s/\lambda_p) \nabla R_p, \rho)} \left( B^*_{||,p} F_p \right) \cdot \frac{\partial}{\partial \rho} \left( B^*_{||,p} F_p \rho \right) + \frac{\partial}{\partial \mu_p} \left( B^*_{||,p} F_p \mu_{p,1} \right) \right\} + O(\epsilon_s^4).
\]

Note that terms that seemed to be of second order in \( \epsilon_s \) are in reality first order contributions. Similarly, terms that seem to be of third order and are neglected in (174) are in reality of second order. To obtain a gyrokinetic Poisson’s equation correct to order \( \epsilon_s^2 \) it is necessary to carry (172) to an order higher

\[
\pi^r(T_{p,\rho}(R_p, u_p, \mu_p, \theta_p, t)) = R_p + \frac{\epsilon_s^2}{\lambda_p^2} \rho + \frac{\epsilon_s^3}{\lambda_p^3} \left( R_{p,2} + \mu_{p,1} \frac{\partial \rho}{\partial \mu_p} + \theta_{p,1} \frac{\partial \rho}{\partial \theta_p} \right) + \mu_{p,1} \theta_{p,1} \left( \frac{\partial^2 \rho}{\partial \mu_p \partial \theta_p} + \frac{1}{2} \theta_{p,1}^2 \frac{\partial^2 \rho}{\partial \theta_p^2} \right) + O(\epsilon_s^4).
\]

All the terms entering this equation are computable from the results found in section 3. We leave this for future work.

By integrating over the delta function in (178) we make the gyrokinetic Poisson’s equation more explicit,

\[
- \frac{\epsilon_s^2 \lambda_p}{L^2} \nabla^2 \varphi(r, t) = \sum_p Z_p \int du_p \, d\rho_p \, d\theta_p \frac{1}{\det(I + (\epsilon_s/\lambda_p) \nabla R_p, \rho)} \left\{ B^*_{||,p} F_p - \frac{\epsilon_s^3}{\lambda_p^3} \frac{1}{\det(I + (\epsilon_s/\lambda_p) \nabla R_p, \rho)} \left( B^*_{||,p} F_p \right) \cdot \frac{\partial}{\partial \rho} \left( B^*_{||,p} F_p \rho \right) + \frac{\partial}{\partial \mu_p} \left( B^*_{||,p} F_p \mu_{p,1} \right) \right\} |_{R_p = \hat{R}(r, \mu_p, \theta_p, \epsilon_s/\lambda_p)} + O(\epsilon_s^4),
\]

where \( \hat{R}(r, \mu_p, \theta_p, \epsilon_s/\lambda_p) \) is the function defined by solving for \( R_p \) the equation \( R_p + (\epsilon_s/\lambda_p) \rho(R_p, \mu_p, \theta_p) = r \).

Finally, we would like to point out that usually equation (180) can be simplified even more. The following discussion is not meant to be an exhaustive review of all possible orderings and simplifications, but a brief comment on the most typical approach. Usually \( F_p \) has contributions with different characteristic scales of variation, ranging from the shortest scale \( \rho_s \) to the background profile variation scale \( L \). At the same time \( \nabla R_p, F_p \sim 1 \) because the saturation amplitude of the short scale fluctuation \( F_{p,\rho_s} \) is sufficiently small, i.e. \( F_{p,\rho_s}(R_{p,\perp}/\epsilon_s, r_p, u_p, \mu_p, t) \sim \epsilon_s \varphi(r_{\perp}/\epsilon_s, t) \), where \( F_{p,\rho_s} \) is a function of order unity with derivatives with respect to its arguments of order unity. For \( \nabla R_p, F_p \sim 1 \), we can use the approximate expression \( \hat{R}(r, \mu_p, \theta_p, \epsilon_s/\lambda_p) = r - (\epsilon_s/\lambda_p) \rho(r, \mu_p, \theta_p) + O(\epsilon_s^2) \) because \( F_p(\hat{R}(r, \mu_p, \theta_p, \epsilon_s/\lambda_p), u_p, \mu_p, \theta_p, t) = F_p(r - (\epsilon_s/\lambda_p) \rho(r, \mu_p, \theta_p), u_p, \mu_p, \theta_p, t) + O(\epsilon_s^2) \). In addition, the term containing \( R_{p,2} \) in (180) becomes of next order because \( \nabla R_p, F_p \sim \nabla R_p, B^*_{||,p} \sim \nabla R_p \cdot R_{p,2} \sim 1 \). To simplify even more,
we use \( \det (\hat{I} + (\epsilon_s/\lambda_p) \nabla R_p \rho) = 1 + (\epsilon_s/\lambda_p) \nabla R_p \cdot \rho + O(\epsilon_s^2) \). As a result,

\[
- \frac{\epsilon_s \lambda_2}{L^2} \nabla^2 \varphi(r, t) = \sum_p Z_p \int d\mu_p d\theta_p d\theta_p \left\{ B_{\| p}^* F_p - \frac{\epsilon_s}{\lambda_p} \left[ B_{\| p}^* F_p (\nabla R_p \cdot \rho) \right] \right. \\
+ \frac{\partial}{\partial \mu_p} (B_{\| p}^* F_p \mu_{p,1}) + \frac{\partial}{\partial \theta_p} (B_{\| p}^* F_p \theta_{p,1}) \right\} \Bigg|_{R_p = r - (\epsilon_s/\lambda_p) \rho(r, \rho, \theta_p)} + O(\epsilon_s^2).
\]

(181)

This is the simplest Poisson’s equation correct to first order in \( \epsilon_s \). As we have already pointed out, it is necessary to include the corrections \( \bar{R}_{p,3} \), \( \bar{\mu}_{p,2} \) and \( \bar{\theta}_{p,2} \) to find the contributions of order \( \epsilon_s^2 \). We finish by checking that the assumption \( \nabla R_p F_p \sim 1 \) is consistent with the gyrokinetic system of equations. Consider the short scale pieces of the Vlasov equation (168) and of the gyrokinetic Poisson’s equation (181). The result will be a typical \( \delta f \) formulation [3–7], demonstrating that it is possible to find a closed non-linear system of equations to determine the short scale fluctuations \( F_p, \rho, \theta \) and \( \varphi \). This does not mean that \( \delta f \) formulations are always valid, but it indicates that the assumption \( \nabla R_p F_p \sim 1 \) is consistent with the gyrokinetic system of equations.

6. Gyrokinetic field theory

Gyrokinetic field theory is the formulation of gyrokinetics as a classical field theory and is defined by an action functional \( \Sigma \) that depends on the trajectories of the particles and the electromagnetic field. A closed system of equations coupling the electromagnetic field and the gyrokinetic distribution function is obtained by finding the stationary points of \( \Sigma \). The first application of field theory to plasma physics is the work by Low in [45] that was later extended to gyrokinetic theory in the seminal papers by Sugama [46] and Brizard [47]. The development of gyrokinetic field theory in the last decade has been motivated mostly by the fact that it allows us to identify in a systematic way conservation laws from symmetries of the Lagrangian. For the sake of completeness we briefly recast now our results in field theory language following the Lagrangian formulation of Sugama [46].

In section 6.1 we show how to obtain the gyrokinetic equations of motion and the gyrokinetic Poisson’s equation from a variational principle. The advantage of this procedure is that it ensures that an energy-like invariant is conserved. We obtain the invariant and prove that it is conserved in section 6.2.

6.1. Equations via variational principle

Using the abbreviated notation \( Z_p = [R_p, u_p, \mu_p, \theta_p] \), the action functional is

\[
\Sigma[Z_p(\bullet, \cdot, \cdot), \varphi(\bullet, \cdot, \cdot)] = \frac{\lambda_2^2 \epsilon_s^2}{2L^2} \int_{t_0}^{t_f} dt \int d^3 r |\nabla \varphi(r, t)|^2 \\
+ \sum_p \int_{t_0}^{t_f} dt \int d^3 r_0 B_{\| p}^* (Z_{p\|}^0) F_{p\|}^0 (Z_{p\|}^0) \\
\times \bar{Z}_p (Z_p(Z_{p\|}^0, t_0; t), \bar{Z}_p (Z_{p\|}^0, t_0; t), \varphi(\bullet, t)),
\]

(182)

where \( F_{p\|}^0 \) is the distribution function of species \( p \) at time \( t_0 \), and \( Z_p(Z_{p\|}^0, t_0; t) \) is the trajectory in phase-space of a particle of species \( p \) satisfying the initial condition \( Z_p(Z_{p\|}^0, t_0; t_0) = Z_{p\|}^0 \) at \( t = t_0 \). The first term on the right side of (182) is the action for the electric field (recall that we are considering a static magnetic field) and the second term is the sum of the actions of the
particles. The gyrokinetic Lagrangian $\mathcal{L}_p$ of the species $p$ is defined in (163) (recall that we have dropped the underlining for normalized variables). Its third argument stresses that $\mathcal{L}_p$ depends as a functional on the electrostatic potential.

The Euler–Lagrange equations for $\Sigma$ are obtained by finding its stationary points under infinitesimal variations of the maps $Z_p(Z_{p0}, t_0; t)$ and $\varphi(r, t)$. The allowed infinitesimal perturbations to $\varphi(r, t)$ vanish at the boundary of the spatial domain of interest, and the perturbations to both $Z_p(Z_{p0}, t_0; t)$ and $\varphi(r, t)$ must be zero at $t = t_0$ and $t = t_1$. The calculation of the variation with respect to $Z_p(Z_{p0}, t_0; t)$ gives the gyrokinetic equations of motion and is a repetition of that leading to equations (143), (144), (145) and (146) (recall the discussion on the phase-space Lagrangian methodology in section 3.1). The distribution function at time $t$ is

$$F_p(Z_p, t) := F_{p0}(Z_{p0}(Z_p, t; t_0)), \quad (183)$$

where $Z_{p0}(Z_p, t; t_0)$ is the inverse of the map $Z_p(Z_{p0}, t; t_0)$, i.e., $Z_{p0}(Z_p(Z_{p0}, t_0; t), t; t_0) \equiv Z_{p0}$. Note that $F_p$ automatically satisfies (168). Observe also that from condition (153), which is automatically satisfied by (143), (144), (145) and (146), we obtain

$$d^6Z_p B_{\parallel p}^*(Z_p(Z_{p0}, t_0; t)) = d^6Z_p \left( \frac{B_{\parallel p}^*(Z_p(Z_{p0}, t_0; t))}{|\det(J_{Z_{p0} \rightarrow Z_p}(Z_{p0}(Z_p, t; t_0)))|} \right) = d^6Z_p B_{\parallel p}^*(Z_p). \quad (184)$$

where $(J_{Z_{p0} \rightarrow Z_p})_a^b = \partial Z^a_p / \partial Z^b_{p0}$ is the Jacobian of the map $Z_p(Z_{p0}, t_0; t)$, and $\det(J_{Z_{p0} \rightarrow Z_p}(Z_{p0})) = B_{\parallel p}^*(Z_p(Z_{p0}, t_0; t))$ is its determinant. This property is proven in appendix H for completeness.

The stationary points of $\Sigma$ under variations of $\varphi(r, t)$ are given by $\delta \varphi \Sigma = 0$, with

$$\delta \varphi \Sigma = \frac{\lambda^2 \delta \epsilon_p^2}{L^2} \int d^3r \nabla \delta \varphi(r, t) \cdot \nabla \varphi(r, t)$$

$$- \sum_p Z_p \delta \epsilon_p \int d^3r Z_p B_{\parallel p}^*(Z_{p0}) F_{p0}(Z_{p0}) \delta \varphi_p(Z_p(Z_{p0}, t_0; t), t)$$

$$- \sum_p Z_p \lambda_p \epsilon_p^2 \int d^3r Z_{p0} B_{\parallel p}^*(Z_{p0}) F_{p0}(Z_{p0}) \delta \varphi_p \Psi_p^{(2)}(Z_p(Z_{p0}, t_0; t), t)$$

$$- \sum_p Z_p \lambda_p^{2} \int d^3r Z_{p0} B_{\parallel p}^*(Z_{p0}) F_{p0}(Z_{p0}) \delta \varphi_p \Psi_p^{(2)}(Z_p(Z_{p0}, t_0; t), t). \quad (185)$$

In appendix I we have evaluated all the terms. The final result is

$$\delta \psi \Sigma = \epsilon_p \int d^3r \delta \varphi(r, t) \mathcal{P}(r, t), \quad (186)$$

where

$$\mathcal{P}(r, t) = -\frac{\lambda^2 \delta \epsilon_p^2}{L^2} \nabla^2 \psi(r, t) - \sum_p Z_p \int d^3R_p du_p d\mu_p d\theta_p \frac{B_{\parallel p}^*(R_p, u_p, \mu_p)}{\lambda_p^2}$$

$$\times F_p(R_p, u_p, \mu_p, t) \left[ \delta \left( R_p + \frac{\epsilon_p}{\lambda_p} \rho - r \right) \right.$$

$$+ \epsilon_p \lambda_p \left[ \frac{\rho}{\lambda_p} R_p 2 \nabla R_p \delta \left( R_p + \frac{\epsilon_p}{\lambda_p} \rho - r \right) + \mu_p \lambda_p \frac{\partial}{\partial \mu_p} \delta \left( R_p + \frac{\epsilon_p}{\lambda_p} \rho - r \right) \right.$$

$$\left. + \theta_p \lambda_p \frac{\partial}{\partial \theta_p} \delta \left( R_p + \frac{\epsilon_p}{\lambda_p} \rho - r \right) \right]. \quad (187)$$

$31$
By imposing that $\delta_p \Sigma = 0$ for any $\delta \varphi(r, t)$, we find that $\mathcal{P}(r, t) = 0$. This is Poisson’s equation. By integrating by parts to leave the delta function undifferentiated, and then integrating over the delta function, we recover (180). Note that even though the Hamiltonian is obtained to second order in $\epsilon_s$, and hence the Vlasov equation is also known to second order, the gyrokinetic Poisson’s equation that we have found using the variational principle is only correct to first order in $\epsilon_s$. It coincides to first order with the first order equation that we found independently in section 5.2.

Observe that equation (187) keeps some second-order terms that could have been neglected, as done in (181). These terms are important if we want to take advantage of the field theory formulation of gyrokinetics. The action (182) is invariant under time translations, so Noether’s theorem automatically provides a conserved quantity, interpreted as the total energy of the system, which we denote by $\mathcal{H}(t)$ because it is indeed the field theory Hamiltonian. However, $\mathcal{H}(t) = 0$ on the equations of motion of $\Sigma$, that is, on solutions of (143), (144), (145), (146) and (187), without neglecting any terms. We show this in the next subsection.

### 6.2. Conservation of energy

In this subsection we prove that if the equations of motion are obtained via a variational principle of the action (182), there is an energy-like invariant $\mathcal{H}(t)$. This is an application of Noether’s theorem, and it is based on the fact that in the action (182) the only time dependence is through the functions $Z_p(Z_{p0}, t_0; t)$ and $\varphi(r, t)$. To obtain the energy-like invariant, we use the equations of motion derived from the variational principle in section 6.1.

In order to find the conserved quantity we first perform the change of variable $t = t' + \delta t$ in every integral on the right side of (182), which, of course, does not change the value of the action, giving

$$
\Sigma[Z_p(\bullet, \cdot, \cdot), \varphi(\bullet, \cdot)] = \int^{t_0 - \delta t}_{t_0 - \delta t} dt' \left[ \frac{\lambda^2_{De} \epsilon_s^2}{2L^2} \int d^3r |\nabla \varphi(r, t' + \delta t)|^2 ight. \\
+ \sum_p \int d^6Z_{p0} B^*_p(Z_{p0}) F_{p0}(Z_{p0}) Z_p |_{t = t' + \delta t} ,
$$

where

$$
Z_p |_{t = t' + \delta t} = Z_p(Z_{p0}(Z_{p00}, t_0; t' + \delta t), Z_p(Z_{p00}, t_0; t' + \delta t), \varphi(\bullet, t' + \delta t)).
$$

We expand the right side of the previous equation up to first order in $\delta t$ to find

$$
\Sigma[Z_p(\bullet, \cdot, \cdot), \varphi(\bullet, \cdot)] = \Sigma[Z_p(\bullet, \cdot, \cdot), \varphi(\bullet, \cdot)] + \delta_p \Sigma[Z_p(\bullet, \cdot, \cdot), \varphi(\bullet, \cdot)] |_{\delta \varphi = \delta t (\partial \varphi / \partial t)} \\
+ \delta_p \sum_p \int d^6Z_{p0} B^*_p(Z_{p0}) F_{p0}(Z_{p0}) Z_p |_{t = t'},
$$

where $\delta_p \Sigma[Z_p(\bullet, \cdot, \cdot), \varphi(\bullet, \cdot)] |_{\delta \varphi = \delta t (\partial \varphi / \partial t)}$ and $\delta_Z \Sigma[Z_p(\bullet, \cdot, \cdot), \varphi(\bullet, \cdot)] |_{\delta \varphi = \delta t (\partial \varphi / \partial t)}$ are the variations of $\Sigma$ under perturbations of both $\varphi(r, t)$ and $Z_p(Z_{p0}, t_0; t)$ with the specific form $\delta \varphi = \delta t (\partial \varphi / \partial t)$ and $\delta Z_p = \delta t \dot{Z}_p$. For the
variations with respect to $Z_p$, we obtain
\[
\delta Z_p \Sigma = \int_0^{t_1} dt' \int d^6 Z_{\rho 0} B_{[\cdot, p]}^*(Z_{\rho 0}) F_{\rho 0}(Z_{\rho 0}) \delta Z_p \Sigma \bigg|_{t=t'}
\]
\[
= \int_0^{t_1} dt' \int d^6 Z_{\rho 0} B_{[\cdot, p]}^*(Z_{\rho 0}) F_{\rho 0}(Z_{\rho 0}) \sum_{\alpha=1}^{6} \left( \delta Z_p^\alpha \frac{\partial \Sigma}{\partial \Sigma} + \delta \Sigma^\alpha \frac{\partial \Sigma}{\partial \Sigma} \right) \bigg|_{t=t'}
\]
\[
= \left[ \int d^6 Z_{\rho 0} B_{[\cdot, p]}^*(Z_{\rho 0}) F_{\rho 0}(Z_{\rho 0}) \sum_{\alpha=1}^{6} \delta Z_p^\alpha \frac{\partial \Sigma}{\partial \Sigma} \right]_{t=t'},
\]
where to obtain this last equality we have integrated by parts in $t$ and we have used the equations of motion
\[
\frac{d}{dt} \left( \frac{\partial \Sigma}{\partial \Sigma} \right) = \frac{\partial \Sigma}{\partial \Sigma} \alpha = 1, 2, \ldots, 6.
\]
Applying $\delta Z_p = \delta t \dot{Z}_p$ we find
\[
\delta Z_p \Sigma |_{\delta Z_p=\delta t \dot{Z}_p} = \delta t \left[ \int d^6 Z_{\rho 0} B_{[\cdot, p]}^*(Z_{\rho 0}) F_{\rho 0}(Z_{\rho 0}) \sum_{\alpha=1}^{6} \dot{Z}_p^\alpha \frac{\partial \Sigma}{\partial \Sigma} \right]_{t=t'},
\]
For the variations with respect to $\delta \varphi$ we obtain that
\[
\delta \varphi \Sigma |_{\delta \varphi=\delta t \partial \varphi/\partial t} = 0,
\]
where we have employed (186) and Poisson’s equation $\mathcal{P}(r, t) = 0$. To obtain (186) we assumed that $\delta \varphi$ vanishes at the boundaries of the domain, and this is not necessarily the case for $\delta \varphi = \delta t \partial \varphi/\partial t$. We avoid this problem by assuming either that the domain extends to infinity, where $\varphi = 0$, or that we are in a periodic box and the contribution from one half of the boundary cancels with the contribution from the other half. In either case, we are assuming that there is no net energy flux through the boundary.

By substituting (193) and (194) into (190) and observing that the term linear in $\delta t$ in (190) has to be identically zero, that is, the sum of the second, third and fourth terms in the right side must vanish, we find
\[
\delta t \left[ -\frac{\lambda^2 \Delta \Sigma}{2L^2} \int d^3 r \left| \nabla \varphi(r, t') \right|^2 \right.
\]
\[
+ \sum_p \int d^6 Z_{\rho 0} B_{[\cdot, p]}^*(Z_{\rho 0}) F_{\rho 0}(Z_{\rho 0}) \left( \sum_{\alpha=1}^{6} \dot{Z}_p^\alpha \frac{\partial \Sigma}{\partial \Sigma} - \Sigma \right) \bigg|_{t=t'} \bigg]_{t=t'} = 0.
\]
Then,
\[
\left[ \mathcal{H}(t') \right]_{t'=t_1} = 0
\]
for solutions of the equations of motion, where
\[
\mathcal{H}(t) = \sum_p \int d^6 Z_{\rho 0} B_{[\cdot, p]}^*(Z_{\rho 0}) F_{\rho 0}(Z_{\rho 0}) \overline{\mathcal{H}}(Z_p(Z_{\rho 0}, t_1; t), \varphi(\bullet, t))
\]
\[
-\frac{\lambda^2 \Delta \Sigma}{2L^2} \int d^3 r \left| \nabla \varphi(r, t) \right|^2.
\]

33
Since this must hold for every interval \([t_0, t_1]\), we deduce that
\[
\frac{dH(t)}{dt} = 0
\] (198)
on the equations of motion. Finally, using equations (183) and (184), we can write the integrals over \(Z_p\) in (197) as integrals over \(Z_p\), giving
\[
\mathcal{H}(t) = \sum_p \int d^6z_p \, \mathcal{B}^*_{\mu p}(Z_p) F_p(Z_p, t) \bar{\mathcal{H}}_{\mu}(Z_p, \varphi(\bullet, t)) - \lambda \frac{\varepsilon_s^2}{2L^2} \int d^3r \, |\nabla \varphi(r, t)|^2.
\] (199)

The conservation of this energy-like invariant is only satisfied on the equations of motion, that is, the variational equations of motion and the variational Poisson’s equation obtained in section 6.1 must be used. It is important to keep all the given terms, even if they are higher order than first (the order to which Poisson’s equation is correct). For example, it is necessary to obtain the exact function \(\hat{R}(\mathbf{r})\) in (180) to have exact conservation of \(\mathcal{H}\).

7. Conclusions and further work

The gyrokinetic ordering in a static magnetic field consists of the ordering assumptions (1), defined by a single parameter \(\epsilon = \rho/L = \omega/\Omega\). In this paper we have strictly implemented the gyrokinetic ordering in the phase-space Lagrangian to obtain explicitly the gyrokinetic Lagrangian to order \(\epsilon^2\) for general magnetic geometry.

Our approach differs from previous phase-space Lagrangian (or Hamiltonian) derivations of gyrokinetics. In previous work [25] the calculation is performed in two steps. First, with zero fluctuating electrostatic potential, an expansion in powers of \(\epsilon = \rho/L\) is performed and a gyrophase independent guiding-center Lagrangian is determined to order \(\epsilon\). Then, the electrostatic fluctuations, whose size is given by a new expansion parameter \(\epsilon_\varphi = Ze\varphi/Mv^2\), are switched-on, reintroducing a gyrophase dependence that is removed order by order in \(\epsilon_\varphi\) yielding the final gyrokinetic Lagrangian, usually computed up to order \(\epsilon^2_\varphi\). When the expansion in \(\epsilon_\varphi\) is performed, the fact that there has been a previous expansion in \(\epsilon\) is ignored and the terms of order \(\epsilon\epsilon_\varphi\) are never calculated. Thus, the final Lagrangian is missing relevant terms of order \(\epsilon\epsilon_\varphi\) and \(\epsilon^2\).

The novelty of our work can be easily understood by examining the explicit expression of the second-order gyrokinetic Hamiltonian, \(\bar{\mathcal{H}}^{(2)}\), given in equations (132), (133), (134) and (135). It shows in a transparent way that gyrokinetic theory ties together geometry and turbulence, so that no splitting between guiding-center and gyrokinetic dynamics is possible. See, for example, \(\Psi^{(2)}_{\mu\mu}\) in (134) where magnetic geometry and electrostatic potential appear together. This is the first time that the electrostatic gyrokinetic equations in general geometry are fully computed to order \(\epsilon^2\) and the calculations are pursued to the point of reaching formulae, like the one for \(\bar{\mathcal{H}}^{(2)}\), that can be straightforwardly implemented in a computer code.

From the new phase-space Lagrangian in (136) and the new Hamiltonian in (137) we obtain a new Vlasov equation and a new gyrokinetic Poisson’s equation. The Vlasov equation is correct to second order in the expansion parameter \(\epsilon\), and it is to our knowledge the highest order full \(f\) gyrokinetic equation available in the literature for general geometry. In the limit where the electrostatic potential has a scale of variation much larger than the gyroradius of the species of interest, this equation is also the highest order drift kinetic equation that we are aware of.

The gyrokinetic Poisson’s equation derived from the new Lagrangian is, however, only correct to first order in \(\epsilon\). We have calculated Poisson’s equation employing two methods. In
section 5, in the integrals in velocity space that enter in Poisson’s equation we have simply changed from the coordinates \([r, v]\) to the gyrokinetic variables. In section 6 we have used a variational formalism. The equation obtained with the variational formalism is set by the form of the Hamiltonian and it is only correct to first order, although it contains terms that are higher order. These higher order terms do not add accuracy, but they are necessary to have an exact energy-like invariant. By directly changing from \([r, v]\) to the gyrokinetic variables in the integrals that enter Poisson’s equation, it is possible to obtain a higher order Poisson’s equation if the higher order corrections \(\tilde{R}_3, \tilde{\mu}_2\) and \(\tilde{\theta}_2\) are calculated. It is not necessary to calculate the Hamiltonian to next order. The calculation of \(\tilde{R}_3, \tilde{\mu}_2\) and \(\tilde{\theta}_2\) is, however, very tedious and is left for future work. If this procedure is followed and the third order Hamiltonian is not obtained, the conservation of the energy-like invariant will not be exact.

Some natural steps following this work are the extension to the electromagnetic case and the introduction of external flows. We will also investigate the implications of the new terms in the Hamiltonian for the transport of toroidal angular momentum in tokamaks.

Acknowledgments

The authors are indebted to the programme of visits to TJ-II at Laboratorio Nacional de Fusión de CIEMAT (Spain) and to the summer programme ‘Gyrokinetics in Laboratory and Astrophysical Plasmas’ at the Isaac Newton Institute for Mathematical Sciences, without which this work would not have been possible.

This research was partially supported by the Post-doctoral programme of the Engineering and Physical Sciences Research Council of the UK, by the Junior Research Fellowship programme of Christ Church at University of Oxford, and by grant ENE2009-07247, Ministerio de Ciencia e Innovación (Spain).

Appendix A. Proof that the inverse of the matrix in (20) defines a Poisson bracket

Let \(L_{\alpha\beta}(Z)\) be an invertible, antisymmetric matrix of dimension \(2n\) whose components are functions defined on a region \(\mathcal{U} \subseteq \mathbb{R}^{2n}\), and such that
\[
\frac{\partial L_{\alpha\beta}}{\partial Z^\gamma} + \frac{\partial L_{\gamma\alpha}}{\partial Z^\beta} + \frac{\partial L_{\beta\gamma}}{\partial Z^\alpha} = 0, \quad 1 \leq \alpha, \beta, \gamma \leq 2n. \tag{A.1}
\]
Denote by \(P^{\alpha\beta}(Z)\) the inverse of \(L_{\alpha\beta}(Z)\), \(P^{\alpha\beta}(Z) = (L^{-1}(Z))^{\alpha\beta}\). Then, \(P^{\alpha\beta}\) defines a Poisson bracket by contraction with the differentials of pairs of functions on \(\mathcal{U}\), i.e.
\[
\{F, G\} = \sum_{\alpha, \beta = 1}^{2n} P^{\alpha\beta} \frac{\partial F}{\partial Z^\alpha} \frac{\partial G}{\partial Z^\beta}. \tag{A.2}
\]
We must prove that (A.2) satisfies skew-symmetry, given in (24), the Leibniz rule, given in (25), and the Jacobi identity, given in (26). Skew-symmetry is satisfied because the inverse of an antisymmetric matrix is antisymmetric, i.e. \(P^{\alpha\beta} = -P^{\beta\alpha}\). The Leibniz rule is trivial to check. As for the Jacobi identity, it is immediate to see that it is equivalent to
\[
\sum_{\delta = 1}^{2n} P^{\alpha\delta} \frac{\partial P^{\delta\gamma}}{\partial Z^\delta} + \sum_{\delta = 1}^{2n} P^{\beta\delta} \frac{\partial P^{\gamma\alpha}}{\partial Z^\delta} + \sum_{\delta = 1}^{2n} P^{\gamma\delta} \frac{\partial P^{\alpha\beta}}{\partial Z^\delta} = 0, \quad 1 \leq \alpha, \beta, \gamma \leq 2n. \tag{A.3}
\]
Showing that (A.3) is equivalent to (A.1) is a simple exercise of application of the formula of the derivative of the inverse of a matrix,
\[
\frac{\partial(L^{-1})^{\alpha\beta}}{\partial Z^\gamma} = -\sum_{\delta, \rho = 1}^{2n} (L^{-1})^{\alpha\delta} \frac{\partial L_{\delta\rho}}{\partial Z^\gamma} (L^{-1})^{\rho\beta}. \tag{A.4}
\]
Appendix B. Calculation of the Lagrangian after the non-perturbative change of variables

In this appendix we prove equations (83) and (84). For equation (83), we use that according to (81)

\[
\nabla R_g S_{NP} = -\int_0^{\mu_g} \frac{d\mu_g'}{2\mu_g} \left[ \nabla A(R_g + \epsilon \rho') \cdot \rho' + \epsilon \nabla R_g \rho' \cdot \nabla A(R_g + \epsilon \rho') \cdot \rho' + \nabla R_g \rho' \cdot A(R_g + \epsilon \rho') \right].
\]

(B.1)

Recall that the prime here indicates that the function depends on \( \mu_g' \). Employing

\[
\frac{1}{2\mu_g} \nabla A(R_g + \epsilon \rho') \cdot \rho' = \frac{1}{2\mu_g} \rho' \cdot \nabla A(R_g + \epsilon \rho') + \frac{1}{2\mu_g} \rho' \times \left[ \nabla \times A(R_g + \epsilon \rho') \right]
\]

\[
= \frac{1}{\epsilon \partial \mu_g} \left[ A(R_g + \epsilon \rho') \right] + \frac{1}{2\mu_g} \rho' \times B(R_g + \epsilon \rho'),
\]

(B.2)

\[
\frac{\epsilon}{2\mu_g} \nabla_R \rho' \cdot \nabla A(R_g + \epsilon \rho') = \frac{\epsilon}{2\mu_g} \nabla_R \rho' \cdot \left[ \rho' \cdot \nabla A(R_g + \epsilon \rho') + \rho' \times B(R_g + \epsilon \rho') \right]
\]

\[
= \nabla_R \rho' \cdot \left[ \frac{\partial}{\partial \mu_g} \left[ A(R_g + \epsilon \rho') \right] + \frac{\epsilon}{2\mu_g} \nabla_R \rho' \cdot \left[ \rho' \times B(R_g + \epsilon \rho') \right] \right],
\]

(B.3)

and

\[
\frac{1}{2\mu_g} \nabla R_g \rho' \cdot A(R_g + \epsilon \rho') = \nabla R_g \left( \frac{\partial \rho'}{\partial \mu_g} \right) \cdot A(R_g + \epsilon \rho'),
\]

(B.4)

equation (B.1) becomes

\[
\nabla R_g S_{NP} = -\frac{1}{\epsilon} A(R_g + \epsilon \rho) + \frac{1}{\epsilon} A_x - \nabla R_g \rho \cdot A(R_g + \epsilon \rho)
\]

\[
- \int_0^{\mu_g} \frac{d\mu_g'}{2\mu_g} \left[ \rho' \times B(R_g + \epsilon \rho') + \epsilon \nabla R_g \rho' \cdot (\rho' \times B(R_g + \epsilon \rho')) \right].
\]

(B.5)

To obtain equation (83) from this equation we use (78) to write

\[
\int_0^{\mu_g} \frac{d\mu_g'}{2\mu_g} \rho' \times B_g = \rho \times B_g,
\]

(B.6)

and we employ (77) to get

\[
\nabla_R \rho' \cdot (\rho' \times B(R_g + \epsilon \rho')) = [\rho' \times \hat{b}_g] \cdot B(R_g + \epsilon \rho') \nabla_R \hat{b}_g \cdot \rho'
\]

\[
+ \frac{2\mu_g'}{B_g} \frac{B_g'}{B_g} \nabla_R (\hat{b}_g \cdot B(R_g + \epsilon \rho' \nabla_R \hat{e}_g \cdot \hat{e}_g).
\]

(B.7)

For equation (84), we use that according to (81)

\[
\frac{\partial S_{NP}}{\partial \theta_g} = -\int_0^{\mu_g} \frac{d\mu_g'}{2\mu_g} \left[ \epsilon \frac{\partial \rho'}{\partial \theta_g} \cdot \nabla A(R_g + \epsilon \rho') \cdot \rho' + \frac{\partial \rho'}{\partial \theta_g} \cdot A(R_g + \epsilon \rho') \right].
\]

(B.8)
Employing

\[ \frac{\epsilon}{2\mu'_g} \frac{\partial \rho'}{\partial \theta_g} \cdot \nabla A(R_g + \epsilon \rho') \cdot \rho' = \frac{\epsilon}{2\mu'_g} \frac{\partial \rho'}{\partial \theta_g} \cdot \left[ \rho' \cdot \nabla A(R_g + \epsilon \rho') \right] + \frac{\epsilon}{2\mu'_g} \frac{\partial \rho'}{\partial \mu'_g} \cdot \left[ \rho' \times B(R_g + \epsilon \rho') \right] \]

\[ = \frac{\partial \rho'}{\partial \theta_g} \cdot \frac{\partial}{\partial \mu'_g} \left[ A(R_g + \epsilon \rho') \right] + \frac{\epsilon}{2\mu'_g} \frac{\partial \rho'}{\partial \mu'_g} \cdot \left[ \rho' \times B(R_g + \epsilon \rho') \right] \]  

(B.9)

and

\[ \frac{1}{2\mu'_g} \frac{\partial \rho'}{\partial \theta_g} \cdot A(R_g + \epsilon \rho') = \frac{\partial^2 \rho'}{\partial \theta_g \partial \mu'_g} \cdot A(R_g + \epsilon \rho'), \]  

(B.10)

equation (B.8) becomes equation (84). To obtain the final form of the equation we have also used (79) to write

\[ \frac{\partial \rho'}{\partial \theta_g} \cdot \left[ \rho' \times B(R_g + \epsilon \rho') \right] = \frac{2\mu'_g}{B_g} \hat{b}_g \cdot B(R_g + \epsilon \rho'). \]  

(B.11)

**Appendix C. Comparison to first order with the results in [17]**

In this appendix we compare the gyrokinetic variable transformation obtained in this paper with the variables found in [17].

To be able to compare with the results in [17], given in the form \( R_{PC}(r, v, t), E_{PC}(r, v, t), \mu_{PC}(r, v, t) \) and \( \theta_{PC}(r, v, t) \), we will use the transformation \( (r, v) = T_r(R, u, \mu, \theta, t) \) to write them as \( R_{PC}(R, u, \mu, t), E_{PC}(R, u, \mu, t), \mu_{PC}(R, u, \mu, t) \) and \( \theta_{PC}(R, u, \mu, \theta, t) \). After doing so, we will see that the gyrokinetic variables \( R_{PC}, E_{PC} \) and \( \mu_{PC} \) in [17] are gyrophase independent quantities up to the order that they are defined. This is a property that must be satisfied because otherwise the new gyrokinetic variables would have fast time dependence through the gyrophase. The variable \( \mu_{PC} \) must be a function of \( \mu \) only because there is only one adiabatic invariant associated with the gyromotion.

The gyrokinetic variables of [17] are the gyrocenter position

\[ R_{PC} = r + \frac{\epsilon}{B} v \times \hat{b} + \frac{\epsilon^2}{B} \left[ \left( v_\perp \hat{b} + \frac{1}{4} v_\perp \right) (v \times \hat{b}) + (v \times \hat{b}) \left( v_\perp \hat{b} + \frac{1}{4} v_\perp \right) \right] \times \nabla \left( \frac{\hat{b}}{B} \right) \]

\[ \quad + \frac{\epsilon^2 v_\parallel}{B^2} v_\perp \cdot \nabla \hat{b} + \frac{\epsilon^2 v_\parallel}{B^2} \hat{b} \cdot \nabla \hat{b} \cdot v_\perp + \frac{\epsilon^2}{8B^2} \hat{b} \times v_\perp \times (v \times \hat{b}) \times (v \times \hat{b}) \cdot \nabla \hat{b} \]

\[- \frac{\Lambda \epsilon^2}{\lambda B^2} \nabla_{(R, \mu)} \phi \times \hat{b} + O(\epsilon^3), \]  

(C.1)

where \( a \times \vec{M} = \alpha \times (b \cdot \vec{M}) \); the gyrokinetic kinetic energy

\[ E_{PC} = \frac{v^2}{2} + \Lambda \epsilon \phi + O(\epsilon^2); \]  

(C.2)

the magnetic moment

\[ \mu_{PC} = \frac{v_\perp^2}{2B} - \frac{\epsilon v_\perp^2}{2B^3} (v \times \hat{b}) \cdot \nabla B - \frac{\epsilon v_\parallel^2}{B^3} \hat{b} \cdot \nabla \hat{b} \cdot (v \times \hat{b}) - \frac{\epsilon v_\parallel^2}{2B^2} \hat{b} \cdot \nabla \times \hat{b} \]

\[- \frac{\epsilon v_\parallel}{4B^2} [v_\perp (v \times \hat{b}) + (v \times \hat{b}) v_\perp] \cdot \nabla \hat{b} + \frac{\Lambda \epsilon}{B} \phi + O(\epsilon^3); \]  

(C.3)
and the gyrophase
\[ \theta_{\text{PC}} = \arctan \left( \frac{v \cdot \hat{e}_2}{v \cdot \hat{e}_1} \right) = -\frac{e v}{B} \nabla B - \frac{e v^2}{B^2} \hat{b} \cdot \nabla v + \frac{e}{B} (v \times \hat{b}) \cdot \nabla \hat{e}_2 \cdot \hat{e}_1 \]
\[ -\frac{e v}{4v^2} [v \times v] - (v \times \hat{b})(v \times \hat{b}) : \nabla \hat{b} = \frac{\Lambda e}{\lambda B^2} \partial \Phi + O(\epsilon^2). \quad (C.4) \]

To these variables we apply the gyrokinetic transformation. First we apply the non-perturbative change of variables \( T_{\text{NP}} \). For the variable \( R_{\text{PC}} \) we find
\[ R_{\text{PC}} = R_{\theta} - e^2 \rho \cdot \nabla R_{\theta} \left( \frac{b_x}{B_x} \right) \times (v_{||} \hat{b}_x + \rho \times B_x) + e^2 v_{||} \rho \cdot \nabla R_{\theta} \left( \frac{b_x}{B_x} \right) \times (v_{||} \hat{b}_x + \rho \times B_x) + O(\epsilon^2) \]
\[ = -\rho - e^2 \rho \cdot \nabla R_{\theta} \left( \frac{b_x}{B_x} \right) \times (v_{||} \hat{b}_x + \rho \times B_x) + O(\epsilon^2). \quad (C.6) \]

We then apply the transformation \( R_{\theta}, v_{||}, \mu, \theta, t) = T_{\theta, \epsilon}(R, u, \mu, \theta, t) \) by using \( R_{\theta} = R + e^2 R_2 + O(\epsilon^3) \), with \( R_2 \) given in (110), \( v_{||} = u + O(\epsilon), \mu = \mu + O(\epsilon) \) and \( \theta = \theta + O(\epsilon) \).

As a result we find the expression
\[ R_{\text{PC}} = R + \frac{\epsilon^2}{2} \rho [\rho \times (\rho \times \hat{b})(\rho \times \hat{b})] : \nabla R \hat{b} + \frac{\epsilon^2}{4} \rho [\rho \times (\rho \times \hat{b})(\rho \times \hat{b})] : \nabla R B \]
\[ + \frac{\epsilon^2}{B^2} \rho \times (\rho \times \hat{b}) \cdot \nabla R \hat{b} + \epsilon^2 \rho \cdot \nabla R \frac{\hat{b}}{B} \times \rho - \frac{\epsilon^2}{2} \rho \hat{b} \times \nabla R \hat{b} \cdot \rho + O(\epsilon^3), \quad (C.7) \]

where \( B \) and \( \hat{b} \) depend on \( R \). To simplify expression (C.7) we use
\[ \rho \rho + (\rho \times \hat{b})(\rho \times \hat{b}) = 2 \frac{\mu}{B} (1 - \hat{b} \hat{b}) \]
\[ \quad (C.8) \]
and
\[ (\rho \times \hat{b}) \cdot \nabla R \hat{b} - \hat{b} \times \nabla R \hat{b} \cdot \rho = (\hat{b} \times \nabla R) \times \hat{b} \times \rho = (\nabla R \cdot \hat{b})(\rho \times \hat{b}). \quad (C.9) \]

We finally obtain
\[ R_{\text{PC}} = R - \frac{\epsilon^2 \mu}{B^2} \hat{b} \cdot \nabla R B + \frac{\epsilon^2 \mu}{2B^2} \nabla R B + O(\epsilon^3). \quad (C.10) \]

To find this result we have used \( \nabla R \cdot \hat{b} = -B^{-1} \hat{b} \cdot \nabla R B \).
Applying the non-perturbative change of variables to $E_{PC}$ we find

$$E_{PC} = \frac{1}{2} v_{\|}^2 + \frac{1}{2} \rho \times B g \cdot B g + \Lambda \epsilon \phi + O(\epsilon^2) = \frac{1}{2} v_{\|}^2 + \mu_1 B g + \Lambda \epsilon \phi + O(\epsilon^2).$$

(C.11)

To transform to the variables $R, u, \mu$ and $\theta$ we use $v_{\|} = u + \epsilon u_1 + O(\epsilon^2), \mu_g = \mu + \epsilon \mu_1 + O(\epsilon^2)$, with $u_1$ and $\mu_1$ given in (111) and (112), $R, u = R + \epsilon u_1 + O(\epsilon^2)$ and $\theta_g = \theta + \epsilon \theta_1$, giving

$$E_{PC} = \frac{1}{2} u^2 + \mu B(R) + O(\epsilon^2).$$

(C.12)

For $\mu_g$, we find that the transformation $(r, v) = T_{NP, \epsilon}(R, v_{\|}, \mu_g, \theta_g)$ gives

$$\mu_g = \mu_g - \epsilon v_{\|} \rho \cdot \nabla R g \cdot (\rho \times \hat{b}_g) + \frac{\epsilon v_{\|}^2}{B g} \cdot \nabla R_g \cdot \hat{b}_g \cdot \rho - \frac{\epsilon v_{\|} \mu_g}{B g} \cdot \nabla R_g \times \hat{b}_g$$

$$+ \frac{\epsilon v_{\|}}{4} ((\rho \times \hat{b}_g) \rho + \rho (\rho \times \hat{b}_g)) : \nabla R_g \hat{b}_g + \frac{\Lambda \epsilon \phi}{B g} + O(\epsilon^2),$$

(C.13)

where we have used $v_{\|} = v_{\|} + O(\epsilon)$, $v_{\|} = \rho \times B g + O(\epsilon)$, and

$$\frac{v_{\|}^2}{2 B r(r)} = \frac{1}{2 B g} [v^2 - (v \cdot \hat{b}(r))^2] - \frac{\mu_g}{B g} \rho \cdot \nabla R_g B g + O(\epsilon^2).$$

(C.14)

and

$$v^2 - (v \cdot \hat{b}(r))^2 = v_{\|}^2 + 2 \mu_g B g - [(v_{\|} \hat{b}_g + \rho \times B g) \cdot \hat{b}(r))^2 + O(\epsilon^2)$$

$$= 2 \mu_g B g - 2 \epsilon v_{\|} \rho \cdot \nabla R_g \hat{b}_g \cdot (\rho \times B g) + O(\epsilon^2).$$

(C.15)

Then employing $\mu_g = \mu + \epsilon \mu_1 + O(\epsilon^2)$, with $\mu_1$ given in (112), $R, u = R + \epsilon u_1 + O(\epsilon)$ and $\theta_g = \theta + O(\epsilon)$, we find

$$\mu_g = \mu - \frac{\epsilon v_{\|} \mu_g}{B g} \cdot \nabla R_g \cdot \hat{b}_g + \frac{\epsilon v_{\|}^2}{2} (\rho \times \hat{b}_g - (\rho \times \hat{b}_g) \rho) : \nabla R_g \hat{b}_g + O(\epsilon^2)$$

$$= \mu + O(\epsilon^2),$$

(C.16)

where to obtain the final equality we have used (C.8).

Finally, for $\theta_g$ we obtain that the non-perturbative transformation gives

$$\theta_g = \theta_g - \frac{\epsilon v_{\|} \rho}{2 \mu_g} \cdot \nabla R_g \hat{b}_g \cdot \rho - \frac{\epsilon}{B g} (\rho \times \hat{b}_g) \cdot \nabla R_g B g - \frac{\epsilon v_{\|}^2}{2 \mu_g B g} \cdot \nabla R_g \hat{b}_g \cdot (\rho \times \hat{b}_g)$$

$$+ \frac{\epsilon v_{\|}}{8 \mu_g} [(\rho \rho - (\rho \times \hat{b}_g) (\rho \times \hat{b}_g)) : \nabla R_g \hat{b}_g - \frac{\Lambda \epsilon}{\Lambda^2 B g} \frac{\partial \phi}{\partial (\mu/\Lambda^2)} + O(\epsilon^2),$$

(C.17)

where we have used

$$\text{arctan} \left( \frac{v \cdot \hat{e}_2(r)}{v \cdot \hat{e}_1(r)} \right) = \text{arctan} \left[ \frac{(v_{\|} \hat{b}_g + \rho \times B g) \cdot \hat{e}_2(r)}{(v_{\|} \hat{b}_g + \rho \times B g) \cdot \hat{e}_1(r)} \right]$$

$$= \theta_g + \frac{\epsilon}{2 \mu_g B g} [(\rho \times B g) \cdot \hat{e}_1] \rho \cdot \nabla R_g \hat{e}_{2g} \cdot (v_{\|} \hat{b}_g + \rho \times B g)$$

$$- \frac{\epsilon}{2 \mu_g B g} [(\rho \times B g) \cdot \hat{e}_{2g}] \rho \cdot \nabla R_g \hat{e}_{1g} \cdot (v_{\|} \hat{b}_g + \rho \times B g) + O(\epsilon^2)$$

$$= \theta_g - \frac{\epsilon v_{\|} \rho}{2 \mu_g} \cdot \nabla R_g \hat{b}_g \cdot \rho + \epsilon \rho \cdot \nabla R_g \hat{e}_{2g} \cdot \hat{e}_{1g} + O(\epsilon^2).$$

(C.18)

Here we have employed $\nabla R_g \hat{e}_{1g} = - (\nabla R_g \hat{b}_g \cdot \hat{e}_{1g}) \hat{b}_g - (\nabla R g \hat{e}_{2g} \cdot \hat{e}_{1g}) \hat{e}_{2g}$ and $\nabla R g \hat{e}_{2g} = - (\nabla R g \hat{b}_g \cdot \hat{e}_{2g}) \hat{b}_g + (\nabla R g \hat{e}_{2g} \cdot \hat{e}_{1g}) \hat{e}_{1g}$ to write the last equality. Equation (C.17) is now rewritten.
by using $\theta_g = \theta + \epsilon \theta_1 + O(\epsilon^2)$, with $\theta_1$ given in (113), $R_g = R + O(\epsilon^2)$, $v_{\|g} = u + O(\epsilon)$ and $\mu_g = \mu + O(\epsilon)$, giving

$$\theta_{PC} = \theta - \frac{e u}{4 \mu} [\rho \rho + (\rho \times \dot{b})(\rho \times \dot{b})] : \nabla_R \dot{b} + O(\epsilon^2) = \theta + \frac{e u}{2 B^2} \dot{b} \cdot \nabla_R B + O(\epsilon^2),$$

(C.19)

where we have employed (C.8) and $\nabla_R \cdot \dot{b} = -B^{-1} \dot{b} \cdot \nabla_R B$.

From equations (C.10), (C.12), (C.16) and (C.19) we see that the gyrokinetic variables defined in [17] are simple functions of the gyrokinetic variables in this paper. Note that $R_{PC}$ and $E_{PC}$ do not depend on the gyrophase and that $\mu_{PC}$ is equal to $\mu$ to the order of interest, as expected.

Appendix D. Calculation of the second-order gyrokinetic Hamiltonian

In this appendix we show how to obtain (132) from equation (130).

First, by employing equations (91), (93), (108), (110) and (123), we find that $(\Gamma^{(2)}_R) = 0$, $(\nabla R_S^{(2)}) = 0$, $(\langle \delta S^{(2)}_{\mu} / \partial t \rangle) = 0$,

$$\langle R_2 \rangle = -\frac{\mu}{2 B^2} \nabla R \cdot B$$

(D.1)

and

$$\langle \Gamma^{(2)}_{\theta} \rangle = \frac{\mu^2}{4 B^2} (I - \dot{b} \dot{b}) : \nabla_R \cdot B \cdot \dot{b}.$$

(D.2)

Here we have used equation (102) repeatedly. These results give that equation (130) is equal to

$$\tilde{H}^{(2)} = -\frac{\mu^2}{2 B^2} \hat{b} \cdot \nabla_R \hat{b} \cdot \nabla_R B + \frac{\mu^2}{4 B^2} (I - \dot{b} \dot{b}) : \nabla_R \nabla_R B \cdot \hat{b} = -\frac{\mu^2}{2 B^2} (\nabla_R B)^2 + \langle h_{\theta} \rangle^2$$

(D.3)

where we have used $(\nabla_R \times \dot{b}) \times \dot{b} = \hat{b} \cdot \nabla_R \dot{b}$ to write

$$\hat{b} \cdot [(\nabla_R \times \dot{b}) \times \langle R_2 \rangle] = \frac{\mu}{2 B^2} [\nabla_R \nabla_R B] \cdot \hat{b} \cdot \nabla_R B = \frac{\mu}{2 B^2} \hat{b} \cdot \nabla_R \dot{b} \cdot \nabla_R B.$$

(D.4)

The terms $T_1$, $T_2$, $T_3$ and $T_4$ in equation (D.3) are

$$T_1 = B \left( \langle \Gamma^{(1)}_{R \perp} + R_2 \cdot \nabla R A \rangle \cdot \frac{\partial R_2}{\partial \theta} \right),$$

(D.5)

$$T_2 = B \left( \langle \Gamma^{(1)}_{R \theta} \hat{b} + u_1 \hat{b} \cdot \frac{\partial R_2}{\partial \theta} \rangle \right),$$

(D.6)

$$T_3 = B \left( \langle \Gamma^{(1)}_{\theta} - \mu_1 \frac{\partial \theta_1}{\partial \theta} \rangle \right)$$

(D.7)

and

$$T_4 = -u \hat{b} \cdot \left( u_1 \frac{\partial \Gamma^{(1)}_{\mu}}{\partial u} + \mu_1 \frac{\partial \Gamma^{(1)}_{\mu}}{\partial \mu} + \theta_1 \frac{\partial \Gamma^{(1)}_{\mu}}{\partial \theta} \right) + B \left( \mu_1 \frac{\partial \Gamma^{(1)}_{\theta}}{\partial \mu} + \theta_1 \frac{\partial \Gamma^{(1)}_{\theta}}{\partial \theta} \right) +$$

$$+ \left( \mu_1 \frac{\partial \hat{H}^{(1)}}{\partial \mu} + \theta_1 \frac{\partial \hat{H}^{(1)}}{\partial \theta} \right).$$

(D.8)

We proceed to write these terms in more convenient forms.
The term $T_1$ in (D.5) is rewritten using

$$R_2 \cdot \nabla R A \cdot \frac{\partial R_2}{\partial \theta} = \frac{1}{2} \left( R_2 \cdot \nabla R A \cdot \frac{\partial R_2}{\partial \theta} - \frac{\partial R_2}{\partial \theta} \cdot \nabla R A \cdot R_2 \right) + \frac{1}{2} \frac{\partial}{\partial \theta} \left( R_2 \cdot \nabla R A \cdot R_2 \right)$$

(90)

$$+ \frac{1}{2} (B \times R_2) \cdot \frac{\partial R_2}{\partial \theta} + \frac{1}{2} \frac{\partial}{\partial \theta} (R_2 \cdot \nabla R A \cdot R_2)$$

(D.9)

to obtain

$$T_1 = B \left( \left( \Gamma_{R \perp}^{(1)} + \frac{1}{2} B \times R_2 \right) \cdot \frac{\partial R_2}{\partial \theta} \right)$$

(91)

Employing (47), this last result becomes

$$T_1 = \frac{1}{2} \left( \left( \begin{array}{c} \Gamma_R^{(1)} \rangle - \nabla R_{\perp i,j} S_p^{(2)} \rangle \cdot \left( \hat{b} \times \left( \frac{\partial \Gamma_R^{(1)}}{\partial \theta} + \nabla R_{\perp i,j} \left( \frac{\partial S_p^{(2)}}{\partial \theta} \right) \right) \right) \right) \right).$$

(111)

Realizing that

$$\Gamma_R^{(1)} \cdot \left( \hat{b} \times \nabla R_{\perp i,j} \left( \frac{\partial S_p^{(2)}}{\partial \theta} \right) \right) - \nabla R_{\perp i,j} S_p^{(2)} \cdot \left( \hat{b} \times \frac{\partial \Gamma_R^{(1)}}{\partial \theta} \right)$$

$$= \frac{\partial}{\partial \theta} \left( \Gamma_R^{(1)} \cdot (\hat{b} \times \nabla R_{\perp i,j} S_p^{(2)}) \right),$$

(D.12)

we finally obtain

$$T_1 = \frac{1}{2} \left( \left( \begin{array}{c} \Gamma_R^{(1)} \rangle \cdot \left( \hat{b} \times \left( \frac{\partial \Gamma_R^{(1)}}{\partial \theta} + \nabla R_{\perp i,j} \left( \frac{\partial S_p^{(2)}}{\partial \theta} \right) \right) \right) \right) \right).$$

(113)

These two terms can be expressed in terms of magnetic and electric field gradients. Using equation (90) and $(rpp)$ from (102), we find

$$\left( \begin{array}{c} \Gamma_R^{(1)} \rangle \cdot \left( \hat{b} \times \frac{\partial \Gamma_R^{(1)}}{\partial \theta} \right) \right) = \frac{u^2 \mu}{B} \left( \hat{b} \times \nabla R \hat{b} \times \hat{b} \right) : (\nabla R \hat{b}) = -\frac{\mu^2}{2B^2} |\nabla R \times B|^2.$$  

(D.14)

where $M^\perp$ is the transpose of the matrix $M$. Employing that $\hat{b} \times \nabla R \hat{b} \times \hat{b} = (\nabla R \hat{b})^T - (\nabla R \cdot \hat{b})(\hat{b}^T - \hat{b} \hat{b})$, we finally find

$$\left( \begin{array}{c} \Gamma_R^{(1)} \rangle \cdot \left( \hat{b} \times \frac{\partial \Gamma_R^{(1)}}{\partial \theta} \right) \right) = \frac{u^2 \mu}{B} \nabla R \hat{b} : \nabla R \hat{b} = \frac{u^2 \mu}{B} \nabla R \cdot \hat{b}^2 - \frac{\mu^2}{2B^2} |\nabla R \times B|^2.$$  

(D.15)

Substituting this expression and the definition of $S_p^{(2)}$ in (108) into equation (D.13) gives

$$T_1 = \frac{u^2 \mu}{2B} \nabla R \hat{b} : \nabla R \hat{b} - \frac{u^2 \mu}{2B} (\nabla R \cdot \hat{b})^2 - \frac{\mu^2}{4B^2} |\nabla R \times B|^2$$

$$- \frac{\Lambda^2}{2B^2} \left( \nabla R_{\perp i,j} \Phi \cdot \left( \hat{b} \times \nabla R_{\perp i,j} \Phi \right) \right).$$  

(D.16)

Using equations (47) and (48), the term $T_2$ in (D.6) becomes

$$T_2 = B \cdot \left( \Gamma_R^{(1)} \rangle \left( \frac{\partial S_p^{(2)}}{\partial \theta} \right) \right) = 0.$$  

(D.17)

Employing equations (49) and (50) the term $T_3$ in (D.7) is written as

$$T_3 = B \left( \frac{\partial S_p^{(2)}}{\partial \theta} \frac{\partial S_p^{(2)}}{\partial \mu} \right) = B \left( \frac{\partial S_p^{(2)}}{\partial \theta} \right)^2 \left( \frac{\partial S_p^{(2)}}{\partial \theta} \right)^2.$$  

(D.18)
Finally, the term $T_4$ in (D.8) is manipulated to yield

$$
T_4 = - \left( uu_1 \hat{b} \cdot \frac{\partial \Gamma_R^{(1)}}{\partial u} \right) + \left( \mu_1 \frac{\partial^{2} \phi}{\partial \mu \partial \theta} + \theta_1 \frac{\partial^{2} \phi}{\partial \theta^{2}} \right) \left( H^{(1)} - u \hat{b} \cdot \Gamma_R^{(1)} + B \Gamma_{\phi}^{(1)} \right). 
$$

(D.19)

Using equation (53) to write $H^{(1)} - u \hat{b} \cdot \Gamma_R^{(1)} + B \Gamma_{\phi}^{(1)} = -B(\partial S_p^{(2)}/\partial \theta) + (H^{(1)} - u \hat{b} \cdot \Gamma_R^{(1)} + B \Gamma_{\phi}^{(1)})$ and equation (48) to write $\hat{b} \cdot \Gamma_R^{(1)} = -u_1 + \hat{b} \cdot \Gamma_R^{(1)}$, and employing that $\langle u_1 \rangle = 0$ and $\langle \mu_1 \rangle = 0$, we obtain

$$
T_4 = \frac{u}{2} \frac{\partial^{2} S_p^{(2)}}{\partial u} - B \left( \frac{\partial^{2} S_p^{(2)}}{\partial \mu \partial \theta} + \theta_1 \frac{\partial^{2} S_p^{(2)}}{\partial \theta^{2}} \right). 
$$

(D.20)

Using equations (49) and (50), we find

$$
T_4 = \frac{u}{2} \frac{\partial^{2} S_p^{(2)}}{\partial u} - B \left( \frac{\partial^{2} S_p^{(2)}}{\partial \mu \partial \theta} \right) - B \left( \frac{\partial^{2} S_p^{(2)}}{\partial \theta^{2}} \right)^2. 
$$

(D.21)

Here we have used

$$
\left\{ \frac{\partial^{2} S_p^{(2)}}{\partial \theta} \frac{\partial^{2} S_p^{(2)}}{\partial \mu \partial \theta} \right\} = 2 \left( \frac{\partial^{2} S_p^{(2)}}{\partial \theta} \frac{\partial^{2} S_p^{(2)}}{\partial \mu \partial \theta} \right) = \frac{\partial}{\partial \mu} \left( \frac{\partial^{2} S_p^{(2)}}{\partial \theta} \right)^2. 
$$

(D.22)

where we have integrated by parts in $\theta$ to obtain the first equality.

Substituting the results $H^{(1)} = \Lambda \langle \phi \rangle + \Lambda \tilde{\phi}$, (D.16), (D.17), (D.18) and (D.21) into equation (D.3) gives

$$
\tilde{H}^{(2)} = \frac{-u^{2} \mu}{2B^{2}} b \cdot \nabla_{R} b \cdot \nabla_{R} b + \frac{\mu^{2}}{4B^{2}} (I - \hat{b} \hat{b}) \cdot \nabla_{R} \nabla_{R} B \cdot \hat{b} - \frac{3\mu^{2}}{4B^{2}} |\nabla_{R} B|^{2}
$$

$$
+ \frac{u^{2} \mu}{2B} \nabla_{R} \hat{b} \cdot \nabla_{R} b - \frac{u^{2} \mu}{2B} (\nabla_{R} \cdot \hat{b})^{2} + \frac{\Lambda}{\lambda} (\nabla_{R} \cdot \nabla_{R} \phi) + \frac{\Lambda}{\lambda} (\nabla_{R} \cdot \nabla_{R} \phi) + \frac{\Lambda}{\lambda} (\nabla_{R} \cdot \nabla_{R} \phi)
$$

$$
= \frac{\mu^{2}}{2B^{2}} \nabla_{R} b \cdot \nabla_{R} \phi + \frac{\mu^{2}}{2B^{2}} \nabla_{R} \phi \cdot \nabla_{R} B
$$

$$
- \frac{\Lambda \mu}{\lambda} B^{2} \left( \frac{\partial^{2} S_p^{(2)}}{\partial \mu \partial \theta} \right) - \frac{\Lambda \mu}{\lambda} B^{2} \left( \frac{\partial^{2} S_p^{(2)}}{\partial \theta^{2}} \right)^2. 
$$

(D.23)

Employing equations (49), (51), (52) and (69) to write

$$
B \frac{\partial^{2} S_p^{(2)}}{\partial \phi} = -\Lambda \tilde{\phi} - u u_1 - B \Gamma_{\phi}^{(1)},
$$

(D.24)

we find

$$
- B \left( \frac{\partial^{2} S_p^{(2)}}{\partial \mu \partial \theta} \right) - \frac{B}{2} \frac{\partial}{\partial \mu} \left( \frac{\partial^{2} S_p^{(2)}}{\partial \theta} \right)^2 = \frac{\Lambda^{2}}{\lambda^{2} B} \frac{\partial}{\partial \mu} \left( \frac{\partial \phi}{\partial \mu} \right)
$$

$$
- \frac{\Lambda u}{\lambda^{2} B} \left( \frac{\partial \phi}{\partial \mu} \right) u_1 - \frac{\Lambda u}{B} \left( \frac{\partial \phi}{\partial \mu} \right) = \Lambda \left( \frac{\partial \phi}{\partial \mu} \right) - u \left( \frac{\partial \phi}{\partial \mu} \right). 
$$

(D.25)
Substituting this result into (D.23) gives equation (132) with $\Psi^{(2)}_\theta$ and $\Psi^{(2)}_{\phi B}$ as in (133) and (134), and $\Psi^{(2)}_B$ given by

$$
\Psi^{(2)}_B = -\frac{u^2\mu}{2B^2} \hat{b} \cdot \nabla_R \hat{b} \cdot \nabla_R B + \frac{\mu^2}{4B^2} (\hat{b} \cdot \nabla_R \hat{b}) \cdot \nabla_R \nabla_R B \cdot \hat{b} - \frac{3\mu^2}{4B^2} |\nabla_R \nabla_R B|^2 + \frac{u^2\mu}{2B} \nabla_R \hat{b} \cdot \frac{u^2\mu}{2B} (\nabla_R \hat{b}^2) + \frac{u}{2} \frac{\partial (u_i^2)}{\partial \mu} - \frac{u^2}{2B^2} \frac{\partial (u_i^2)}{\partial \mu},
$$

To obtain equation (135) from equation (D.26) we only need to calculate $\langle u_i^2 \rangle$ and $\langle u_i (\partial \Gamma_i^{(1)} / \partial \mu) \rangle$. The gyroaverage of $u_i^2$ is

$$
\langle u_i^2 \rangle = \frac{u^2\mu}{B} \hat{b} \cdot \nabla_R \hat{b}^2 + \frac{B^2}{4} (\langle (\rho \times \hat{b}) \cdot \nabla_R \hat{b} \cdot \rho \rangle^2) - \frac{\mu^2}{4} (\hat{b} \cdot \nabla_R \hat{b})^2,
$$

where we have used the definition of $u_1$ in (111), we have taken the gyroaverage $\langle \rho \rho \rangle$ from (102), and we have employed equation (107) to write $[\rho (\rho \times \hat{b}) + (\rho \times \hat{b}) \rho] : \nabla_R \hat{b} = 2(\rho \times \hat{b}) \cdot \nabla_R \hat{b} - 2 \mu / B b \cdot \nabla_R \hat{b}$ and hence obtain

$$
\langle [\rho (\rho \times \hat{b}) + (\rho \times \hat{b}) \rho] : \nabla_R \hat{b}^2 \rangle = 4 \langle (\rho \times \hat{b}) \cdot \nabla_R \hat{b} \cdot \rho \rangle^2 - \frac{4\mu^2}{B^2} (\hat{b} \cdot \nabla_R \hat{b})^2.
$$

The gyroaverage of the second term in (D.27) is given by

$$
\langle [\rho \times \hat{b}] \cdot \nabla_R \hat{b} \cdot \rho \rangle^2 = \frac{\mu^2}{2B^2} (\hat{b} \cdot \nabla_R \hat{b})^2 + \frac{\mu^2}{2B^2} \nabla_R \hat{b} : (\nabla_R \hat{b})^T + \frac{\mu^2}{2B^2} \nabla_R \hat{b} : (\nabla_R \hat{b})^T = \frac{\mu^2}{2B^2} (\hat{b} \cdot \nabla_R \hat{b})^2 + \frac{\mu^2}{2B^2} (\hat{b} \cdot \nabla_R \hat{b})^2 + \frac{\mu^2}{2B^2} \nabla_R \hat{b} : (\nabla_R \hat{b})^T
$$

where we have used

$$
\langle \rho \rho \rangle \rho \rho \rangle = \frac{\mu^2}{2B^2} [\delta_{ij} - \hat{b}_i \hat{b}_j] (\delta_{kl} - \hat{b}_k \hat{b}_l) + (\delta_{ij} - \hat{b}_i \hat{b}_j) (\delta_{kl} - \hat{b}_k \hat{b}_l) + (\delta_{ij} - \hat{b}_i \hat{b}_j) (\delta_{kl} - \hat{b}_k \hat{b}_l).
$$

Here $\delta_{ij}$ is the Kronecker delta. We have employed $\hat{b} \times \nabla_R \hat{b} \times \hat{b} = (\nabla_R \hat{b})^T - (\nabla_R \hat{b}) (\hat{b} \times \nabla_R \hat{b})^T - (\hat{b} \times \nabla_R \hat{b}) (\hat{b} \times \nabla_R \hat{b})^T$ to obtain the second equality in (D.29). Substituting equation (D.29) into equation (D.27) gives

$$
\langle u_i^2 \rangle = \frac{u^2\mu}{B} \hat{b} \cdot \nabla_R \hat{b}^2 + \frac{\mu^2}{4} \nabla_R \hat{b} : (\nabla_R \hat{b})^T - \frac{\mu^2}{8} (\nabla_R \hat{b})^2 - \frac{\mu^2}{8} (\hat{b} \cdot \nabla_R \hat{b})^2.
$$

The gyroaverage of $u_1 (\partial \Gamma_1^{(1)} / \partial \mu)$ is

$$
\langle u_1 (\partial \Gamma_1^{(1)} / \partial \mu) \rangle = \frac{u^2\mu}{B} \hat{b} \cdot \nabla_R \hat{b} \cdot \nabla_R B.
$$

Finally, substituting equations (D.31) and (D.32) into (D.26) gives (135).
Appendix E. Poisson bracket

In this appendix we prove that the Poisson bracket that corresponds to the gyrokinetic Lagrangian in (136) is (139). Since the symplectic part\(^{10}\) of the gyrokinetic Lagrangian (136) is exactly the same as in [25], the Poisson bracket in gyrokinetic coordinates will also be.

As explained in section 3.1, to obtain the Poisson bracket, given in (22), we need to calculate the inverse of the matrix \(L\) in (20). We explicitly write this matrix by writing the gyrokinetic coordinates as \(\{Z_α\}_{α=1}^6\), with \((Z_1, Z_2, Z_3) = \{R, Z^4 = u, Z^5 = µ\) and \(Z_6 = \theta\). The gyrokinetic Lagrangian (136) is written as in (29), with \((\bar{Γ}_1, \bar{Γ}_2, \bar{Γ}_3) = \epsilon^{-1} A(R) + u \hat{b}(R) - \epsilon µ K(R), \bar{Γ}_4 = 0, \bar{Γ}_5 = 0\) and \(\bar{Γ}_6 = -\epsilon µ\). Then, using (20) for the gyrokinetic Lagrangian, we find that the matrix \(L\) is given by

\[
L_{αβ} := \frac{\partial Γ_β}{\partial Z_α} - \frac{\partial Γ_α}{\partial Z_β},
\]

or in matrix form

\[
L = \left(\begin{array}{cccc}
-\epsilon^{-1} B^* \times \hat{I} & \hat{b} & \epsilon K & 0 \\
-\epsilon^{-1} B^* \times \hat{I} & \hat{b} & \epsilon K & 0 \\
\epsilon^{-1} K & 0 & 0 & 0 \\
0 & 0 & 0 & \epsilon^{-1} \end{array}\right),
\]

where \(B^*\) is defined in (140). Its inverse is given by

\[
P = L^{-1} = \frac{1}{B^*_||} \left(\begin{array}{cccc}
\epsilon \hat{b} \times \hat{I} & B^* & 0 & \epsilon K \times \hat{b} \\
\epsilon \hat{b} \times \hat{I} & B^* & 0 & \epsilon K \times \hat{b} \\
-\epsilon K \times \hat{b} & 0 & 0 & B^* \cdot \hat{K} \\
-\epsilon K \times \hat{b} & 0 & 0 & B^* \cdot \hat{K} \end{array}\right),
\]

where \(B^*_||\) is defined in (141). It is easy to check by direct calculation that \(P\) is the inverse of \(L\).

The Poisson bracket of two functions \(F(Z)\) and \(G(Z)\) is then given by equation (22) that can be compactly rewritten as (139).

Appendix F. Calculation of the Jacobian

In this appendix we show that the determinant of the Jacobian matrix of the gyrokinetic transformation is \(B^*_||\), defined in (141). This result coincides with the results in [25] because of our choice for the final form of the Lagrangian (136).

To obtain the Jacobian of the gyrokinetic transformation, we use the matrix \(L\), defined in (20). This matrix is defined for both the original coordinates \(\{r, v\}\) and the new gyrokinetic coordinates \(\{R, u, µ, \theta\}\). The matrix \(L\) in the original phase space and the matrix \(L\) in the new gyrokinetic phase space are related by the Jacobian matrix of the gyrokinetic transformation

\(^{10}\) It is common to call the symplectic part of a phase-space Lagrangian to the piece linear in the time derivatives of the coordinates.
\( T_\epsilon \). It is possible to use this relation to calculate the determinant of the Jacobian matrix by calculating the matrix \( L \) in both coordinate systems.

We denote the original coordinates by \( \{X^\alpha\}_{\alpha=1}^6 \) with \( (X^1, X^2, X^3) = r \) and \( (X^4, X^5, X^6) = v \), and the gyrokinetic coordinates by \( \{Z^\alpha\}_{\alpha=1}^6 \) with \( (Z^1, Z^2, Z^3) = R \), \( Z^4 = u \), \( Z^5 = \mu \) and \( Z^6 = \theta \). The Jacobian matrix of the transformation is given by

\[
(J_{T_\epsilon})^\alpha_\beta(Z) = \frac{\partial X^\alpha(Z)}{\partial Z^\beta}.
\]

We write the Lagrangian in the coordinates \( X \) as

\[
L^X = \sum_{\alpha=1}^6 \gamma_\alpha(X) \frac{dX^\alpha}{dt} - H^X(X, t),
\]

with \( (\gamma_1, \gamma_2, \gamma_3) = \epsilon^{-1} A(r) + v, \gamma_4 = 0, \gamma_5 = 0 \) and \( \gamma_6 = 0 \). We write the Lagrangian \( \overline{L} \) in gyrokinetic coordinates \( Z \) as in (29), with \( (\Gamma_1, \Gamma_2, \Gamma_3) = \epsilon^{-1} A(R) + \hat{u} \hat{b}(R) - \epsilon \mu K(R), \Gamma_4 = 0, \Gamma_5 = 0 \) and \( \Gamma_6 = -\epsilon \mu \).

From the Lagrangians \( L^X \) and \( \overline{L} \), we obtain the matrix \( L \), defined in (20), in both coordinate systems, given by (E.1) for the gyrokinetic coordinates, and by

\[
l_{\alpha\beta} := \frac{\partial \gamma_\beta}{\partial X^\alpha} - \frac{\partial \gamma_\alpha}{\partial X^\beta}
\]

for the original coordinates. It is immediate to check that \( L = J_{T_\epsilon}^T l_{T_\epsilon} \), with the superscript \( T \) standing for matrix transposition. It is then obvious that

\[
det(J_{T_\epsilon}) = \sqrt{\det(L) \det(l)},
\]

where we have used that the Jacobian of \( T_\epsilon \) is positive at \( \epsilon = 0 \) to determine the sign in front of the square root. Then, to calculate the Jacobian is enough to calculate the determinants of the matrices \( L \) and \( l \).

The matrix \( l \) is

\[
l = \begin{pmatrix}
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
-\epsilon^{-1} B \times \hat{I} & -\epsilon \hat{I} & I & 0 & \vdots & \vdots \\
I & 0 & \epsilon \hat{b} & \hat{b} & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{pmatrix},
\]

and the matrix \( L \) was given in (E.2). The determinant of \( l \) is

\[
det(l) = 1.
\]

As for \( L \), given in (E.2), we have that

\[
det(L) = \epsilon^2 \begin{vmatrix}
\vdots & \vdots & \vdots & \vdots & \vdots \\
-\epsilon^{-1} B \times \hat{I} & -\epsilon \hat{I} & \hat{b} & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\hat{b} & \vdots & \vdots & \vdots & 0
\end{vmatrix}
\]

\[45\]
Writing the matrix in the reference system $\{\hat{e}_1, \hat{e}_2, \hat{b}\}$, where $\hat{b} = (0, 0, 1)$, we find that

$$\det(L) = \epsilon^2 \begin{vmatrix} 0 & \epsilon^{-1} B^* \cdot \hat{b} & -\epsilon^{-1} B^* \cdot \hat{e}_2 & 0 \\ -\epsilon^{-1} B^* \cdot \hat{b} & 0 & \epsilon^{-1} B^* \cdot \hat{e}_1 & 0 \\ \epsilon^{-1} B^* \cdot \hat{e}_2 & -\epsilon^{-1} B^* \cdot \hat{e}_1 & 0 & 1 \\ 0 & 0 & 0 & \epsilon^{-1} B^* \parallel \end{vmatrix} = \epsilon^2 \begin{vmatrix} 0 & 0 & \epsilon^{-1} B^* \parallel \\ 0 & 0 & 0 \\ \epsilon^{-1} B^* \parallel & 0 & 0 \\ 0 & 0 & \epsilon^{-1} B^* \parallel \end{vmatrix} = (B^* \parallel)^2.$$ (F.8)

Substituting this result and (F.6) into (F.4), we finally obtain

$$\det(J_{\epsilon} L) = B^* \parallel.$$ (F.9)

**Appendix G. Conservation of phase-space volume**

In this appendix we prove that

$$\sum_{\alpha=1}^{6} \frac{\partial}{\partial Z^\alpha} \left[ \det(J_{\epsilon} L) \frac{dZ^\alpha}{dt} \right] = 0.$$ (G.1)

This equation is satisfied by any gyrokinetic Lagrangian $L$ with a symplectic part $\Gamma_{1\alpha}$ that is independent of time, as is in our Lagrangian (136). Relation (G.1) gives equation (153).

To prove (G.1) we use equations (21), (F.4) and (F.6). From (F.4) and (F.6) we find

$$\sum_{\alpha=1}^{6} \frac{\partial}{\partial Z^\alpha} \left[ \det(J_{\epsilon} L) \frac{dZ^\alpha}{dt} \right] = \frac{1}{2\sqrt{\det(L)}} \left\{ \sum_{\alpha=1}^{6} \frac{dZ^\alpha}{dt} \frac{\partial}{\partial Z^\alpha} \left[ \det(L) \right] + 2 \det(L) \sum_{\alpha=1}^{6} \frac{\partial}{\partial Z^\alpha} \left( \frac{dZ^\alpha}{dt} \right) \right\},$$ (G.2)

and using (21) for the gyrokinetic Lagrangian $L$, we obtain

$$\frac{dZ^\alpha}{dt} = \sum_{\beta=1}^{6} (L^{-1})^{\alpha\beta} \frac{\partial H}{\partial Z^\beta}, \quad \alpha = 1, 2, \ldots, 6.$$ (G.3)

Employing that $(L^{-1})^{\alpha\beta} = -(L^{-1})^{\beta\alpha}$, we find that this equation leads to

$$\sum_{\alpha=1}^{6} \frac{\partial}{\partial Z^\alpha} \left( \frac{dZ^\alpha}{dt} \right) = \sum_{\alpha, \beta=1}^{6} \frac{\partial (L^{-1})^{\alpha\beta}}{\partial Z^\alpha} \frac{\partial H}{\partial Z^\beta}.$$ (G.4)

Equation (G.4) can be further simplified by using (A.4), giving

$$\sum_{\alpha=1}^{6} \frac{\partial}{\partial Z^\alpha} \left( \frac{dZ^\alpha}{dt} \right) = - \sum_{\alpha, \beta, \gamma=1}^{6} (L^{-1})^{\alpha\beta} \frac{\partial L_{\beta\gamma}}{\partial Z^\alpha} \frac{dZ^\gamma}{dt},$$ (G.5)

where to obtain the last equality we have used that $L_{\alpha\beta} = -L_{\beta\alpha}$ and $(L^{-1})^{\alpha\beta} = -(L^{-1})^{\beta\alpha}$.

Substituting (G.5) into (G.2) and using that the derivatives of a determinant are

$$\frac{\partial}{\partial Z^\alpha} \det(L) = \det(L) \sum_{\beta, \gamma=1}^{6} (L^{-1})^{\alpha\beta} \frac{\partial L_{\gamma\beta}}{\partial Z^\alpha},$$ (G.6)

finally gives (G.1). To prove that all the terms cancel we just need to use that, as trivially deduced from its definition (20), $L_{\alpha\beta}$ satisfies (A.1) with $n = 3$.}

46
Appendix H. Proof of equation (184)

In this appendix we prove (184) by showing that \( \det(J_{Z_0\to Z_p}(Z_p)) = B^*_{||,p}(Z_p)/B^*_{\|,p}(Z_p, t_0; t) \). First we evaluate the time derivative of \( \det(J_{Z_0\to Z_p}(Z_p)) \), given by

\[
\frac{d}{dt}[\det(J_{Z_0\to Z_p}(Z_p))] = \det(J_{Z_0\to Z_p}(Z_p)) \sum_{\beta,\gamma=1}^6 \frac{\partial Z_\beta^p}{\partial Z_{\gamma}^0} = \det(J_{Z_0\to Z_p}(Z_p)) \sum_{\beta=1}^6 \frac{\partial Z_\beta^p}{\partial Z_{\beta}^0},
\]

where we have employed the formula for the derivative of a determinant in (G.6) and omitted the arguments hoping that no confusion will be caused. Using (153), or its equivalent (G.1), we obtain that

\[
\frac{1}{\det(J_{Z_0\to Z_p}(Z_p))} \frac{d}{dt}[\det(J_{Z_0\to Z_p}(Z_p))] = - \frac{1}{B^*_{||,p}(Z_p(Z_p(t_0), t_0; t))} \frac{d}{dt}[B^*_{||,p}(Z_p(Z_p(t_0), t_0; t))],
\]

that is, the product \( B^*_{||,p}(Z_p(Z_p(t_0), t_0; t)) \) \( \det(J_{Z_0\to Z_p}(Z_p)) \) is constant in time. Since for \( t = t_0 \) the map \( Z_p(Z_p(t_0), t_0; t) \) is the identity, giving \( \det(J_{Z_0\to Z_p}(Z_p)) = 1 \), we find that the constant is \( B^*_{||,p}(Z_p(t_0)) \), implying that \( \det(J_{Z_0\to Z_p}(Z_p(t_0))) = B^*_{||,p}(Z_p(t_0))/B^*_{||,p}(Z_p(Z_p(t_0), t_0; t)) \).

Appendix I. Manipulations leading to equation (186)

In this appendix we obtain equation (186) from (185).

First, we evaluate the variations with respect to \( \psi(r, t) \) in (185) term by term. For the first term in (185), we find

\[
\int dt \, d^3r \nabla \psi(r, t) \cdot \nabla \psi(r, t) = - \int dt \, d^3r \, \delta \psi(r, t) \nabla^2 \psi(r, t),
\]

where we have integrated by parts and we have taken into account that \( \delta \psi = 0 \) at the boundary. Using relation (70) and equation (162) to write

\[
\phi_p(R_p, \mu_p, \theta_p, t) = \int d^3r \, \delta \left( R_p + \frac{\epsilon_s}{\lambda_p} \rho(R_p, \mu_p, \theta_p) - r \right) \psi(r, t),
\]

we find

\[
\delta_s(\phi_p(R_p, \mu_p, \theta_p, t)) = \int d^3r \, \delta \psi(r, t) \left( \delta \left( R_p + \frac{\epsilon_s}{\lambda_p} \rho - r \right) \right).
\]

Using relation (1.2) again and the identity

\[
\frac{1}{2} \delta \psi(\nabla (R_{p,i} / \epsilon_s) \tilde{\Phi}_p) \cdot (\dot{b} \times \nabla (R_{p,i} / \epsilon_s) \tilde{\Phi}_p) = \frac{1}{2} \nabla (R_{p,i} / \epsilon_s)(\delta \Phi_2) \cdot (\dot{b} \times \nabla (R_{p,i} / \epsilon_s) \tilde{\Phi}_p)
\]

\[
- \frac{1}{2} \nabla (R_{p,i} / \epsilon_s) \delta \Phi_2 \cdot (\dot{b} \times \nabla (R_{p,i} / \epsilon_s) \tilde{\Phi}_p)
\]

\[
= - \nabla (R_{p,i} / \epsilon_s)(\delta \Phi_2) \cdot (\dot{b} \times \nabla (R_{p,i} / \epsilon_s) \tilde{\Phi}_p),
\]

where to obtain the last equality we have integrated by parts in \( \theta_p \) and we have used that \( \partial \Phi_2 / \partial \theta_p = \tilde{\Phi}_p \) and that \( \partial (\delta \Phi_2) / \partial \theta_p = \delta \phi_p \), we obtain

\[
\Lambda p \delta \phi_2 \Psi_\Phi^{(2)}(R_p, \mu_p, \theta_p, t) + \delta \phi_2 \Psi_\Phi^{(2)}_{B_B}(R_p, u_p, \mu_p, \theta_p, t)
\]

\[
= \int d^3r \, \delta \psi(r, t) \left( \frac{1}{\lambda_p} \nabla R_p(R_p, u_p, \mu_p, \theta_p, t) \cdot \nabla (R_{p,i} / \epsilon_s) \delta \left( R_p + \frac{\epsilon_s}{\lambda_p} \rho - r \right) \right)
\]

\[
+ V_{\mu_p}(R_p, u_p, \mu_p, \theta_p, t) \frac{\partial}{\partial \mu_p} \left[ \delta \left( R_p + \frac{\epsilon_s}{\lambda_p} \rho - r \right) \right]
\]

\[
+ V_{\theta_p}(R_p, u_p, \mu_p, \theta_p, t) \delta \left( R_p + \frac{\epsilon_s}{\lambda_p} \rho - r \right).
\]
In this equation we have separated the different terms into three types: the terms that contain the gradient of the delta function, the terms that contain the derivative with respect to $\mu_p$ of the delta function, and the terms that contain the delta function. The coefficient multiplying the gradient of the delta function is

$$V_{\mu_p}(R_p, u_p, \mu_p, \theta_p, t) = -\frac{\Lambda_p}{\lambda_p^2 B^2} \hat{b} \times \nabla (R_p) \frac{\delta}{\delta \phi_p} - \frac{u_p}{B} \hat{b} \times \nabla R_p \hat{b} \cdot \rho - \frac{\mu_p}{2B^2} \nabla R_p \hat{b}$$

the coefficient multiplying the derivative with respect to $\mu_p$ of the delta function is

$$V_{\mu_p}(R_p, u_p, \mu_p, \theta_p, t) = -\frac{\Lambda_p}{B} \frac{\delta}{\delta \phi_p} - \frac{u_p}{B} \hat{b} \cdot \nabla R_p \hat{b} \cdot \rho + \frac{\mu_p}{4} \nabla R_p \hat{b} : (\rho (\rho \times \hat{b}) + (\rho \times \hat{b}) \rho)$$

and the coefficient multiplying the delta function is

$$V_{\theta_p}(R_p, u_p, \mu_p, \theta_p, t) = -\frac{\Lambda_p}{\lambda_p^2 B} \frac{\delta}{\delta \phi_p} - \frac{1}{B} \nabla R_p \hat{b} \cdot \rho - \frac{u_p^2}{2\mu_p B} \hat{b} \cdot \nabla R_p \hat{b} \cdot \rho + \frac{\mu_p}{4\mu_p} \nabla R_p \hat{b} : (\rho (\rho \times \hat{b}) + (\rho \times \hat{b}) \rho).$$

Substituting (1.1), (1.3) and (1.5) into (185), and employing (183) and (184) to write the integrals as integrals over $Z_p$ and not over $Z_{\rho\theta}$, we find that the variations are of the form given in (186) with

$$\mathcal{P}(r, t) = -\frac{\alpha^2}{L^2} \nabla^2 \psi(r, t)$$

Note that we did not need to keep the gyroaverages in (1.3) and (1.5) because there is an overall integral in $\theta_p$ and neither $B_{11, p}$ nor $F_p$ depend on $\theta_p$.

We have only left to prove that (1.9) is equal to (187). Using the result in (110) and employing (102) and (107) to write

$$\frac{\mu_p}{B} \nabla R_p \hat{b} + \frac{1}{2} [\rho \rho - (\rho \times \hat{b})(\rho \times \hat{b})] \cdot \nabla R_p \hat{b} = \rho \rho \cdot \nabla R_p \hat{b},$$

it is clear that $V_{R_p}$ is exactly the perpendicular component of $R_p$. Using (112), we find that $V_{\mu_p} = \mu_{p, 1}$. It only remains to use that $V_{\theta_p} = -\partial \theta_{p, 1} / \partial \theta_p$, with $\theta_{p, 1}$ given in (113). To obtain this identity we have used $\rho = \delta (\rho \times \hat{b}) / \partial \theta_p$, $\rho (\rho \times \hat{b}) + (\rho \times \hat{b}) \rho = -1/2 \delta [\rho \rho - (\rho \times \hat{b})(\rho \times \hat{b})] / \partial \theta_p$, and $\hat{\phi}_p = \delta \hat{\Phi}_p / \partial \theta_p$. Using $V_{\theta_p} = -\partial \theta_{p, 1} / \partial \theta_p$, and integrating by parts in $\theta_p$ in (1.9) gives the final form (187).

References

[1] Catto P J 1978 Plasma Phys. 20 719
[2] Dimits A M, Williams T J, Byers J A and Cohen B I 1996 Phys. Rev. Lett. 77 71
[3] Dorland W, Jenko F, Kotschenreuther M and Rogers B N 2000 Phys. Rev. Lett. 85 5579
[4] Dannert T and Jenko F 2005 Phys. Plasmas 12 072309
[5] Candy J and Waltz R E 2003 J. Comput. Phys. 186 545
[6] Chen Y and Parker S E 2003 J. Comput. Phys. 189 463
[7] Peeters A G and Strinzi D 2004 Phys. Plasmas 11 3748
[8] Romanelli F 1989 Phys. Fluids B 1 1018
[9] Dimits A M, LoDestr o L L and Dubin DHE 1992 Phys. Fluids B 4 274
[10] Dimits A M 2010 Phys. Plasmas 17 055901
[11] Hazeltine R D 1973 Plasma Phys. 15 77
[12] Hinton F L and Hazeltine R D 1976 Rev. Mod. Phys. 48 239
[13] Helander P and Sigmar D J 2002 Collisional Transport in Magnetized Plasmas (Cambridge Monographs on Plasma Physics) ed Haines M G et al (Cambridge: Cambridge University Press)
[14] Lee X S, Myra J R and Catto P J 1983 Phys. Fluids 26 223
[15] Lee W W 1983 Phys. Fluids 26 556
[16] Bernstein I B and Catto P J 1985 Phys. Fluids 28 1342
[17] Parra F I and Catto P J 2008 Plasma Phys. Control. Fusion 50 065014
[18] Littlejohn R G 1981 Phys. Fluids 24 1730
[19] Cary J R 1981 Phys. Rep. 79 129
[20] Cary J R and Littlejohn R G 1983 Ann. Phys. 151 1
[21] Littlejohn R G 1983 J. Plasma Phys. 29 111
[22] Littlejohn R G 1985 Phys. Fluids 28 2015
[23] Dubin D H E, Krommes J A, Oberman C and Lee W W 1983 Phys. Fluids 26 3524
[24] Hahn T S 1988 Phys. Fluids 31 2670
[25] Brizard A J and Hahn T S 2007 Rev. Mod. Phys. 79 421
[26] Heikkinen J A, Jauhunen S J, Kiviniemi T P and Ogando F 2008 J. Comput. Phys. 227 5582
[27] Grandgirard V et al 2006 J. Comput. Phys. 217 395
[28] Xu Q et al 2007 Nucl. Fusion 47 809
[29] Chang C S and Ku S 2008 Phys. Plasmas 15 062510
[30] Parra F I and Catto P J 2009 Plasma Phys. Control. Fusion 51 095008
[31] Lee W W and Kolesnikov R A 2009 Phys. Plasmas 16 044506
[32] Parra F I and Catto P J 2009 Phys. Plasmas 16 124701
[33] Parra F I and Catto P J 2010 Phys. Plasmas 17 056106
[34] Parra F I and Catto P J 2010 Plasma Phys. Control. Fusion 52 085011
[35] Parra F I and Catto P J 2010 Plasma Phys. Control. Fusion 52 045004
[36] Parra F I and Catto P J 2010 Plasma Phys. Control. Fusion 52 059801
[37] Peeters A G and Angioni C 2005 Phys. Plasmas 12 072515
[38] Parra F I, Barnes M and Peeters A G 2011 Symmetry arguments for turbulent transport of toroidal angular momentum in local δf gyrokinetic models Phys. Plasmas submitted
[39] Goldstein H, Poole C P and Safko J L 2002 Classical Mechanics 3rd edn (San Francisco, CA: Addison-Wesley) p 353
[40] Sugiyama L E 2008 Phys. Plasmas 15 092112
[41] Krommes J A 2009 Phys. Plasmas 16 084701
[42] Sugiyama L E 2009 Phys. Plasmas 16 084702
[43] Littlejohn R G 1988 Phys. Rev. A 38 6034
[44] Baños A 1967 J. Plasma Phys. 1 305
[45] Brizard A J 2004 Phys. Plasmas 11 4429
[46] Low F E 1958 Proc. R. Soc. Lond. A 248 282
[47] Sugama H 2000 Phys. Plasmas 7 466
[48] Brizard A J 2000 Phys. Plasmas 7 4816