Gyrokinetics

Xavier Garbet, Maxime Lesur

Gyrokinetics

X. Garbet[1,2] and M. Lesur[3]

[1]CEA, IRFM, Saint-Paul-lez-Durance, F-13108, France
[2]School of Physical and Mathematical Sciences, Nanyang Technological University, 637371 Singapore
[3]Institut Jean Lamour, UMR 7198 CNRS-Université de Lorraine, 54000 Nancy, France

February 6, 2023

Abstract

Gyrokinetics is a self-consistent kinetic model of magnetised plasmas that applies to dynamical systems characterised by typical frequencies lower than the cyclotron frequency. Any gyrokinetic theory proceeds in two steps. The first one is the derivation of a gyrokinetic Vlasov equation for each charged species. This is done by building a new adiabatic invariant of motion, the magnetic moment, associated with a virtual particle, the gyrocentre, slightly shifted from the particle guiding-centre. It relies on a near-identity change of variables in a 8D extended phase space. This change of variables is not unique. Several options are discussed in this lecture note. The second part is the derivation of particle charge and current densities that enter the Maxwell equations, knowing the gyrocentre distribution functions. It appears that a magnetised plasma behaves as a medium that is both electrically polarised and magnetised. The resulting model encompasses one kinetic equation per species and the Maxwell equations. It can be used to address any self-consistent electromagnetic problem in magnetised plasmas, in particular instabilities and turbulent transport.

Sections labelled with a star "*" can be skipped in a first reading. Notations can be found in Appendix A.

1 Introduction

This note aims at deriving a theory of low frequency electromagnetic fluctuations in a magnetised plasma. Let us remind that 3 invariants of motion can be built in a magnetised plasma under stringent conditions. These invariants are the particle energy if the field is static, magnetic moment for a strong enough magnetic field, and canonical momentum whenever the configuration is left invariant by a rotational symmetry. However dynamical invariance breaks down in presence of fluctuations. This is a consequence of symmetry losses. For instance time variations invalidates energy conservation. A loss of axisymmetry in a tokamak implies that the kinetic toroidal momentum is no longer an invariant of motion. Also the particle magnetic moment is no longer an invariant whenever field perturbations oscillate on time scales commensurate with the cyclotron period. Losses of invariants of motion imply that particle trajectories are generically chaotic. Predicting properties of a chaotic dynamics is a formidable task since the dynamics dimensionality is intrinsically 6D in the phase space. Fortunately, some simplification is possible since most MHD instabilities, or micro-instabilities that drive turbulence in laboratory magnetic configurations, are characterised by frequencies that are lower than the cyclotron frequency.

A new invariant of motion, close to the magnetic moment, can then be constructed, hence reducing the number of degrees of freedom. Nevertheless the other coordinates must also be modified in this process. This new set of coordinates describes the motion of a virtual particle called gyrocentre, combined with a distorted cyclotron motion (see Fig.1). The change of coordinates, from guiding-centre to gyrocentre, is called a gyrocentre transform. A gyrocentre stays close to the guiding-centre position, provided the amplitude of fluctuations is small, say of the order of the expansion parameter ϵ_B (defined in subsection 2.3) that is used to derive the guiding-centre equations of motion. This assumption is fulfilled in most practical cases[1]. Technically speaking, this means that a near identity change of variables, conjugate or not, must be performed. It turns out that powerful methods were developed in celestial mechanics to perform effectively near identity changes of variables, at all orders in the expansion parameter. Kinetic equations that rule the particle distribution functions can be computed based on this new set of variables.

However the task is not over. Indeed solving kinetic equations provides all necessary information on moments, i.e. density, fluid velocity and temperature, and transport equations, but it does not ensure self-consistency, i.e. coherence of the Maxwell equations with the plasma response to the electromagnetic field. More precisely, electromagnetic fluctuations must be solutions of the Maxwell equations, with charge and currents densities that depend on the electromagnetic field itself via the particle distribution functions. Here a major difficulty emerges. Maxwell equations live in the physical space, not in the space of gyrocentre coordinates. Hence a second part of the gyrokinetic theory consists in relating the charge and current densities in the physical space to the moments of the gyrocentre distribution function. It turns out they differ via terms that can be seen as polarisation and magnetisation terms. This second step is conveniently achieved by using a variational method. The gyrokinetic theory was initiated by Frieman and Chen in 1982 [1], and was then largely expanded and improved upon. This lecture note derives from an extensive overview on gyrokinetics [2], and the papers [3, 4] for the specific "Hamiltonian approach". Readers interested in advanced issues should read them in depth. Numerical implementation and results of simulations are not addressed in this note, but some details can be found in the review [5].

2 Choice of coordinates and ordering

2.1 Canonical coordinates

The simplest[2] set of Hamiltonian canonical (conjugate) variables is the set of position/momentum $\mathbf{z} = (\mathbf{x}, \mathbf{p})$. The equations of equilibrium motion expressed in canonical variables bear the Hamiltonian form

$$\begin{aligned} \frac{d\mathbf{x}}{dt} &= \frac{\partial h_{eq}}{\partial \mathbf{p}} \\ \frac{d\mathbf{p}}{dt} &= -\frac{\partial h_{eq}}{\partial \mathbf{x}} \end{aligned} \tag{1}$$

[1]Exceptions exist however like turbulence in the edge of a magnetic configuration, where the turbulence intensity is high, or large amplitude MHD instabilities.

[2]An alternate set of Hamiltonian canonical variables can be built, which is suitable to a toroidal magnetic configuration. The first set of variables are $\frac{m_a}{e_a}\mu$ (m the mass, e the charge and μ the magnetic moment), and the gyroangle ς - these two variables are conjugate. Using Boozer coordinates, a canonical poloidal momentum is defined that is conjugate to the poloidal angle. Finally the toroidal angle is conjugate to the canonical toroidal momentum. If 3 invariants of the equilibrium motion can be found, and trajectories are bounded, a set of angle/action variables can be constructed. This procedure [6] is not detailed here.

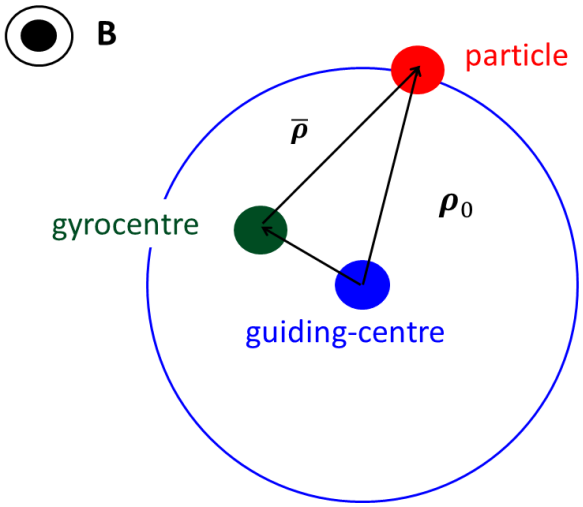


Figure 1: Schematic trajectory separated in a guiding-centre $\mathbf{X}$ and cyclotron motion $\boldsymbol{\rho}$, taken circular. In presence of a perturbed electromagnetic field, the magnetic moment is no longer an (adiabatic) invariant of motion. To recover an invariant of motion, the guiding-centre position is slightly shifted to become a "gyrocentre". The particle position $\mathbf{x}$ is the sum of the gyrocentre position $\bar{\mathbf{X}}$ and a gyrovector $\bar{\boldsymbol{\rho}}$. The later is the sum of a circular cyclotron motion plus a deformation due to higher harmonics and also a mean part responsible for plasma polarisation and magnetisation effects.

where $h_{eq}(\mathbf{x},\mathbf{p})$ is the unperturbed Hamiltonian[3]. A convenient simplification is the use of an *extended set of conjugate variables*, by adding a pair of conjugate variables $\left(x^0 = t, p_0 = -E\right)$, where E is the particle energy[4]. The new Hamiltonian is $\mathtt{h}_{eq}(\boldsymbol{\mathfrak{z}}) = h_{eq}\left(\mathbf{x},\mathbf{p}\right) - E$, where $\boldsymbol{\mathfrak{z}} = (x^\mu, p_\mu)$ with $\mu = (0,1,2,3)$, x^μ are contravariant components, and p_μ are covariant components[5]. The new flow "time" (similar to a curvilinear coordinate for the orbit in extended phase-space) is noted τ, and the motion is constrained to take place on the manifold $E = h_{eq}\left(\mathbf{x},\mathbf{p}\right)$. The equilibrium evolution equations for the new variables (x^0, p_0) are

$$\begin{aligned}\frac{dx^0}{d\tau} &= \frac{\partial \mathtt{h}_{eq}}{\partial p_0} = -\frac{\partial \mathtt{h}_{eq}}{\partial E} = 1\\ \frac{dp_0}{d\tau} &= -\frac{\partial \mathtt{h}_{eq}}{\partial x^0} = -\frac{\partial \mathtt{h}_{eq}}{\partial t} = 0\end{aligned}$$

The first equation imposes that the coordinate t coincides with τ, up to an initial condition. The last equation is the expected time evolution of the Hamiltonian $\frac{dh}{dt} = \frac{\partial h}{\partial t}$, which states that energy is conserved for an Hamiltonian that is time independent. Hence this extended set of conjugate variables verifies the same canonical Hamilton equations as Eq.(1), but applied to 8 conjugate variables, and with t replaced by τ. An alternative expression of the Hamiltonian equations of motion is

$$\frac{d\boldsymbol{\mathfrak{z}}}{d\tau} = \mathtt{J}\cdot\frac{\partial \mathtt{h}}{\partial \boldsymbol{\mathfrak{z}}}$$

[3]Notations are such that small letters are used for fields in the $(\mathbf{x},\mathbf{p})$ space. Capital letters are used for the guiding-centre space, and over-lined capital letters for gyrocentres. See details in Appendix A

[4]This extension is useful for treating a time-dependent electromagnetic field A^μ. Indeed the variation of energy, $dE/dt = e_a v_\mu \partial_t A^\mu$, then appears naturally as part of Hamilton's equations in the extended set of conjugate variables.

[5]Although the notations are similar to those of relativity, relativistic effects are not addressed here.

where $\mathtt{J}$ is called "symplectic" matrix. For a set of canonical variables, the symplectic matrix reads

$$\mathtt{J} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

where I is the $N = 4$ identity matrix[6].

2.2 Non-canonical coordinates

Let us consider a toroidal equilibrium magnetic configuration. A set of guiding-centre coordinates can be constructed if the magnetic field is strong enough. This represents a first change of coordinates from the canonical set of position/momentum $\mathbf{z} = (\mathbf{x}, \mathbf{p})$ to a new set of guiding-centre variables $\mathbf{Z}$. The latter is usually built as a set of non canonical variables, for instance $(\mathbf{X}, u_\parallel, \mu, \varsigma)$, where $\mathbf{X}$ is the position of the guiding-centre, $u_\parallel$ its parallel velocity, μ the magnetic moment and ς the gyroangle (or cyclotron angle). This set of 6 coordinates can be split into a *reduced set* of "slow" guiding-centre coordinates, noted $\mathbf{Z}^* = (\mathbf{X}, u_\parallel, \mu)$, and the gyroangle ς. "Slow" means that these coordinates characterise the part of the motion that evolves on time scales much longer than a cyclotron period, in contrast with the gyroangle ς, dubbed here "fast" coordinate. The corresponding extended set of variables is $\boldsymbol{\mathcal{Z}} = \big(\mathbf{X}, u_\parallel, \mu, \varsigma, E, t\big)$. For such non-conjugate variables, the symplectic matrix contains many more non zero elements. The calculation of $\mathtt{J}$ is an important part of the derivation of gyrokinetic equations. The matrix $\mathtt{J}$ appears in the equations of motion (as shown in appendix B),

$$\frac{d\boldsymbol{\mathcal{Z}}}{d\tau} = \mathtt{J} \cdot \frac{\partial \mathcal{H}}{\partial \boldsymbol{\mathcal{Z}}}$$

and in the Vlasov equation (as shown in subsection 3.4). Each element $\mathtt{J}^{\mu\nu}$ is in fact a Poisson bracket $\{\mathcal{Z}^\mu, \mathcal{Z}^\nu\}$[7]. The best procedure to compute these elements is to devise a least action principle for the equations of motion, and find its extremum with respect to all variations near the actual particle trajectory (see note on particle trajectories). This procedure is detailed in section 3.

It is useful to familiarize oneself with the concepts of phase-space Lagrangian, Lagrange matrix ω and Poisson matrix J, and variational formulation in non-canonical coordinates. Appendix B is a introduction to these concepts, and a simple example which does not involve cyclotron motion, nor magnetic field, nor extended phase-space.

2.3 Ordering

Let us recall that the usual small parameter for the guiding-centre theory is $\epsilon_B = \rho_c / L_B$, where ρ_c is the particle gyroradius and L_B is the magnetic field scale length. The perturbed electromagnetic electric and vector potentials $(\tilde{\phi}, \tilde{\mathbf{A}})$ are characterised by typical amplitude, pulsation ω and wave number $\boldsymbol{k}$. The latter is split into wave numbers $k_\parallel$ and $k_\perp$, where $k_\parallel$ is the wave number projected along the unperturbed magnetic field, while $k_\perp$ is the modulus of the perpendicular wave number. This procedure is adopted for all vector fields, in particular the vector potential that is split in a parallel and perpendicular components $(\tilde{A}_\parallel, \tilde{\mathbf{A}}_\perp)$. Gyrokinetics relies on the assumption of a pulsation ω much

[6] A variant of the phase space extension is to use non conjugate variables $(t, +E)$ – note the plus sign. The identity matrix is then replaced by the Minkowski space metric $diag(-1, 1, 1, 1)$.

[7] Poisson brackets are conveniently defined by using the original set of conjugate variables $\{f, g\}_{\mathfrak{z}} = \frac{\partial f}{\partial \mathbf{x}} \cdot \frac{\partial g}{\partial \mathbf{p}} - \frac{\partial f}{\partial \mathbf{p}} \cdot \frac{\partial g}{\partial \mathbf{x}}$. The label $\mathfrak{z}$ can be omitted as long as it designates a set of canonical variables since a Poisson bracket can be shown to be invariant under a change of conjugate of variables. However the use of a Poisson bracket becomes trickier whenever non-conjugate variables are used - in this case it is better to keep the index to designate the set of variables that is being used to compute the bracket. The knowledge of the Poisson brackets $\{\mathcal{Z}^\mu, \mathcal{Z}^\nu\}$ *as functions of the variables* $\mathcal{Z}_\mu$ allows the computation of any Poisson bracket of two functions $\mathcal{F}(\boldsymbol{\mathcal{Z}})$ and $\mathcal{G}(\boldsymbol{\mathcal{Z}})$ since $\{\mathcal{F}, \mathcal{G}\}_{\boldsymbol{\mathcal{Z}}} = \partial_{\mathcal{Z}_\mu} \mathcal{F} \{\mathcal{Z}^\mu, \mathcal{Z}^\nu\} \partial_{\mathcal{Z}_\nu} \mathcal{G}$.

smaller than the cyclotron frequency $\Omega_c = e_a B_{eq}/m$ – this is mandatory to build a new magnetic moment that is an adiabatic invariant of motion, – hence

$$\epsilon_\omega = \frac{\omega}{\Omega_c} \ll 1$$

Regarding the ordering on wave numbers, let us remind first that a major objective of the gyrokinetic theory is to avoid an assumption of large field scale length $\lambda_\perp \sim 1/k_\perp$ compared with a Larmor radius ρ_c. This is to be contrasted with the guiding-centre theory where the scale of the equilibrium magnetic field is assumed much larger than a gyroradius. Two dimensionless numbers are introduced to characterise the perpendicular and parallel wave numbers. The first one $\epsilon_\perp = k_\perp \rho_c$ is the perpendicular wave number normalised to the gyroradius. The second one $\epsilon_\parallel = k_\parallel v_T/\omega$, is a parallel transit frequency built with a thermal velocity $v_T = \sqrt{T_{eq}/m_a}$ (T_{eq} the species temperature, m the particle mass) normalised to the frequency. Two emblematic limit cases can be identified[8]:

- Small scale kinetic turbulence is characterised by perpendicular wavelength of the order of the gyroradius ρ_c, i.e.

  $$\epsilon_\perp = k_\perp \rho_c \sim 1$$

  This ordering is called "gyroBohm" scaling. Resonant wave-particle interaction is optimum when a longitudinal Landau resonance takes place $\omega \sim k_\parallel u_\parallel$, and $u_\parallel \sim v_T$. This implies $\epsilon_\parallel \sim 1$, and

  $$\frac{k_\parallel}{k_\perp} \sim \epsilon_\omega \ll 1$$

  Hence parallel wave numbers are small compared to their perpendicular counterpart, a feature that is well verified experimentally and numerically. This is a so-called "flute-like" structure of fluctuations, referring to flute modes, i.e. perturbations that are aligned with the mean magnetic field ($k_\parallel = 0$).

- MHD instabilities and turbulence in the hydrodynamic limit. The hydrodynamic approach is appropriate whenever $\epsilon_\parallel \ll 1$, so that Landau resonances per se are avoided[9]. The perpendicular wavelength of MHD instabilities is usually larger than the gyroradius, so that $\epsilon_\perp \ll 1$ as well.

The scaling of the perturbation amplitude requires some care. Let $\tilde{\phi}$ and $\tilde{A}_\parallel$ denote the perturbed electric and parallel vector potentials - we ignore for now the perpendicular component of the vector potential [10]. We want to build a dimensionless small parameter to characterise the amplitude of the perturbed field. This can be done by defining a parameter ϵ_δ that measures the perturbed Hamiltonian $\tilde{h} = e_a(\tilde{\phi} - u_\parallel \tilde{A}_\parallel)$ (see subsection 3.1 for the justification of this structure of the perturbed Hamiltonian), normalised to the unperturbed temperature T_{eq}. An electromagnetic perturbation such that the electric component $\tilde{\phi}$ of the perturbed Hamiltonian is of the same order of magnitude as its magnetic component $v_\parallel \tilde{A}_\parallel$ satisfies

$$\frac{e_a \tilde{\phi}}{T_{eq}} \sim \frac{e_a v_T \tilde{A}_\parallel}{T_{eq}} \sim \epsilon_\delta$$

[8]This classification is indicative. The reader should not take it too seriously: numerous flavours and variants exist that depart from these orderings. Strictly speaking, one could add a third limit case that is a neoclassical collisional equilibrium $\omega \sim 0$, $k_\parallel q R_0 \sim 1$. This one obeys $\epsilon_\perp \to 0$, which corresponds to the drift-kinetic limit. In other words a gyrokinetic theory is not requested to compute collisional transport, though nothing opposes using it.

[9]This does not forbid however closure schemes where Landau damping is mimicked via appropriate additional terms in fluid equations.

[10]Unperturbed fields are labelled with an index "eq", while perturbed fields are written with a tilde.

Other choices are possible. For instance an electrostatic perturbation is such that $\tilde{A}_\parallel = 0$. A Boltzmann response[11], also called adiabatic response, then implies

$$\frac{\tilde{F}}{F_{eq}} \sim \frac{\widetilde{H}}{T_{eq}} \sim \epsilon_\delta$$

where $\tilde{F}$ is the perturbed distribution function. In the turbulence case, another path is to use a "critical balance" argument, which assumes that all terms in the Vlasov drift-kinetic equation (written in guiding-centre variables) are of the same order of magnitude once a statistical equilibrium is reached. In the electrostatic case, this implies[12]

$$\frac{\partial F}{\partial t} \quad \sim \quad u_\parallel \nabla_\parallel F \quad \sim \quad \mathbf{v}_E \cdot \nabla F \quad \sim \quad \frac{e_a E_\parallel}{m_a} \frac{\partial F}{\partial u_\parallel}$$

where

$$\mathbf{v}_E = \frac{\mathbf{E} \times \mathbf{B}}{B^2}$$

is the electric drift velocity. The first equality reproduces the argument based on the longitudinal Landau resonance. The electric drift velocity can be split in unperturbed and perturbed parts. The unperturbed piece scales as a diamagnetic velocity $T_{eq}/e_a B_{eq} L_p$, where L_p is a plasma gradient length. Equating $\tilde{\mathbf{v}}_E \cdot \nabla F_{eq}$ with $\mathbf{v}_{E,eq} \cdot \nabla \tilde{F}$ yields the relation

$$\frac{\tilde{F}}{F_{eq}} \sim \epsilon_\perp \frac{e_a \tilde{\phi}}{T_{eq}}$$

If a scaling $\tilde{F}/F_{eq} \sim \epsilon_\delta \ll 1$ is imposed, and a "gyroBohm"gyrokinetic turbulence considered $\epsilon_\perp \sim 1$, the ordering of the mode amplitude is found consistent with our previous result

$$\frac{e_a \tilde{\phi}}{T_{eq}} \sim \epsilon_\delta \tag{2}$$

The next step is to establish a link between ϵ_δ and the other dimensionless parameters $\epsilon_B, \epsilon_\parallel, \epsilon_\perp$. The analysis is restricted to the "gyroBohm" framework $\epsilon_\perp \sim 1$. In this regime, the scaling Eq.(2) is consistent with $e_a \tilde{\phi}/T_{eq} \sim \epsilon_\delta$. It is also consistent with the electroneutrality condition in the same limit. In addition, the perturbed drift velocity is such that $|\tilde{\mathbf{v}}_E| / v_T \sim \epsilon_\delta$. Let us bring the parallel vector potential $\tilde{A}_\parallel \simeq \tilde{\phi}/v_T$ back in the game. It is responsible for a perturbed perpendicular magnetic field $\tilde{B}_\perp \simeq k_\perp \tilde{A}_\parallel$, thus such that $\tilde{B}_\perp / B_{eq} \sim \epsilon_\delta$. The critical balance condition also imposes that $\tilde{\mathbf{v}}_E \cdot \nabla F_{eq} \sim \mathbf{v}_{E,eq} \cdot \nabla \tilde{F}$, i.e. $\epsilon_\delta \sim \rho_c / L_p$, which is close to the expansion parameter for the guiding-centre theory $\epsilon_B = \rho_c / L_B$, since the plasma gradient length L_p, magnetic field variation scale L_B and plasma size a are all commensurate. The parameter ρ_c/a is traditionally noted ρ_*. A "gyroBohm" scaling is therefore consistent with $\epsilon_\perp \sim 1$ and $\epsilon_\delta \sim \rho_*$. Still pursuing with a critical balance model, equating $\frac{\partial \tilde{F}}{\partial t}$ with $\tilde{\mathbf{v}}_E \cdot \nabla F_{eq}$ yields $\epsilon_\omega \sim \rho_*$. Hence one arrives to a quite simple "gyroBohm" ordering

$$\frac{\omega}{\Omega_c} \quad \sim \quad \frac{\lambda_\perp}{\lambda_\parallel} \quad \sim \quad \frac{\tilde{F}}{F_{eq}} \quad \sim \quad \frac{\widetilde{H}}{T_{eq}} \quad \sim \quad \rho_*$$

[11] A Boltzmann response should be understood as a special case where the distribution is the one expected for a local thermodynamical equilibrium, i.e. $F = C \exp(-H/T_{eq})$, where C is a constant. An adiabatic transform is such that whenever the Hamiltonian changes from H_{eq} to an new Hamiltonian $H = H_{eq} + \widetilde{H}$, the new distribution function remains given by $F = C \exp(-H/T_{eq})$ at all times. A Taylor development at first order in the perturbed Hamiltonian yields $\tilde{F}/F_{eq} = -\widetilde{H}/T_{eq}$.

[12] The form of the Vlasov equation is not demonstrated here, as it is rather intuitive - its precise structure will be given as one of the main outcomes of the gyrokinetic model. Critical balance is not always satisfied. It is however a reasonable working hypothesis for turbulence.

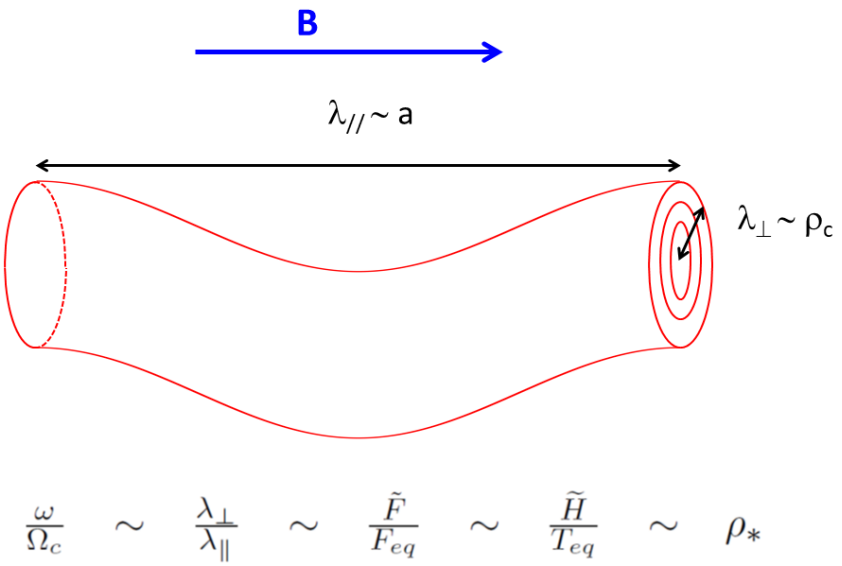


Figure 2: Schematic 3D view of the contour lines of the perturbed Hamiltonian, and corresponding "gyroBohm" scaling on the perpendicular and parallel wavelengths.

where $\lambda_\perp \sim 1/k_\perp$ and $\lambda_\parallel \sim 1/k_\parallel$ are perpendicular and parallel wavelengths (see Figs. 2 and 3). This scaling goes together with $\lambda_\perp \sim \rho_c$ and[13] $\lambda_\parallel \sim a$. Again it is stressed that this scaling holds for turbulence problems (not all in fact), but not for MHD instabilities, nor for collisional transport. The main results in this lecture note do not rely on this specific ordering.

Finally the perpendicular component of the vector potential is mainly responsible for a compressional component of the magnetic field $\tilde{B}_\parallel = \mathbf{e}_\parallel \cdot \nabla \times \tilde{\mathbf{A}}_\perp$. It will be seen that it scales as $\beta\epsilon_\delta$, where β is the ratio of kinetic to magnetic pressure. Hence at low values of β, the perturbed electromagnetic field can be represented by two scalar fields only, the electric potential and the parallel projection of the vector potential. It is called "$(\phi, A_\parallel)$ representation". At higher values of β, the compressional component of the magnetic field must be kept - this limit is dubbed "$(\phi, A_\parallel, B_\parallel)$ representation".

3 The gyrokinetic Vlasov equation

3.1 Breakdown of dynamical invariance

Let us remind a few results regarding the guiding-centre description of a charged particle motion in a strong magnetic field. The particle position is decomposed as the sum of a guiding-centre position $\mathbf{X}$ and a cyclotron displacement $\boldsymbol{\rho}$ (see Fig.4). The displacement vector $\boldsymbol{\rho}$ can be expanded in powers of the small parameter ϵ_B, i.e. $\boldsymbol{\rho} = \boldsymbol{\rho}_0 + \boldsymbol{\rho}_1$, where $\boldsymbol{\rho}_0$ is a circular cyclotron displacement while $\boldsymbol{\rho}_1$ is responsible for plasma polarisation effects. The latter effect can be incorporated in the gyrocentre transform as long as the mean field is small enough[14]. To avoid unnecessary complications, this option is adopted here, even if in practice the mean electric field is not that small. So in the following $\boldsymbol{\rho}_1 = 0$. Hence

[13] In toroidal fusion devices, an additional dimensionless number is the inverse aspect ratio, which allows a distinction between the plasma minor and major radius a and R_0.

[14] Strictly speaking there is also a contribution that comes from magnetic field gradient and line curvature - this tiny effect is ignored as well.

a guiding-centre is a virtual particle localised at the centre of the cyclotron motion

$$\mathbf{x} = \mathbf{X} + \boldsymbol{\rho}_0(\mathbf{X}, \mu, \varsigma) \tag{3}$$

The particle velocity reads

$$\mathbf{v} = \dot{\mathbf{X}} + \dot{\boldsymbol{\rho}}_0(\mathbf{X}, \mu, \varsigma) \tag{4}$$

where $\dot{\mathbf{X}}$ is the guiding-centre velocity and $\dot{\boldsymbol{\rho}}_0 = -\Omega_c(\mathbf{X})\mathbf{e}_\parallel \times \boldsymbol{\rho}_0$ the cyclotron velocity (dots denote time derivative). The simplest version of cyclotron motion is adopted here

$$\begin{aligned} \boldsymbol{\rho}_0(\mathbf{X}, \mu, \varsigma) &= \rho_c \left(\cos\varsigma\, \mathbf{e}_1 - \sin\varsigma\, \mathbf{e}_2\right) \\ \dot{\boldsymbol{\rho}}_0(\mathbf{X}, \mu, \varsigma) &= v_\perp \left(-\sin\varsigma\, \mathbf{e}_1 - \cos\varsigma\, \mathbf{e}_2\right) \end{aligned} \tag{5}$$

where $\mathbf{e}_1(\mathbf{X}), \mathbf{e}_2(\mathbf{X})$ are two unit vectors locally orthogonal to the unperturbed magnetic field $\mathbf{B}_{eq}$, such that $\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_\parallel$. Here, ς is the cyclotron phase[15], $v_\perp = \sqrt{2\mu B_{eq}(\mathbf{X})/m_a}$ is the modulus of the cyclotron velocity (μ is the magnetic moment), $\Omega_c = e_a B_{eq}(\mathbf{X})/m$ is the cyclotron angular frequency, and $\rho_c = m_a v_\perp / e_a B_{eq}$ is the gyroradius. Note that both Ω_c and ρ_c take the sign of e. Eqs. (3) and (4) can be applied to change variables from the extended set of particle coordinates $\boldsymbol{\mathfrak{z}}$ to the extended set of guiding-centre coordinates $\boldsymbol{\mathcal{Z}} = (\boldsymbol{\mathcal{Z}}^*, \varsigma)$ described in subsection 2.2. Here we have noted the 7 slow guiding-centre coordinates as $\boldsymbol{\mathcal{Z}}^* = (\mathbf{X}, u_\parallel, \mu, E, t)$. Any dynamical function of $\boldsymbol{\mathcal{Z}}$ can be expanded as a periodic function of ς, with coefficients that depend on $\boldsymbol{\mathcal{Z}}^*$. The operation that consists in transforming a function of $\boldsymbol{\mathfrak{z}}$ in a function of guiding-centre coordinates $\boldsymbol{\mathcal{Z}}$ is called a guiding-centre transform.

The equilibrium electromagnetic field is inhomogeneous, but assumed static to simplify the analysis. Deriving the guiding-centre equations of motion is still challenging, but can be done efficiently by using a principle of least action. This procedure is detailed in Appendix C. To summarize, the equilibrium Lagrangian $L_{eq} = \mathbf{p} \cdot \dot{\mathbf{x}} - H_{eq}$ is transformed into a guiding-centre Lagrangian, which, up to the zeroth order in ϵ_B, is written as

$$L_{eq}\left(\mathbf{Z}, \dot{\mathbf{Z}}\right) = e_a \mathbf{A}^*_{eq} \cdot \dot{\mathbf{X}} + \frac{m_a}{e_a}\mu\dot{\varsigma} - H_{eq}\left(\mathbf{X}, u_\parallel, \mu\right) \tag{6}$$

where the Hamiltonian is

$$H_{eq}\left(\mathbf{X}, u_\parallel, \mu\right) = \frac{1}{2} m_a u_\parallel^2 + \mu B_{eq}(\mathbf{X}) + e_a \phi_{eq}(\mathbf{X})$$

and ϕ_{eq} is the unperturbed electric potential. The modified vector potential is

$$\mathbf{A}^*_{eq}\left(\mathbf{X}, u_\parallel\right) = \mathbf{A}_{eq}(\mathbf{X}) + \frac{m_a}{e_a} u_\parallel \mathbf{e}_\parallel(\mathbf{X})$$

The principle of least action yields the Poisson brackets associated with the guiding-centre coordinates Eq.(112) in Appendix C, and the equations of motion Eqs.(113)-(116). Since the Lagrangian does not depend on ς, the corresponding Euler-Lagrange equation yields

$$\frac{\partial L_{eq}}{\partial \varsigma} = \frac{d}{dt}\left(\frac{\partial L_{eq}}{\partial \dot{\varsigma}}\right) = \frac{m_a}{e_a}\dot{\mu} = 0$$

Hence the adiabatic invariance of the magnetic moment μ is recovered.

Let us now introduce a perturbation of the electromagnetic field. The total vector and electric potentials read

$$\begin{aligned} \mathbf{A}(\mathbf{x}, t) &= \mathbf{A}_{eq}(\mathbf{x}) + \tilde{\mathbf{A}}(\mathbf{x}, t) \\ \phi(\mathbf{x}, t) &= \phi_{eq}(\mathbf{x}) + \tilde{\phi}(\mathbf{x}, t) \end{aligned}$$

[15]The convention in the gyrokinetic literature is that ς is defined as the angle from $\boldsymbol{\rho}_0$ to $\mathbf{e}_1$ for ions, and from $\boldsymbol{\rho}_0$ to $-\mathbf{e}_1$ for electrons.

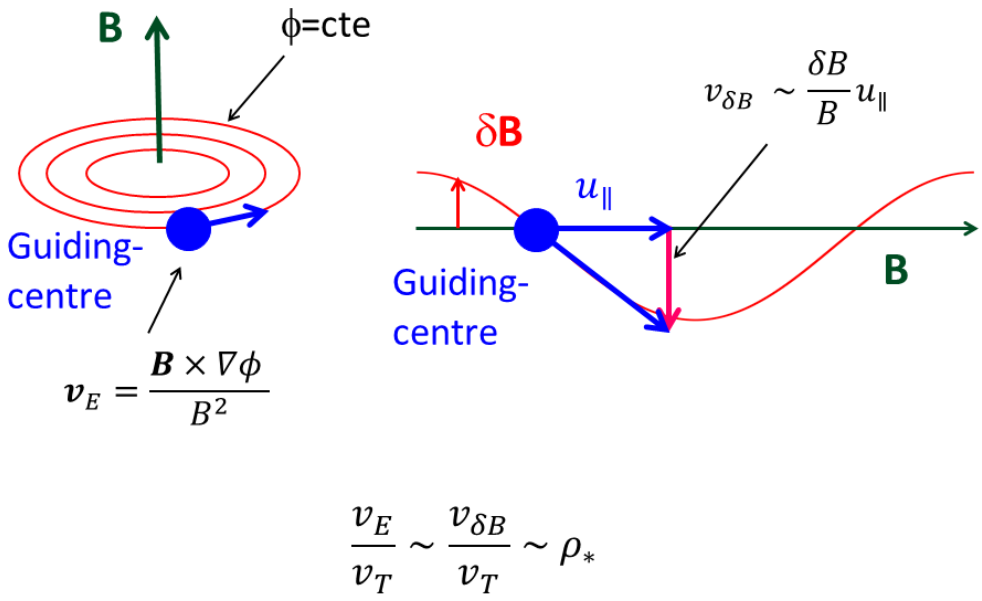


Figure 3: Schematic effect of perturbed electro-magnetic field on a guiding-centre velocities. Left: contour lines of the perturbed electric potential. A guiding-centre is submitted to an $E \times B$ drift velocity that is tangent to the contour lines as a result of its expression $\mathbf{v}_E = \frac{\mathbf{B}\times\nabla\phi}{B^2}$, i.e. normal to the magnetic field and the potential gradient $\nabla\phi$. Right: unperturbed field line (here a straight line) that is deformed by a perturbed magnetic field δB. A guiding-centre with a parallel velocity $u_\parallel$ along the field lines acquires a velocity transverse to the initial field line $v_{\delta B} \sim \frac{\delta B}{B} u_\parallel$. GyroBohm scaling corresponds to $\frac{|\mathbf{v}_E|}{v_T} \sim \frac{v_{\delta B}}{v_T} \sim \rho_*$. This picture applies as well to gyrocentres.

The perturbed vector potential is separated in components parallel and perpendicular to the unperturbed magnetic field

$$\tilde{\mathbf{A}}(\mathbf{x},t) = \tilde{A}_\parallel(\mathbf{x},t)\mathbf{e}_\parallel(\mathbf{x}) + \tilde{\mathbf{A}}_\perp(\mathbf{x},t)$$

The perturbed fields $(\tilde{\mathbf{A}}, \tilde{\phi})$ are of order 1 in the expansion parameter ϵ_δ once properly normalised, so will be sometimes noted $(\mathbf{A}_1, \phi_1)$. The Lagrangian in the original phase space $\mathbf{z} = (\mathbf{x}, \mathbf{v})$ reads

$$L(\mathbf{z}, \dot{\mathbf{z}}, t) = (e_a \mathbf{A}(\mathbf{x},t) + m_a \mathbf{v}) \cdot \dot{\mathbf{x}} - h(\mathbf{x}, \mathbf{v}, t) \tag{7}$$

where $h = \frac{1}{2} m v^2 + e_a \phi(\mathbf{x}, t)$ is the "Hamiltonian"[16]. This Lagrangian can be as well written in guiding-centre coordinates

$$\begin{aligned} L\left(\mathbf{Z}, \dot{\mathbf{Z}}, t\right) &= e_a \mathbf{A}^*_{eq}\left(\mathbf{X}, u_\parallel\right) \cdot \dot{\mathbf{X}} + \frac{m_a}{e_a} \mu \dot{\varsigma} - H_{eq}\left(\mathbf{X}, u_\parallel, \mu\right) \\ &+ e_a \tilde{\mathbf{A}}(\mathbf{X} + \boldsymbol{\rho}_0, t) \cdot \left(\dot{\mathbf{X}} + \dot{\boldsymbol{\rho}}_0\right) - e_a \tilde{\phi}(\mathbf{X} + \boldsymbol{\rho}_0, t) \end{aligned} \tag{8}$$

The structure of the new Lagrangian Eq.(8) is deeply modified. Thus all Poisson brackets should be recalculated. However this difficult task can be circumvented with the following trick. Let us go back to the initial Lagrangian Eq.(7) and introduce the shifted velocity

$$\mathbf{v}'(\mathbf{x}, \mathbf{v}, t) = \mathbf{v} + \frac{e_a}{m_a} \tilde{\mathbf{A}}(\mathbf{x}, t)$$

[16]The name "Hamiltonian" is a bit stretched as Hamilton arrived after Lagrange, but appears to be convenient

The new Lagrangian reads

$$L\left(\mathbf{z},\dot{\mathbf{z}},t\right) \quad = \quad \left(e_a\mathbf{A}_{eq}(\mathbf{x},t)+m_a\mathbf{v}'\right)\cdot\dot{\mathbf{x}}-h^*(\mathbf{x},\mathbf{v}',t)$$

where the new Hamiltonian is

$$h^*(\mathbf{x},\mathbf{v}',t)=\frac{1}{2}m_a\mathbf{v}'^2+e_a\phi_{eq}(\mathbf{x},t)+e_a\left(\tilde{\phi}(\mathbf{x},t)-\mathbf{v}'\cdot\tilde{\mathbf{A}}(\mathbf{x},t)\right)+\frac{1}{2}\frac{e_a^2}{m_a}\tilde{\mathbf{A}}^2(\mathbf{x},t) \qquad (9)$$

What is the gain? It appears that the coefficient of $\dot{\mathbf{x}}$ in the Lagrangian is the same as

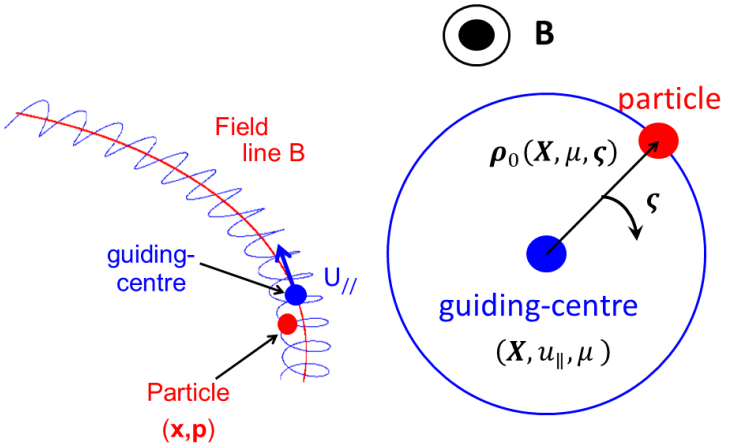


Figure 4: Decomposition of a charged particle motion into a guiding-centre position $\mathbf{X}$ and a cyclotron motion $\boldsymbol{\rho}$. Strictly speaking, the cyclotron motion can be itself be separated into a periodic component, restricted here to a circular gyration $\boldsymbol{\rho}_0$, and a mean value, neglected here.

the one computed without the perturbed field. The immediate consequence is the identity of the Poisson brackets with those computed for the unperturbed system. Hence when moving to gyrocentre coordinates, the Poisson brackets remain the same. Let us note however that the new Hamiltonian h^* is now second order in the expansion parameter ϵ_δ. This adds some complexity to the problem at hand, but is the price to pay for not modifying the Poisson brackets. The shifted velocity $\mathbf{v}'$ can be decomposed in parallel and perpendicular components to the equilibrium magnetic field. The perpendicular component participates in the cyclotron motion. When moving to guiding-centre coordinates, the velocity reads

$$\dot{\mathbf{X}}=p_{\|}\mathbf{e}_{\|}(\mathbf{X})+\dot{\mathbf{X}}_{\perp}$$

where $p_{\|}$ is proportional to the parallel momentum[17]

$$p_{\|}\left(\boldsymbol{\mathcal{Z}}\right)=u_{\|}+\frac{e_a}{m_a}\tilde{A}_{\|}(\mathbf{X}+\boldsymbol{\rho}_0,t)$$

The total Lagrangian can be recast as

$$L\left(\mathbf{Z},\dot{\mathbf{Z}},t\right)=e_a\mathbf{A}^*\cdot\dot{\mathbf{X}}+\frac{m_a}{e_a}\mu\dot{\varsigma}-\left(H^*_{eq}+\widetilde{H}\right) \qquad (10)$$

[17]The parallel momentum is in fact $m_a p_{\|}$. $p_{\|}$ scales as a velocity, which may be disconcerting given the notation. However this in line with the fact that most codes compute a parallel "velocity", whereas it is in fact a momentum. This notation is not universal. There is no use of the variable $p_{\|}$ per se in the approach [2], where $p_{\|}$ designates in fact $m_a u_{\|}$ (but the variable $\bar{p}_{\|}$ that is built later on is the same). In the references [3, 4], $p_{\|}$ is noted p_z, and $p_{\|}=m_a u_{\|}$.

The new generalised vector potential reads

$$\mathbf{A}^* \left(\mathbf{X}, p_\parallel\right) = \mathbf{A}^*_{eq} + \tilde{A}_\parallel \mathbf{e}_\parallel = \mathbf{A}_{eq}\left(\mathbf{X}\right) + \frac{m_a}{e_a} p_\parallel \mathbf{e}_\parallel\left(\mathbf{X}\right)$$

The unperturbed and perturbed Hamiltonians read

$$H^*_{eq}\left(\mathbf{X}, p_\parallel, \mu\right) = \frac{1}{2} m_a p_\parallel^2 + \mu B_{eq}\left(\mathbf{X}\right) + e_a \phi_{eq}\left(\mathbf{X}\right)$$

and

$$\begin{aligned} \widetilde{H}\left(\mathbf{X}, p_\parallel, \mu, \varsigma, t\right) &= e_a \tilde{\phi}(\mathbf{X} + \boldsymbol{\rho}_0, t) - e_a p_\parallel \tilde{A}_\parallel(\mathbf{X} + \boldsymbol{\rho}_0, t) \\ &- e_a \left(\dot{\mathbf{X}}_\perp + \dot{\boldsymbol{\rho}}_0\right) \cdot \tilde{\mathbf{A}}_\perp(\mathbf{X} + \boldsymbol{\rho}_0, t) + \frac{1}{2} \frac{e_a^2}{m_a} \tilde{\mathbf{A}}^2(\mathbf{X} + \boldsymbol{\rho}_0, t) \end{aligned} \quad (11)$$

The structure of the Lagrangian Eq.(10) is very similar to the unperturbed one Eq.(6) : the "symplectic" part (the bit proportional to $\dot{\mathbf{X}}$) of the Lagrangian remains structurally the same, but the Hamiltonian H_{eq} is replaced by $H = H^*_{eq} + \widetilde{H}$. Note that this is not the sole strategy: the perturbed vector potential could be inserted in the "symplectic" part of the Lagrangian, i.e. added to $\mathbf{A}_{eq}$, hence leading to another definition of the velocity $p_\parallel$ and Hamiltonian H. The latter approach is called "symplectic", while the former is dubbed "Hamiltonian". A more systematic derivation will be given in section 5. In the Hamiltonian approach, the symplectic part of the Lagrangian is left unchanged compared with the standard guiding-centre description. Therefore the Poisson brackets are the same, and the Jacobian as well. Hence the trajectory equations are the same with H_{eq} replaced by $H = H^*_{eq} + \widetilde{H}$. Since the equilibrium Hamiltonian has been redefined, we also redefine $\mathcal{H}_{eq} = H^*_{eq} - E$ (instead of $H_{eq} - E$). The Hamiltonian perturbation is small, i.e. first order in an expansion parameter $\epsilon_\delta \ll 1$, opening the path to a perturbative approach in ϵ_δ. As seen in section 2.3, the expansion parameter ϵ_δ characterises the perturbed Hamiltonian normalised to the equilibrium one, itself of the order of the plasma temperature T_{eq}, hence $\epsilon_\delta \sim \widetilde{H}/T_{eq}$.

However, a disturbing feature of the Lagrangian Eq.(10) is the dependence of the new Hamiltonian on the gyroangle ς, which implies that the magnetic moment is no longer an invariant of motion. This is most unfortunate since a breakdown of dynamical invariance usually implies an onset of chaos, and therefore a loss of confinement in the context of fusion devices. There is nevertheless one case where an adiabatic invariant can be rebuilt efficiently, which is the case where the typical time scale of the Hamiltonian perturbation is much longer than the cyclotron period, provided the amplitude of the perturbed Hamiltonian is small enough. To recover an adiabatic invariant, we impose a change of coordinates such that the new Hamiltonian does not depend on the gyroangle, or more exactly its gyrocentre counterpart. This is a situation of interest in laboratory devices, where many instabilities evolve slowly compared with a cyclotron time scale. Because the new Hamiltonian does not depend on the gyroangle, it allows treating the Vlasov/Maxwell system as a 4D dimensional problem parametrized by the new adiabatic invariant. This invariant is now going to be built, proceeding step by step.

3.2 Building a new adiabatic invariant

Solving the Maxwell equations requires distribution functions expressed in the initial set of variables $(\mathbf{x}, \mathbf{p})$ since they involve charge and current densities of "true" particles, not virtual guiding-centres. Moreover, it is desirable to build new variables, called "gyrocentres", close to guiding-centres, and such that the modified magnetic moment remains an adiabatic invariant of motion. The corresponding coordinates can be quite far from the original set $(\mathbf{x}, \mathbf{p})$. It is therefore important to keep track of this double change of variables. The choice of gyrocentre variables is in fact quite broad. However not all choices

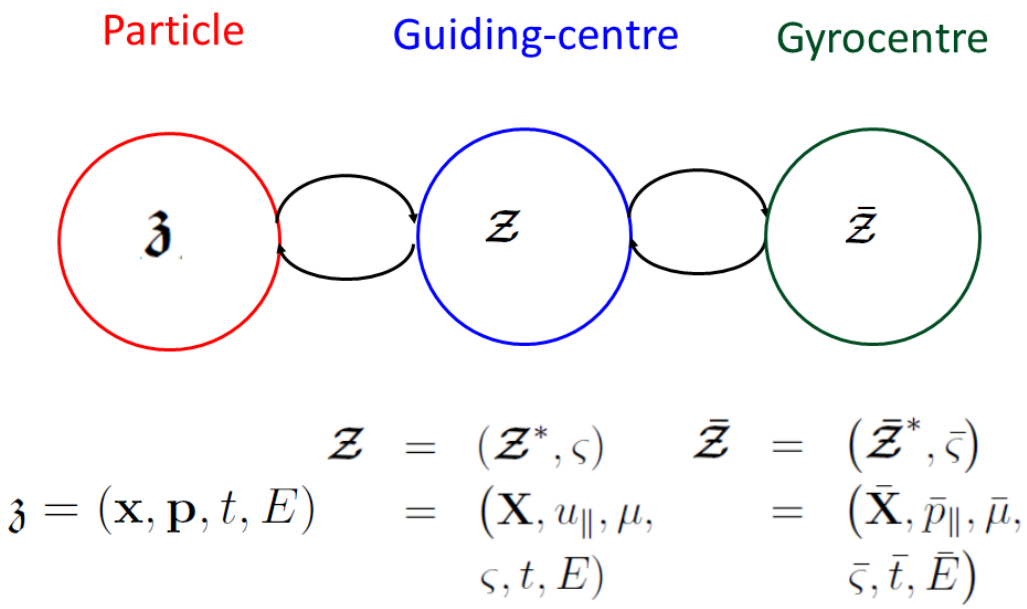


Figure 5: Sequence of transforms from particle to guiding-centres, then to gyrocentres. A complete gyrokinetic theory must provide the two corresponding changes of variables, to move back and forth from one representation to the others.

are suitable to compute the charge and current densities. The optimum procedure can then be summarised as follows. It appears that useful coordinates in presence of the perturbed field are the non conjugate guiding-centre variables $\mathbf{Z} = (\mathbf{X}, p_\parallel, \mu, \varsigma)$, where $\mathbf{X}$ is the guiding-centre position, $p_\parallel$ its parallel momentum (constructed from $u_\parallel$ by adding $e_a \tilde{A}_\parallel / m_a$), μ the magnetic moment and ς the gyroangle. The distribution function of guiding-centres is solution of a Vlasov equation, as shown by Littlejohn [7] (see also a previous note on trajectories). However this set of variables is not good enough because, as mentioned above, the magnetic moment is not conserved, a consequence of the dependence of the perturbed electromagnetic field on the gyroangle ς. It is then desirable to build a set of gyrocentre coordinates, noted $\bar{\mathbf{Z}} = (\bar{\varsigma}, \bar{\mu}, \bar{\mathbf{X}}, \bar{p}_\parallel)$, such that $\bar{\mu}$ is an invariant of motion. In the extended phase space, these two changes of variables are first a transform from the particle coordinates, noted $\mathfrak{z} = (t, -E, \mathbf{x}, \mathbf{p})$, to the guiding-centre coordinates $\mathcal{Z} = (\mathbf{X}, p_\parallel, \mu, \varsigma, t, E)$, and second a transform from $\mathcal{Z}$ to the gyrocentre coordinates $\bar{\mathcal{Z}} = (\bar{\mathbf{X}}, \bar{p}_\parallel, \bar{\mu}, \bar{\varsigma}, \bar{t}, \bar{E})$ (see Fig. 5). It will be shown in the next section that charge and current densities in real space can be suitably calculated knowing the distribution function expressed with this set of gyrocentre coordinates. A systematic way to take this step is to use a principle of least action, as for guiding-centre variables. This is a rather technical methodology that will be detailed in the section 5. For the moment, we show that the gyrocentre distribution function is solution of a Vlasov equation expressed in non-canonical variables.

In the Hamiltonian approach, the new set of variables $\bar{\mathcal{Z}}$ is related to the old set $\mathcal{Z}$ by using a near identity transform that is called a *contact transformation* [8, 9]

$$\mathcal{Z} = \bar{\mathcal{Z}} - \epsilon_\delta \left\{ S_1, \bar{\mathcal{Z}} \right\}_{\bar{\mathcal{Z}}} + o(\epsilon_\delta^2) \tag{12}$$

where $\{\ ,\ \}_{\bar{\mathcal{Z}}}$ is the Poisson bracket written in the new variables $\bar{\mathcal{Z}}$, and S_1 is a generating

function. It can be shown that this change of variables preserve the structure of the Poisson brackets. In other words, the Lagrangian is the same in both sets of coordinates $\boldsymbol{\mathcal{Z}}$ and $\bar{\boldsymbol{\mathcal{Z}}}$, i.e.

$$\begin{aligned}\{F,G\}_{\bar{\mathbf{Z}}} &= \frac{e_a}{m_a}\left(\frac{\partial F}{\partial\bar{\varsigma}}\frac{\partial G}{\partial\bar{\mu}}-\frac{\partial F}{\partial\bar{\mu}}\frac{\partial G}{\partial\bar{\varsigma}}\right)\\ &+\frac{1}{m_a}\frac{\mathbf{B}^*}{B^*_{||}}\cdot\left(\frac{\partial F}{\partial\bar{\mathbf{X}}}\frac{\partial G}{\partial\bar{p}_{||}}-\frac{\partial F}{\partial\bar{p}_{||}}\frac{\partial G}{\partial\bar{\mathbf{X}}}\right)\\ &-\frac{\mathbf{e}_{||}}{e_aB^*_{||}}\cdot\left(\frac{\partial F}{\partial\bar{\mathbf{X}}}\times\frac{\partial G}{\partial\bar{\mathbf{X}}}\right)\\ &+\frac{\partial F}{\partial\bar{E}}\frac{\partial G}{\partial\bar{t}}-\frac{\partial F}{\partial\bar{t}}\frac{\partial G}{\partial\bar{E}}\end{aligned} \tag{13}$$

where

$$\mathbf{B}^*\left(\bar{\mathbf{X}},\bar{p}_{||}\right)=\nabla\times\mathbf{A}^*\left(\bar{\mathbf{X}},\bar{p}_{||}\right)=\mathbf{B}_{eq}\left(\bar{\mathbf{X}}\right)+\frac{m_a\bar{p}_{||}}{e_a}\nabla\times\mathbf{e}_{||}$$

and

$$B^*_{||}\left(\bar{\mathbf{X}},\bar{p}_{||}\right)=\mathbf{e}_{||}\cdot\mathbf{B}^*\left(\bar{\mathbf{X}},\bar{p}_{||}\right)=B_{eq}\left(\bar{\mathbf{X}}\right)+\frac{m_a\bar{p}_{||}}{e_a}\mathbf{e}_{||}\cdot\nabla\times\mathbf{e}_{||}$$

is the Jacobian of the guiding-centre transform (by convention $\mathbf{e}_{||}$ is calculated at $\bar{\mathbf{X}}$ and all gradients are derivative with respect to $\bar{\mathbf{X}}$). The Poisson structure of the change of variables Eq.(12), that involves a Poisson bracket, is easy to demonstrate when using conjugate variables (see Appendix D). It was shown by Cary [10] that it remains valid for non conjugate variables. This transform is essential in gyrokinetic theory, in its Hamiltonian version [11, 12, 13]. A general demonstration that applies to all changes of variables (i.e. not only Hamiltonian) can be found in the overview [2] - a summary is given in Appendix H. At first order in ϵ_δ, the near identity change of variables Eq.(12) can as well be written

$$\bar{\boldsymbol{\mathcal{Z}}}=\boldsymbol{\mathcal{Z}}+\epsilon_\delta\left\{S_1,\boldsymbol{\mathcal{Z}}\right\}_{\boldsymbol{\mathcal{Z}}}+o(\epsilon_\delta^2) \tag{14}$$

Hence the gyrocentre change of variables can be seen as a translation, where the shift is the Poisson bracket $\{S_1,\boldsymbol{\mathcal{Z}}\}_{\boldsymbol{\mathcal{Z}}}$. The following relation holds[18]

$$\bar{\mathcal{H}}(\bar{\boldsymbol{\mathcal{Z}}})=\mathcal{H}(\boldsymbol{\mathcal{Z}}) \tag{15}$$

A Taylor development of $\mathcal{H}(\boldsymbol{\mathcal{Z}})$ in Eq.(15) implies that at lowest order in ϵ_δ

$$\bar{\mathcal{H}}_{eq}\left(\bar{\boldsymbol{\mathcal{Z}}}\right)=\mathcal{H}_{eq}(\boldsymbol{\mathcal{Z}})$$

The next order in ϵ_δ provides the evolution equation of the generating function

$$-\left\{\mathcal{H}_{eq},S_1\right\}_{\bar{\boldsymbol{\mathcal{Z}}}}=\mathcal{H}_1(\bar{\boldsymbol{\mathcal{Z}}})-\bar{\mathcal{H}}_1(\bar{\boldsymbol{\mathcal{Z}}}) \tag{16}$$

which is written in *gyrocentre coordinates* $\bar{\boldsymbol{\mathcal{Z}}}$ *exclusively*. Let us now recall our initial purpose: the new Hamiltonian should be independent of the new gyroangle $\bar{\varsigma}$, assuming that all typical frequencies of the perturbed Hamiltonian are smaller than the cyclotron pulsation by at least a factor ϵ_δ. Let us remind that the perturbed Hamiltonian is supposed to exhibit low frequencies compared with the cyclotron frequency Ω_c. This is the key assumption of gyrokinetic theory. The left hand side of Eq.(16) reads

$$-\left\{\mathcal{H}_{eq},S_1\right\}_{\bar{\boldsymbol{\mathcal{Z}}}}=\frac{\partial S_1}{\partial\bar{t}}+\Omega_c\frac{\partial S_1}{\partial\bar{\varsigma}}+\bar{p}_{||}\left(\mathbf{e}_{||}\cdot\nabla\right)S_1+\left(\dot{\bar{\mathbf{X}}}_\perp\cdot\nabla\right)S_1$$

[18]This a consequence of the independence of the generating function on the extended time τ.

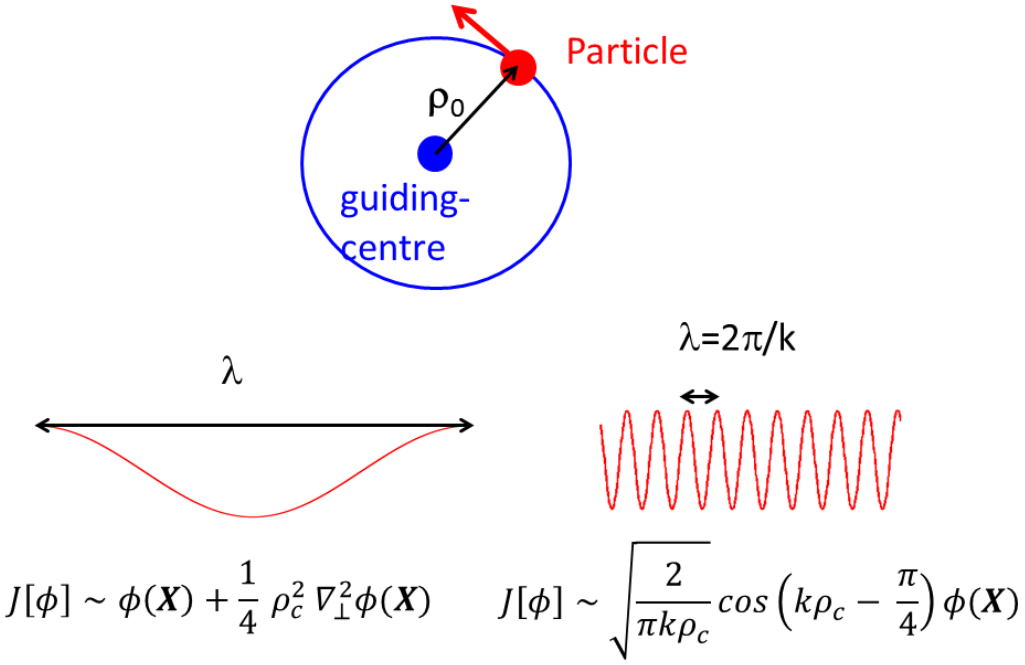


Figure 6: Gyroaverage operator applied to a field $\phi(\mathbf{x})$ in two limits. On the left, case of a field with a typical variation scale λ that is much larger than a gyroradius ρ_c, $\lambda \gg \rho_c$. The asymptotic limit is obtained via a Taylor development of $\phi\left(\mathbf{X}+\boldsymbol{\rho}_0\right)$, followed by an average over the gyrophase. The gyroaverage operator is close to $\mathcal{J}\left[\phi\right] \sim \phi\left(\mathbf{X}\right) + \frac{1}{4}\rho_c^2\nabla_\perp^2\phi\left(\mathbf{X}\right) + o\left(\rho_c/\lambda\right)^4$. An average of the field over 4 points on the cyclotron circle leads to a finite difference discretisation of the same operator. The right panel corresponds to the opposite limit of a scale length much smaller than a gyroradius $\lambda \ll \rho_c$. A simple case corresponds to a periodic variation of the field. The gyroaverage operator consists in multiplying the field by a Bessel function of argument $k_\perp\rho_c$ in the Fourier space (see Appendix E). Hence the gyroaverage field tends to zero like $\sqrt{\lambda/\rho_c}$ in the limit of small wavelengths. A similar operation can be defined on gyrocentres.

Because of the gyrokinetic ordering, the second term, related to the cyclotron motion, is $\epsilon_\omega^{-1} \sim \epsilon_\delta^{-1}$ larger than the others. This property allows restricting the left hand side of Eq.(16) to a derivative in $\bar{\varsigma}$ only, i.e.

$$\Omega_c\left(\bar{\boldsymbol{\mathcal{Z}}}^*\right)\frac{\partial S_1}{\partial\bar{\varsigma}} = \mathcal{H}_1(\bar{\boldsymbol{\mathcal{Z}}}) - \bar{\mathcal{H}}_1(\bar{\boldsymbol{\mathcal{Z}}}) \tag{17}$$

The average of Eq.(17) over the gyroangle $\bar{\varsigma}$, a solubility constraint, provides the new Hamiltonian

$$\bar{\mathcal{H}}_1(\bar{\boldsymbol{\mathcal{Z}}}^*) = \left\langle \mathcal{H}_1(\bar{\boldsymbol{\mathcal{Z}}})\right\rangle_{\bar{\varsigma}}$$

where the bracket $\langle\ \ \rangle_{\bar{\varsigma}}$ indicates an average over the gyroangle associated with the gyrocentre

$$\langle\mathcal{H}_1\rangle_{\bar{\varsigma}}(\bar{\boldsymbol{\mathcal{Z}}}^*) = \int_0^{2\pi}\frac{d\bar{\varsigma}}{2\pi}\mathcal{H}_1(\bar{\boldsymbol{\mathcal{Z}}}^*,\bar{\varsigma})$$

An explicit form of the generating function is therefore

$$S_1\left(\bar{\boldsymbol{\mathcal{Z}}}^*,\bar{\varsigma}\right) = \frac{1}{\Omega_c\left(\bar{\boldsymbol{\mathcal{Z}}}^*\right)}\int^{\bar{\varsigma}} d\varsigma'\left(\mathcal{H}_1(\bar{\boldsymbol{\mathcal{Z}}}^*,\varsigma') - \bar{\mathcal{H}}_1(\bar{\boldsymbol{\mathcal{Z}}}^*)\right)$$

It appears to be a periodic function of $\bar{\varsigma}$ with zero mean. Here $\bar{\boldsymbol{\mathcal{Z}}}^*$ designates the slow gyrocentre coordinates $\bar{\boldsymbol{\mathcal{Z}}}^* = (\bar{\mathbf{X}}, \bar{p}_\parallel, \bar{\mu}, \bar{t}, \bar{E})$, and $\bar{\boldsymbol{\mathcal{Z}}} = (\bar{\boldsymbol{\mathcal{Z}}}^*, \bar{\varsigma})$ is the full set of gyrocentre coordinates. The average over a gyroangle is called "gyroaverage". Let us insist

on the mixed nature of the integrand $\mathcal{H}_1(\bar{\boldsymbol{\mathcal{Z}}})$, an Hamiltonian originally expressed in guiding-centre variables $\boldsymbol{\mathcal{Z}}$, but calculated with gyrocentre variables $\bar{\boldsymbol{\mathcal{Z}}}$. The perturbed Hamiltonian Eq.(11) reads at order 1

$$\begin{aligned}\mathcal{H}_1(\bar{\mathbf{X}},\bar{p}_\parallel,\bar{\mu},\bar{\varsigma},t) &= e_a\tilde{\phi}\left(\bar{\mathbf{X}}+\boldsymbol{\rho}_0(\bar{\mathbf{X}},\bar{\mu},\bar{\varsigma}),t\right)-e_a p_\parallel\tilde{A}_\parallel\left(\bar{\mathbf{X}}+\boldsymbol{\rho}_0(\bar{\mathbf{X}},\bar{\mu},\bar{\varsigma}),t\right)\\ &- e_a\left(\dot{\bar{\mathbf{X}}}_\perp+\dot{\boldsymbol{\rho}}_0(\bar{\mathbf{X}},\bar{\mu},\bar{\varsigma})\right)\cdot\tilde{\mathbf{A}}_\perp\left(\mathbf{X}+\boldsymbol{\rho}_0(\bar{\mathbf{X}},\bar{\mu},\bar{\varsigma}),t\right)\end{aligned}\tag{18}$$

The gyroaverage over $\bar{\varsigma}$ is an average over a circular cyclotron motion $(\boldsymbol{\rho}_0,\dot{\boldsymbol{\rho}}_0)(\bar{\mathbf{X}},\bar{\mu},\bar{\varsigma})$ *centred on the gyrocentre position and velocity* , noted[19]

$$\bar{\mathcal{H}}_1(\bar{\boldsymbol{\mathcal{Z}}}^*)=\mathcal{J}[\mathcal{H}_1](\bar{\boldsymbol{\mathcal{Z}}}^*)=\left\langle\mathcal{H}_1(\bar{\mathbf{X}},\bar{p}_\parallel,\bar{\mu},\bar{\varsigma})\right\rangle_{\bar{\varsigma}}$$

A similar operation will be needed in guiding-centre variables, i.e.

$$\bar{\mathcal{H}}_1(\boldsymbol{\mathcal{Z}}^*)=\mathcal{J}[\mathcal{H}_1](\boldsymbol{\mathcal{Z}}^*)=\left\langle\mathcal{H}_1(\mathbf{X},p_\parallel,\mu,\varsigma)\right\rangle_{\varsigma}$$

Here the average is performed over a circular cyclotron motion $(\boldsymbol{\rho}_0,\dot{\boldsymbol{\rho}}_0)(\mathbf{X},\mu,\varsigma)$ *centred on the guiding-centre position.* A gyroaveraged field is roughly speaking the field "felt" by a gyrocentre (or guiding-centre), a virtual particle that moves slowly compared with the cyclotron motion. If the scale length of the perturbed field is much larger than the gyroradius, the gyroaverage field is close to the field itself. On the contrary, if the field exhibits fast spatial scales compared with the gyroradius, its gyroaverage tends to zero, as a result of a summation over plus and minus values during a cyclotron motion (see Fig. 6).

The variable $\bar{\mu}$ is an invariant of motion since the new Hamiltonian does not depend on its conjugate variable[20] $\bar{\varsigma}$. Hence the objective is met: the *new Hamiltonian does not depend on the gyrocentre gyroangle* $\bar{\varsigma}$, which immediately imposes that $\bar{\mu}$ *is an invariant of motion.* It must be stressed here that no information is lost here: the gyrocentre transform that changes the coordinates $\boldsymbol{\mathcal{Z}}$ to $\bar{\boldsymbol{\mathcal{Z}}}$ is fully defined by Eq.(12) - there is no reduction of dimensionality, even if the new Hamiltonian is the gyroaverage of the old one. Once the generating function S_1 is known, the change of variables Eq.(12) is fully determined.

At this stage, we have computed a first order Hamiltonian $\bar{\mathcal{H}}_1$ that does not depend on the gyroangle. However we saw that the perturbed Hamiltonian is in fact of second order in ϵ_δ. Hence the calculation should be done at second order, a formidable task that will handled later on (an impatient reader may give a look at the Appendix G). Let us note moreover that most gyrokinetic codes solve the equations of motion at first order in ϵ_δ. So at this point, we will just admit that there exists a second order Hamiltonian $\bar{\mathcal{H}}_2(\boldsymbol{\mathcal{Z}}^*)$ that will not be detailed for now.

3.3 Plasma polarisation

It is useful to clarify the relationship between particle, guiding-centre and gyrocentre positions. The link between the particle and the guiding-centre position is given by Eq.(3). The relationship between the gyrocentre and guiding-centre coordinates imposes that

$$\mathbf{x}=\mathbf{X}+\boldsymbol{\rho}_0(\mathbf{X},\mu,\varsigma)=\bar{\mathbf{X}}+\bar{\boldsymbol{\rho}}(\bar{\mathbf{X}},\bar{p}_\parallel,\bar{\mu},\bar{\varsigma})$$

where $\bar{\boldsymbol{\rho}}$ can be expanded as

$$\bar{\boldsymbol{\rho}}(\bar{\boldsymbol{\mathcal{Z}}})=\bar{\boldsymbol{\rho}}_0(\bar{\boldsymbol{\mathcal{Z}}})+\bar{\boldsymbol{\rho}}_1(\bar{\boldsymbol{\mathcal{Z}}})$$

[19] We use a bracket [ ] to isolate the argument of the gyroaverage operator, thus avoiding ambiguities. It should be differentiated from the notation [,] that is used for Lagrange brackets.

[20] This is a direct consequence of the equation of motion $\frac{d\bar{\mu}}{dt}=-\frac{e_a}{m_a}\frac{\partial\bar{\mathcal{H}}}{\partial\bar{\varsigma}}$.

with $\bar{\boldsymbol{\rho}}_0(\bar{\boldsymbol{\mathcal{Z}}}) = \boldsymbol{\rho}_0(\bar{\boldsymbol{\mathcal{Z}}})$, and

$$\bar{\boldsymbol{\rho}}_1(\bar{\boldsymbol{\mathcal{Z}}}) = -\left\{S_1, \mathbf{X} + \boldsymbol{\rho}_0\right\}_{\bar{\boldsymbol{\mathcal{Z}}}}$$

It appears that $\bar{\boldsymbol{\rho}}$ does not average out to zero when expressed in the gyroangle $\bar{\varsigma}$. Indeed the average of $\{S_1, \mathbf{X} + \boldsymbol{\rho}_0\}_{\bar{\boldsymbol{\mathcal{Z}}}}$ has no reason to vanish, and indeed it does not. The gyroaverage of the cyclotron deformation $\bar{\boldsymbol{\rho}}_1$ is equal to

$$\langle \bar{\boldsymbol{\rho}}_1 \rangle_{\bar{\varsigma}} (\bar{\boldsymbol{\mathcal{Z}}}^*) = -\int \frac{d\bar{\varsigma}}{2\pi} \left\{S_1, \bar{\mathbf{X}} + \boldsymbol{\rho}_0\right\}$$

Using the expression Eq.(17) of the generating function S_1, one gets

$$\langle \bar{\boldsymbol{\rho}}_1 \rangle_{\bar{\varsigma}} (\bar{\boldsymbol{\mathcal{Z}}}^*) = -\frac{1}{B(\bar{\mathbf{X}})} \frac{\partial}{\partial \bar{\mu}} \int \frac{d\bar{\varsigma}}{2\pi} \boldsymbol{\rho}_0(\bar{\boldsymbol{\mathcal{Z}}}^*, \bar{\varsigma}) \left\{\mathcal{H}_1(\bar{\boldsymbol{\mathcal{Z}}}^*, \bar{\varsigma}) - \mathcal{J}[\mathcal{H}_1](\bar{\boldsymbol{\mathcal{Z}}}^*)\right\}$$

This quantity can be explicitly calculated in the special case where only the electric potential and the parallel components of the vector potential are considered ($\tilde{\phi}, \tilde{A}_\parallel$ representation). It is then equal to

$$\langle \boldsymbol{\rho}_1 \rangle_{\bar{\varsigma}} (\bar{\boldsymbol{\mathcal{Z}}}^*) = -\frac{m_a}{e_a^2 B^2(\bar{\mathbf{X}})} \frac{\partial \mathcal{J}[\mathcal{H}_1]}{\partial \bar{\mathbf{X}}_\perp} \tag{19}$$

In the long wavelength electrostatic limit, a simple expression is found

$$\langle \boldsymbol{\rho}_1 \rangle_{\bar{\varsigma}} (\bar{\boldsymbol{\mathcal{Z}}}^*) = \frac{m_a}{e_a B^2(\bar{\mathbf{X}})} \mathbf{E}_\perp(\bar{\mathbf{X}}_\perp) \tag{20}$$

The distortion $\langle \boldsymbol{\rho}_1 \rangle_{\bar{\varsigma}}$ corresponds to a mean displacement of the gyrocentre compared with the particle position, and is thus assimilable to an electric dipole responsible for plasma polarisation effects [14] (see section D.6 and Fig.7). Indeed, the polarisation field $\mathbf{P}(\mathbf{x})$ is the sum over the unperturbed distribution of gyrocentres of this displacement times the particle charge e, calculated at the particle position in the physical space

$$\mathbf{P}(\mathbf{x}) \simeq \frac{N_{eq} m_a}{B_{eq}^2(\mathbf{x})} \mathbf{E}_\perp(\mathbf{x})$$

The medium permittivity is then $\epsilon_{pol} \simeq \sum_{species} \frac{N_{eq} m_a}{B_{eq}^2}$.

3.4 Gyrokinetic Vlasov equation

Let us start from the original set of coordinates $\mathbf{z} = (\mathbf{x}, \mathbf{p})$. The distribution function of particles is some function $f(\mathbf{z}, t)$ solution of the Vlasov equation

$$\frac{\partial f}{\partial t} - \{h, f\}_{\mathbf{z}} = 0$$

which can be written in the extended phase space as

$$\{\mathtt{h}, \mathtt{f}\}_{\mathfrak{z}} = 0$$

where $\mathtt{h}$ and $\mathtt{f}$ are the Hamiltonian and the particle distribution function both in the 8D extended phase space. The change of coordinates from the original set $\mathfrak{z}$ to the new set $\boldsymbol{\mathcal{Z}} = (\mathbf{X}, u_\parallel, \mu, \varsigma, t, E)$ preserves the structure of the Vlasov equation [7], which thus reads

$$\{\mathcal{H}, \mathcal{F}\}_{\boldsymbol{\mathcal{Z}}} = 0$$

with $\mathtt{f}(\mathfrak{z}) = \mathcal{F}(\boldsymbol{\mathcal{Z}})$ and $\mathtt{h}(\mathfrak{z}) = \mathcal{H}(\boldsymbol{\mathcal{Z}})$. After this first step, the Vlasov equation should be written in the gyrocentre phase space. The gyrocentre distribution function is related to the old distribution function via the relation

$$\bar{\mathcal{F}}(\bar{\boldsymbol{\mathcal{Z}}}) = \mathcal{F}(\boldsymbol{\mathcal{Z}}) \tag{21}$$

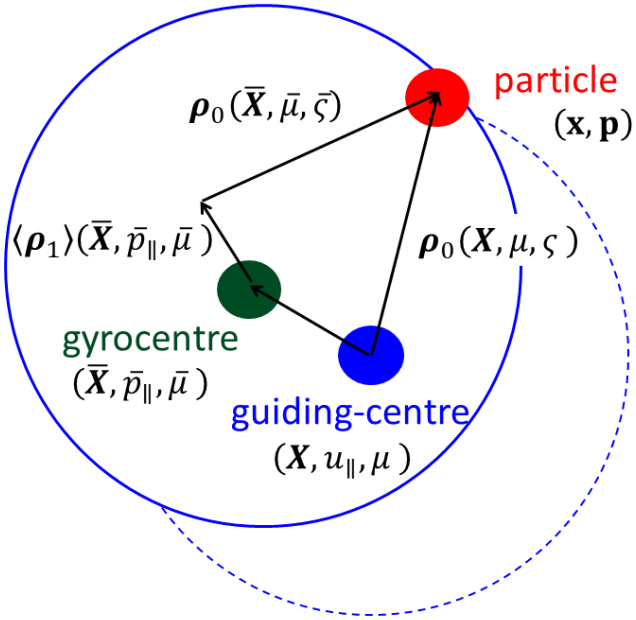


Figure 7: Schematic construction of the gyrocentre position from the particle position. The displacement between the particle position and the gyrocentre is the sum of a cyclotron motion $\boldsymbol{\rho}_0$ and a smaller component $\boldsymbol{\rho}_1$. The later contains a part that oscillates with the gyroangle (not shown on the figure) and a mean part $\langle \boldsymbol{\rho}_1 \rangle_{\bar{\varsigma}}$ that is responsible for plasma polarisation. This displacement results from the distortion of the cyclotron motion under the effect of the electromagnetic field.

where the gyrocentre coordinates $\bar{\boldsymbol{\mathcal{Z}}}$ are functions of the guiding-centre coordinates $\boldsymbol{\mathcal{Z}}$ - this relationship is requested to be invertible. The meaning of this definition is not as intuitive as it seems, in particular because the functions $\bar{\mathcal{F}}(\bar{\boldsymbol{\mathcal{Z}}})$ and $\mathcal{F}(\boldsymbol{\mathcal{Z}})$ are expressed by using *two different set of variables* (see discussion in Appendix D). It is useful to express the distribution functions $\mathcal{F}$ and $\bar{\mathcal{F}}$ *with the same set of variables*. An appropriate choice for the section to come on the Maxwell equation is the set of guiding-centre coordinates $\boldsymbol{\mathcal{Z}}$. Let us note first that at the requested order, Eq.(16) can be written as well in the guiding-centre coordinates

$$-\{\mathcal{H}_{eq}, S_1\}_{\boldsymbol{\mathcal{Z}}} = \mathcal{H}_1(\boldsymbol{\mathcal{Z}}^*, \varsigma) - \bar{\mathcal{H}}_1(\boldsymbol{\mathcal{Z}}^*)$$

Using again the relation $\bar{\boldsymbol{\mathcal{Z}}} = \boldsymbol{\mathcal{Z}} + \epsilon_\delta \{S_1, \boldsymbol{\mathcal{Z}}\}$ and a Taylor development on $\bar{\mathcal{F}}$, one gets the relation

$$\mathcal{F}(\boldsymbol{\mathcal{Z}}^*, \varsigma) = \bar{\mathcal{F}}(\boldsymbol{\mathcal{Z}}^*, \varsigma) + \mathcal{F}'_{eq}(\boldsymbol{\mathcal{Z}}^*) \left(\mathcal{H}_1(\boldsymbol{\mathcal{Z}}^*, \varsigma) - \bar{\mathcal{H}}_1(\boldsymbol{\mathcal{Z}}^*)\right) + o\left(\epsilon_\delta^2\right) \tag{22}$$

where $\mathcal{F}_{eq}(\boldsymbol{\mathcal{Z}}^*)$ is the solution of the unperturbed Vlasov equation, hence does not depend on the gyroangle ς[21], and

$$\mathcal{F}'_{eq}(\boldsymbol{\mathcal{Z}}^*) = \frac{1}{B_{eq}(\mathbf{X})} \frac{\partial \mathcal{F}_{eq}(\boldsymbol{\mathcal{Z}}^*)}{\partial \mu}$$

An inspection of the solution at zeroth order in ϵ implies that $\bar{\mathcal{F}}_{eq}(\bar{\boldsymbol{\mathcal{Z}}}^*) = \mathcal{F}_{eq}(\boldsymbol{\mathcal{Z}}^*)$. Fields in Eq.(22) that do not depend on the cyclotron angle are highlighted by their dependence on $\boldsymbol{\mathcal{Z}}^*$. It is quite important to understand the difference between the relations Eq.(21) and Eq.(22). In Eq.(21), the guiding-centre distribution function calculated at $\boldsymbol{\mathcal{Z}}$ is equal to the gyrocentre distribution computed at $\bar{\boldsymbol{\mathcal{Z}}}$. In Eq.(22), *both distribution functions are calculated at the same guiding-centre position* $\boldsymbol{\mathcal{Z}}$. The difference between $\mathcal{F}$ and $\bar{\mathcal{F}}$ comes

[21] A consequence of the independence of the unperturbed Hamiltonian on ς.

from the displacement between the guiding-centre and the gyrocentre. As discussed in section 3.3, this displacement is responsible for a plasma electric polarisation and magnetisation, and it is therefore essential to keep it fully (see Fig. 8).

The Poisson bracket $\{\mathcal{H},\mathcal{F}\}$ is left invariant by a gyrocentre change of coordinates (see Appendix D), so that $\{\mathcal{H},\mathcal{F}\}_{\boldsymbol{\mathcal{Z}}}=0$ implies the compact, but essential, result

$$\left\{\bar{\mathcal{H}},\bar{\mathcal{F}}\right\}_{\bar{\boldsymbol{\mathcal{Z}}}}=0$$

The symplectic part of the Lagrangian that rules the particle action is the same in the new and old set of variables within the context of the "Hamiltonian" approach - hence the Poisson bracket bears the same expression in the variables $\mathbf{Z}$ and $\bar{\mathbf{Z}}$ (and obviously in the variables $\boldsymbol{\mathcal{Z}}$ and $\bar{\boldsymbol{\mathcal{Z}}}$). It is useful for numerical applications to write the Vlasov equation as an evolution equation in time t in the conventional $6D$ phase space. Since the gyrocentre Hamiltonian does not depend on the gyroangle $\bar{\varsigma}$, the distribution function does not depend on $\bar{\varsigma}$ either. In other words the distribution function of gyrocentres depends on slow variables $\boldsymbol{\mathcal{Z}}^*$ only. Let us recall that another consequence of $\partial_{\bar{\varsigma}}\bar{\mathcal{H}}=0$ is the invariance of the magnetic moment $\bar{\mu}$. Hence the gyrocentre distribution functions depends only on 4 dynamical variables $(\bar{\mathbf{X}},\bar{p}_{\parallel})$, and time t, while $\bar{\mu}$ plays the role of a parameter. The gyrokinetic Vlasov equation reads[22]

$$\frac{\partial\bar{F}}{\partial t}-\left\{\bar{H},\bar{F}\right\}_{\bar{\mathbf{Z}}}=0 \tag{23}$$

A more explicit form is

$$\frac{\partial}{\partial t}\left(B_{||}^{*}\bar{F}\right)+\frac{\partial}{\partial\bar{\mathbf{X}}}\cdot\left(B_{||}^{*}\bar{F}\frac{d\bar{\mathbf{X}}}{dt}\right)+\frac{\partial}{\partial\bar{p}_{\parallel}}\left(B_{||}^{*}\bar{F}\frac{d\bar{p}_{\parallel}}{dt}\right)=0$$

where

$$\frac{d\bar{\mathbf{X}}}{dt} = -\left\{\bar{H},\bar{\mathbf{X}}\right\}=\frac{1}{m_a}\frac{d\bar{H}}{dp_{\parallel}}\frac{\mathbf{B}^{*}}{B_{||}^{*}}+\frac{1}{e_aB_{||}^{*}}\mathbf{e}_{\parallel}\times\nabla\bar{H} \tag{24}$$

$$m_a\frac{d\bar{p}_{\parallel}}{dt} = -\left\{\bar{H},m_a\bar{p}_{\parallel}\right\}=-\left(\frac{\mathbf{B}^{*}}{B_{||}^{*}}\cdot\nabla\right)\bar{H} \tag{25}$$

Here the Hamiltonian is[23]

$$\bar{H}(\bar{\boldsymbol{\mathcal{Z}}}^{*})=\frac{1}{2}m_a\bar{p}_{\parallel}^{2}+\mu B_{eq}(\bar{\mathbf{X}})+e_a\phi_{eq}(\bar{\mathbf{X}})+\mathcal{J}[\mathcal{H}_1]\left(\bar{\mathbf{X}},\bar{p}_{\parallel},\bar{\mu},t\right)+\bar{H}_2(\bar{\boldsymbol{\mathcal{Z}}}^{*})$$

where

$$\begin{aligned}\mathcal{J}[\mathcal{H}_1](\bar{\mathbf{X}},\bar{p}_{\parallel},\bar{\mu},t) = & \int\frac{d\bar{\varsigma}}{2\pi}\Big\{e_a\tilde{\phi}\left(\bar{\mathbf{X}}+\boldsymbol{\rho}_0(\bar{\mathbf{X}},\bar{\mu},\bar{\varsigma}),t\right)-e_ap_{\parallel}\tilde{A}_{\parallel}\left(\bar{\mathbf{X}}+\boldsymbol{\rho}_0(\bar{\mathbf{X}},\bar{\mu},\bar{\varsigma}),t\right)\\ & -e_a\left(\dot{\bar{\mathbf{X}}}_{\perp}+\dot{\boldsymbol{\rho}}_0(\bar{\mathbf{X}},\bar{\mu},\bar{\varsigma})\right)\cdot\tilde{\mathbf{A}}_{\perp}\left(\mathbf{X}+\boldsymbol{\rho}_0(\bar{\mathbf{X}},\bar{\mu},\bar{\varsigma}),t\right)\Big\}\end{aligned}$$

The gyrocentre equations motion Eqs.(24,25) contains the information on the unperturbed motion (guiding-centre - see note on trajectories), plus the effect of the perturbed electromagnetic field. The r.h.s. of Eq.(24) covers both the perturbed $E\times B$ electric drift and magnetic flutter, as described in Fig.3. Equation Eq.(25) yields the parallel force,

[22]Strictly speaking, time should be noted $\bar{t}$ in the equations below. However $\bar{t}=t$ since the generating function does not depend on the energy. Hence we will simplify the notations by changing safely $\bar{t}$ in t.

[23]The full gyrocentre Hamiltonian is formally kept here, i.e. second order, whereas the derivation is so far first order. This does not do any harm as these equations are conservative for any Hamiltonian. We will see that a more rigorous derivation, consistent with the symplectic derivation, actually brings naturally this second order term.

essentially the effect of the perturbed parallel electric field. Note that the force does not contain an induction term $\partial_t \tilde{A}_\parallel$, as would be expected from the traditional equations of motion in presence of an electromagnetic field. This is a consequence of the Hamiltonian approach adopted here. This important point will be discussed in detail in section 5.2.

In summary, a gyrokinetic kinetic equation has been derived that rules the distribution function of gyrocentres $\bar{\mathcal{F}}(\bar{\mathbf{X}}, \bar{p}_\parallel, \bar{\mu}, t)$, where $\bar{\mu}$ is an invariant of motion. This distribution function is related to the distribution function expressed in guiding-centres coordinates $\mathcal{F}(\mathbf{X}, p_\parallel, \mu, \varsigma, t)$, using the relation $\bar{\mathcal{F}}(\bar{\boldsymbol{\mathcal{Z}}}) = \mathcal{F}(\boldsymbol{\mathcal{Z}})$, or equivalently Eq.(22). It then remains to express the charge and current densities as moments of the gyrocentre distribution function $\bar{\mathcal{F}}(\bar{\mathbf{X}}, \bar{p}_\parallel, \bar{\mu}, t)$.

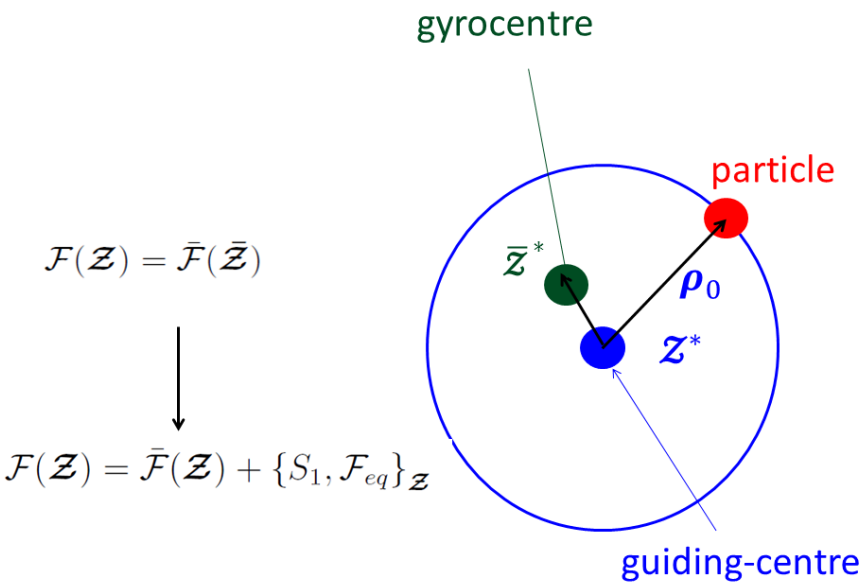


Figure 8: Rationale of a first order gyrokinetic approach. The guiding-centre distribution function $\mathcal{F}$ in the corresponding set of variables $\boldsymbol{\mathcal{Z}}$ is equal to the new gyrocentre distribution $\bar{\mathcal{F}}$ expressed in the gyrocentre variables $\bar{\boldsymbol{\mathcal{Z}}}$. The gyrocentre and guiding-centre coordinates are shifted by a displacement $\{S_1, \boldsymbol{\mathcal{Z}}\}_{\boldsymbol{\mathcal{Z}}}$. The old and new distribution functions $\mathcal{F}$ and $\bar{\mathcal{F}}$ can be expressed at first order in the same set of coordinates $\boldsymbol{\mathcal{Z}}$.

4 Maxwell equations

Solving the Maxwell equations is a delicate step in gyrokinetics. Indeed the Gauss and Ampère equations involve charge and current densities in the physical space, whereas the gyrokinetic Vlasov equation provides the distribution function in the gyrocentre coordinate space, and therefore yields moments that will be called "charge and current densities of gyrocentres". There is no reason why gyrocentre and particle charge and current densities should coincide, and in fact, they do not (see Fig.9). There are two distinct reasons for this difference.

First a plasma can be seen as a collection of gyrocentres (which play the role of free charges) in a medium that is both polarised and magnetised [14]. Indeed the cyclotron motion is deformed under the effect of an electric field in presence of a strong magnetic field (see section 3.3). This deformation contains a mean displacement between the guiding-centre and gyrocentre positions, and an oscillating part (see Fig. 10). This charge shift is

assimilated to an electric dipole that produces a polarisation field. At long wavelengths, the displacement $\langle\bar{\boldsymbol{\rho}}_1\rangle_{\bar{\varsigma}}$ given by Eq.(20) is equal to $(m_a/e_aB_{eq}^2)\mathbf{E}_\perp$ for each species (m is the particle mass, and $\mathbf{E}_\perp$ the transverse electric field). It corresponds to a polarisation field [14]

$$\mathbf{P} = \sum_{species} N_{eq} e_a \langle\bar{\boldsymbol{\rho}}_1\rangle_{\bar{\varsigma}} = \sum_{species} \frac{N_{eq} m_a}{B_{eq}^2}\mathbf{E}_\perp$$

where N_{eq} is the equilibrium density for a given species and

$$\epsilon_{pol} = \sum_{species} \frac{N_{eq} m_a}{B_{eq}^2}$$

is an effective permittivity, which is larger than the vacuum permittivity ϵ_0 in most magnetised fusion plasmas. The time derivative of the corresponding polarisation field is the polarisation current $\frac{\partial \mathbf{P}}{\partial t}$. Magnetisation comes from the current produced by the cyclotron motion. The particle behaves as a small magnet centred on the gyrocentre, with a magnetic moment [14]

$$\mathbf{m} = \frac{1}{2} e_a \langle\boldsymbol{\rho}_0 \times \dot{\boldsymbol{\rho}}_0\rangle_{\bar{\varsigma}} = -\bar{\mu}\mathbf{e}_\parallel$$

Assuming a Maxwellian distribution functions of gyrocentres, the linearised magnetisation field is (see Fig.11)

$$\mathbf{M} = -\sum_{species} \frac{N_{eq} T_{eq}}{B^2}\mathbf{B}$$

This mechanism is responsible for a plasma diamagnetic response, since the magnetisation is opposed to the external field $\mathbf{B}$. The order magnitude of the ratio $\mu_0 M/B$ is the plasma beta parameter, defined as[24]

$$\beta = -\sum_{species} \frac{\mu_0 N_{eq} T_{eq}}{B_{eq}^2}$$

The corresponding diamagnetic current is $\mathbf{J} = \nabla \times \mathbf{M}$. Polarisation and magnetisation effects are additive, so that the total current is

$$\mathbf{J} = \mathbf{J}_{gy} + \frac{\partial \mathbf{P}}{\partial t} + \nabla \times \mathbf{M}$$

where $\mathbf{J}_{gy}$ is the current carried by gyrocentres[25].

The other reason that explains the difference between particle and gyrocentre charge and current densities has to do with particle counting. A density is a number of particles divided by a volume calculated in the limit of vanishing volume size. In an inhomogeneous plasma, densities of particles and gyrocentres differ because some particles have their gyrocentre counterpart outside the volume, and vice-versa, as illustrated in Fig.12. These are called finite Larmor radius (FLR) effects. The same effect occurs to the current density.

These delicate points are detailed in the next section. We use here a straightforward approach that consists in moving back from the gyrocentre to the guiding-centre phase space. In this approach, the Vlasov equation is solved in gyrocentre coordinates. The resulting distribution function is subsequently expressed in guiding-centre coordinates, using Eq.(22). This step is readily followed by a (pull-back) transform to the physical

[24]Strictly speaking a factor 2 should be incorporated in the definition of β, as it usually defined as the ratio of the kinetic to magnetic pressures. As a general rule, numerical factors will be avoided in the definition of the dimensionless parameters.

[25]These effects are already present for guiding-centres alone (just replace $\bar{\mu}$ by μ for instance). The expression of the current is related to the so-called Spitzer paradox: the guiding-centre current is very different from the "true" current - the difference is due to the magnetisation.

space to solve the Maxwell equations (see Fig.13). The advantage of this approach is that the expression of second order gyrocentre Hamiltonian is not needed. The last section will be dedicated to another approach where one moves directly from the gyrocentre to the physical space. The latter requires an explicit computation of the second order gyrocentre Hamiltonian, a difficult step that is postponed for now.

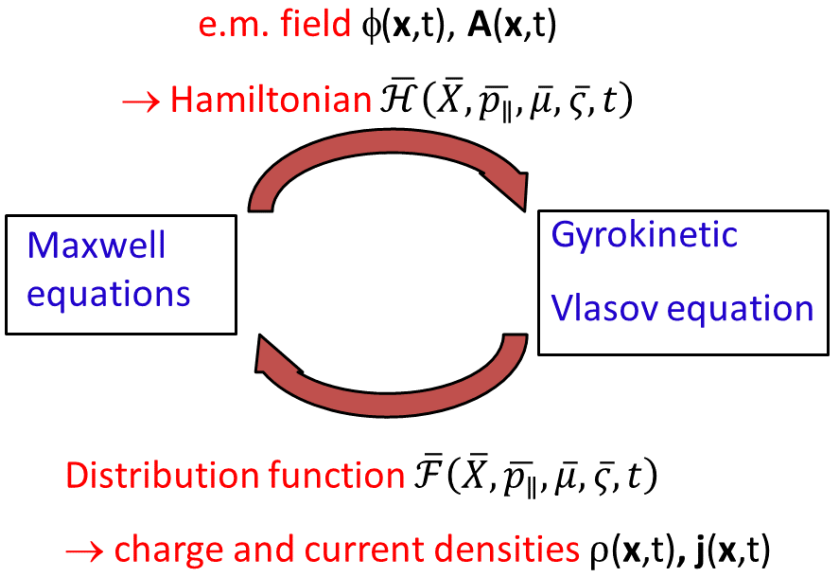


Figure 9: Schematics of a self-consistent gyrokinetic model. The electromagnetic field produces an Hamiltonian that is written in gyrocentre coordinates to solve the gyrokinetic equations. The resulting distribution function is used to compute particle charge and current densities, which are then inserted in the Maxwell equations in the physical space to produce an updated electromagnetic field.

4.1 Preamble: field action principle

Before presenting the electromagnetic action in its full beauty, let us remind how an action principle works for a scalar field. The idea is the same as for computing the equations of motion of a particle. Let us consider a field ϕ that depends on the 4 coordinates $x^\mu = (t, \mathbf{x})$. A field action is a functional of the field and its derivatives, and reads

$$\mathcal{A} = \int_{t_1}^{t_2} dt \int_{\mathbf{x}_1}^{\mathbf{x}_2} d^3\mathbf{x}\, \mathcal{L}\left(\phi, \partial_\mu \phi\right) \tag{26}$$

where $\mathcal{L}$ is the field Lagrangian, and the notation $\partial_\mu \phi$ means $\frac{\partial \phi}{\partial x^\mu}$. The equations that rule the field ϕ are obtained by expressing that $\mathcal{A}$ is an extremum for variations of the field $\delta\phi(x^\mu)$ such that $\delta\phi = 0$ at the boundary of the integration volume (in a 4D space). An integration by parts provides readily the following expression of the action variation

$$\begin{aligned} \delta\mathcal{A} &= \int_{t_1}^{t_2} dt \int_{\mathbf{x}_1}^{\mathbf{x}_2} d^3\mathbf{x} \left[\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\right)\right] \delta\phi \\ &+ \int_{t_1}^{t_2} dt \int_{\mathbf{x}_1}^{\mathbf{x}_2} d^3\mathbf{x}\, \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta\phi\right) \end{aligned}$$

The last term vanishes. Indeed, it is the integral of the divergence of a "current" $J^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta\phi$. The Gauss-Ostrogradsky theorem implies that it is equal to the current calculated at the surface that bounds the integration volume. Since $\delta\phi = 0$ on this

surface, this piece can just be ignored. It will be seen however that it has its utility when using the Noether's theorem to derive the law of energy conservation. Imposing that $\delta\mathcal{A} = 0$ for all variations of $\delta\phi$, one gets the field equations

$$\frac{\partial \mathcal{L}}{\partial \phi} = \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right)$$

These equations are called Euler-Lagrangian field equations, and are generalisations of the Lagrangian equations that rule the motion of a particle. An example for a simple 1D case is illustrated on Fig.14. The action for the electromagnetic field works on the same principle, but with a scalar field ϕ replaced by a 4-field $(\phi, \mathbf{A})$, where ϕ and $\mathbf{A}$ are the electric and vector potentials.

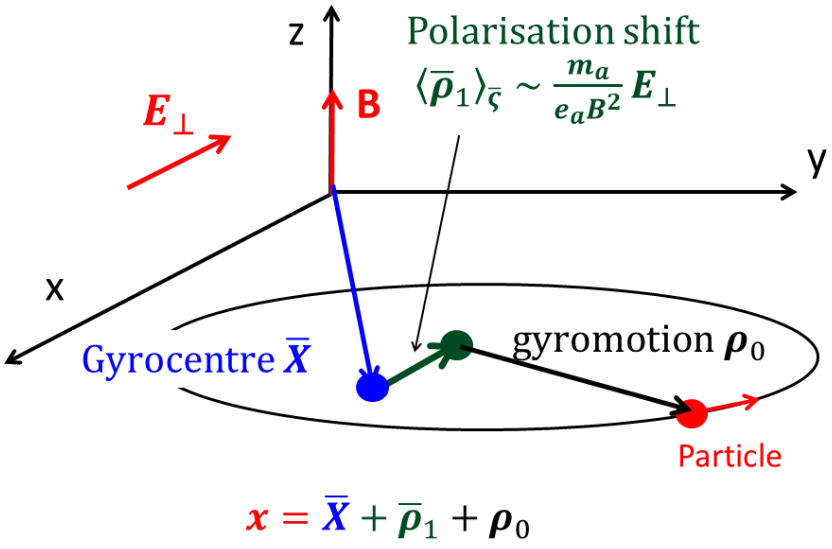


Figure 10: The displacement vector between a charged particle position $\mathbf{x}$ and its gyrocentre $\bar{\mathbf{X}}$ is the sum of a cyclotron motion $\boldsymbol{\rho}_0(\bar{\mathbf{X}}, \bar{\mu}, \bar{\varsigma})$ and a lowest order component $\bar{\boldsymbol{\rho}}_1(\bar{\mathbf{X}}, \bar{p}_\parallel, \bar{\mu}, \bar{\varsigma})$. In the electrostatic case, the mean part of the later $\langle \bar{\boldsymbol{\rho}}_1 \rangle (\bar{\mathbf{X}}, \bar{\mu})$ is co-linear with the perpendicular electric field $\mathbf{E}_\perp$, and roughly of the order of $\frac{m_a}{B_{eq}^2(\mathbf{x})} \mathbf{E}(\mathbf{x})$. This behaviour is responsible for a polarisation field $\mathbf{P} \simeq \sum_{species} N_{eq} e_a \langle \bar{\boldsymbol{\rho}}_1 \rangle_{\bar{\varsigma}} = \sum_{species} \frac{N_{eq} m_a}{B_{eq}^2} \mathbf{E}_\perp$.

4.2 Electromagnetic action principle

The theory of electromagnetic fields [15] states that the field-particle action $\mathcal{A}$ is a sum of an action of the electromagnetic field $\mathcal{A}_{em}$

$$\mathcal{A}_{em} = \frac{1}{2} \int dt\, d^3\mathbf{x} \left(\epsilon_0 \mathbf{E}^2 (\mathbf{x}, t) - \frac{1}{\mu_0} \mathbf{B}^2 (\mathbf{x}, t) \right) \tag{27}$$

and a action $\mathcal{A}_{fp}$ that rules field-particle interactions

$$\mathcal{A}_{fp} = \sum_{species} \int d^3\mathbf{x} dt \left\{ \mathbf{J} (\mathbf{x}, t) \cdot \mathbf{A} (\mathbf{x}, t) - \varrho (\mathbf{x}, t) \phi (\mathbf{x}, t) \right\} \tag{28}$$

where $\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \nabla\phi$ is the electric field, and $\mathbf{B} = \nabla \times \mathbf{A}$ the magnetic field. Zeroing the functional derivatives of the action $\mathcal{A} = \mathcal{A}_{em} + \mathcal{A}_{fp}$ with respect to $(\phi, \mathbf{A})$ leads to

the Maxwell equations, more precisely

$$\begin{aligned} \frac{\delta \mathcal{A}}{\delta \phi} = 0 \quad &\rightarrow \quad \nabla \cdot \mathbf{E} = \frac{\varrho}{\epsilon_0} \\ \frac{\delta \mathcal{A}}{\delta \mathbf{A}} = 0 \quad &\rightarrow \quad \frac{1}{\mu_0} \nabla \times \mathbf{B} = \mathbf{J} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \end{aligned} \tag{29}$$

The charge and current densities are functional derivatives of the field-particle functional,

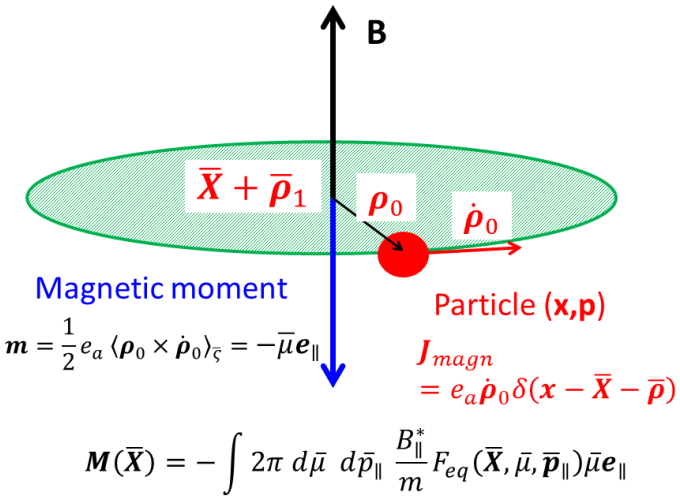


Figure 11: A particle that moves fast on a circle, a cyclotron motion, behaves as a small magnet with amplitude $\bar{\mu}$, with a direction opposite to the magnetic field. This process is responsible for diamagnetism in a magnetised plasma.

i.e. formally

$$\begin{aligned} \varrho &= \frac{\delta \mathcal{A}_{fp}}{\delta \phi} \\ \mathbf{J} &= \frac{\delta \mathcal{A}_{fp}}{\delta \mathbf{A}} \end{aligned} \tag{30}$$

Let us stress that these derivatives can be performed only if the field-particle action $\mathcal{A}_{fp}$ is of the form Eq.(28). It will be seen in the following that the action $\mathcal{A}_{fp}$, once expressed as a function of the gyrocentre distribution function, rather involves operators that apply on the electric and vector potentials. A major difficulty is to reshape it as an action of the form Eq.(28). From a physical standpoint, this is equivalent to compute charge and current densities $(\varrho(\mathbf{x}, t), \mathbf{J}(\mathbf{x}, t))$ in the physical 3D space of particle positions, knowing the distribution of guiding centres (resp. gyrocentres), known as a function of guiding-centre $\mathbf{X}$ (resp. gyrocentre $\bar{\mathbf{X}}$) positions. This point is essential. It is easily verified that the wave-particle action can be replaced by

$$\mathcal{A}_{fp} = - \sum_{species} \int d\gamma \mathtt{f}(\mathfrak{z}) \mathtt{h}(\mathfrak{z}) \tag{31}$$

Using the expression of the particle Hamiltonian

$$h(\mathbf{x}, \mathbf{p}, t) = \frac{1}{2m_a} \left[\mathbf{p} - e_a \mathbf{A}(\mathbf{x}, t) \right]^2 + e_a \phi(\mathbf{x}, t)$$

where

$$d\gamma = d^3\mathbf{x} d^3\mathbf{p} dt dE \, \delta \left(E - h_{eq}(\mathbf{x}, \mathbf{p}) \right) \tag{32}$$

is the volume element in the particle extended phase space. Maxwell equations found in this way are consistent with the classical expressions of the charge and current densities

$$\begin{aligned}\varrho\left(\mathbf{x},t\right) &= e_a \int d^3\mathbf{p} f\left(\mathbf{x},\mathbf{p},t\right) \\ \mathbf{J}\left(\mathbf{x},t\right) &= e_a \int d^3\mathbf{p} f\left(\mathbf{x},\mathbf{p}\right)\mathbf{v}\left(\mathbf{x},\mathbf{p},t\right)\end{aligned}$$

Quite interestingly, if all possible variations of the distribution functions are constrained to be of the form $\delta f = \{S, f\}$ [12, 13] (for an overview, see [2] and references therein)[26], the action principle allows recovering the Vlasov equation in the extended phase space[27]

$$\{h, f\}_{\mathfrak{z}} = 0$$

The action Eq.(31) can be written in the set of guiding-centre or gyrocentre coordinates $\boldsymbol{\mathcal{Z}}$ or $\bar{\boldsymbol{\mathcal{Z}}}$, e.g.

$$\mathcal{A}_{fp} = -\sum_{species}\int d\Gamma \mathcal{F}(\boldsymbol{\mathcal{Z}})\mathcal{H}(\boldsymbol{\mathcal{Z}}) = -\sum_{species}\int d\bar{\Gamma}\bar{\mathcal{F}}(\bar{\boldsymbol{\mathcal{Z}}})\bar{\mathcal{H}}(\bar{\boldsymbol{\mathcal{Z}}}) \tag{33}$$

where $d\Gamma$ (resp. $d\bar{\Gamma}$) is the volume element in the guiding-centre (resp. gyrocentre) phase space. The difficulty now appears quite clearly. As said, the Maxwell equations are retrieved from the action principle in the physical space $\mathbf{x}$. On the other hand solving a gyrokinetic Vlasov equation provides $\bar{\mathcal{F}}(\bar{\boldsymbol{\mathcal{Z}}})$, hence a function that lives in the gyrocentre phase space.

So far, the action principle applies to the full electromagnetic field. Since the equilibrium magnetic field is supposed given, it is used from now on to compute the perturbed electromagnetic field, including the unperturbed electric field. The unperturbed electric potential is treated on the same ground as the perturbed one. A shifted velocity $\mathbf{v}'$ is introduced, that was already used to compute the gyrocentre equations of motion. Indeed the velocity $\mathbf{v}$ is a dynamical variable, i.e. a function of $(\mathbf{x},\mathbf{p},t)$. It can be split in unperturbed $\mathbf{v}_{eq}\left(\mathbf{x},\mathbf{p}\right)$ and perturbed $\mathbf{v}_1\left(\mathbf{x},\mathbf{p},t\right)$ components

$$\begin{aligned}\mathbf{v}_{eq}\left(\mathbf{x},\mathbf{p}\right) &= \frac{1}{m_a}\left(\mathbf{p} - e_a\mathbf{A}_{eq}(\mathbf{x})\right) \\ \mathbf{v}_1\left(\mathbf{x},\mathbf{p},t\right) &= -\frac{e_a}{m_a}\tilde{\mathbf{A}}(\mathbf{x},t)\end{aligned}$$

It then appears readily that $\mathbf{v}_{eq} = \mathbf{v} - \mathbf{v}_1$ is the same as $\mathbf{v}'$[28]. The new Hamiltonian becomes h^* as given by Eq.(9). The field-particle Lagrangian has to be computed at second order in ϵ_δ^2. It can be recast as

$$\begin{aligned}\mathcal{A}_{fp} &= -\sum_{species}\int d\gamma f_1(\mathfrak{z})\mathtt{h}_1(\mathfrak{z}) \\ &+ \frac{1}{2}\sum_{species}\int d\gamma\, \mathcal{F}_{eq}(\mathbf{x},\mathbf{p})\,\frac{e_a^2}{m_a}\tilde{\mathbf{A}}^2(\mathbf{x},t)\end{aligned} \tag{34}$$

The volume element $d\gamma$ is the same as Eq.(32) with h_{eq} replaced by h^*_{eq}. The 1st order perturbed Hamiltonian is

$$\mathtt{h}_1(\mathbf{x},\mathbf{p},t) = e_a\tilde{\phi}(\mathbf{x},t) - e_a\mathbf{v}_{eq}(\mathbf{x},\mathbf{p})\cdot\tilde{\mathbf{A}}(\mathbf{x},t)$$

[26]This can be roughly understood as follows. For any canonical change of variables $\bar{\mathfrak{z}} = \mathfrak{z} + \{S, \mathfrak{z}\}$ with small S, the variation $\delta f = f(\bar{\mathfrak{z}}) - f(\mathfrak{z}) \simeq \frac{\partial f}{\partial \mathfrak{z}}\cdot(\bar{\mathfrak{z}} - \mathfrak{z})$ is equal to $\delta f = \{S, f\}_{\mathfrak{z}}$.

[27]This is a consequence of the identity $\int d\gamma\,\{S, f\}\,h = \int d\gamma S\,\{f, h\}$ up to a divergence term that will be discussed in the section on energy conservation. Hence $\delta\mathcal{A}_{fp} = 0$ for any S implies Eq.(33)

[28]Note however that $\mathbf{v}'$ was calculated in the $(\mathbf{x},\mathbf{v})$ variables, while $\mathbf{v}_{eq}$ is a function of $(\mathbf{x},\mathbf{p})$. The notations will be therefore kept different

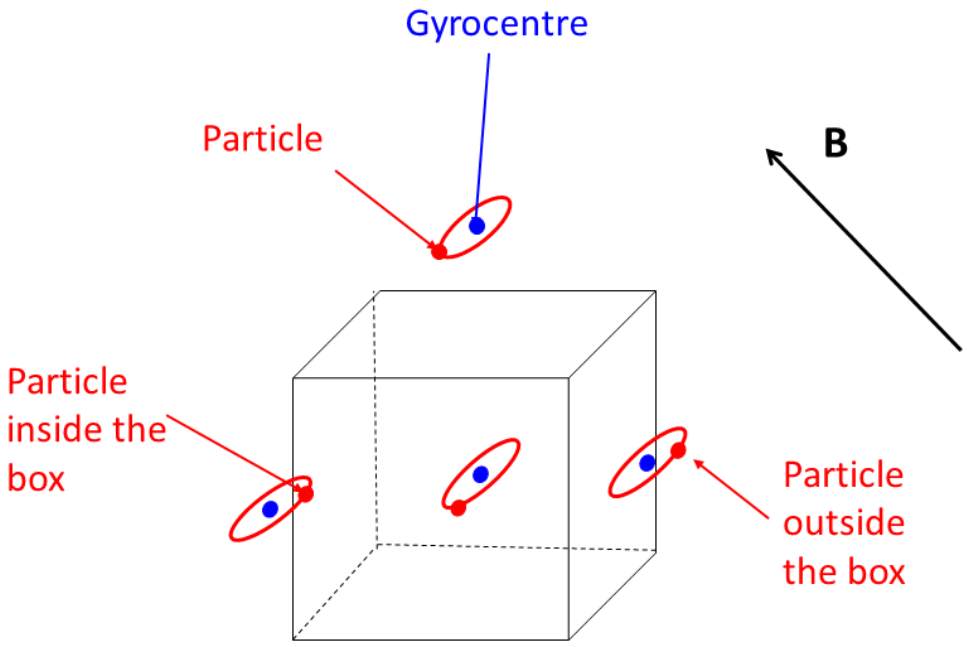


Figure 12: Calculation of the particle density knowing the distribution of gyrocentres. When counting particles in a box, some particles have their gyrocentres outside the box, while the opposite may also appear. When the plasma is inhomogeneous, this leads to differences between the densities of particles and gyrocentres. A detailed calculation leads to a particle density in accordance with Eq.(49).

The reader may be somewhat disconcerted by the second line of Eq.(34), which is dubbed "skin depth action", reason being that it represents physically a plasma skin depth effect[29]. Its appearance is a direct consequence of the Hamiltonian approach that is chosen here, i.e. the fact that the velocity is a dynamical variable, hence a function of $(\mathbf{x}, \mathbf{p}, t)$, not a coordinate. This skin depth action deserves some comments. Let us note first the pre-factor 1/2. It must be recalled that the current is derived as a functional derivative of the action with respect to $\tilde{\mathbf{A}}$. However, if the current is itself proportional to the vector potential, a factor 1/2 must be introduced since the functional derivative of $\int d^3\mathbf{x}dt\tilde{\mathbf{A}}^2$ is $2\tilde{\mathbf{A}}$, not $\tilde{\mathbf{A}}$. Forgetting the pre-factor 1/2 would lead to a current that is twice too large. We will meet again this subtle point when treating the polarisation/magnetisation terms. It occurs whenever the proportionality of the plasma response to the field is made explicit. Second, integrating the skin depth functional over the momenta $\mathbf{p}$ seems attractive. This is not a smart move since it will be seen that unperturbed distribution function that is considered as given is the guiding-centre equilibrium distribution function. Hence the unperturbed density N_{eq} is a density of guiding-centres, not of particles.

A contribution $f_2 h^*_{eq}$ was omitted here. It does not play any role in the calculation of the perturbed field. This term represents physically a transport equation that rules the long time scale evolution of the distribution function. It is of course included in the

[29]The plasma skin depth is the penetration length of a high frequency wave that propagates into a collisionless plasma. The skin depth is defined as $d_e = \sqrt{m_e/\mu_0 N_{eq} e_a^2}$, where N_{eq} is the equilibrium density, or alternatively as $d_e = c/\omega_{pe}$, where ω_{pe} is the plasma frequency, and c the speed of light. The corresponding current can be recast as $-\tilde{A}_\parallel/\mu_0 d_e^2$.

gyrokinetic equation, and can be added at the very end in the full action principle for completion.

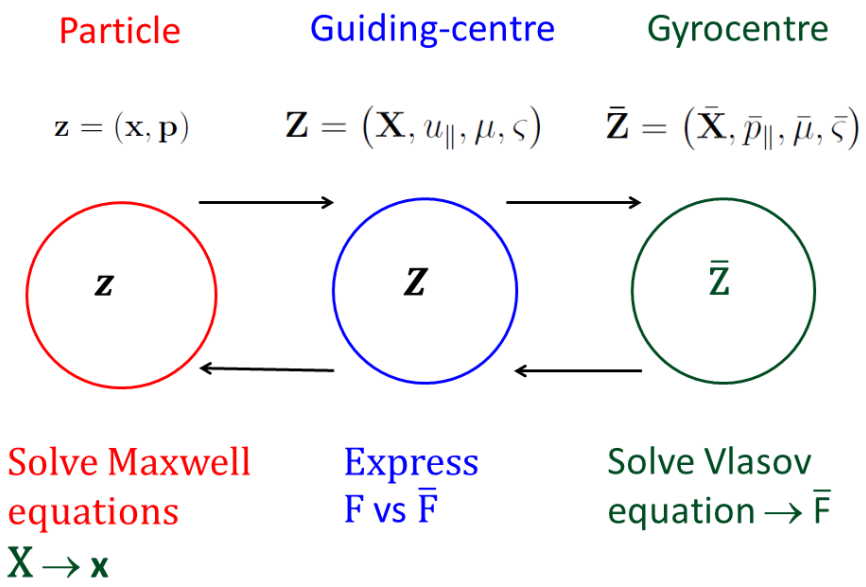


Figure 13: Strategy of 1st order gyrokinetic theory. The Vlasov equation is solved in gyrocentre coordinates. It is then expressed in guiding-centre coordinates, followed by a pull-back to the physical space to solve the Maxwell equations. Polarisation effects are contained in the difference between $F(\boldsymbol{\mathcal{Z}})$ and $\bar{F}(\boldsymbol{\mathcal{Z}})$

4.3 Variational principle in guiding-centre variables

The advantage of a variational approach is that the integrand of the field-particle Lagrangian Eq.(34) lives naturally in the extended phase space. The set of coordinates $\boldsymbol{\mathcal{Z}}$ is chosen here as an extended set of *non conjugate* guiding-centre coordinates $\boldsymbol{\mathcal{Z}} = (\boldsymbol{\mathcal{Z}}^*, \varsigma)$, where $\boldsymbol{\mathcal{Z}}^*$ is a group of 7 "slow" guiding-centre coordinates, conveniently chosen here as $\boldsymbol{\mathcal{Z}}^* = (\mathbf{X}, p_\parallel, \mu, E, t)$, while ς is the "fast" gyroangle. Here $p_\parallel$ is the parallel component of the guiding-centre velocity. The energy coordinate E is eliminated via a straightforward integration over the Kronecker delta function $\delta\left(E - H_{eq}(\boldsymbol{\mathcal{Z}})\right)$. The element of volume integration in the phase space is now $d\Gamma = d\Gamma^* \frac{d\varsigma}{2\pi}$, where[30]

$$\Gamma^* = d^3\mathbf{X} dt d^3\mathbf{V}$$

is a 5D guiding-centre phase space volume element and

$$d^3\mathbf{V} = \frac{2\pi}{m_a} B^*_\parallel dp_\parallel d\mu$$

is the element of volume in the guiding-centre velocity space. The unperturbed distribution function of guiding-centres, noted $\mathcal{F}_{eq}$, depends on the slow variables $\boldsymbol{\mathcal{Z}}^*$ only, since it must be solution of the Vlasov equation $\{\mathcal{H}_{eq}, \mathcal{F}_{eq}\} = 0$.

The action Eq.(34) keeps the same form with $\mathfrak{z}$ replaced by $\boldsymbol{\mathcal{Z}}$, and $d\gamma$ replaced by $d\Gamma$. The unperturbed particle velocity reads

$$\mathbf{v}_{eq} = p_\parallel \mathbf{e}_\parallel + \dot{\mathbf{X}}_\perp + \dot{\boldsymbol{\rho}}_0$$

[30] $B^*_\parallel(\mathbf{X}, p_\parallel)$ is the Jacobian of the guiding-centre set of coordinates, hence its appearance in the velocity element of integration $d^3\mathbf{V}$.

Hence the functional $\mathcal{A}_{fp}$ can be split in two parts, a "skin depth" contribution, which reads for one species

$$\mathcal{A}_{skin} = -\frac{e_a^2}{2m_a} \int d\Gamma \, \mathcal{F}_{eq}(\boldsymbol{\mathcal{Z}}^*) \tilde{\mathbf{A}}^2(\mathbf{X} + \boldsymbol{\rho}_0, t) \tag{35}$$

and a modified field-particle functional expressed as an integral in the extended phase space

$$\mathcal{A}_{part} = -\int d\Gamma \, \mathcal{F}_1 \left(\boldsymbol{\mathcal{Z}}\right) \mathcal{H}_1 \left(\boldsymbol{\mathcal{Z}}\right) \tag{36}$$

where the order 1 perturbed Hamiltonian H_1 is

$$\begin{aligned} \mathcal{H}_1 &= e_a \tilde{\phi}(\mathbf{X} + \boldsymbol{\rho}_0, t) - e_a p_\parallel \tilde{A}_\parallel(\mathbf{X} + \boldsymbol{\rho}_0, t) \\ &- e_a \left(\dot{\mathbf{X}}_\perp + \dot{\boldsymbol{\rho}}_0\right) \cdot \tilde{\mathbf{A}}_\perp(\mathbf{X} + \boldsymbol{\rho}_0, t) \end{aligned}$$

Here the cyclotron motion $(\boldsymbol{\rho}_0, \dot{\boldsymbol{\rho}}_0)$ is expressed as a function of guiding-centre coordinates $(\mathbf{X}, \mu, \varsigma)$[31]. The total functional thus reads

$$\mathcal{A} = \mathcal{A}_{em} + \mathcal{A}_{skin} + \mathcal{A}_{part} \tag{37}$$

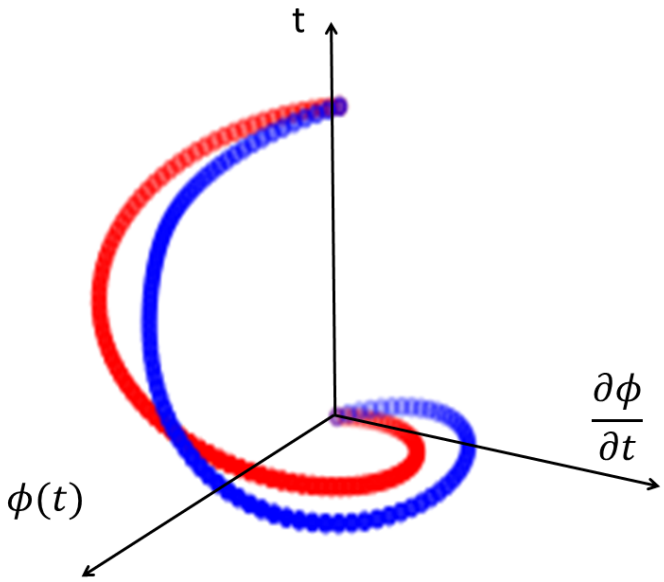


Figure 14: Extremum of the the action $\mathcal{A} = \int_{t_1}^{t_2} dt \mathcal{L}\left(\phi, \partial_t \phi\right)$ where the Lagrangian $\mathcal{L}\left(\phi, \partial_t \phi\right) = \frac{1}{2}\left(\partial_t \phi\right)^2 - \frac{1}{2}\phi^2$. The Euler-Lagrange equation is $\partial_{tt}\phi + \phi = 0$. The solution that satisfies the boundary conditions $\phi = 0$, $\partial_t \phi = \phi_0$ at $t = t_1$ is $\phi(t) = \phi_0 \sin(t - t_1)$. It is plotted in red for the special choice $t_1 = 0$ and $t_2 = 2\pi$. The action is $\mathcal{A} = 0$ for this solution. A neighbouring function is shown in blue that bears the same values at $t = 0$ and $t_2 = 2\pi$, in this peculiar example $\phi(t) = \phi_0 \sin t(1 + \epsilon \sin(t))$. The corresponding action is non zero, namely $\mathcal{A} = \frac{\pi}{8}\epsilon^2 \phi_0^2$.

The functional $\mathcal{A}_{part}$ must now be detailed. Since it is written in guiding-centre coordinates, the gyrocentre distribution function must be expressed in the same set of variables. This is done by using the relationship Eq.(22) between $\mathcal{F}$ and $\bar{\mathcal{F}}$ in the guiding-centre

[31] There no dependence on the parallel momentum $p_\parallel$ since $\boldsymbol{\rho}_0$ is restricted here to a circular cyclotron motion that depends on $(\mathbf{X}, \mu, \varsigma)$ ony.

phase space. One has to be extra-careful at this step. Charge and current densities are computed by finding an extremum of the action *at constant distribution function*. However Eq.(22) contains terms, related to plasma polarisation and magnetisation, that are proportional to the field. When building the action, the contribution must be multiplying by a factor 1/2 to counter-balance the factor 2 that pops-out when calculating the extremum[32]. The new action is a functional that involves the perturbed gyrocentre distribution function $\bar{\mathcal{F}}_1$, namely

$$\begin{aligned}\mathcal{A}_{part} &= -\int d\Gamma \bar{\mathcal{F}}_1(\boldsymbol{\mathcal{Z}}^*)\mathcal{H}_1(\boldsymbol{\mathcal{Z}}) \\ &- \frac{1}{2}\int d\Gamma \mathcal{F}'_{eq}(\boldsymbol{\mathcal{Z}}^*)\left[\mathcal{H}_1(\boldsymbol{\mathcal{Z}}) - \bar{\mathcal{H}}_1(\boldsymbol{\mathcal{Z}}^*)\right]\mathcal{H}_1(\boldsymbol{\mathcal{Z}})\end{aligned} \tag{38}$$

This expression differs from a naive translation of the particle functional Eq.(36) to a functional in the guiding-centre phase space that would only be the first part in the r.h.s. of Eq.(38). The second line of Eq.(38) involves the difference between the perturbed Hamiltonian and its gyroaverage $\mathcal{H}_1(\boldsymbol{\mathcal{Z}}) - \bar{\mathcal{H}}_1(\boldsymbol{\mathcal{Z}}^*)$. It is responsible for polarisation and magnetisation contributions that play an important role in the field dynamics. An alternative expression to Eq.(38), expressed in "slow" guiding-centre coordinates, is

$$\begin{aligned}\mathcal{A}_{part} = &- \int d\Gamma^* \bar{\mathcal{F}}_1(\boldsymbol{\mathcal{Z}}^*)\mathcal{J}[\mathcal{H}_1](\boldsymbol{\mathcal{Z}}^*) \\ &- \frac{1}{2}\int d\Gamma^* \mathcal{F}'_{eq}(\boldsymbol{\mathcal{Z}}^*)\left[\mathcal{J}[\mathcal{H}_1^2] - (\mathcal{J}[\mathcal{H}_1])^2\right](\boldsymbol{\mathcal{Z}}^*)\end{aligned} \tag{39}$$

where the relation $d\Gamma = d\Gamma^* \frac{d\varsigma}{2\pi}$ has been used, and the gyroaverage operator in guiding-centre variables $\mathcal{J}$ is defined by Eq.(19). A similar treatment can be done for the skin depth functional $\mathcal{A}_{skin}$ Eq.(35). This leads to the alternative expression

$$\mathcal{A}_{skin} = -\sum_{species}\frac{e_a^2}{2m_a}\int d\Gamma^*\, \mathcal{F}_{eq}\left(\boldsymbol{\mathcal{Z}}^*\right)\mathcal{J}[\tilde{\mathbf{A}}^2](\boldsymbol{\mathcal{Z}}^*) \tag{40}$$

The functionals Eq.(39) and Eq.(40) are still not fully tractable. Indeed the electric and vector potentials appear via their gyroaverage, i.e. operators. These operators must be inverted to find the extremum of these functionals with respect to $(\tilde{\phi}, \tilde{\mathbf{A}})$. There are several ways to proceed. If the gyroaverage operator is known as a differential operator, typically via an expansion in gyroradius, then an integration by parts provides the requested form. This is an efficient procedure when the expansion is stopped at low order: typically second order in Larmor radius, so that operators are Laplacians. It becomes rapidly cumbersome when moving to higher orders. A second approach consists in using the definition Eq.(19) of the gyroaverage operator $\mathcal{J}$, and build the adjoint operator. The latter approach will be adopted here. Both techniques are equivalent and consists in deriving charge and current densities, knowing their gyroaverage. Hence a last technical step is needed before properly expressing the Maxwell equations. We will proceed here step by step, starting with the $(\phi, A_\parallel)$ representation.

4.4 The $(\phi, A_\parallel)$ representation

4.4.1 Current and charge densities in the $(\phi, A_\parallel)$ representation

An extremum of the functional Eq.(37) cannot be found easily, because test fields that appear in the functionals Eqs.(39,40) are gyroaveraged. In other words, the functionals

[32] A simple example is given by the action $\mathcal{A} = \int d^3\mathbf{x}\varrho\phi$. The charge density ϱ is given by a functional derivative with respect to ϕ. Suppose now that the charge density is proportional to the potential (adiabatic response), i.e. $\rho = \epsilon\phi$, then a simple replacement would give $\mathcal{A} = \int d^3\mathbf{x}\epsilon\phi^2$. A functional derivative yields a wrong answer, that is a charge density that is twice the right number. Hence the correct answer is $\mathcal{A} = 1/2\int d^3\mathbf{x}\epsilon\phi^2$.

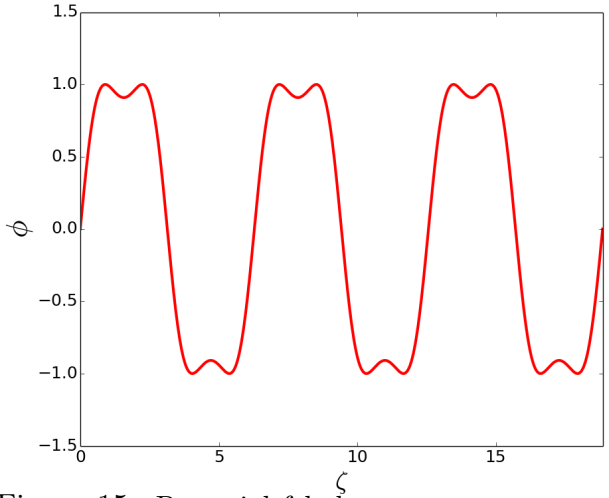


Figure 15: Potential felt by a gyrocentre - example of a potential $\phi(x) = \cos(kx)$ and $x = \rho_c \sin(\varsigma)$ with $k\rho_c = 2$.

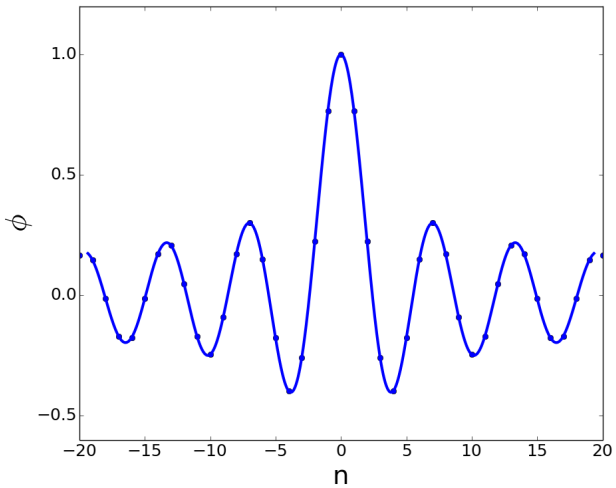


Figure 16: Fourier transform in ζ of the potential ϕ. Values of the discrete Fourier spectrum are shown in blue circles. The blue solid line shows the function $J_0(k\rho_c)$.

Eqs.(39,40) live in the guiding-centre space, whereas a functional must be expressed in the physical space to compute its extremum. To do this, the perturbed Hamiltonian must be scrutinised. A staged approach is adopted here, to avoid excessive technicalities. The analysis is restricted first to a perturbed field that derives from an electric potential $\tilde{\phi}$ and a vector potential $\tilde{A}_{\|}\mathbf{e}_{\|}$ aligned with the equilibrium magnetic field. The perpendicular component of the vector potential is neglected in this first step. This approximation allows treating a large variety of physical problems, including turbulent transport due to electrostatic drift waves (zero plasma β), or shear Alfvén waves (low values of $\beta \ll 1$). In fact many gyrokinetic codes operate in this framework. The parallel component of the perturbed magnetic field is discarded in this approximation, which becomes questionable for increasing values of the plasma β. In particular this framework is not appropriate to address large scale MHD modes such as interchange or kink modes. The corresponding perturbed Hamiltonian is calculated from Eq.(11) and reads

$$\mathcal{H}_1\left(\boldsymbol{\mathcal{Z}}\right) = e_a\tilde{\phi}\left(\mathbf{X} + \boldsymbol{\rho}_0(\mathbf{X},\mu,\varsigma),t\right) - e_a p_{\|}\tilde{A}_{\|}\left(\mathbf{X} + \boldsymbol{\rho}_0(\mathbf{X},\mu,\varsigma),t\right)$$

Its gyroaverage is

$$\mathcal{J}[\mathcal{H}_1] = e_a\mathcal{J}[\tilde{\phi}] - e_a p_{\|}\mathcal{J}[\tilde{A}_{\|}]$$

For any field $\Lambda(\mathbf{x},t)$ that depends on the physical space coordinates $\mathbf{x}$ and time t, the gyroaverage operator reads

$$\mathcal{J}[\Lambda]\left(\boldsymbol{\mathcal{Z}}^*\right) = \int_0^{2\pi}\frac{d\varsigma}{2\pi}\Lambda\left(\mathbf{X} + \boldsymbol{\rho}_0\left(\mathbf{X},\mu,\varsigma\right),t\right)$$

As discussed earlier, a gyroaverage potential is the potential "felt" by a gyrocentre, a virtual particle. A simple example is shown in Fig.15 and Fig.16. The functional Eq.(39) is the sum of 3 contributions, which can be translated to the physical space (details can be found in Appendix D.4)

$$\begin{aligned}
\mathcal{A}_{part} &= -\int d\gamma^*\,\frac{1}{B^*_{||}(\boldsymbol{\mathfrak{z}}^*)}\mathcal{J}^t\left[B^*_{||}\bar{\mathcal{F}}_1\right](\boldsymbol{\mathfrak{z}}^*)\mathcal{H}_1(\boldsymbol{\mathfrak{z}}^*) \\
&- \frac{1}{2}\int d\gamma^*\,\frac{1}{B^*_{||}(\boldsymbol{\mathfrak{z}}^*)}\mathcal{J}^t\left[B^*_{||}\mathcal{F}'_{eq}\right](\boldsymbol{\mathfrak{z}}^*)\mathcal{H}_1^2(\boldsymbol{\mathfrak{z}}^*) \\
&+ \frac{1}{2}\int d\gamma^*\,\frac{1}{B^*_{||}(\boldsymbol{\mathfrak{z}}^*)}\mathcal{J}^t\left[B^*_{||}\mathcal{F}'_{eq}\mathcal{J}[\mathcal{H}_1]\right](\boldsymbol{\mathfrak{z}}^*)\mathcal{H}_1(\boldsymbol{\mathfrak{z}}^*) \qquad (41)
\end{aligned}$$

Similarly the functional Eq.(40) can be rewritten as

$$\mathcal{A}_{skin} = -\frac{e_a^2}{2m_a}\int d\gamma^* \, \frac{1}{B_{||}^*(\mathfrak{z}^*)}\mathcal{J}^t\left[B_{||}^*\mathcal{F}_{eq}\right](\mathfrak{z}^*)\tilde{A}_{||}^2(\mathbf{x},t) \tag{42}$$

where $\mathfrak{z}^*$ designates the *hybrid* set of variables $\left(\mathbf{x},t,p_{||},\mu\right)$, which mixes position coordinates in the physical space and guiding-centre velocity coordinates. The corresponding volume element is

$$d\gamma^* = d^3\mathbf{x}dtd^3\mathbf{V}$$

The operator $\mathcal{J}^t$ is defined for any field $\Lambda(\mathbf{X},t,p_{||},\mu)$ as

$$\mathcal{J}^t[\Lambda]\left(\mathbf{x},t,p_{||},\mu\right) = \int_0^{2\pi}\frac{d\varsigma}{2\pi}\Lambda\left(\mathbf{x}-\boldsymbol{\rho}_0\left(\mathbf{x},\mu,\varsigma\right),t,p_{||},\mu\right) \tag{43}$$

The operator $\mathcal{J}^t$ hence transforms a function of the guiding-centre position coordinates in a function that depends on the particle position in the physical space, but also on the guiding-centre velocity coordinates $p_{||}$ and μ (see Fig.17). Its structure is very close to the gyroaverage operator.

Let us insist on the hybrid character of this formulation. All functions involved in the l.h.s. of Eqs.(41,42) depend on $\mathfrak{z}^* = \left(\mathbf{x},t,p_{||},\mu\right)$. This is the result that was looked after: the functional lives in the physical space, so that an extremum can be properly computed, but the distribution function is the one that is known, i.e. the gyrocentre distribution function, with $\bar{\mathbf{X}}$ replaced by $\mathbf{x}$, and $(\bar{p}_{||},\bar{\mu})$ replaced by $(p_{||},\mu)$. A step further can be taken by noting that the Jacobian $B_{||}^*$ and the unperturbed distribution function $\mathcal{F}_{eq}$ vary slowly in space, so that

$$\frac{1}{B_{||}^*}\mathcal{J}^t\left[B_{||}^*\mathcal{F}'_{eq}\right] = \mathcal{F}'_{eq} + o(\epsilon_B^2)$$

In this limit the functional bears a compact form

$$\begin{aligned}\mathcal{A}_{part} &= -\int d\gamma^*\mathcal{J}^t[\bar{\mathcal{F}}_1](\mathfrak{z}^*)\mathcal{H}_1(\mathfrak{z}^*)\\ &- \frac{1}{2}\int d\gamma^*\,\mathcal{F}'_{eq}(\mathfrak{z}^*)\left(\mathcal{H}_1 - \mathcal{J}^t\left[\mathcal{J}[\mathcal{H}_1]\right]\right)(\mathfrak{z}^*)\mathcal{H}_1(\mathfrak{z}^*)\end{aligned} \tag{44}$$

$$\mathcal{A}_{skin} = -\frac{e_a^2}{2m_a}\int d\gamma^*\mathcal{J}^t[\mathcal{F}_{eq}](\mathfrak{z}^*)\tilde{A}_{||}^2(\mathbf{x},t)$$

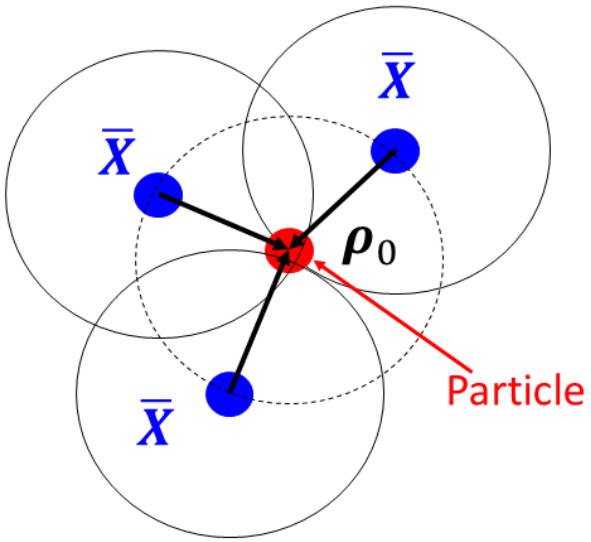
The operator $\mathcal{J}^t$ is in this limit the adjoint of the gyroaverage operator $\mathcal{J}$. Their expressions in the Fourier space are the same, reason why the combination $\mathcal{J}^t\mathcal{J}$ is often replaced in the literature by $\mathcal{J}^2$ - this confusion will be avoided[33]. After integration over the velocity variables $(p_{||},\mu)$, the functional $\mathcal{A}_{part}$ appears as an integral over $d^3\mathbf{x}dt$ of fields that depend on $(\mathbf{x},t)$. Hence an extremum can be found in the physical space and charge and current densities become available.

4.4.2 Gyrokinetic Vlasov equation in the $(\phi, A_{||})$ representation

The Vlasov gyrokinetic equation is given by Eq.(24) where the equations of motion are set by Eqs.(24,25) with

$$\mathcal{J}[\widetilde{H}] = e_a\mathcal{J}[\tilde{\phi}] - e_a\bar{p}_{||}\mathcal{J}[\tilde{A}_{||}]$$

[33]This would actually obscure the meaning of this "double average" procedure. One operator represents the gyroaverage operation of a field, i.e. of a field "seen" by a gyrocentre. The second step consists in moving back to the physical space.



$$N(\boldsymbol{x},t) = \int d\bar{\varsigma}\, d\bar{\mu} d\bar{p}_{\|} \frac{B^*_{\|}}{m} \bar{F}(\boldsymbol{x} - \boldsymbol{\rho}_0, \bar{\mu}, \bar{p}_{\|})$$

Figure 17: Construction of a particle density from a gyrocentre distribution $\bar{F}$ computed at the guiding-centre position $\mathbf{X} = \mathbf{x} - \boldsymbol{\rho}_0$. The density at one point $\mathbf{x}$ is a sum of contributions from guiding-centres $\mathbf{X}$ located on a cyclotron circle centred on the particle position.

and

$$\frac{1}{m_a}\frac{d\bar{H}}{d\bar{p}_{\|}} = \bar{p}_{\|} - \frac{e_a}{m_a}\mathcal{J}[\tilde{A}_{\|}]\left(\bar{\mathbf{X}}, \bar{p}_{\|}, \bar{\mu}, t\right)$$

The new effective field $\mathbf{B}^*$ is given by the relation

$$\mathbf{B}^*\left(\bar{\mathbf{X}}, \bar{p}_{\|}\right) = \mathbf{B}_{eq}\left(\bar{\mathbf{X}}\right) + \frac{m_a \bar{p}_{\|}}{e_a}\nabla \times \mathbf{e}_{\|}$$

and $B^*_{||}$ is the Jacobian of the gyrocentre coordinate transform

$$B^*_{||}\left(\bar{\mathbf{X}}, \bar{p}_{\|}\right) = B_{eq}\left(\bar{\mathbf{X}}\right) + \frac{m_a \bar{p}_{\|}}{e_a}\mathbf{e}_{\|} \cdot \left(\nabla \times \mathbf{e}_{\|}\right)$$

The Jacobian $B^*_{||}$ plays a central role here. The gyrokinetic derivation ensures that the flow in incompressible in the gyrocentre phase space, i.e.

$$\frac{\partial B^*_{||}}{\partial t} + \frac{\partial}{\partial \bar{\mathbf{X}}} \cdot \left(B^*_{||}\frac{d\bar{\mathbf{X}}}{dt}\right) + \frac{\partial}{\partial \bar{p}_{\|}}\left(B^*_{||}\frac{d\bar{p}_{\|}}{dt}\right) = 0$$

so that the Vlasov equation can be written as well in an advective form

$$\frac{\partial \bar{F}}{\partial t} + \frac{d\bar{\mathbf{X}}}{dt} \cdot \frac{\partial \bar{F}}{\partial \bar{\mathbf{X}}} + \frac{d\bar{p}_{\|}}{dt}\frac{\partial \bar{F}}{\partial \bar{p}_{\|}} = 0$$

Both conservative and advective forms of the Vlasov equation are useful depending on the purpose, and also the numerical scheme chosen to solve this equation. Let us stress that this is a 4D equation parametrised by the magnetic moment $\bar{\mu}$.

4.4.3 Poisson and Ampère equations in the $(\phi, A_\parallel)$ representation

These equations allow solving the gyrokinetic Vlasov equation. The gyrokinetic model becomes self-consistent by solving the Poisson equation

$$\begin{aligned} -\epsilon_0 \nabla^2 \tilde{\phi} &= \sum_{species} e_a \int d^3\mathbf{V} \mathcal{J}^t \left[\bar{\mathcal{F}}_1\right] \\ &+ \sum_{species} e_a \int d^3\mathbf{V} \mathcal{F}'_{eq} \left(\mathcal{H}_1 - \mathcal{J}^t \left[\mathcal{J}[\mathcal{H}_1]\right]\right) \end{aligned} \tag{45}$$

and the Ampère equation

$$\begin{aligned} -\frac{1}{\mu_0}\nabla_\perp^2 \tilde{A}_\parallel + \sum_{species} \frac{N_{eq} e_a^2}{m_a} \tilde{A}_\parallel &= \sum_{species} e_a \int d^3\mathbf{V} p_\parallel \mathcal{J}^t \left[\bar{\mathcal{F}}_1\right] \\ &+ \sum_{species} e_a \int d^3\mathbf{V} \mathcal{F}'_{eq} p_\parallel \left(\mathcal{H}_1 - \mathcal{J}^t \left[\mathcal{J}[\mathcal{H}_1]\right]\right) \end{aligned} \tag{46}$$

where

$$N_{eq}\left(\mathbf{x}\right) = \int d^3\mathbf{V} \mathcal{F}_{eq}(\mathfrak{z}^*)$$

is the unperturbed guiding-centre density, and $d^3\mathbf{V} = \frac{2\pi}{m_a} B^*_{||} dp_\parallel d\mu$. The electric and vector potentials $\tilde{\phi}$ and $\tilde{A}_\parallel$ live in the physical space, i.e. depend on $(\mathbf{x}, t)$, while all functions in the integrands depend on $(\mathbf{x}, t, p_\parallel, \mu)$. Some further simplifications met in the literature are the following. Let us impose an equilibrium distribution function of guiding-centres that is a Maxwellian

$$\mathcal{F}_{eq}(\mathbf{X}, p_\parallel, \mu) = N_{eq}(\mathbf{X}) \left(\frac{m_a}{2\pi T_{eq}(\mathbf{X})}\right)^{3/2} \exp\left\{-\frac{m(p_\parallel - V_{\parallel eq}(\mathbf{X}))^2}{2T_{eq}(\mathbf{X})} - \frac{\mu B_{eq}(\mathbf{X})}{T_{eq}(\mathbf{X})}\right\} \tag{47}$$

where the unperturbed field B_{eq}, density N_{eq}, parallel velocity $V_{\parallel eq}$ and temperature T_{eq} depend on $\mathbf{X}$ only. The function $\mathcal{F}'_{eq}$ then becomes

$$\frac{1}{B_{eq}} \frac{\partial \mathcal{F}_{eq}}{\partial \mu} = -\frac{\mathcal{F}_{eq}}{T_{eq}} \tag{48}$$

The second common approximation consists in developing the gyroaverage operator at second order in the gyroradius (see Appendix D). The Poisson equation becomes

$$-\epsilon_0 \nabla^2 \tilde{\phi} - \sum_{species} \nabla \cdot \left(\frac{N_{eq} m_a}{B_{eq}^2} \nabla_\perp \tilde{\phi}\right) = \sum_{species} e_a \int d^3\mathbf{V} \mathcal{J}^t \left[\bar{\mathcal{F}}_1\right] \tag{49}$$

and the Ampère equation

$$\begin{aligned} -\frac{1}{\mu_0}\nabla_\perp^2 \tilde{A}_\parallel + \sum_{species} \frac{N_{eq} e_a^2}{m_a} \tilde{A}_\parallel &+ \sum_{species} \nabla \cdot \left(\frac{N_{eq} T_{eq}}{B_{eq}^2} \nabla_\perp \tilde{A}_\parallel\right) \\ &= \sum_{species} e_a \int d^3\mathbf{V} p_\parallel \mathcal{J}^t \left[\bar{\mathcal{F}}_1\right] \end{aligned} \tag{50}$$

The ratio of permittivities $(\sum_{species} N_{eq} m / B_{eq}^2)/\epsilon_0$ is dominated by the ion contribution. It can also be reformulated as ρ_i^2/λ_D^2, where ρ_i is the ion thermal gyroradius, and λ_D the Debye length. In fusion magnetised plasmas, this number is much greater than one. Hence in most practical cases $N_{eq} m_i / B_{eq}^2$ is much larger than ϵ_0 so that the first term in

the l.h.s of Eq.(49) can be neglected. This is equivalent to consider an electrically quasi-neutral plasma. The physical meaning of the second term in Eq.(49) is a polarisation term associated with the cyclotron motion of charged particles around their gyrocentres. The mechanism has been commented above and is described in Fig.10. The r.h.s. of Eq.(49) is a density of particles constructed from the distribution of gyrocentres, as illustrated in Fig.17. It incorporates finite Larmor radius (FLR) effects. As explained in the introduction of section 4, another way to understand this term is to count particles in a box to compute their density.

The Ampère equation Eq.(50) deserves some attention too. The second term is in essence a skin depth effect - when balanced with the first term (Laplacian), exponential solutions are found, with a decay length equal to the skin depth d_e. This effect is responsible for the weak penetration of a pulsating magnetic field in a collisionless plasma (see Fig.18). However this description holds only for some range of frequencies. In particular, a low frequency field penetrates very well in a plasma (fortunately for magnetic fusion). This means that a compensating current develops in the r.h.s. of Eq.(50). This phenomenon is called "magnetic cancellation" and is source of numerical difficulties when solving this equation. The third term in the l.h.s. is a magnetisation term. Indeed, it was seen that from the stand point of a guiding-centre, the gyrating particle produces a magnetic moment $-\mu\mathbf{e}_\parallel$. The perturbed vector potential $\tilde{A}_\parallel$ produces a transverse magnetic field that bends field lines, and therefore changes slightly $-\mu\mathbf{e}_\parallel$. The magnetisation increment is $-\mu\nabla\times(\tilde{A}_\parallel\mathbf{e}_\parallel)/B_{eq}$ at the particle level, and $-\beta/\mu_0\nabla\times(\tilde{A}_\parallel\mathbf{e}_\parallel)$ after integration over the equilibrium distribution function and summation over all species, where β is the beta parameter defined in Eq.(26) - see Fig.(19). This increment is responsible for a diamagnetic modification of the Ampère equation. Its amplitude goes like the total β, which measures the diamagnetism in a magnetised plasma. Finally the gyroaverage operator that appears in the r.h.s. of Eq.(50) (gyrocentre current) contains the finite Larmor radius (FLR) effects on the current, similar to those found for the charge density.

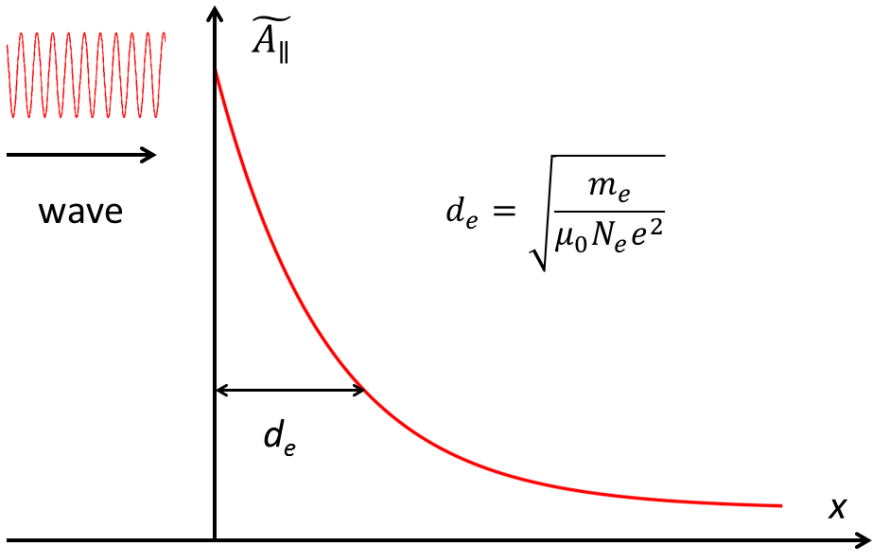


Figure 18: Illustration of the plasma skin depth effect. The amplitude $\tilde{A}_\parallel$ of a high frequency electromagnetic wave that comes from the left decays exponentially when entering the plasma at $x = 0$. The decay length is the collisionless skin depth d_e.

4.5 The $(\phi, A_\parallel, B_\parallel)$ representation*.

The $(\phi, A_\parallel)$ representation bears many advantages. However limitations are met for high values of β. The perturbed perpendicular wave vector $\tilde{\mathbf{A}}_\perp$ must then be accounted for. Since a gauge choice is always possible, only one component of $\tilde{\mathbf{A}}_\perp$ can be considered, or any scalar related to $\tilde{\mathbf{A}}_\perp$. An attractive choice, motivated by the theory of MHD instabilities, is to compute the projection $\tilde{B}_\parallel$ of the perturbed magnetic field along its unperturbed component. The calculations are lengthy and cumbersome. The analysis is restricted to the case $k_\perp \rho_c = o(1)$. This is a rather restrictive assumption since it appears that $\tilde{B}_\parallel$ impacts strongly large scale MHD modes [16], but it is the price to get something practical. Hence this section addresses the gyrokinetic equations that govern turbulence and small scale MHD modes. In this limit $\mathcal{F}_{eq}$ and its derivative $\mathcal{F}'_{eq}$ are large scale functions, so that the gyroaverage operator applied on these functions is just the identity. We will also see that the perturbed Hamiltonian is explicit in $\tilde{B}_\parallel$ in this case.

4.5.1 Perturbed Hamiltonian and functional in the $(\phi, A_\parallel, B_\parallel)$ representation

The perpendicular component of the vector potential is now added. The gyroaverage Hamiltonian reads (see Appendix E)[34]

$$\mathcal{J}[\widetilde{\mathcal{H}}] = e_a \mathcal{J}[\tilde{\phi}] - e_a p_\parallel \mathcal{J}[\tilde{A}_\parallel] + \mu \mathcal{B}[\mathbf{e}_\parallel \cdot \nabla \times \tilde{\mathbf{A}}_\perp] - e_a \dot{\mathbf{X}}_\perp \cdot \mathcal{J}[\tilde{\mathbf{A}}_\perp] + \frac{e_a^2}{2m_a} \mathcal{J}[\tilde{\mathbf{A}}^2]$$

The two last terms are of second order in $\epsilon_B \sim \epsilon_\delta$[35]. Therefore the perturbed Hamiltonian is at first order

$$\mathcal{J}[\mathcal{H}_1] = e_a \mathcal{J}[\tilde{\phi}] - e_a p_\parallel \mathcal{J}[\tilde{A}_\parallel] + \mu \mathcal{B}[\mathbf{e}_\parallel \cdot \nabla \times \tilde{\mathbf{A}}_\perp] \tag{51}$$

The last term of Eq.(51) calls for some comments. The operator $\mathcal{B}$ is defined in the Fourier space by multiplying the Fourier component $\boldsymbol{k}$ of the perpendicular potential vector $\tilde{\mathbf{A}}_\perp$ by $2J_1(k_\perp \rho_c)/k_\perp \rho_c$, where J_1 is the Bessel function of index 1, $k_\perp$ the modulus of the perpendicular wave number, and $\rho_c = \sqrt{2\mu/e_a \Omega_c}$ the gyroradius (see Appendix E). The operator $\mathcal{J}$ consists in a similar operation, where the field Fourier component is multiplied by $J_0(k_\perp \rho_c)$. Both operators $\mathcal{J}$ and $\mathcal{B}$ come near identity when $k_\perp \rho_c \to 0$. The scalar $\mathbf{e}_\parallel \cdot \nabla \times \tilde{\mathbf{A}}_\perp$ is close to the parallel component of the perturbed magnetic field $\tilde{B}_\parallel$ for small scale fluctuations, more precisely

$$\tilde{B}_\parallel = \mathbf{e}_\parallel \cdot \tilde{\mathbf{B}} = \mathbf{e}_\parallel \cdot \nabla \times \tilde{\mathbf{A}}_\perp + \tilde{A}_\parallel \frac{\mu_0 J_{\parallel eq}}{B_{eq}} \tag{52}$$

The quantity $\mu_0 J_{\parallel eq}/B_{eq}$ is of the order of the inverse of the plasma size, while gradients of perturbed quantities scale as an inverse of fluctuation scale length. For small scale perturbations (i.e. turbulent, but not global MHD modes) the second term in the r.h.s. of Eq.(52) can be neglected against the first term, so that $\tilde{B}_\parallel = \mathbf{e}_\parallel \cdot \nabla \times \tilde{\mathbf{A}}_\perp + o(B_{eq}\epsilon_\delta^2)$. The field $\tilde{B}_\parallel$ also coincides with the 1st order perturbation of the magnetic field[36]. Hence the physical meaning of the last term in the perturbed Hamiltonian Eq.(51) is clear : it corresponds to a contribution $\mu \tilde{B}_\parallel$ to the Hamiltonian that modifies its unperturbed counterpart μB_{eq}. As such, it can be seen as a perturbation of the magnetisation field $\mathbf{M}$ (see Appendix F). This approximation will not be used for now, to keep some generality.

[34]The reader should be warned that $\mathcal{J}[\widetilde{\mathcal{H}}]$ is not the gyrocentre Hamiltonian $\bar{\mathcal{H}}$ - this is only true at first order in ϵ_δ.

[35]The perpendicular velocity $\dot{\tilde{\mathbf{X}}}_\perp$ is of order ϵ_B compared with the parallel velocity. Assuming $\epsilon_B \sim \epsilon_\delta$, it appears that the piece $\dot{\tilde{\mathbf{X}}}_\perp \cdot \mathcal{J}[\tilde{\mathbf{A}}_\perp]$ is second order compared with the 3 first terms and can be neglected in a first order calculation.

[36]Indeed $B^2 = (\mathbf{B}_{eq} + \tilde{\mathbf{B}})^2 = B_{eq}(1 + 2\tilde{B}_\parallel/B_{eq} + o(\epsilon_\delta^2))$ and therefore $B = B_{eq} + \tilde{B}_\parallel + o(\epsilon_\delta^2 B_{eq})$

It is further discussed in the section F.4 and will be used to derive the Maxwell equations in section 4.5.3. The total functional reads

$$\mathcal{A} = \mathcal{A}_{em} + \mathcal{A}_{part} + \mathcal{A}_{skin}$$

where

$$\mathcal{A}_{em} = \frac{1}{2}\int d^3\mathbf{x}\, dt\, \left(\epsilon_0 \tilde{\mathbf{E}}^2(\mathbf{x},t) - \frac{1}{\mu_0}\tilde{\mathbf{B}}^2(\mathbf{x},t)\right) \tag{53}$$

$$\begin{aligned}\mathcal{A}_{part} &= -\sum_{species}\int d\Gamma^* \bar{\mathcal{F}}_1(\boldsymbol{\mathcal{Z}}^*)\mathcal{J}[\mathcal{H}_1](\boldsymbol{\mathcal{Z}}^*) \\ &- \frac{1}{2}\sum_{species}\int d\Gamma^* \mathcal{F}'_{eq}(\boldsymbol{\mathcal{Z}}^*)\left\{\mathcal{J}[\mathcal{H}_1^2] - (\mathcal{J}[\mathcal{H}_1])^2\right\}(\boldsymbol{\mathcal{Z}}^*)\end{aligned}$$

$$\mathcal{A}_{skin} = -\sum_{species}\frac{e_a^2}{2m_a}\int d\Gamma^*\, \mathcal{F}_{eq}\,(\boldsymbol{\mathcal{Z}}^*)\,\mathcal{J}[\tilde{\mathbf{A}}^2](\boldsymbol{\mathcal{Z}}^*) \tag{54}$$

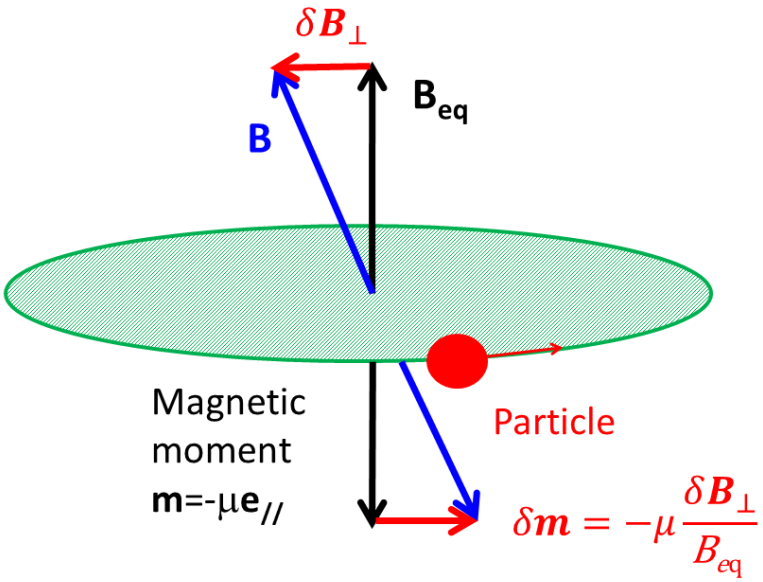


Figure 19: Diamagnetic effect associated with a perpendicular perturbation of the magnetic field. A perturbed parallel potential vector $\tilde{A}_\parallel$ produces a perturbed field $\delta\mathbf{B}_\perp = \nabla\times(\tilde{A}_\parallel\mathbf{e}_\parallel) \simeq -\mathbf{e}_\parallel\times\nabla\tilde{A}_\parallel$ that is essentially perpendicular to the equilibrium field $\mathbf{B}_{eq}$. Field line bending is responsible for a change of direction of the magnetic moment associated with the cyclotron motion of a charged particle, equal to $-\bar{\mu}\delta\mathbf{B}_\perp/B_{eq}$. The corresponding change of magnetisation density is $-\int d^3\bar{\mathbf{v}}\bar{\mathcal{F}}_{eq}\bar{\mu}\delta\mathbf{B}_\perp/B_{eq}$. For a Maxwellian distribution function, the perturbed magnetisation density is $\delta\mathbf{M}_\perp = -(N_{eq}T_{eq}/B_{eq}^2)\delta\mathbf{B}_\perp$, which opposes to $\delta\mathbf{B}_\perp$. The ratio $\mu_0\delta M_\perp/\delta B_\perp$ is of the order of the beta parameter given by Eq.(26). This effect can be seen as a diamagnetic effect on the perturbed perpendicular magnetic field.

4.5.2 Skin depth term and magnetisation

The derivation is the same as for the representation $(\phi, A_\parallel)$ (see details in Appendix F). The unperturbed guiding-centre distribution function is again supposed Maxwellian

Eq.(47). The function $\mathcal{F}'_{eq}$ is then given by Eq.(48). The second term of the functional $\mathcal{A}$, physically related to the adiabatic response of particles, and defined for each species as

$$\mathcal{A}_{ad} = -\int d\Gamma^* \mathcal{F}'_{eq}(\boldsymbol{\mathcal{Z}}^*)\mathcal{J}[\mathcal{H}_1^2](\boldsymbol{\mathcal{Z}}^*)$$

raises some difficulty due to the presence of the perpendicular vector potential. A detailed calculation can be found in Appendix F where the following identity is demonstrated

$$\begin{aligned}\mathcal{A}_{ad} &= \frac{e_a^2}{2}\int d\gamma^* \frac{F_{eq}(\mathbf{x},p_\parallel,\mu)}{T_{eq}(\mathbf{x})}\left(\tilde{\phi}^2(\mathbf{x},t) - 2p_\parallel\tilde{\phi}(\mathbf{x},t)\tilde{A}_\parallel(\mathbf{x},t)\right)\\ &+ \frac{e_a^2}{2m_a}\int d^3\mathbf{x}dt\, N_{eq}(\mathbf{x})\tilde{\mathbf{A}}^2(\mathbf{x},t)\end{aligned}$$

in the limit where the Jacobian $B^*_{||}$ and the equilibrium distribution function $\mathcal{F}_{eq}$ exhibit large scales compared with the gyroradius. The last term of this functional compensates the skin depth term in the functional of the electromagnetic field $\mathcal{A}_{field}$. The total functional reads

$$\begin{aligned}\mathcal{A} &= \frac{\epsilon_0}{2}\int d^3\mathbf{x}dt\,\tilde{\mathbf{E}}^2(\mathbf{x},t) - \frac{1}{2\mu_0}\int d^3\mathbf{x}dt\,\tilde{\mathbf{B}}^2(\mathbf{x},t)\\ &+ \sum_{species}\frac{e_a^2}{2}\int d\gamma^*\frac{F_{eq}(\mathbf{x},p_\parallel,\mu)}{T_{eq}(\mathbf{x})}\left(\tilde{\phi}^2(\mathbf{x},t) - 2p_\parallel\tilde{A}_\parallel(\mathbf{x},t)\tilde{\phi}(\mathbf{x},t)\right)\\ &- \frac{1}{2}\sum_{species}\int d\Gamma^*\frac{\mathcal{F}_{eq}}{T_{eq}}(\mathcal{J}[\mathcal{H}_1])^2(\boldsymbol{\mathcal{Z}}^*) - \sum_{species}\int d\Gamma^*\bar{\mathcal{F}}_1(\boldsymbol{\mathcal{Z}}^*)\mathcal{J}[\mathcal{H}_1](\boldsymbol{\mathcal{Z}}^*)\end{aligned}$$

The last line represents the interaction between particles and the electromagnetic field.

4.5.3 Maxwell equations in the $(\phi, A_\parallel, B_\parallel)$ representation

The calculation of the charge and current densities in $(\phi, A_\parallel, B_\parallel)$ representation is somewhat technical. It is detailed in the Appendix F. Expressions of the charge and current densities become much simpler under some reasonable assumptions. In most calculation of instabilities, the current density current carried by ions is small compared with the electron current. Conversely, finite Larmor radius effects, which are important for ions, can usually be neglected for electrons since their gyroradius is often much smaller than wavelengths of interest. This suppress "cross-terms", i.e. terms involving the vector potential in the charge density, or involving the electric potential in the current density. The Poisson and Ampère equations become

$$\begin{aligned}\sum_{species}\int d^3\mathbf{V}\frac{F_{eq}e_a^2}{T_{eq}}\left(1-\mathcal{J}^t\mathcal{J}\right)[\tilde{\phi}] &= \sum_{species}\int d^3\mathbf{V}e_a\mathcal{J}^t[\bar{\mathcal{F}}_1]\\ &+ \sum_{species}\int d^3\mathbf{V}\frac{F_{eq}e_a}{T_{eq}}\mu\mathcal{J}^t\mathcal{B}[\tilde{B}_\parallel] \qquad (55)\end{aligned}$$

$$-\frac{1}{\mu_0}\nabla_\perp^2\tilde{A}_\parallel + \sum_{species}\int d^3\mathbf{V}\frac{F_{eq}e_a^2}{m_a}\mathcal{J}^t\mathcal{J}[\tilde{A}_\parallel] = \sum_{species}e_a\int d^3\mathbf{V}p_\parallel\mathcal{J}^t[\bar{\mathcal{F}}_1] \qquad (56)$$

Up to a cross-term in the Poisson equation that involves $\tilde{B}_\parallel$, the equations over $(\tilde{\phi}, \tilde{A}_\parallel)$ are the same as without $\tilde{B}_\parallel$. Since $\tilde{B}_\parallel$ is itself known as a function of the perturbed gyrocentre distribution function and the perturbed potential $\tilde{\phi}$ (see below), this means that Eqs.(55,56) are two decoupled equations that rule separately $\tilde{\phi}$ and $\tilde{A}_\parallel$.

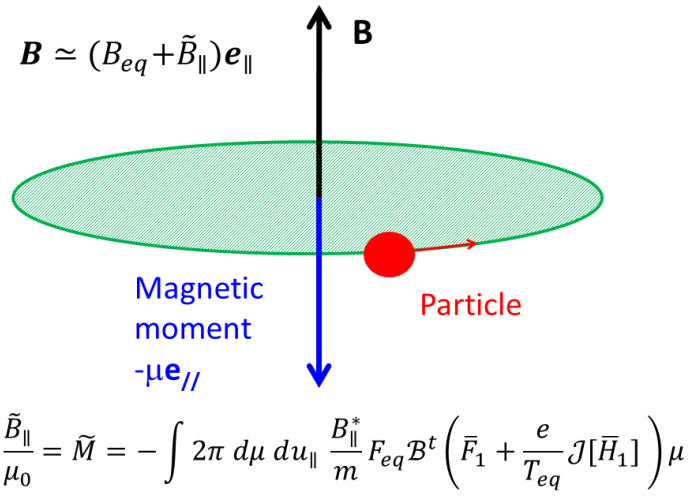
The compressional component $\tilde{B}_\parallel$ is given by the equation

$$\frac{1}{\mu_0}\tilde{B}_\parallel + \sum_{species}\int d^3\mathbf{V}\frac{F_{eq}e}{T_{eq}}\mu^2\mathcal{B}^t\mathcal{B}[\tilde{B}_\parallel] = -\sum_{species}\int d^3\mathbf{V}\mu\mathcal{B}^t[\bar{\mathcal{F}}_1 + \frac{e_a}{T_{eq}}\mathcal{J}[\tilde{\phi}]] \tag{57}$$

The equations Eqs.(55,56,57) are close to those given in [17]. In fact the choice of Maxwellian distribution functions allows an explicit integration of the gyroaverage operators over the velocity space - details can be found in the Appendix F.6. The equation Eq.(57) is reminiscent of the MHD constraint $B_{eq}\tilde{B}_\parallel + \mu_0\tilde{P}_\perp = 0$, which can be derived using fluid equations, thus a different path, but similar assumptions. A simple physical explanation can be offered for this equation. When submitted to a perturbed perpendicular vector potential, gyrocentres produce a perturbed magnetisation field

$$\tilde{\mathbf{M}} = -\int d^3\mathbf{V}\mu\mathcal{B}^t[\bar{\mathcal{F}}_1 + \frac{F_{eq}e_a}{T_{eq}}\mathcal{J}[\tilde{\phi}]]\mathbf{e}_\parallel$$

Eq.(57) is equivalent to state that $\tilde{B}_\parallel = \mu_0\tilde{M}$ (see figure 20). This formulation is reasonably accurate for low β plasmas and short wavelengths, but should not be used for some classes of long wavelength instabilities. It is autonomous in the sense that the knowledge of the perturbed electric potential $\tilde{\phi}$ and the gyrocentre distribution function yields the value of $\tilde{B}_\parallel$. Another advantage is that no Laplacian operator needs be inverted.



$$\boldsymbol{B} \simeq (B_{eq}+\tilde{B}_\parallel)\boldsymbol{e}_\parallel$$

$$\frac{\tilde{B}_\parallel}{\mu_0} = \tilde{M} = -\int 2\pi\, d\mu\, du_\parallel \frac{B^*_\parallel}{m} F_{eq}\mathcal{B}^t\left(\bar{F}_1 + \frac{e}{T_{eq}}\mathcal{J}[\bar{H}_1]\right)\mu$$

Figure 20: When submitted to a perturbed perpendicular vector potential, gyrocentres produce a perturbed magnetisation field $\tilde{\mathbf{M}} = \tilde{M}\mathbf{e}_\parallel$. In first approximation the compressional component of the magnetic field $\tilde{B}_\parallel$ is proportional to the perturbed magnetisation $\tilde{B}_\parallel = \mu_0\tilde{M}$, which yields an explicit expression of $\tilde{B}_\parallel$ versus a moment of the gyrocentre distribution function.

5 Gyrokinetic theory - a general formulation*

The theory above suffers from some drawbacks. It is restricted to canonical changes of variables of the form Eq.(12), i.e. to the so-called Hamiltonian version of the gyrokinetic theory. Moreover the gyrocentre Hamiltonian was calculated at first order only, which may appear somewhat artificial whenever a full-F model is to be built. In the previous approach, a full-F model can be reconstructed by defining the full F distribution function as the sum of the unperturbed distribution function, allowed to vary slowly, and the first order perturbed distribution function, that evolves on fast time scales. Moreover it

appears that polarisation effects requires a second order theory when fully expressed in gyrocentre coordinates. In fact it is possible to build an action principle which does not need a reconstruction of the full distribution function. However it requires handling an Hamiltonian at second order in ϵ_δ. Hence a theory is needed that is valid at second order in the ordering parameter, and handles non canonically conjugate variables [2, 3, 4]. This is a rather formidable task, which necessitates some know-how in differential geometry. This part is not indispensable in a first reading.

5.1 Canonical change of variables at second order

As mentioned before, one has to move at some point to the hybrid set of variables $(\mathbf{x}, p_\parallel, \mu)$. The expression Eq.(11) shows that the Hamiltonian is second order in the parameter ϵ_δ. Hence the gyrocentre coordinates must be built at second order in the order parameter. It is stressed here that solving the problem at second order in ϵ_δ is challenging when allowing all possible choices of coordinates. So it is wise to focus first the analysis on canonical changes of variables of the form Eq.(12) ("Hamiltonian approach"), but at all orders. Then we will see that the calculation can be generalised to arbitrary changes of variables. Near identity changes of canonical coordinates at all orders were studied extensively in the context of celestial mechanics. These are generalisations at all order of Eq.(12), hence of the form

$$\boldsymbol{\mathcal{Z}} = \bar{\boldsymbol{\mathcal{Z}}} - \left\{S, \bar{\boldsymbol{\mathcal{Z}}}\right\}_{\bar{\boldsymbol{\mathcal{Z}}}} \tag{58}$$

where S is a generating function. Both S and the Hamiltonian are expanded in Fourier series

$$S(\bar{\mathcal{Z}}) = \sum_{n=0}^{+\infty} S_n(\bar{\mathcal{Z}})\epsilon_\delta^n$$

$$\begin{aligned}\mathcal{H}(\boldsymbol{\mathcal{Z}}) &= \sum_{n=0}^{+\infty} \mathcal{H}_n(\boldsymbol{\mathcal{Z}})\epsilon_\delta^n \\ \bar{\mathcal{H}}(\bar{\boldsymbol{\mathcal{Z}}}) &= \sum_{n=0}^{+\infty} \bar{\mathcal{H}}(\bar{\boldsymbol{\mathcal{Z}}})\epsilon_\delta^n\end{aligned}$$

A powerful approach to calculate the generating function is based on the Deprit theory [9, 18], detailed in Appendix G. Some of these works were devised for conjugate variables, but can be extended to non conjugate variables by using the change of variables Eq.(58). So "canonical change of variables" must be understood in this spirit. The second order Hamiltonian is found to be of the form

$$\bar{\mathcal{H}}_2 = \mathcal{J}[\mathcal{H}_2] - \frac{1}{2}\mathcal{J}\left[\left\{S_1, \mathcal{H}_1 + \bar{\mathcal{H}}_1\right\}\right]$$

where the generating function S_1 is given by the equation Eq.(17). A refined calculation (see Appendix G) provides the following expression

$$\begin{aligned}\bar{\mathcal{H}}_2 &= \langle\mathcal{H}_2\rangle - \frac{1}{2B_{eq}}\frac{\partial}{\partial\bar{\mu}}\left\langle\left(\mathcal{H}_1 - \mathcal{J}[\mathcal{H}_1]\right)^2\right\rangle \\ &+ \frac{1}{2\Omega_c}\left\langle\frac{\mathbf{e}_\parallel}{e_a B_\parallel^*}\cdot\left(\frac{\partial}{\partial\bar{\mathbf{X}}}\left(\frac{m_a}{e_a B_{eq}(\bar{\mathbf{X}})}\int^\varsigma d\varsigma'\widetilde{\mathcal{H}}_1\right)\times\frac{\partial\widetilde{\mathcal{H}}_1}{\partial\bar{\mathbf{X}}}\right)\right\rangle\end{aligned} \tag{59}$$

The last term of Eq.(59) is sometimes dubbed "ponderomotive"[37]. It is quite small in practical cases (see Appendix G). It can actually be neglected rigorously if the sub-ordering $\epsilon_\omega \sim \epsilon_\delta^2$ holds [3, 4]. If it is ignored, the second order Hamiltonian reduces

[37] A ponderomotive force is the average force exerted by an oscillating field on a particle. It can be understood with a simple example. Consider a particle submitted to a 1D electric field perturbation E (no B field), its position $x(t)$ is solution of the equation $m_a\ddot{x} = eE(x)\cos(\omega t)$. The position x can be decomposed in an average

to

$$\bar{\mathcal{H}}_2 = \mathcal{J}[\mathcal{H}_2] - \frac{1}{2B_{eq}}\frac{\partial}{\partial\bar{\mu}}\left(\mathcal{J}[\mathcal{H}_1^2] - (\mathcal{J}[\mathcal{H}_1])^2\right) \tag{60}$$

Introducing this gyrocentre Hamiltonian in the field-particle action Eq.(33) leads to the expression

$$\mathcal{A}_{fp} = -\sum_{species}\int d\bar{\Gamma}\left(\bar{\mathcal{F}}_0(\bar{\mathcal{Z}})\bar{\mathcal{H}}_2(\bar{\mathcal{Z}}) + \bar{\mathcal{F}}_1(\bar{\mathcal{Z}})\bar{\mathcal{H}}_1(\bar{\mathcal{Z}}) + \bar{\mathcal{F}}_2(\bar{\mathcal{Z}})\bar{\mathcal{H}}_0(\bar{\mathcal{Z}})\right)$$

The term that involves $\bar{\mathcal{F}}_2$ is multiplied by $\bar{\mathcal{H}}_0$, which is a function of the invariants of motion. Hence it addresses only the mean part of the distribution function, that is its average over the unperturbed periodic motion of particles. This piece evolves on a transport time scale, i.e. terms of order ϵ_δ^3, and is part of the Vlasov equation. It does not participate in the determination of the Maxwell equations, except for the mean flow. It is left apart from now on. An integration by part yields

$$\mathcal{A}_{fp} = -\sum_{species}\int d\bar{\Gamma}\left(\bar{\mathcal{F}}_0\mathcal{J}[\mathcal{H}_2] + \frac{1}{2B_{eq}}\frac{\partial}{\partial\bar{\mu}}\bar{\mathcal{F}}_0\left(\mathcal{J}[\mathcal{H}_1^2] - (\mathcal{J}[\mathcal{H}_1])^2\right) + \bar{\mathcal{F}}_1\mathcal{J}[\mathcal{H}_1]\right) \tag{61}$$

This functional is the same as Eq.(39) obtained previously in the first order theory, keeping in mind that the term $\bar{\mathcal{F}}_0\mathcal{J}[\mathcal{H}_2]$ is the skin depth term that was included in the field action. In other words the Maxwell equations are the same in this approach. Also there is no test field here - the functional should be made extremum with all variations of the electric and magnetic potentials - hence the extra factor $\frac{1}{2}$ in Eq.(61) that was not apparent in Eq.(39).

It is quite remarkable that the polarisation term, proportional to $(\mathcal{J}[\mathcal{H}_1^2] - (\mathcal{J}[\mathcal{H}_1])^2)$, comes here from the second order Hamiltonian. One may wonder why. The answer lies in the expression of the mean displacement of the gyrocentre with respect to the particle position Eq.(19). This displacement produces contributions which are second order in the expansion parameter ϵ_δ, and hence belong to the category of ponderomotive forces. A feeling of this important point can be grasped by noting that in the limit of long wavelengths, the Hamiltonian $\bar{\mathcal{H}}_1(\bar{\mathbf{X}} + \boldsymbol{\rho}_0 + \bar{\boldsymbol{\rho}}_1)$ "seen" by a guiding centre can be Taylor developed to give an average difference with $\mathcal{H}_1(\bar{\mathbf{X}} + \boldsymbol{\rho}_0)$ equal to

$$\langle\bar{\boldsymbol{\rho}}_1\rangle_{\bar{\varsigma}}\cdot\frac{\partial\mathcal{H}_1(\bar{\mathbf{X}})}{\partial\bar{\mathbf{X}}_\perp} = -\frac{m_a}{e_a^2B_{eq}^2(\bar{\mathbf{X}})}\left[\frac{\partial\mathcal{H}_1}{\partial\bar{\mathbf{X}}_\perp}\right]^2 \tag{62}$$

This is also what is found when developing the second term in the r.h.s. of Eq.(60). In first order gyrokinetic theory, it is incorporated in the difference between the guiding-centre and gyrocentre distribution functions when moving back to the guiding-centre phase space. The polarisation effects are therefore contained in the difference between $F(\boldsymbol{\mathcal{Z}})$ and $\bar{\mathcal{F}}(\boldsymbol{\mathcal{Z}})$. Differences between the two approaches are illustrated by Fig.13 and Fig.21.

The last step consists in moving back to the physical space. This is done via a change of variables from the gyrocentre position variables $\bar{\mathbf{X}}$ to the particle position $\mathbf{x} = \bar{\mathbf{X}} + \bar{\boldsymbol{\rho}}$. An apparent difficulty is the occurrence of $\bar{\boldsymbol{\rho}}$ which is in principle not just a simple cyclotron motion. In particular it contains the polarisation physics via $\bar{\boldsymbol{\rho}}_1$. However the polarisation shift $\bar{\boldsymbol{\rho}}_1$ introduces higher order corrections. This is consistent with our previous remark: in this approach, polarisation effects are included in the second order gyrocentre Hamiltonian and do not need further treatment - otherwise that would mean double counting.

value $\bar{x}$ and and a variation $\tilde{x}(t)$. The field can be Taylor development around $\bar{x}$, $E(x) = E(\bar{x}) + \tilde{x}\,\partial_x E|_{x=\bar{x}}$. Obviously $\tilde{x}(t) = -\frac{eE(\bar{x})}{m_a\omega^2}\cos(\omega t)$, so that the average force is $-\frac{e_a^2}{2m_a\omega^2}E(\bar{x})\,\partial_x E|_{x=\bar{x}}$. The effect of a strong magnetic guide field can be mimicked by imposing $\omega = e_aB/m_a$, thus giving a force $-\frac{1}{2}\frac{m_a}{B^2}E(\bar{x})\,\partial_x E|_{x=\bar{x}}$, similar to Eq.(62), since it corresponds to an effective potential $-\frac{1}{4}\frac{m_a}{B^2}\left[\frac{d\phi}{d\bar{x}}\right]^2$.

This means that in fact the change of variables is just $\mathbf{x} = \bar{\mathbf{X}} + \boldsymbol{\rho}_0$, i.e. similar to the one for a guiding-centre. Performing this change of variables provide the same action as the one found previously by pulling back the distribution function from the gyrocentre to the guiding-centre position space.

Hence the expected result is recovered by using a second order Hamiltonian theory. However this does not tell us how this should be managed when a non-canonical change of variables is needed at second order in ϵ_δ. Hence the Deprit performance must repeated in the non-canonical case. This is done in Appendix H.

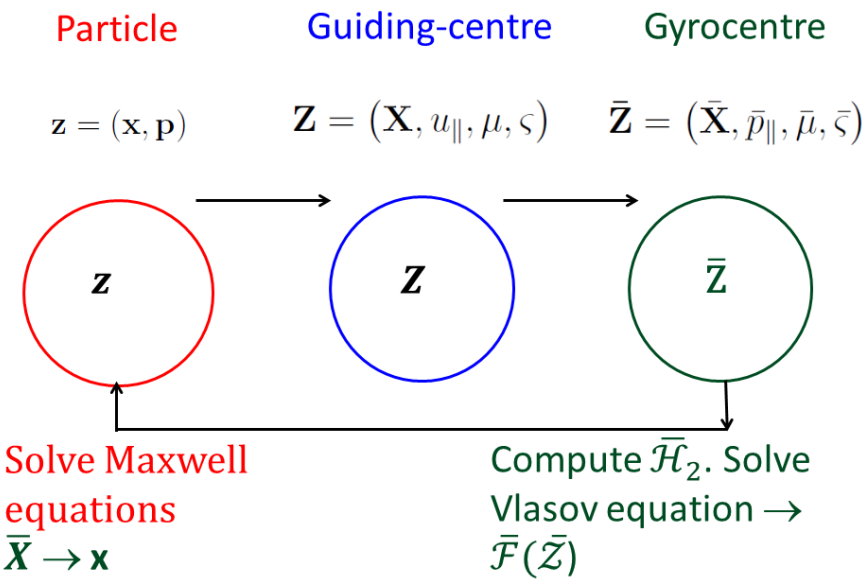


Figure 21: Strategy of a gyrokinetic theory using a second order gyrocentre Hamiltonian. The Vlasov equation is solved in gyrocentre coordinates with no separation of perturbed and unperturbed distribution functions. This procedure requires a second order Hamiltonian. A pull-back from the gyrocentre space to the physical space is then done to solve the Maxwell equations. Polarisation effects are included in the second order Hamiltonian.

5.2 Non canonical change of variables

Calculations in the non canonical case are treacherous (see Appendix H). The result is shown here for the $(\phi, A_\parallel)$ representation. The parallel velocity is then given by the expression

$$\bar{p}_\parallel = u_\parallel + \frac{e_a}{m_a}\left(\tilde{A}_\parallel\left(\mathbf{X} + \boldsymbol{\rho}_0(\mathbf{X}, \mu, \varsigma)\right) - \epsilon_s \mathcal{J}[\tilde{A}_\parallel](\mathbf{X}, \mu)\right)$$

where $\epsilon_s = 0$ in the Hamiltonian case and $\epsilon_s = 1$ in the symplectic case. For the Hamiltonian case $\epsilon_s = 0$, the velocity $\bar{p}_\parallel$ coincides with $p_\parallel$. The situation is quite different in the symplectic case since the symplectic parallel "velocity" $\bar{p}_\parallel$ comes near the guiding-centre velocity $u_\parallel$ for long wavelength fields $\mathcal{J}[\tilde{A}_\parallel] \sim \tilde{A}_\parallel$. The equations of motion of the

	Hamiltonian	Symplectic
Skin depth term	yes	no
Induction in parallel force	no	yes
$\bar{p}_\parallel = u_\parallel$ at long wavelengths	no	yes
Jacobian depends on perturbed field	no	yes

Table 1: Some properties of the "Hamiltonian" and "Symplectic" gyrokinetic models

gyrocentre are[38]

$$\begin{aligned} \frac{d\bar{\mathbf{X}}}{dt} &= \bar{p}_\parallel \mathbf{b}^* + \frac{\mathbf{e}_\parallel}{e_a B^*_\parallel} \times \nabla \left(\bar{\mu} B_{eq} + e_a \mathcal{J}[\tilde{\phi}] \right) \\ m_a \frac{d\bar{p}_\parallel}{dt} &= -\mathbf{b}^* \cdot \left[\nabla \left(\bar{\mu} B_{eq} + e_a \mathcal{J}[\tilde{\phi}] \right) + e_a \frac{\partial \mathcal{J}[\tilde{A}_\parallel]}{\partial t} \mathbf{e}_\parallel \right] \end{aligned}$$

where $\mathbf{b}^* = \mathbf{B}^*/B^*_\parallel$. The vector $\mathbf{B}^*$ is given by the relation

$$\mathbf{B}^* = \mathbf{B}_{eq} + \frac{m_a \bar{p}_\parallel}{e_a} \nabla \times \mathbf{e}_\parallel + \nabla \times \left(\mathcal{J}[\tilde{A}_\parallel] \mathbf{e}_\parallel \right)$$

and

$$B^*_\parallel = \mathbf{e}_\parallel \cdot \mathbf{B}^* = B_{eq} + \frac{m_a \bar{p}_\parallel}{e_a} \mathbf{e}_\parallel \cdot \nabla \times \mathbf{e}_\parallel + \mathbf{e}_\parallel \cdot \nabla \times \mathcal{J}[\tilde{A}_\parallel] \mathbf{e}_\parallel$$

6 Energy conservation*

Since an Hamiltonian dynamics underlies any gyrokinetic theory, an energy conservation theorem is expected to hold. There are various ways to construct an energy balance equation. The most elegant, but technical, approach is based on the Noether's theorem combined with a generalised action principle [11, 12, 13, 19]. It can be used to build both energy and momentum conservation equations. Its use is restricted here to energy conservation, which appears to be a consequence of the field Lagrangian invariance under time translation. Energy conservation is useful to identify populations responsible for kinetic instabilities by computing the energy transfer between particles and field. It is also useful for code verification.

6.1 Preamble: field action and Noether's theorem

Before deriving a gyrokinetic equation of energy conservation, it is probably useful to start with an example based on the simplified action used in section 4.1. Let us note first that the Lagrangian in Eq.(26) does not to depend explicitly on time - only via the field and its derivatives. As a consequence, the field ϕ solution of the Euler-Lagrange equations must remain the same if time is translated by some small amount δt. Such a dynamical system is said to be invariant under time translation symmetry. The idea behind the Noether's theorem is to make use of this type of symmetry invariance. It can be formalised as follows in the specific case of time translation. Introducing a new time variable $\bar{t} = t + \delta t$ and a field $\bar{\phi}$ such that $\bar{\phi}(\bar{t}) = \phi(t)$, the invariance under time translation symmetry implies that the action

$$\bar{\mathcal{A}} = \int_{\bar{t}_1}^{\bar{t}_2} d\bar{t} \int_{\bar{\mathbf{x}}_1}^{\bar{\mathbf{x}}_2} d^3\bar{\mathbf{x}}\, \mathcal{L}\left(\bar{\phi}, \partial_{\bar{\mu}} \bar{\phi}\right)$$

[38] Since the vector potential is directed along the unperturbed magnetic field, the corresponding inductive electric field does not contribute to the $E \times B$ drift

must be equal to the action $\mathcal{A}$. Hence the variation $\delta\mathcal{A} = \bar{\mathcal{A}} - \mathcal{A}$ must vanish. It can be calculated explicitly for a small time variation δt. Since $\bar{\phi}(\bar{t}) = \phi(\bar{t} - \delta t)$, it appears that $\delta\phi = \bar{\phi}(\bar{t}) - \phi(\bar{t}) = -\partial_{\bar{t}}\phi(\bar{t})\delta t$. This relation allows the computation $\bar{\mathcal{A}}$ as a functional of ϕ. Note in particular that once this operation is done, $(\bar{\mathbf{x}}, \bar{t})$ are dummy variables in the integral $\bar{\mathcal{A}}$ and can be replaced by $(\mathbf{x}, t)$. One has to be careful that the integral time bounds are changed since $\bar{t}_1 = t_1 + \delta t$ and $\bar{t}_2 = t_2 + \delta t$. Hence the variation $\delta\mathcal{A}$ contains two contributions: one due to the changes in integration bounds, noted $\delta\mathcal{A}_{bound}$, and the other one due to the field variation $\delta\phi = -\partial_{\bar{t}}\phi(\bar{t})\delta t$ in the "bulk" of the integral. This variation is noted $\delta\mathcal{A}_{bulk}$. Obviously

$$\delta\mathcal{A}_{bound} = \delta t \int_{\mathbf{x}_1}^{\mathbf{x}_2} d^3\mathbf{x}\ (\mathcal{L}_2 - \mathcal{L}_1) = \int_{t_1}^{t_2} dt \int_{\mathbf{x}_1}^{\mathbf{x}_2} d^3\mathbf{x}\ \partial_t\mathcal{L}\left(\phi, \partial_\mu\phi\right)$$

where $\mathcal{L}_1$ (resp. $\mathcal{L}_2$) is the Lagrangian calculated with the field ϕ solution of the Euler-Lagrange equations and computed at time $t = t_1$ (resp. $t = t_2$). The variation $\delta\mathcal{A}_{bulk}$ due to the field variation $\delta\phi$ is calculated by using the action variation Eq.(27). The first part vanishes since ϕ is solution of the Euler-Lagrange equations. Hence only the current divergence term contributes, so that

$$\delta\mathcal{A}_{bulk} = -\delta t \int_{t_1}^{t_2} dt \int_{\mathbf{x}_1}^{\mathbf{x}_2} d^3\mathbf{x}\ \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_t\phi\right)$$

The sum of these two contributions $\delta\mathcal{A}_{bound} + \delta\mathcal{A}_{bulk}$ must vanish *whatever the choice of the integral lower and upper bounds* $(t_1, t_2, \mathbf{x}_1, \mathbf{x}_2)$. This constraint provides a local conservation law

$$\partial_t\mathcal{E} + \partial_\alpha\Gamma^\alpha_{\mathcal{E}} = 0$$

where the index α runs from 1 to 3,

$$\mathcal{E} = \mathcal{L} - \frac{\partial\mathcal{L}}{\partial(\partial_t\phi)}\partial_t\phi$$

is the energy density, and

$$\Gamma^\alpha_{\mathcal{E}} = \frac{\partial\mathcal{L}}{\partial(\partial_\alpha\phi)}\partial_t\phi$$

is an energy flux. An example is shown on Fig.22. When applied to the action in absence of charge and current densities, it can be verified that the energy density is $\mathcal{E}_{em} = \frac{\epsilon_0}{2}\mathbf{E}^2 + \frac{1}{2\mu_0}\mathbf{B}^2$, and the energy flux is the Poynting flux $\frac{\mathbf{E}\times\mathbf{B}}{\mu_0}$. This recipe should now be applied to the gyrokinetic field action.

6.2 Noether's theorem in gyrokinetics

6.2.1 Variation of the gyrokinetic Lagrangian

Keeping in mind that the right variables are $(t, \mathbf{x})$, we use the version of the action $\mathcal{A} = \mathcal{A}_{field} + \mathcal{A}_{part}$ that is split in field and particle functionals Eqs.(27, 33). These are repeated here to better grasp the structure

$$\mathcal{A}_{em} = \frac{1}{2}\int dt\, d^3\mathbf{x}\ \left(\epsilon_0\mathbf{E}^2\left(\mathbf{x}, t\right) - \frac{1}{\mu_0}\mathbf{B}^2\left(\mathbf{x}, t\right)\right)$$

$$\mathcal{A}_{fp} = -\sum_{species}\int d\bar{\Gamma}\bar{\mathcal{F}}(\bar{\boldsymbol{\mathcal{Z}}})\bar{\mathcal{H}}(\bar{\boldsymbol{\mathcal{Z}}})$$

Let us remind that the Maxwell equations are obtained by finding an extremum with respect to variations of the electric and vector potential $\delta\phi$ and $\delta\mathbf{A}$.

$$\begin{aligned}\delta\mathcal{A}_{em} &= \int dt\, d^3\mathbf{x}\ \left[\epsilon_0(\nabla\cdot\mathbf{E})\delta\phi + \left(\epsilon_0\frac{\partial\mathbf{E}}{\partial t} - \frac{1}{\mu_0}\nabla\times\mathbf{B}\right)\cdot\delta\mathbf{A}\right] \\ &+ \int dt\, d^3\mathbf{x}\ \left[-\epsilon_0\frac{\partial}{\partial t}\left(\delta\mathbf{A}\cdot\mathbf{E}\right) + \nabla\cdot\left(\frac{\mathbf{B}\times\delta\mathbf{A}}{\mu_0} - \epsilon_0\mathbf{E}\delta\phi\right)\right]\end{aligned}$$

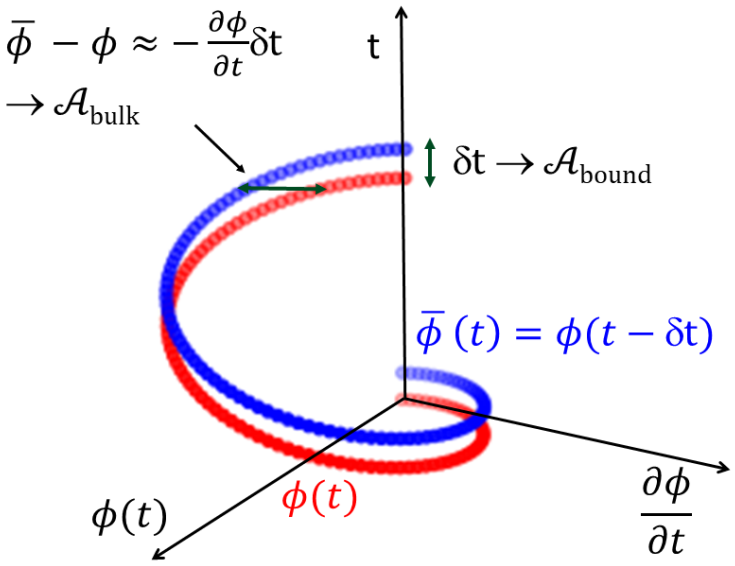


Figure 22: The example Fig.14 is reconsidered. The action is $\mathcal{A} = \int_{t_1}^{t_2} dt \mathcal{L}\left(\phi, \partial_t \phi\right)$ where the Lagrangian $\mathcal{L}\left(\phi, \partial_t \phi\right) = \frac{1}{2}\left(\partial_t \phi\right)^2 - \frac{1}{2}\phi^2$. The action calculated for the solution $\phi(t) = \phi_0 \sin(t - t_1)$ is $\mathcal{A}(t) = \frac{1}{4}\phi_0^2 \sin(t_2 - t_1)$. It is left invariant by a time shift δt. The considered solution is plotted in red for the choice $t_1 = 0$ and $t_2 = 2\pi$. The function $\bar{\phi}(t)$ shown in blue is defined as $\bar{\phi}(t + \delta t) = \phi(t)$. The shift of the time bounds (t_1, t_2) provides a first contribution to the action variation. The difference $\bar{\phi}(t) - \phi(t) \simeq -\frac{\partial \phi}{\partial t}\delta t$ provides a second contribution. The total variation must cancel whatever the choice of (t_1, t_2). In this particular case, it provides the energy conservation equation $\partial_t \mathcal{E} = 0$, where $\mathcal{E} = \frac{1}{2}(\partial_t \phi)^2 + \frac{1}{2}\phi^2$.

Adding the charged particle functional $\mathcal{A}_{fp}$ Eq.(28) and computing the extremum of the full electromagnetic action with respect to $(\delta\phi, \delta\mathbf{A})$ provides the Maxwell equations Eqs.(29). The functional $\mathcal{A}_{fp}$ can be rewritten as a functional of the gyrocentre distribution function. Using the recipes described in the previous sections provide the gyrokinetic Maxwell equations discussed previously.

The gyrokinetic Vlasov equation is obtained via an extremum with a gauge function $\delta\mathcal{S}$ such that $\delta\bar{\mathcal{F}} = \left\{\delta\mathcal{S}, \bar{\mathcal{F}}\right\}$ that results from a change of variable $\bar{\mathcal{Z}} = \mathcal{Z} + \{\delta\mathcal{S}, \mathcal{Z}\}$. It is useful to make use of the Jacobi identity, a property of Poisson brackets, which states that

$$\left\{\delta\mathcal{S}, \bar{\mathcal{F}}\right\}\bar{\mathcal{H}} + \left\{\bar{\mathcal{H}}, \delta\mathcal{S}\right\}\bar{\mathcal{F}} + \left\{\bar{\mathcal{F}}, \bar{\mathcal{H}}\right\}\delta\mathcal{S} = 0$$

so that

$$\delta\mathcal{A}_{fp} = -\sum_{species}\int d\bar{\Gamma}\left\{\bar{\mathcal{H}}, \bar{\mathcal{F}}\right\}\delta\mathcal{S} + \sum_{species}\int d\bar{\Gamma}\left\{\bar{\mathcal{H}}, \delta\mathcal{S}\right\}\bar{\mathcal{F}}$$

Using the relations $\left\{\bar{\mathcal{H}}, \delta\mathcal{S}\right\} = \dot{\bar{\mathcal{Z}}}\partial_{\bar{\mathcal{Z}}}\delta\mathcal{S}$ and $\partial_{\bar{\mathcal{Z}}}(\dot{\bar{\mathcal{Z}}}\mathcal{F}) = 0$, and ignoring the partial derivative in parallel gyrocentre momentum that does not contribute to the Noether's current, one finds

$$\delta\mathcal{A}_{fp} = -\sum_{species}\int d\bar{\Gamma}\left\{\bar{\mathcal{H}}, \bar{\mathcal{F}}\right\}\delta\mathcal{S} + \sum_{species}\int d\bar{\Gamma}\left[\frac{\partial}{\partial\bar{t}}\left(\bar{\mathcal{F}}\delta\mathcal{S}\right) + \frac{\partial}{\partial\bar{\mathbf{X}}}\left(\dot{\bar{\mathbf{X}}}\bar{\mathcal{F}}\delta\mathcal{S}\right)\right]$$

The Euler-Lagrange equation is the gyrokinetic Vlasov equation $\left\{\bar{\mathcal{H}}, \bar{\mathcal{F}}\right\} = 0$, and the gauge current is $\dot{\bar{\mathbf{X}}}\bar{\mathcal{F}}\delta\mathcal{S}$.

6.2.2 Noether's energy and flux

The recipe described in section 6.1 is now applied to the gyrokinetic Lagrangian. The Noether's field variations are

$$\delta\phi = -\frac{\partial\phi}{\partial t}\delta t \qquad \delta\mathbf{A} = -\frac{\partial\mathbf{A}}{\partial t}\delta t \qquad \delta\mathcal{S} = -E\delta t$$

The last relation is justified by the change of variable $\bar{\mathcal{Z}} = \mathcal{Z} + \{\delta\mathcal{S}, \mathcal{Z}\}$, and therefore $\delta t = \{\delta\mathcal{S}, t\} = -\partial_E\delta\mathcal{S}$. We start with the electromagnetic action Eq.(63). The boundary contribution is

$$\delta\mathcal{A}_{em,bound} = \frac{1}{2}\delta t \int dt\, d^3\mathbf{x}\, \frac{\partial}{\partial t}\left[\epsilon_0\mathbf{E}^2 - \frac{1}{\mu_0}\mathbf{B}^2\right]$$

Regarding the bulk contribution, it was seen that only the Noether's current contributes since the field is solution of the Euler-Lagrange equations. Using $\mathbf{E} = -\partial_t\mathbf{A} - \nabla\phi$, the bulk action variation is

$$\begin{aligned}\delta\mathcal{A}_{em,bulk} &= \delta t \int dt\, d^3\mathbf{x}\, \frac{\partial}{\partial t}\left[-\epsilon_0\mathbf{E}^2 - \epsilon_0\mathbf{E}\cdot\nabla\phi\right] \\ &- \int dt\, d^3\mathbf{x}\, \nabla\cdot\left[\frac{\mathbf{E}\times\mathbf{B}}{\mu_0} + \frac{\nabla\phi\times\mathbf{B}}{\mu_0} - \epsilon_0\mathbf{E}\frac{\partial\phi}{\partial t}\right]\end{aligned}$$

Terms can be reorganised by using the relationships

$$\begin{aligned}\nabla\cdot(\nabla\phi\times\mathbf{B}) &= -(\nabla\times\mathbf{B})\cdot\nabla\phi \\ \nabla\cdot\left(\mathbf{E}\frac{\partial\phi}{\partial t}\right) &= (\nabla\cdot\mathbf{E})\frac{\partial\phi}{\partial t} + \mathbf{E}\cdot\frac{\partial\nabla\phi}{\partial t}\end{aligned}$$

In addition, since the field is solution of the Maxwell equations, and also $\partial_t\varrho + \nabla\cdot\mathbf{J} = 0$, one gets finally

$$\begin{aligned}\delta\mathcal{A}_{em} &= -\delta t \int dt\, d^3\mathbf{x}\, \frac{\partial}{\partial t}\left[\frac{\epsilon_0\mathbf{E}^2}{2} + \frac{\mathbf{B}^2}{2\mu_0} - \varrho\phi\right] \\ &- \delta t \int dt\, d^3\mathbf{x}\, \nabla\cdot\left[\frac{\mathbf{E}\times\mathbf{B}}{\mu_0} - \mathbf{J}\phi\right]\end{aligned}$$

A similar calculation shows that the variation of the field-particle interaction action is

$$\delta\mathcal{A}_{fp} = -\delta t \sum_{species} \int d\bar{\Gamma}\left[\frac{\partial}{\partial\bar{t}}\left(\bar{\mathcal{F}}E\right) + \frac{\partial}{\partial\bar{\mathbf{X}}}\cdot\left(\dot{\bar{\mathbf{X}}}\bar{\mathcal{F}}E\right)\right]$$

6.2.3 Energy in the physical space

We are now facing the same problem as usual : moving the action written in the gyrocentre position space back to the physical space. However the task is much simpler now. Indeed there is no operator to invert here. Hence the dummy variable $\bar{\mathbf{X}}$ can just be replaced by $\mathbf{x}$. To illustrate this point, we proceed step by step. Let us note that

$$d\bar{t}\int d^3\bar{\mathbf{X}}... = dt\int d^3\mathbf{x}\int d^3\bar{\mathbf{X}}\delta(\mathbf{x}-\bar{\mathbf{X}})\delta(t-\bar{t})...$$

It is reminded that

$$d\bar{\Gamma} = d\bar{t}\, d^3\bar{\mathbf{X}}\, d\bar{p}_\parallel\, d\bar{\mu}\, d\varsigma\, dE\, B^*_{||}(\bar{\mathbf{X}}, \bar{p}_\parallel)\delta(E-\bar{H})$$

So that finally

$$\delta\mathcal{A}_{fp} = -\delta t \sum_{species}\int dt\, d^3\mathbf{x}\int d^3\bar{\mathbf{X}}d^3\bar{\mathbf{V}}\left[\frac{\partial}{\partial t}\left(\bar{\mathcal{F}}\bar{H}\right) + \frac{\partial}{\partial\bar{\mathbf{X}}}\cdot\left(\dot{\bar{\mathbf{X}}}\bar{\mathcal{F}}\bar{H}\right)\right]\delta(\bar{\mathbf{X}}-\mathbf{x})$$

where $d^3\bar{\mathbf{V}} = \frac{2\pi}{m_a} B_{||}^* d\bar{p}_\parallel d\bar{\mu}$. So finally the energy conservation that is looked for is

$$\frac{\partial \mathtt{E}}{\partial t} + \nabla \cdot \mathbf{\Gamma}_{\mathtt{E}} = 0$$

where

$$\begin{aligned} \mathtt{E} &= \frac{\epsilon_0}{2}\mathbf{E}^2(\mathbf{x},t) + \frac{1}{2\mu_0}\mathbf{B}^2(\mathbf{x},t) - \varrho(\mathbf{x},t)\phi(\mathbf{x},t) \\ &+ \sum_{species} \int d^3\bar{\mathbf{V}} \left(\bar{\mathcal{F}}\bar{H}\right)(\mathbf{x},\bar{\mu},\bar{p}_\parallel,t) \end{aligned}$$

is the energy density, and

$$\begin{aligned} \mathbf{\Gamma}_{\mathtt{E}} &= \frac{1}{\mu_0}\left(\mathbf{E}(\mathbf{x},t)\times\mathbf{B}(\mathbf{x},t)\right) - \mathbf{J}(\mathbf{x},t)\phi(\mathbf{x},t) \\ &+ \sum_{species} \int d^3\bar{\mathbf{V}} \left(\dot{\bar{\mathbf{X}}}\bar{\mathcal{F}}\bar{H}\right)(\mathbf{x},\bar{\mu},\bar{p}_\parallel,t) \end{aligned} \tag{63}$$

the energy flux. The contribution of $-\rho\phi$ can be reworked to be regrouped with the particle energy. Indeed, the following identity can be shown (see Appendix I)

$$\int dt\, d^3\mathbf{x}\, \varrho(\mathbf{x},t)\phi(\mathbf{x},t) = \sum_{species} e_a \int dt\, d^3\mathbf{x}\, d^3\bar{\mathbf{V}}\, \bar{\mathcal{F}}(\mathbf{x},\bar{\mu},\bar{p}_\parallel,t)\mathcal{J}[\phi](\mathbf{x},\bar{\mu},t) \tag{64}$$

A similar operation can be done for the current. This transformation allows recovering the energy conservation law derived in [13].

6.3 Global energy conservation

6.3.1 Expression of the total energy

The local law of energy conservation is elegant, but difficult to use in practice. Building a global version, i.e., an equation over the volume integrated energy, is rewarding. The analysis is further simplified by assuming a vacuum permittivity ϵ_0 that is much smaller than the polarisation permittivity. In this limit, the plasma is quasi-neutral, i.e. $\varrho = 0$. The electric energy term $\epsilon_0\mathbf{E}^2/2$ can then be neglected. The analysis is restricted to the perturbed electromagnetic field - this is not mandatory, but consistent with a given guide magnetic field. The electromagnetic energy is the classical one

$$\mathcal{E}_{em} = \frac{1}{2\mu_0}\int d^3\mathbf{x}\, \tilde{\mathbf{B}}^2(\mathbf{x},t)$$

The particle energy is

$$\mathcal{E}_{part} = \sum_{species} \int d^3\mathbf{x} \int d^3\bar{\mathbf{V}}\bar{\mathcal{F}}\bar{\mathcal{H}}$$

where $\bar{\mathcal{H}} = \bar{\mathcal{H}}_{eq} + \bar{\mathcal{H}}_1 + \bar{\mathcal{H}}_2$. The second order Hamiltonian is given by Eq.(60). Each order of the Hamiltonian reads

$$\bar{\mathcal{H}}_{eq} = \frac{m_a}{2}\bar{p}_\parallel^2 + \bar{\mu}B_{eq}$$

$$\bar{\mathcal{H}}_1 = \mathcal{J}[\mathcal{H}_1]$$

$$\bar{\mathcal{H}}_2 = \frac{e_a^2}{2m_a}\mathcal{J}[\tilde{\mathbf{A}}^2] - \frac{1}{2B_{eq}}\frac{\partial}{\partial\bar{\mu}}\left(\mathcal{J}[\mathcal{H}_1^2] - (\mathcal{J}[\mathcal{H}_1])^2\right)$$

The particle energy for ons species only then reads

$$\begin{aligned}
\mathcal{E}_{part} &= \int d^3\mathbf{x}dt \int d^3\bar{\mathbf{V}}\bar{\mathcal{F}}\bar{\mathcal{H}}_{eq} \\
&+ \int d^3\mathbf{x}dt \int d^3\bar{\mathbf{V}}\bar{\mathcal{F}}\mathcal{J}[\mathcal{H}_1] \\
&+ \frac{1}{2}\int d^3\mathbf{x}dt \int d^3\bar{\mathbf{V}}\mathcal{F}'_{eq}\left(\mathcal{J}[\mathcal{H}_1^2] - (\mathcal{J}[\mathcal{H}_1])^2\right) \\
&+ \int d^3\mathbf{x}dt \int d^3\bar{\mathbf{V}}\mathcal{F}_{eq}\frac{e_a^2}{2m_a}\mathcal{J}[\tilde{\mathbf{A}}^2]
\end{aligned}$$

Since $(\bar{p}_\parallel, \bar{\mu})$ are dummy variables in the integral above, they can be safely replaced by $(p_\parallel, \mu)$. The element of integration in the velocity space then becomes $d^3\mathbf{V} = \frac{2\pi}{m_a}B^*_{||}dp_\parallel d\mu$. Hence the total energy is

$$\begin{aligned}
\mathcal{E} &= \int d^3\mathbf{x}\,\frac{\tilde{\mathbf{B}}^2}{2\mu_0} \\
&+ \int d^3\mathbf{x}\,d^3\mathbf{V}\bar{\mathcal{F}}\bar{\mathcal{H}}_{eq} \\
&+ \sum_{species}\int d^3\mathbf{x}\,d^3\mathbf{V}\bar{\mathcal{F}}\mathcal{J}[\mathcal{H}_1] \\
&+ \frac{1}{2}\sum_{species} d^3\mathbf{x}\int d^3\mathbf{V}\mathcal{F}'_{eq}\left(\mathcal{J}[\mathcal{H}_1^2] - (\mathcal{J}[\mathcal{H}_1])^2\right) \\
&+ \sum_{species}\int d^3\mathbf{x}\,d^3\mathbf{V}\mathcal{F}_{eq}\frac{e_a^2}{2m_a}\mathcal{J}[\tilde{\mathbf{A}}^2] \qquad (65)
\end{aligned}$$

This result calls for some comments. First, the energy contains clearly a vacuum field energy (the 1st line of Eq.(65)), a kinetic energy term (the 2nd line). However the physical nature of the 3 other contributions is less clear. For instance, the two last ones apparently belong to the category "field energy" since they involve the electric and potential vectors squared, and still they come from the "particle energy" contribution. In fact the 4th line corresponds to polarisation and magnetisation field energy that one would be tempted to include in the electromagnetic energy. Same remark for the skin-depth term (5th line). Finally the status of the field-particle interaction term (3rd line of Eq.(65)) is the only one that remains fuzzy at this point.

Second, the sign of the field energy is not well defined. Obviously a positive energy is preferable. For instance the polarisation term (4th line) is usually negative[39]. The same difficulty occurs with the field-particle interaction term (3rd line), whose sign is undetermined. To make some progress, it it necessary to scrutinise the Maxwell equations to enable a clear separation between kinetic and field energy densities.

6.3.2 Revisiting the kinetic and field energies

The Maxwell equations in the gyrokinetic framework are derived by having the action $\mathcal{A} = \mathcal{A}_{field} + \mathcal{A}_{part} + \mathcal{A}_{skin}$, given by Eqs.(53,54,54), extremum with respect to $(\tilde{\phi}, \tilde{\mathbf{A}})$. Hence the electromagnetic potentials $(\tilde{\phi}, \tilde{\mathbf{A}})$ are given by linear operators applied to charge and current densities. This is just the consequence of the linearity of the Maxwell equations

[39]The term $\mathcal{J}[\mathcal{H}_1^2] - (\mathcal{J}[\mathcal{H}_1])^2$ can be written $\left\langle \left(\mathcal{H}_1 - \langle\mathcal{H}_1\rangle_\zeta\right)^2 \right\rangle_\zeta$ and is thus positive. Most equilibrium distribution function verifies $\mathcal{F}'_{eq} < 0$, so that the overall integral is negative. Note that the term $\tilde{\mathbf{A}}^2_\perp$ in the last line is balanced by an equivalent term in the $\mathcal{J}[\mathcal{H}_1^2]$ contribution to the polarisation functional (the third line) so that the polarisation energy does not depend explicitly on $\tilde{\mathbf{A}}_\perp$.

that relate fields to sources, i.e. charge and current densities. An extra condition is obtained by integrating the Maxwell equations over $(\mathbf{x}, t)$. The following identity is derived

$$\begin{aligned}\sum_{species} \int d^3\mathbf{x}\, d^3\mathbf{V} \bar{\mathcal{F}} \mathcal{J}[\mathcal{H}_1] &= -\int d^3\mathbf{x}\, \frac{\tilde{\mathbf{B}}^2}{\mu_0} \\ &- \sum_{species} \int d^3\mathbf{x}\, d^3\mathbf{V} \mathcal{F}'_{eq} \left(\mathcal{J}[\mathcal{H}_1^2] - (\mathcal{J}[\mathcal{H}_1])^2 \right) \\ &- \sum_{species} \int d^3\mathbf{x}\, d^3\mathbf{V} \mathcal{F}_{eq} \frac{e_a^2}{m_a} \mathcal{J}[\tilde{\mathbf{A}}^2] \end{aligned} \tag{66}$$

Combining Eqs.(65,66) yields the conservation equation $\partial_t \mathcal{E} = 0$, with

$$\mathcal{E} = \int d^3\mathbf{x}\, d^3\mathbf{V} \bar{\mathcal{F}} \bar{\mathcal{H}}_{eq} + \frac{1}{2} \sum_{species} \int d^3\mathbf{x}\, d^3\mathbf{V}\, \bar{\mathcal{F}}_1 \mathcal{J}[\mathcal{H}_1] \tag{67}$$

This relation is identical to the one derived in [20]. Note also that it can be derived directly from the gyrokinetic Vlasov equation (see Appendix I). Eq.(67) can then be used to eliminate the field-particle contribution $\bar{\mathcal{F}}_1 \mathcal{J}[\mathcal{H}_1]$, whose sign is undetermined, in the expression of the energy Eq.(65). The latter then becomes

$$\begin{aligned}\mathcal{E} &= -\int d^3\mathbf{x}\, \frac{\tilde{\mathbf{B}}^2}{2\mu_0} + \int d^3\mathbf{x}\, d^3\mathbf{V} \bar{\mathcal{F}} \bar{\mathcal{H}}_{eq} \\ &- \frac{1}{2} \sum_{species} d^3\mathbf{x} \int d^3\mathbf{V} \mathcal{F}'_{eq} \left(\mathcal{J}[\mathcal{H}_1^2] - (\mathcal{J}[\mathcal{H}_1])^2 \right) \\ &- \sum_{species} \frac{1}{2} \int d^3\mathbf{x}\, d^3\mathbf{V} \mathcal{F}_{eq} \frac{e_a^2}{m_a} \mathcal{J}[\tilde{\mathbf{A}}^2] \end{aligned} \tag{68}$$

It appears that unfortunately the magnetic energy appears with a negative sign, and so does the skin depth term. An overall positive sign of the energy is demonstrated as follows. Let us remark first that

$$\int d^3\mathbf{x}\, \frac{\tilde{\mathbf{B}}^2}{\mu_0} = \int d^3\mathbf{x}\, \tilde{\mathbf{J}} \cdot \tilde{\mathbf{A}}$$

assuming $\tilde{\mathbf{B}} \times \tilde{\mathbf{A}} = 0$ at the boundary of the integration volume. Next step follows the same chain of transformations performed to compute the integral of $\varrho\phi$ (see Appendix I)

$$\begin{aligned}\int d^3\mathbf{x}\, \tilde{\mathbf{J}} \cdot \tilde{\mathbf{A}} &= \sum_{species} e_a \int dt\, d^3\mathbf{x}\, d^3\mathbf{V}\, \bar{\mathcal{F}} \left(p_\parallel \mathcal{J}[\tilde{A}_\parallel] - \mu \mathcal{B}[\tilde{B}_\parallel] \right) \\ &- \sum_{species} \int d^3\mathbf{x}\, d^3\mathbf{V}\, \bar{\mathcal{F}}_{eq} \frac{e_a^2}{m_a} \mathcal{J}[\tilde{\mathbf{A}}^2] \end{aligned}$$

This relationship is approximate and is valid under the same conditions as the calculations done in the $(\phi, A_\parallel, B_\parallel)$ representation. Rearranging various terms, it appears that the energy can be split in 2 components $\mathcal{E} = \mathcal{E}_{kin} + \mathcal{E}_{field}$, where $\mathcal{E}_{kin}$ is a particle kinetic energy defined as

$$\mathcal{E}_{kin} = \sum_{species} \int d^3\mathbf{x}\, d^3\mathbf{V}\, \bar{\mathcal{F}} \left\{ \frac{1}{2} m_a \left(p_\parallel - \frac{e_a}{m_a} \mathcal{J}[\tilde{A}_\parallel] \right)^2 + \mu \left(B_{eq} + \mathcal{B}[\tilde{B}_\parallel] \right) \right\}$$

and $\mathcal{E}_{field}$ is a "field energy", itself split in two parts, that is

$$\mathcal{E}_{field} = \mathcal{E}_{em} + \mathcal{E}_{pol}$$

where

$$\mathcal{E}_{em} = \int d^3\mathbf{x}\, \frac{\tilde{\mathbf{B}}^2}{2\mu_0}$$

is the electromagnetic field energy (neglecting the vacuum electric energy), and

$$\begin{aligned}
\mathcal{E}_{pol} = & - \frac{1}{2}\sum_{species}\int d^3\mathbf{x}\, d^3\mathbf{V}\, \mathcal{F}'_{eq}\left(\mathcal{J}[\mathcal{H}_1^2] - (\mathcal{J}[\mathcal{H}_1])^2\right) \\
& + \frac{1}{2}\sum_{species}\int d^3\mathbf{x}\, d^3\mathbf{V}\mathcal{F}_{eq}\frac{e_a^2}{m_a}\mathcal{J}[\tilde{\mathbf{A}}_\perp^2] \\
& + \frac{1}{2}\sum_{species}\int d^3\mathbf{x}\, d^3\mathbf{V}\mathcal{F}_{eq}\frac{e_a^2}{m_a}\left\{\mathcal{J}[\tilde{A}_\parallel^2] - (\mathcal{J}[\tilde{A}_\parallel])^2\right\}
\end{aligned}$$

is the "field polarisation-magnetisation" energy, which is always positive[40]. This expression is consistent with [21] in the long wavelength limit of the $(\phi, A_\parallel)$ representation. Hence a meaningful energy conservation constraint is found

$$\frac{\partial}{\partial t}\left(\mathcal{E}_{kin} + \mathcal{E}_{field}\right) = 0 \tag{69}$$

A close inspection of Eq.(68) is rewarding. As mentioned, the first term $\mathcal{E}_{em}$ correspond to the classical definition of the energy of the electromagnetic field in the vacuum. The second one $\mathcal{E}_{pol}$ is more intriguing. In the electrostatic limit $\tilde{\mathbf{A}} = 0$, long wavelength limit $\mathcal{J} = 1 - \frac{1}{4}\rho_c^2\nabla_\perp^2$ and for a Maxwellian distribution function, the field energy can be recast as

$$\mathcal{E}_{field} = \sum_{species}\int d^3\mathbf{x}\, \frac{N_{eq}m_a}{B_{eq}^2}\left|\nabla_\perp\phi\right|^2$$

This is the energy associated with the polarisation field $\mathbf{P} = \epsilon_{pol}\mathbf{E}_\perp$[41]. In the $(\phi, A_\parallel)$ electromagnetic case, an additional energy term appears. Assuming an unshifted large scale Maxwellian [42], it is of the form

$$\mathcal{E}_{field} = \sum_{species}\int d^3\mathbf{x}\, \frac{N_{eq}m_a}{B_{eq}^2}\left|\nabla_\perp\phi\right|^2 + \int d^3\mathbf{x}\, \frac{1}{\mu_0}\left(1 + \sum_{species}\frac{\mu_0 N_{eq}T_{eq}}{B_{eq}^2}\right)\left|\nabla_\perp\tilde{A}\right|^2$$

In the fully electromagnetic case, the energy $\mathcal{E}_{pol}$ contains also compressional magnetic terms that involve $\tilde{B}_\parallel$. Its full expression is complex and beyond the scope of this overview.

6.4 Exchange power

It is interesting to analyse the energy budget from the stand point of the electromagnetic field. Ignoring the influx of energy in the volume of interest associated with the Poynting vector, the electromagnetic energy balance equation reads

$$\frac{\partial\mathcal{E}_{em}}{\partial t} = -\int d^3\mathbf{x}\, \tilde{\mathbf{J}}\cdot\tilde{\mathbf{E}} \tag{70}$$

[40]Note that for a Maxwellian distribution function, the term in $\tilde{\mathbf{A}}_\perp^2$ is balanced by an equivalent term in the $\mathcal{J}[\mathcal{H}_1^2]$ contribution to the "adiabatic" functional so that the polarisation energy does not depend explicitly on $\tilde{\mathbf{A}}_\perp$. On the other hand, quadratic terms in $\tilde{A}_\parallel$ add up.

[41]It corresponds to the classical definition $\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P}$, where $\mathbf{P}$ is the polarisation field [14]. It is recalled that the polarisation permittivity is $\epsilon_{pol} = \rho_m/B_{eq}^2$, with $\rho_m = \sum_{species} N_{eq}m$ the mass density, which is roughly c^2/v_A^2 larger than ϵ_0, where c is the speed of light and $v_A = \left(B_{eq}^2/\mu_0\rho_m\right)^{1/2}$ the Alfvén velocity.

[42]Unshifted means here no equilibrium current so that cross terms proportional to $\tilde{\phi}\tilde{A}_\parallel$ vanish in the polarisation-magnetisation term.

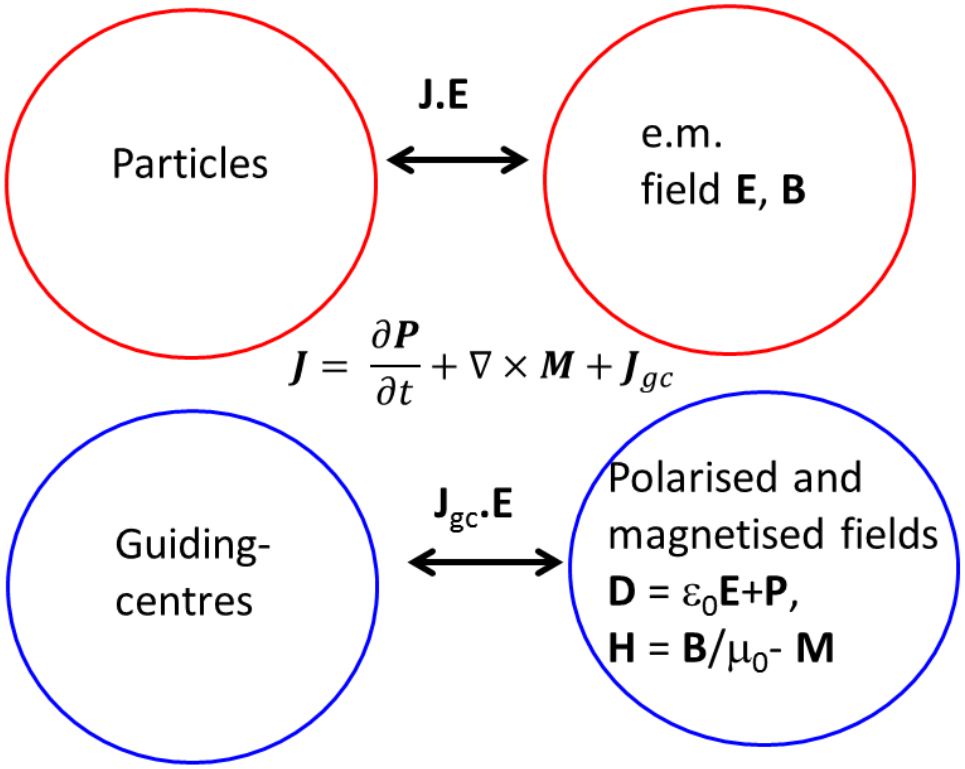


Figure 23: Particle/field energy exchange. Upper panel: particles and electromagnetic field exchange energy via the product $\tilde{\mathbf{J}} \cdot \tilde{\mathbf{E}}$. Lower panel: in the guiding-centre stand point, the kinetic energy is associated with guiding-centres - the polarisation and magnetisation terms enter the electromagnetic energy via the polarised and magnetisation fields.

where

$$\mathcal{E}_{em} = \int d^3\mathbf{x} \left(\frac{\epsilon_0}{2} \tilde{\mathbf{E}}^2 + \frac{1}{2\mu_0} \tilde{\mathbf{B}}^2 \right)$$

is the total classical electromagnetic energy. This equation can be worked out by using the expression of the electric field

$$\tilde{\mathbf{E}} = -\frac{\partial \tilde{\mathbf{A}}}{\partial t} - \nabla \tilde{\phi} \tag{71}$$

and the charge conservation equation

$$\frac{\partial \tilde{\varrho}}{\partial t} + \nabla \cdot \tilde{\mathbf{J}} = 0$$

The new balance equation is

$$\frac{\partial}{\partial t} \int d^3\mathbf{x} \left(\frac{\epsilon_0 \tilde{\mathbf{E}}^2}{2} - \frac{\tilde{\mathbf{B}}^2}{2\mu_0} \right) = \int d^3\mathbf{x} \left(\frac{\partial \tilde{\varrho}}{\partial t} \tilde{\phi} - \frac{\partial \tilde{\mathbf{J}}}{\partial t} \cdot \tilde{\mathbf{A}} \right)$$

up to a surface term that was set to zero [43]. Using the definition of the charge and current densities Eqs.(30), it becomes

$$\frac{\partial}{\partial t} \left(\int d^3\mathbf{x} \left(\frac{\epsilon_0 \tilde{\mathbf{E}}^2}{2} - \frac{\tilde{\mathbf{B}}^2}{2\mu_0} \right) - \sum_{species} \int d^3\mathbf{x}\, d^3\mathbf{V} \mathcal{F}_{eq} \frac{e_a^2}{2m_a} \mathcal{J}[\tilde{\mathbf{A}}^2] \right) = \sum_{species} \int d\Gamma \frac{\partial \mathcal{F}_1}{\partial t} \mathcal{H}_1$$

[43] This term corresponds to a flux of energy $\tilde{\mathbf{J}}\tilde{\phi}$, i.e. a flux of electric potential energy carried by particles, which plays a role similar to the Poynting flux, but for particles, as seen in the energy flux Eq.(63).

The guiding-centre and gyrocentre distribution functions $\mathcal{F}_1$ and $\bar{\mathcal{F}}_1$ are related by Eq.(22). The first term in the latter equation gives rise to the polarisation field energy. When combined with Eqs.(140,141), derived in Appendix I, it appears that the conservation equation $\partial_t \mathcal{E} = 0$ is recovered, with the total energy $\mathcal{E}$ given by Eq.(68). As a consequence, energy conservation law Eq.(69) is also fully recovered. Hence this implies that

$$\frac{\partial}{\partial t}\left(\mathcal{E}_{kin} + \mathcal{E}_{pol}\right) = \int d^3\mathbf{x}\, \tilde{\mathbf{J}} \cdot \tilde{\mathbf{E}} \tag{72}$$

It is satisfactory to find a variation of the electromagnetic energy consistent with its kinetic counterpart. But the real assets are Eqs.(70, 72), which enable diagnosing the power exchange between fields and particles. In particular, when detailed for each species, and even per class of energy, it offers a deep insight in the nature of particles that drive or damp instabilities. It is also interesting to see the polarisation term can be handled in two different ways. On the one hand Eq.(72) clearly indicates that the polarisation term belongs to the particle realm - it comes in fact to a shift of the guiding-centre motion under the effect of the electromagnetic field. On the other hand Eq.(70) suggests that is can be as well included in the field energy via the polarisation field. Both approaches are sound, but this freedom is sometimes source of confusion (see Fig.23).

7 Conclusion

Gyrokinetic theory is commonly considered as technical, and sometimes abstract. Nevertheless it uncovers substantial physical outcomes and prominent features that can be summarized as follows. The first objective of the gyrokinetic theory is to modify the magnetic moment so that it remains an invariant of motion. This is possible as long as as the typical frequencies of electromagnetic perturbations are low compared with the cyclotron frequency. Also the amplitude of fluctuations must be small compared to the equilibrium fields. No constraint on spatial scales is requested, which allows applying the model to small scale turbulence in magnetised plasmas. Beyond its usefulness for tractable analytic calculations, gyrokinetic theory is motivated by one key fact: developing and using 6D kinetic codes is a difficult task given the available computers. 6D codes are barely emerging, and mostly for ions. Gyrokinetic theory thus offers a path to computational modelling of turbulence. The first step in gyrokinetics, namely the reinstatement of the magnetic moment as an invariant of motion, can be seen as a change of variables, from guiding-centre coordinates to the coordinates of a virtual particle, localised nearby the guiding-centre, and called gyrocentre. The starting point is a first change of coordinates, from the "true" particle coordinates to guiding-centre coordinates, a well documented step. Hence a double change of coordinates must be managed.

Not all choices of gyrocentre variables are convenient for the second part of gyrokinetic theory, namely the resolution of the Maxwell equations. Indeed Maxwell equations relate the electric and magnetic vector fields $\mathbf{E}(\mathbf{x}, t), \mathbf{B}(\mathbf{x}, t)$ to charge and current densities. This step requires the computation of charge and current densities as function of the spatial coordinates $\mathbf{x}$ of charged particles in the physical space, and time t. In practice, this requirement enforces the use of non conjugate coordinates, namely 3 positions of gyrocentres, the magnetic moment, a generalised parallel velocity, and a gyroangle - the later does not appear explicitly in the final equations. A change of variables from guiding-centre to gyrocentres coordinates requires some care. Also equations of motion are not the same depending on the choice of generalised parallel velocity. Some are more convenient than others for a subsequent use of the gyrocentre distribution function to compute charge and current densities. Two options were addressed, most commonly mentioned in the literature as "Hamiltonian" and "Symplectic" approaches. In both cases, the Maxwell equations bear a more complex form than their original version, due to the emergence of polarisation and magnetisation fields. This comes physically from the spatial shift between

gyrocentre and particle positions. As in any polarisable medium, the modification of the distance between gyrocentres and particles under the effect of an electromagnetic field introduces a polarisation vector field $\mathbf{P}$. Similarly, the magnetic moment can be seen as a magnet intensity associated with the cyclotron motion of the charged particle around its gyrocentre. A collection of particles is therefore characterised by a magnetisation vector field $\mathbf{M}$. Gyrokinetic Maxwell equations are then similar to those used in classical electromagnetism in continuum media [22, 14]. This formulation involves vector fields $\mathbf{D}, \mathbf{H}$ instead of $\mathbf{E}, \mathbf{B}$, and such that $\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}$, and $\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M})$. The equations of motion are better expressed using scalar fields, like the electric potential, the parallel component of the vector potential, and the perturbed magnetic field modulus. First reason is the Hamiltonian nature of gyrokinetic theory, hence better derived using electric and vector potentials. On the other hand, a $\mathbf{E}, \mathbf{B}$ representation is closer to the traditional MHD description. These considerations explain the intricate aspect of the gyrokinetic Maxwell equations.

The Hamiltonian path is the easiest change of coordinates. It is described in details in the present note. In this case the generalised parallel velocity is close to the parallel canonical momentum. The price to pay however is the appearance in the Ampère equation of a skin depth term, whose amplitude is large, and balances partially the parallel current density. This "magnetic cancellation" raises difficulties in the computation, and has been subject of a number of discussions, and ways to circumvent it. The other path is called "Symplectic" formulation, and involves a gyrocentre parallel velocity close to the particle velocity. The force that appears in the equation of motion of the parallel momentum it the electric parallel electric field, instead of the perturbed Hamiltonian. No magnetic cancellation in this case, and the equations of motion actually look close to the usual ones in the $\mathbf{E}, \mathbf{B}$ representation. Difficulty here is the inductive part of the electric field, which involves the time derivative of the vector potential. This means handling time derivatives in the right hand side of the equations of motion, usually not a good idea in terms of numerical stability. Recent developments suggest the use of hybrid formulations, that combine Hamiltonian and Symplectic approaches [23, 24]. A summary of characteristics for each scheme is given in Table 1.

Ordering is also an issue in gyrokinetic theory. Most results can be understood with a first order theory, where the ordering parameter is the amplitude of the perturbed fields normalised to their unperturbed values. However global codes compute the full distribution function, so that a second order theory turns out attractive. As before, this requires two ingredients: changes of non conjugate variables, and reformulation of Maxwell equations, both at second order. This a formidable task, that is most easily done by using Lie transform methods, and a principle of action extremum. This is still an evolving field of research. Only the main points have been sketched here.

As a last word, it must be understood that gyrokinetic theory is mostly a mathematical tool, in essence a change of variables, but physics guided all along. It remains a necessary step for computation, but also allows a deeper understanding of the physics of magnetised plasmas. It would be illusory to ask more to theory than it can deliver. However, it remains a powerful mean to address many issues related to stability and turbulent transport in plasma physics.

APPENDICES

A Notations

	Coordinates	Hamiltonian	Distribution function
Conjugate variables	$\mathbf{z} = (\mathbf{x}, \mathbf{p})$	$h, h_{eq}, \widetilde{h}$	$f, f_{eq}, \widetilde{f}$
Extended	$\mathfrak{z} = (t, -E, \mathbf{x}, \mathbf{p})$	$\mathtt{h}, \mathtt{h}_{eq}$	$\mathtt{f}$
Guiding-centre	$\mathbf{Z} = (\mathbf{X}, u_\parallel, \mu, \varsigma)$	$H, H_{eq}, \widetilde{H}$	$F, F_{eq}, \tilde{F}$
Extended	$\boldsymbol{\mathcal{Z}} = (\mathbf{X}, u_\parallel, \mu, \varsigma, E, t)$	$\mathcal{H}, \mathcal{H}_{eq}, \mathcal{H}_1, \mathcal{H}_2$	$\mathcal{F}, \mathcal{F}_{eq}, \mathcal{F}_1, \mathcal{F}_2$
Gyrocentres	$\bar{\mathbf{Z}} = (\bar{\mathbf{X}}, \bar{p}_\parallel, \bar{\mu}, \bar{\varsigma})$	$\bar{H}$	$\bar{F}$
Extended	$\bar{\boldsymbol{\mathcal{Z}}} = (\bar{\mathbf{X}}, \bar{p}_\parallel, \bar{\mu}, \bar{\varsigma}, E, t)$	$\bar{\mathcal{H}}$	$\bar{\mathcal{F}}$

Table 2: Notations and conventions - hamiltonian and distribution functions.

	Coordinates	Volume element
Conjugate variables	$\mathbf{z} = (\mathbf{x}, \mathbf{p})$	$d^3\mathbf{x}\, d^3\mathbf{p}$
Extended	$\mathfrak{z} = (t, -E, \mathbf{x}, \mathbf{p})$	$d\gamma = d^3\mathbf{x}\, d^3\mathbf{p}\, dt\, dE\, \delta(E - \mathtt{h}_{eq})$
Guiding-centre	$\mathbf{Z} = (\mathbf{X}, u_\parallel, \mu, \varsigma)$	$B^*_\parallel\, d^3\mathbf{X}\, du_\parallel\, d\mu\, d\varsigma$
Extended	$\boldsymbol{\mathcal{Z}} = (\mathbf{X}, u_\parallel, \mu, \varsigma, E, t)$	$d\Gamma = B^*_\parallel\, d^3\mathbf{X}\, du_\parallel\, d\mu\, d\varsigma\, dE\, dt\, \delta(E - \mathcal{H}_{eq})$
Gyrocentres	$\bar{\mathbf{Z}} = (\bar{\mathbf{X}}, \bar{p}_\parallel, \bar{\mu}, \bar{\varsigma})$	$B^*_\parallel\, d^3\bar{\mathbf{X}}\, d\bar{p}_\parallel\, d\bar{\mu}\, d\bar{\varsigma}$
Extended	$\bar{\boldsymbol{\mathcal{Z}}} = (\bar{\mathbf{X}}, \bar{p}_\parallel, \bar{\mu}, \bar{\varsigma}, E, t)$	$d\bar{\Gamma} = B^*_\parallel\, d^3\bar{\mathbf{X}}\, d\bar{p}_\parallel\, d\bar{\mu}\, d\bar{\varsigma}\, d\bar{E}\, d\bar{t}\, \delta(\bar{E} - \bar{\mathcal{H}}_{eq})$
Hybrid	$(\mathbf{x}, p_\parallel, \mu)$	$d^3\mathbf{V} = \frac{2\pi}{m_a} B^*_\parallel\, dp_\parallel\, d\mu$

Table 3: Notations and conventions - volume element.

B From canonical to non-canonical phase-space

B.1 Poisson brackets for non-canonical variables

We start from a 6D phase-space and a set of canonical variables $\mathbf{z} = (\mathbf{q}, \mathbf{p})$. In this section, for simplicity, we do not extend the phase-space to include time and energy, but the formulation straightforwardly generalises to an 8D extended phase-space.

Compared to the Lagrangian formalism, the Hamiltonian formalism in canonical variables $\mathbf{z} = (\mathbf{q}, \mathbf{p})$ has the advantage of providing straightforward equations of motion in the form of a Poisson bracket,

$$\frac{dz^\mu}{dt} = \{z^\mu, H\} \tag{73}$$

where $\{f, g\} = \partial_\mathbf{q} f \cdot \partial_\mathbf{p} g - \partial_\mathbf{p} f \cdot \partial_\mathbf{q} g$ This is easy to check, since $\{q^i, H\} = \partial H/\partial p_i$ and $\{p^i, H\} = -\partial H/\partial q_i$. However this simple form (73) is a priori limited to canonical variables. Contrarily, the Lagrangian formalism is also valid for non-canonical variables, but the relationship between Euler-Lagrange equations and Poisson brackets may be hidden and very tricky to uncover. In section we detail how a Poisson-bracket-based form of Hamilton's equations can be recovered for non-canonical variables.

It is useful to define a phase-space Lagrangian,

$$L(\mathbf{z}, \dot{\mathbf{z}}, t) = \mathbf{p} \cdot \dot{\mathbf{q}} - H(\mathbf{q}, \mathbf{p}, t) \tag{74}$$

This phase-space Lagrangian equals in value the configuration-space Lagrangian (which depends on $\mathbf{q}$ and $\dot{\mathbf{q}}$ but not explicitly on $\mathbf{p}$ and $\dot{\mathbf{p}}$), but lives in a phase-space with twice higher dimension. The advantage of the phase-space Lagrangian is that Euler-Lagrange equations directly yield Hamilton's equations of motion,

$$\frac{d\mathbf{p}}{dt} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}}\right) = \frac{\partial L}{\partial \mathbf{q}} = -\frac{\partial H}{\partial \mathbf{q}} \tag{75}$$

$$0 = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{p}}}\right) = \frac{\partial L}{\partial \mathbf{p}} = \frac{d\mathbf{q}}{dt} - \frac{\partial H}{\partial \mathbf{p}} \tag{76}$$

whereas Euler-Lagrange equations applied to the configuration-space Lagrangian yield only Eq. (75). Therefore, this can be seen as an alternative Hamiltonian formulation, based on a Lagrangian.

We now consider a change of variables from canonical variables $\mathbf{z}$ to non-canonical variables $\mathbf{Z}$. Note that, when expressed in the canonical variables $\mathbf{z}$, the phase-space Lagrangian L does not depend on $\dot{\mathbf{p}}$, but we will see that for non-canonical variables $\mathbf{Z}$, L may in general depend on all 12 components of $\mathbf{Z}$ and $\dot{\mathbf{Z}}$. To express L in terms of $\mathbf{Z}$, we need to express the functions $q^i(\mathbf{Z})$ and $p^i(\mathbf{Z})$. Here we assume that the change of variables is time-independent for simplicity[44]. Then, $\dot{q^i} = (\partial q^i/\partial Z^\mu)\dot{Z}^\mu$, and

$$L(\mathbf{Z}, \dot{\mathbf{Z}}, t) = \Gamma_\mu \dot{Z}^\mu - H(\mathbf{Z}, \dot{\mathbf{Z}}, t) \tag{77}$$

where Γ_μ is a covariant 6-vector,

$$\Gamma_\mu(\mathbf{Z}) = \mathbf{p} \cdot \frac{\partial \mathbf{q}}{\partial Z^\mu} \tag{78}$$

The first part in Eq. (77) is said to be symplectic because it is linear in $\dot{Z}^\mu$. Then Euler-Lagrange equations yield

$$\frac{d\Gamma_\nu}{dt} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Z}^\nu}\right) = \frac{\partial L}{\partial Z^\nu} = \frac{\partial \Gamma_\mu}{\partial Z^\nu}\dot{Z}^\mu - \frac{\partial H}{\partial Z^\nu} \tag{79}$$

which can be rearranged as

$$\frac{\partial H}{\partial Z^\nu} = \frac{\partial \Gamma_\mu}{\partial Z^\nu}\dot{Z}^\mu - \frac{d\Gamma_\nu}{dt} \tag{80}$$

$$= \frac{\partial \Gamma_\mu}{\partial Z^\nu}\dot{Z}^\mu - \frac{\partial \Gamma_\nu}{\partial Z^\mu}\dot{Z}^\mu \tag{81}$$

$$= \omega_{\mu\nu}\dot{Z}^\mu \tag{82}$$

where we have introduced the Lagrange matrix

$$\omega_{\mu\nu} = \frac{\partial \Gamma_\mu}{\partial Z^\nu} - \frac{\partial \Gamma_\nu}{\partial Z^\mu} \tag{83}$$

Assuming that the Lagrange matrix ω is invertible, we define its inverse J, which is called as the Poisson matrix[45]. Then the equations of motion write simply as

$$\dot{Z}^\mu = J^{\mu\nu}\frac{\partial H}{\partial Z^\nu} \tag{84}$$

Therefore, the equations of motion can finally be expressed in terms of a Poisson bracket,

$$\frac{dZ^\mu}{dt} = \{Z^\mu, H\} \tag{85}$$

[44] If the change of variables is time-dependant, then $q^i(\mathbf{Z}, t)$ and there is an extra term $p_i \partial_t q^i$ in the Lagrangian.

[45] An alternative method to calculate the Poisson matrix J, without calculating and inverting the Lagrange matrix ω, is to use the relationship $J = P \cdot \sigma \cdot P^{-1}$. Here P and σ are matrices defined by $P^\nu_\mu = \partial Z^\nu/\partial z^\mu$ and $\sigma = \begin{pmatrix} 0_3 & I_3 \\ -I_3 & 0_3 \end{pmatrix}$, where 0_3 is the 3x3 zero matrix, and I_3 is the 3x3 identity matrix.

albeit with a generalisation of Poisson brackets to non-canonical variables,

$$\{F, G\} = \frac{\partial F}{\partial Z^\mu} J^{\mu\nu} \frac{\partial G}{\partial Z^\nu} \tag{86}$$

Note that $J^{\mu\nu} = \{Z^\mu, Z^\nu\}$ is an immediate consequence of the latter definition.

B.2 A simple example

Let us illustrate the concepts described in subsection B.1 with a simple example, which does not involve cyclotron motion, nor magnetic field, nor extended phase-space. We consider a particle of mass m and charge e in a prescribed electric field $\phi(\mathbf{q}, t)$. We work in Cartesian coordinates $\mathbf{q} = (x, y, z)$ and start from canonical variables $\mathbf{z} = (\mathbf{q}, \mathbf{p})$ with $\mathbf{p} = (p_x, p_y, p_z) = m_a \dot{\mathbf{q}}$. Since the configuration-space Lagrangian is $m_a \dot{\mathbf{q}}^2/2 - e_a\phi$, the phase-space Lagrangian is

$$L(\mathbf{z}, \dot{\mathbf{z}}, t) = \mathbf{p} \cdot \dot{\mathbf{q}} - H(\mathbf{q}, \mathbf{p}, t) \tag{87}$$

with

$$H(\mathbf{q}, \mathbf{p}, t) = \frac{\mathbf{p}^2}{2m_a} + e_a\phi(\mathbf{q}, t) \tag{88}$$

Let us now change variables from $\mathbf{z}$ to non-canonical $\mathbf{Z} = (x, y, z, p_\perp, \alpha, p_z)$ where $p_\perp$ and α are such that $p_x = p_\perp \cos\alpha$ and $p_y = p_\perp \sin\alpha$. At this point we advice the reader to work through the following steps as an exercise:

1. Calculate the 6 components of the covariant vector $\Gamma_\mu = \mathbf{p} \cdot \frac{\partial \mathbf{q}}{\partial Z^\mu}$ in terms of $\mathbf{Z}$.
2. Calculate the Lagrange matrix $\omega_{\mu\nu} = \frac{\partial \Gamma_\mu}{\partial Z^\nu} - \frac{\partial \Gamma_\nu}{\partial Z^\mu}$. Since this is an anti-symmetric tensor, only 15 components need to be calculated. In this simple example, only 4 components are non-trivial.
3. Calculate the Poisson matrix J by inverting ω.
4. Express the new (generalised) Poisson brackets $\{f, g\} = (\partial f/\partial Z^\mu) J^{\mu\nu} (\partial g/\partial Z^\nu)$.
5. Express H in terms of $\mathbf{Z}$.
6. Deduce from that the equations of motion $\dot{Z}^\mu = \{Z^\mu, H\}$

Solution:

1. One finds $\Gamma_1 = p_\perp \cos\alpha$, $\Gamma_2 = p_\perp \sin\alpha$, $\Gamma_3 = p_z$, and $\Gamma_4 = \Gamma_5 = \Gamma_6 = 0$.
2. The only non-zero elements of the Lagrange matrix are

$$\begin{aligned} &\omega_{4\,1} = -\omega_{1\,4} = \cos\alpha && \omega_{4\,2} = -\omega_{2\,4} = \sin\alpha \\ &\omega_{5\,1} = -\omega_{1\,5} = -p_\perp \sin\alpha && \omega_{5\,2} = -\omega_{2\,5} = p_\perp \cos\alpha \\ &\omega_{6\,3} = -\omega_{3\,6} = 1 \end{aligned}$$

3. The only non-zero elements of the Poisson matrix are

$$\begin{aligned} &J^{1\,4} = -J^{4\,1} = \cos\alpha && J^{2\,4} = -J^{4\,2} = \sin\alpha \\ &J^{1\,5} = -J^{5\,1} = -\sin\alpha/p_\perp && J^{2\,5} = -J^{5\,2} = \cos\alpha/p_\perp \\ &J^{3\,6} = -J^{6\,3} = 1 \end{aligned}$$

4. One immediately obtains

$$\begin{aligned} \{F, G\} \;=\; &\left(\frac{\partial F}{\partial x}\frac{\partial G}{\partial p_\perp} - \frac{\partial F}{\partial p_\perp}\frac{\partial G}{\partial x}\right)\cos\alpha \;+\; \left(\frac{\partial F}{\partial y}\frac{\partial G}{\partial p_\perp} - \frac{\partial F}{\partial p_\perp}\frac{\partial G}{\partial y}\right)\sin\alpha \\ &- \left(\frac{\partial F}{\partial x}\frac{\partial G}{\partial \alpha} - \frac{\partial F}{\partial \alpha}\frac{\partial G}{\partial x}\right)\frac{\sin\alpha}{p_\perp} \;+\; \left(\frac{\partial F}{\partial y}\frac{\partial G}{\partial \alpha} - \frac{\partial F}{\partial \alpha}\frac{\partial G}{\partial y}\right)\frac{\cos\alpha}{p_\perp} \\ &+ \left(\frac{\partial F}{\partial z}\frac{\partial G}{\partial p_z} - \frac{\partial F}{\partial p_z}\frac{\partial G}{\partial z}\right) \end{aligned}$$

5. From $H(\mathbf{z}) = \mathbf{p}^2/(2m_a) + e_a\phi(\mathbf{q},t)$ we obtain $H(\mathbf{Z}) = p_\perp^2/(2m_a) + p_z^2/(2m_a) + e_a\phi(\mathbf{q},t)$.
6. From $\dot{Z}^1 = \{Z^1, H\} = \{x, H\}$, and the derivatives $\frac{\partial H}{\partial p_\perp} = p_\perp/m_a$ and $\frac{\partial H}{\partial \alpha} = 0$, we obtain $\dot{x} = (p_\perp/m_a)\cos\alpha$. Similarly, we obtain the 5 other equations of motion. To summarize,

$$\begin{aligned}
\dot{x} &= \frac{p_\perp}{m_a}\cos\alpha \\
\dot{y} &= \frac{p_\perp}{m_a}\sin\alpha \\
\dot{z} &= \frac{p_z}{m_a} \\
\dot{p_\perp} &= -e_a\cos\alpha\frac{\partial\phi}{\partial x} - e_a\sin\alpha\frac{\partial\phi}{\partial y} \\
\dot{\alpha} &= e_a\frac{\sin\alpha}{p_\perp}\frac{\partial\phi}{\partial x} - e_a\frac{\cos\alpha}{p_\perp}\frac{\partial\phi}{\partial y} \\
\dot{p_z} &= -e_a\frac{\partial\phi}{\partial z}
\end{aligned}$$

It is straightforward to check that these are equivalent to the equations of motion obtained in canonical variables, $\dot{\mathbf{q}} = \mathbf{p}/m_a$ and $\dot{\mathbf{p}} = -e_a\partial\phi/\partial\mathbf{q}$.

C Guiding-center Lagrangian and Poisson brackets

C.1 Guiding-center Lagrangian

Let us derive the equilibrium Lagrangian in guiding-centre coordinates, to the zeroth order[46]. Here we work in the 6-dimensional (non-extended) phase-space.

We start in non-canonical position/velocity variables $\mathbf{z} = (\mathbf{x}, \mathbf{v})$, where the Lagrangian is

$$L_{eq}(\mathbf{x},\mathbf{v}) \;=\; \frac{1}{2}m_a v^2 \;+\; e_a\,\mathbf{A}_{eq}\cdot\mathbf{v} \;-\; e_a\phi_{eq} \tag{89}$$

simply because the Euler-Lagrange then yields the correct Newton equation with the electric and Lorentz forces. We recall that the equilibrium electromagnetic field is assumed static to simplify the analysis at this point. Therefore, in the latter expression, $\mathbf{A}_{eq} = \mathbf{A}_{eq}(\mathbf{x})$ and $\phi_{eq} = \phi_{eq}(\mathbf{x})$.

By definition the conjugate to $\mathbf{x}$ is

$$\mathbf{p}(\mathbf{x},\mathbf{v}) \;=\; \frac{\partial L_{eq}}{\partial\mathbf{v}} \;=\; e_a\,\mathbf{A}_{eq}(\mathbf{x}) \;+\; m_a\mathbf{v} \tag{90}$$

The Legendre transformation yields an alternative expression for the Lagrangian,

$$L_{eq}(\mathbf{x},\mathbf{v}) \;=\; [e_a\,\mathbf{A}_{eq}(\mathbf{x}) \;+\; m_a\mathbf{v}]\cdot\dot{\mathbf{x}} \;-\; H_{eq}\left[\mathbf{x},\mathbf{p}(\mathbf{x},\mathbf{v})\right] \tag{91}$$

with the equilibrium Hamiltonian

$$H_{eq}(\mathbf{x},\mathbf{p}) \;=\; \frac{1}{2}m_a v(\mathbf{x},\mathbf{p})^2 \;+\; e_a\phi_{eq}(\mathbf{x}) \tag{92}$$

The goal now is to obtain the Lagrangian as a function of $\mathbf{Z} = (\mathbf{X}, u_\parallel, \mu, \varsigma)$. The term $\mathbf{A}_{eq}(\mathbf{x})$ can be expressed in terms of $\mathbf{A}_{eq}(\mathbf{X})$ and its derivatives by a Taylor expansion

[46]The method of Lie transforms is much more systematic, and can be used to derive the Lagrangian in an arbitrary set of coordinates, to an arbitrary order. This method was used by Littlejohn to obtain the guiding-centre Lagrangian to higher order [7].

with respect to the small parameter $\epsilon_B = \rho_c/L_B$. Similarly, the relationship between $\mathbf{v}$ and $u_\parallel$ and μ can be expressed as a Taylor series. The main difficulty is to ensure that, in the Lagrangian, all terms up to the order $(\epsilon_B)^0$ – with respect to $m_a v^2$ – are kept (and we wish to keep as few higher order terms as possible in order to keep expressions simple).

We will make use of two sets of right-handed orthonormal basis vectors : the gyroangle-independent set ($\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_\parallel$), as introduced in subsection 3.1, and the gyroangle-dependant set ($\hat{\mathbf{a}}$, $\mathbf{e}_\parallel$, $\hat{\mathbf{c}}$), which is defined by $\hat{\mathbf{a}} = \boldsymbol{\rho}_0/\rho_c$ and $\hat{\mathbf{c}} = \hat{\mathbf{a}} \times \mathbf{e}_\parallel$. Note that all these basis vectors depend on $\mathbf{x}$.

By using the relationships $\mathbf{x} = \mathbf{X} + \boldsymbol{\rho}_0$ and $\mathbf{v} = u_\parallel \mathbf{e}_\parallel + v_\perp(\mu)\,\hat{\mathbf{c}} + O(\epsilon_B)$, we perform the transformation to the new coordinates $(\mathbf{X}, u_\parallel, \zeta, \mu)$. This yields

$$\begin{aligned} L_{eq} \;=\; & \left[\frac{1}{\epsilon_B} e_a\, \mathbf{A}_{eq}(\mathbf{X}) \;+\; e_a\, \boldsymbol{\rho}_0 \cdot \boldsymbol{\nabla}\mathbf{A}_{eq} \;+\; m_a u_\parallel \mathbf{e}_\parallel \;+\; m_a v_\perp \hat{\mathbf{c}} \;+\; O(\epsilon_B)\right] \cdot \left[\dot{\mathbf{X}} \;+\; \dot{\boldsymbol{\rho}}_0\right] \\ & -\; \left[e_a \phi_{eq}(\mathbf{X}) \;+\; \epsilon_B \boldsymbol{\rho}_0 \cdot \boldsymbol{\nabla}\phi_{eq} \;+\; \frac{1}{2} m v^2 \;+\; O(\epsilon_B^2)\right] \end{aligned} \tag{93}$$

where factors ϵ_B have been artificially introduced in front of some terms in order to keep track of the ordering of each term. Note that the first term is of order -1 compared to, for example, $m_a v_\perp$, because $e_a A_{eq}/(m_a v_\perp) \sim e_a L_B B_{eq}/(m_a v_\perp) \sim \epsilon_B^{-1}$. Note also that $\dot{\boldsymbol{\rho}}_0$ is of the same order as $\dot{\mathbf{X}}$ because of the fast cyclotron frequency.

Let us now decompose the Lagrangian into terms of successive order in ϵ_B,

$$L_{eq} \;\equiv\; \frac{1}{\epsilon_B} L_{-1} + L_0 + \epsilon_B L_1 + \ldots - (H_0 + \epsilon_B H_1 + \ldots) \tag{94}$$

with

$$L_{-1} \;=\; e_a\, \mathbf{A}_{eq} \cdot \left[\dot{\mathbf{X}} \;+\; \dot{\boldsymbol{\rho}}_0\right] \tag{95}$$

$$L_0 \;=\; e_a (\boldsymbol{\rho}_0 \cdot \boldsymbol{\nabla})\mathbf{A}_{eq} \cdot \left[\dot{\mathbf{X}} \;+\; \dot{\boldsymbol{\rho}}_0\right] \;+\; \left(m_a u_\parallel \mathbf{e}_\parallel \;+\; m_a v_\perp \hat{\mathbf{c}}\right) \cdot \left[\dot{\mathbf{X}} \;+\; \dot{\boldsymbol{\rho}}_0\right] \tag{96}$$

and

$$H_0 \;=\; e_a \phi_{eq} + \frac{1}{2} m_a v_\perp^2 + \frac{1}{2} m_a u_\parallel^2 \tag{97}$$

The Lagrangian is only defined up to an arbitrary exact differential. It simplifies by adding the total time derivative of the scalar $S = S_0 + \epsilon_B S_1 + \ldots$ such that

$$S_0 \;=\; -e_a\, \mathbf{A}_{eq}(\mathbf{X}) \cdot \boldsymbol{\rho}_0 \tag{98}$$

$$S_1 \;=\; -\frac{e_a}{2}\, \boldsymbol{\rho}_0 \cdot \boldsymbol{\nabla}\mathbf{A}_{eq}(\mathbf{X}) \cdot \boldsymbol{\rho}_0 \tag{99}$$

Since $\dot{\mathbf{A}_{eq}} = (\dot{\mathbf{X}} \cdot \boldsymbol{\nabla})\mathbf{A}_{eq}$, the derivatives $\dot{S}_0$ and $\dot{S}_1$ include terms of the general form $[(\mathbf{V_1} \cdot \boldsymbol{\nabla})\mathbf{A}_{eq}]\mathbf{V_2}$, where $\mathbf{V_1}$ and $\mathbf{V_2}$ are two vectors. It is useful to note that

$$[(\mathbf{V_1} \cdot \boldsymbol{\nabla})\mathbf{A}_{eq}] \cdot \mathbf{V_2} - [(\mathbf{V_2} \cdot \boldsymbol{\nabla})\mathbf{A}_{eq}] \cdot \mathbf{V_1} \;=\; (\mathbf{B}_{eq} \times \mathbf{V_1}) \cdot \mathbf{V_2} \tag{100}$$

The derivative $\dot{S}_0$ can be written as

$$\frac{dS_0}{dt} \;=\; -\frac{1}{\epsilon_B} e_a\, \mathbf{A}_{eq} \cdot \dot{\boldsymbol{\rho}}_0 \;-\; e_a\, [(\dot{\mathbf{X}} \cdot \boldsymbol{\nabla})\mathbf{A}_{eq}] \cdot \boldsymbol{\rho}_0 \tag{101}$$

$$\;=\; -\frac{1}{\epsilon_B} e_a\, \mathbf{A}_{eq} \cdot \dot{\boldsymbol{\rho}}_0 \;-\; e_a\, [(\boldsymbol{\rho}_0 \cdot \boldsymbol{\nabla})\mathbf{A}_{eq}] \cdot \dot{\mathbf{X}} \;+\; e_a\, (\mathbf{B}_{eq} \times \rho_0) \cdot \dot{\mathbf{X}} \tag{102}$$

$$\;=\; -\frac{1}{\epsilon_B} e_a\, \mathbf{A}_{eq} \cdot \dot{\boldsymbol{\rho}}_0 \;-\; e_a\, [(\boldsymbol{\rho}_0 \cdot \boldsymbol{\nabla})\mathbf{A}_{eq}] \cdot \dot{\mathbf{X}} - m_a v_\perp \hat{\mathbf{c}} \cdot \dot{\mathbf{X}} \tag{103}$$

The latter expression of $\dot{S}_0$ comprises a term of order -1 which simplifies L_{-1}, and two terms of order 0 which simplify L_0. Similarly, the derivative $\dot{S}_1$ is such that

$$\epsilon_B \frac{dS_1}{dt} = -\frac{e_a}{2}\left[(\boldsymbol{\rho}_0\cdot\boldsymbol{\nabla})\mathbf{A}_{eq}\right]\cdot\dot{\boldsymbol{\rho}}_0 - \frac{e_a}{2}\left[(\dot{\boldsymbol{\rho}}_0\cdot\boldsymbol{\nabla})\mathbf{A}_{eq}\right]\cdot\boldsymbol{\rho}_0 - \epsilon_B\frac{e_a}{2}\left[(\boldsymbol{\rho}_0\cdot\boldsymbol{\nabla})(\dot{\mathbf{X}}\cdot\mathbf{A}_{eq})\right]\cdot\boldsymbol{\rho}_0 \tag{104}$$

$$= -e_a\left[(\boldsymbol{\rho}_0\cdot\boldsymbol{\nabla})\mathbf{A}_{eq}\right]\cdot\dot{\boldsymbol{\rho}}_0 + \frac{e_a}{2}(\mathbf{B}_{eq}\times\boldsymbol{\rho}_0)\cdot\dot{\boldsymbol{\rho}}_0 + O(\epsilon_B) \tag{105}$$

$$= -e_a\left[(\boldsymbol{\rho}_0\cdot\boldsymbol{\nabla})\mathbf{A}_{eq}\right]\cdot\dot{\boldsymbol{\rho}}_0 - \frac{m_a v_\perp}{2}\hat{\mathbf{c}}\cdot\dot{\boldsymbol{\rho}}_0 + O(\epsilon_B) \tag{106}$$

Therefore, $\epsilon_B\dot{S}_1$ comprises two terms of order 0 which simplifies L_0, and one term of order 1 which we neglect.

Finally, we substitute $\dot{\boldsymbol{\rho}}_0 = \rho_c\dot{\zeta}\hat{\mathbf{c}} + O(\epsilon_B)$, $m_a v_\perp\rho_c = 2(m_a/e_a)\mu$. Then the new lagrangian (after adding $\dot{S}$) writes

$$L_{eq} = \left[\frac{1}{\epsilon_B}e_a\,\mathbf{A}_{eq} + m_a u_\parallel\mathbf{e}_\parallel\right]\cdot\dot{\mathbf{X}} + \frac{m_a}{e_a}\mu\dot{\zeta} - \left[e_a\phi_{eq} + \frac{1}{2}m_a u_\parallel^2 + \frac{1}{2}m_a v_\perp^2\right] + O(\epsilon_B) \tag{107}$$

where all quantitites are evaluated at $\mathbf{X}$.

Note that we have assumed that ϕ_{eq} is of higher order than $\mathbf{A}_{eq}\cdot\mathbf{X}$. Alternatively, one can construct a new guiding-center Lagrangian under the assumption that ϕ_{eq} and $\mathbf{A}_{eq}\cdot\mathbf{X}$ are of the same order [25].

C.2 Guiding-center Poisson brackets

In this subsection we apply the procedure described in appendix B.1 to the unperturbed Lagrangian in guiding-center variables to obtain guiding-center Poisson brackets (and check that we recover the expected guiding-center equations of motion).

We work in non-extended phase-space $\mathbf{Z} = \left(\mathbf{X}, u_\parallel, \mu, \varsigma\right)$. We rewrite the Lagrangian obtained in Eq. (107), keeping only terms up to the zeroth order in ϵ_B,

$$\mathcal{L}_{eq} = e_a\,\mathbf{A}^*_{eq}\cdot\dot{\mathbf{X}} + \frac{m_a}{e_a}\mu\dot{\zeta} - H_{eq} \tag{108}$$

where $\mathbf{A}^*_{eq} = \mathbf{A}_{eq} + (m_a/e_a)u_\parallel\mathbf{e}_\parallel$ and

$$H_{eq} = e_a\phi_{eq} + \frac{1}{2}m_a u_\parallel^2 + \mu B_{eq} \tag{109}$$

Since Eq. (108) is already in the form of Eq. (77), it only remains to identify the components of $\Gamma_\mu = \left(e_a\mathbf{A}^*_{eq}, 0, 0, m_a\mu/e_a\right)$,

Calculating the components of the Lagrange matrix, $\omega_{\mu\nu} = \frac{\partial\Gamma_\mu}{\partial Z^\nu} - \frac{\partial\Gamma_\nu}{\partial Z^\mu}$, is straightforward. Defining $\mathbf{B}^*_{\mathbf{eq}} = \boldsymbol{\nabla}\times\mathbf{A}^*_{eq}$, one obtains the Lagrange matrix,

$$\omega_{\mu\nu} = \begin{pmatrix} 0 & e_a B^{*\,3}_{eq} & -e_a B^{*\,2}_{eq} & -m_a e^1_\parallel & 0 & 0 \\ -e_a B^{*\,3}_{eq} & 0 & e_a B^{*\,1}_{eq} & -m_a e^2_\parallel & 0 & 0 \\ e_a B^{*\,2}_{eq} & -e_a B^{*\,1}_{eq} & 0 & -m_a e^3_\parallel & 0 & 0 \\ m_a e^1_\parallel & m_a e^2_\parallel & m_a e^3_\parallel & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{m_a}{e_a} \\ 0 & 0 & 0 & 0 & -\frac{m_a}{e_a} & 0 \end{pmatrix} \tag{110}$$

This matrix can be inverted by blocks (a 4x4 block, a 2x2 block, and two zero-valued

blocks). This procedure yields the Poisson matrix,

$$J^{\mu\nu} = \begin{pmatrix} 0 & -\frac{1}{e_a}\frac{e_\parallel^3}{B_\parallel^*} & \frac{1}{e_a}\frac{e_\parallel^2}{B_\parallel^*} & -\frac{1}{m_a}\frac{B_{eq}^{*\,1}}{B_\parallel^*} & 0 & 0 \\ \frac{1}{e_a}\frac{e_\parallel^3}{B_\parallel^*} & 0 & -\frac{1}{e_a}\frac{e_\parallel^1}{B_\parallel^*} & \frac{1}{m_a}\frac{B_{eq}^{*\,2}}{B_\parallel^*} & 0 & 0 \\ -\frac{1}{e_a}\frac{e_\parallel^2}{B_\parallel^*} & \frac{1}{e_a}\frac{e_\parallel^1}{B_\parallel^*} & 0 & \frac{1}{m_a}\frac{B_{eq}^{*\,3}}{B_\parallel^*} & 0 & 0 \\ -\frac{1}{m_a}\frac{B_{eq}^{*\,1}}{B_\parallel^*} & -\frac{1}{m_a}\frac{B_{eq}^{*\,2}}{B_\parallel^*} & -\frac{1}{m_a}\frac{B_{eq}^{*\,3}}{B_\parallel^*} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{e_a}{m_a} \\ 0 & 0 & 0 & 0 & \frac{e_a}{m_a} & 0 \end{pmatrix} \tag{111}$$

where $B_\parallel^* = \mathbf{e}_\parallel \cdot \mathbf{B}_{\mathbf{eq}}^*$.

This yields the expression of the equilibrium guiding-center Poisson brackets,

$$\begin{aligned} \{F,G\}_{\mathbf{Z}} &= \frac{e_a}{m_a}\left(\frac{\partial F}{\partial \varsigma}\frac{\partial G}{\partial \mu} - \frac{\partial F}{\partial \mu}\frac{\partial G}{\partial \varsigma}\right) \\ &\quad + \frac{\mathbf{B}_{\mathbf{eq}}^*}{mB_\parallel^*}\cdot\left(\frac{\partial F}{\partial \mathbf{X}}\frac{\partial G}{\partial u_\parallel} - \frac{\partial F}{\partial u_\parallel}\frac{\partial G}{\partial \mathbf{X}}\right) \\ &\quad - \frac{\mathbf{e}_\parallel}{e_a B_\parallel^*}\cdot\left(\frac{\partial F}{\partial \mathbf{X}}\times\frac{\partial G}{\partial \mathbf{X}}\right) \end{aligned} \tag{112}$$

which can be used to compute the guiding-center equations of motion,

$$\dot{\mathbf{X}} = u_\parallel \frac{\mathbf{B}_{\mathbf{eq}}^*}{B_\parallel^*} + \frac{\mathbf{e}_\parallel}{e_a B_\parallel^*}\times\left[e_a\boldsymbol{\nabla}\phi_{eq} + \mu\boldsymbol{\nabla}B_{eq}\right] \tag{113}$$

$$\dot{u_\parallel} = \frac{\mathbf{B}_{\mathbf{eq}}^*}{m_a B_\parallel^*}\times\left[e_a\boldsymbol{\nabla}\phi_{eq} + \mu\boldsymbol{\nabla}B_{eq}\right] \tag{114}$$

$$\dot{\mu} = -\frac{e_a}{m_a}\frac{\partial H_{eq}}{\partial \varsigma} = 0 \tag{115}$$

$$\dot{\varsigma} = \frac{e_a B_{eq}}{m_a} = \omega_c \tag{116}$$

Note that Eq. (113) includes not only the $\mathbf{E}\times\mathbf{B}$ drift (of the form $\mathbf{e}_\parallel\times\boldsymbol{\nabla}\phi_{eq}$) and the $\boldsymbol{\nabla}B$ drift, but also the curvature drift, although in this form it is slightly hidden in the term $u_\parallel\mathbf{B}_{\mathbf{eq}}^*/B_\parallel^*$. Indeed, a Taylor expansion at the lowest order in ϵ_B yields

$$\begin{aligned} \frac{\mathbf{B}_{\mathbf{eq}}^*}{B_\parallel^*} &\approx \mathbf{e}_\parallel + \frac{m_a u_\parallel}{e_a B_{eq}}\left[\boldsymbol{\nabla}\times\mathbf{e}_\parallel\right]_\perp \\ &= \mathbf{e}_\parallel - \frac{m_a u_\parallel}{e_a B_{eq}}\,\mathbf{e}_\parallel\times\left[\frac{1}{2}\boldsymbol{\nabla}\mathbf{e}_\parallel^2 - (\mathbf{e}_\parallel\cdot\boldsymbol{\nabla})\mathbf{e}_\parallel\right]_\perp \\ &= \mathbf{e}_\parallel + \frac{m_a u_\parallel}{e_a B_{eq}}\,\mathbf{e}_\parallel\times\boldsymbol{\nabla}_\parallel\mathbf{e}_\parallel \end{aligned}$$

D Some useful relations

D.1 Contact transformation

We derive here the change of variables that correspond to a contact transformation in the simple case where $\boldsymbol{\mathfrak{z}}$ is a set of conjugate variables $(\mathbf{x},\mathbf{p})$. The new coordinates are $(\bar{\mathbf{x}},\bar{\mathbf{p}})$.

The old and new Hamiltonians $\mathcal{H}$ and $\bar{\mathcal{H}}$ read

$$\begin{aligned}\mathcal{H}(\mathbf{x},\mathbf{p},\tau) &= \mathcal{H}_{eq}(\mathbf{p}) + \epsilon_\delta \mathcal{H}_1(\mathbf{x},\mathbf{p},\tau)\\ \bar{\mathcal{H}}(\bar{\mathbf{x}},\bar{\mathbf{p}},\tau) &= \bar{\mathcal{H}}_{eq}(\bar{\mathbf{x}},\bar{\mathbf{p}},\tau) + \epsilon_\delta \bar{\mathcal{H}}_1(\bar{\mathbf{x}},\bar{\mathbf{p}},\tau)\end{aligned}$$

Since the amplitude of the perturbed Hamiltonian is small, the change of coordinates can be anticipated to be a near-identity transform. The change of canonical coordinates is performed with a methodology based on a least action principle [8, 9], i.e., the action integrals are made equal up to a generating function,

$$\int \mathbf{p}\cdot d\mathbf{x} = \int \bar{\mathbf{p}}\cdot d\bar{\mathbf{x}} - \int dS$$

The generating function S is a function of both new and old variables, for instance

$$S(\bar{\mathbf{x}},\bar{\mathbf{p}},\bar{\mathbf{x}},\bar{\mathbf{p}}) = \epsilon_\delta S_1(\mathbf{x},\bar{\mathbf{p}}) + (\mathbf{x}-\bar{\mathbf{x}})\cdot\bar{\mathbf{p}}$$

Identification of various terms in front of the differential elements $d\mathbf{x}$ and $d\bar{\mathbf{p}}$ provides the requested near identity transform

$$\begin{aligned}p_i &= \bar{p}_i - \epsilon_\delta \frac{\partial S_1(\mathbf{x},\bar{\mathbf{p}})}{\partial x^i} + o(\epsilon_\delta^2)\\ \bar{x}^i &= x^i - \epsilon_\delta \frac{\partial S_1(\mathbf{x},\bar{\mathbf{p}})}{\partial \bar{p}_i} + o(\epsilon_\delta^2)\end{aligned}$$

This relationship provides the expression of $(\bar{\mathbf{x}},\mathbf{p})$, a mix of old and new variables, versus $(\mathbf{x},\bar{\mathbf{p}})$. Moreover one has the relation

$$\bar{\mathcal{H}}(\bar{\mathbf{x}},\bar{\mathbf{p}}) = \mathcal{H}(\mathbf{x},\mathbf{p}) \tag{117}$$

which related the new Hamiltonian to the previous one[47]. One difficulty comes from the mix of new and old variables - a relationship between old and new variables would be easier to handle. This difficulty is easily overcome since at requested order

$$\begin{aligned}p_i &= \bar{p}_i - \epsilon_\delta \frac{\partial S_1(\bar{\mathbf{x}},\bar{\mathbf{p}})}{\partial \bar{x}^i} + o(\epsilon_\delta^2)\\ x^i &= \bar{x}^i + \epsilon_\delta \frac{\partial S_1(\bar{\mathbf{x}},\bar{\mathbf{p}})}{\partial \bar{p}_i} + o(\epsilon_\delta^2)\end{aligned} \tag{118}$$

Note that this relation can be written in a compact form as

$$\boldsymbol{\mathfrak{z}} = \bar{\boldsymbol{\mathfrak{z}}} - \epsilon_\delta \{S_1, \bar{\boldsymbol{\mathfrak{z}}}\}_{\bar{\boldsymbol{\mathfrak{z}}}} + o(\epsilon_\delta^2)$$

where $\{\ ,\ \}_{\bar{\boldsymbol{\mathfrak{z}}}}$ is the Poisson bracket written in the new variables $\bar{\boldsymbol{\mathfrak{z}}} = (\bar{\mathbf{x}},\bar{\mathbf{p}})$, i.e.

$$\{F,G\}_{\bar{\boldsymbol{\mathfrak{z}}}} = \frac{\partial F}{\partial \bar{\mathbf{x}}}\cdot\frac{\partial G}{\partial \bar{\mathbf{p}}} - \frac{\partial F}{\partial \bar{\mathbf{p}}}\cdot\frac{\partial G}{\partial \bar{\mathbf{x}}} = \frac{\partial F}{\partial \bar{\boldsymbol{\mathfrak{z}}}}\cdot \mathtt{J}\cdot\frac{\partial G}{\partial \bar{\boldsymbol{\mathfrak{z}}}}$$

D.2 Invariance of Poisson bracket

Poisson brackets are invariant under a change of variables. The sequence of operations that demonstrate this property is the following

$$\begin{aligned}\{\mathcal{H},\mathcal{F}\}_{\boldsymbol{\mathcal{Z}}} &= \frac{\partial \mathcal{H}}{\partial \mathcal{Z}^k}\mathtt{J}^{kl}\frac{\partial \mathcal{F}}{\partial \mathcal{Z}^l}\\ &= \frac{\partial \bar{\mathcal{H}}}{\partial \bar{\mathcal{Z}}^i}\left(\frac{\partial \bar{\mathcal{Z}}^i}{\partial \mathcal{Z}^k}\mathtt{J}^{kl}\frac{\partial \bar{\mathcal{Z}}^j}{\partial \mathcal{Z}^l}\right)\frac{\partial \bar{\mathcal{F}}}{\partial \bar{\mathcal{Z}}^j}\\ &= \frac{\partial \bar{\mathcal{H}}}{\partial \bar{\mathcal{Z}}^i}\bar{\mathtt{J}}^{ij}\frac{\partial \bar{\mathcal{F}}}{\partial \bar{\mathcal{Z}}^j}\\ &= \{\bar{\mathcal{H}},\bar{\mathcal{F}}\}_{\bar{\boldsymbol{\mathcal{Z}}}}\end{aligned}$$

[47] The index $i=0$ yields the usual evolution equation for the generating function $\epsilon_\delta\frac{\partial S_1(\mathbf{x},\bar{\mathbf{p}},t)}{\partial t} = H(\mathbf{x},\mathbf{p},t) - \bar{H}(\bar{\mathbf{x}},\bar{\mathbf{p}},t)$, where $(\bar{\mathbf{x}},\mathbf{p})$ are functions of $(\mathbf{x},\bar{\mathbf{p}})$, and are now restricted to the usual position and momentum coordinates. The order 0 imposes $\bar{H}_{eq}(\bar{\mathbf{x}},\bar{\mathbf{p}}) = H_{eq}(\mathbf{x},\mathbf{p})$.

The identity $\frac{\partial \bar{\mathcal{Z}}^i}{\partial \mathcal{Z}^k} \mathtt{J}^{kl} \frac{\partial \bar{\mathcal{Z}}^j}{\partial \mathcal{Z}^l} = \bar{\mathtt{J}}^{ij}$ results from the rules of change of variables for a tensor. The new tensor $\bar{\mathtt{J}}^{ij}$ bears the properties of a Poisson bracket for a change of variables of the form $\boldsymbol{\mathcal{Z}} = \bar{\boldsymbol{\mathcal{Z}}} - \{S_1, \bar{\boldsymbol{\mathcal{Z}}}\}_{\bar{\boldsymbol{\mathcal{Z}}}}$.

D.3 Push-forward and pull-back operators

The operator T that transforms $\boldsymbol{\mathcal{Z}}$ in $\bar{\boldsymbol{\mathcal{Z}}}$ is called a push-forward operator, $\bar{\boldsymbol{\mathcal{Z}}} = T\boldsymbol{\mathcal{Z}}$ - here $T = I + \{S, \bullet\}$, I being the identity operator. A "push-forward" operator leads to new coordinates $\bar{\boldsymbol{\mathcal{Z}}}$ from the old ones $\boldsymbol{\mathcal{Z}}$. One can then define a space of functions that operate on $\boldsymbol{\mathcal{Z}}$ variables, and a space of functions that act on the $\bar{\boldsymbol{\mathcal{Z}}}$ variables. The function $\mathcal{F}$ is derived from $\bar{\mathcal{F}}$ via a pull-back operator $\mathcal{T}^*$, i.e. $\mathcal{F} = \mathcal{T}^* \bar{\mathcal{F}}$. It transforms a function that acts in the space of the gyrocentre coordinates into a function that lives in the initial guiding-centre space. Hence the relation $\bar{\mathcal{F}}(\bar{\boldsymbol{\mathcal{Z}}}) = \mathcal{F}(\boldsymbol{\mathcal{Z}})$ can as well be written as $\mathcal{T}^* \bar{\mathcal{F}}(\boldsymbol{\mathcal{Z}}) = \mathcal{F}(\boldsymbol{\mathcal{Z}}) = \bar{\mathcal{F}}(T\boldsymbol{\mathcal{Z}})$ (see Fig.24).

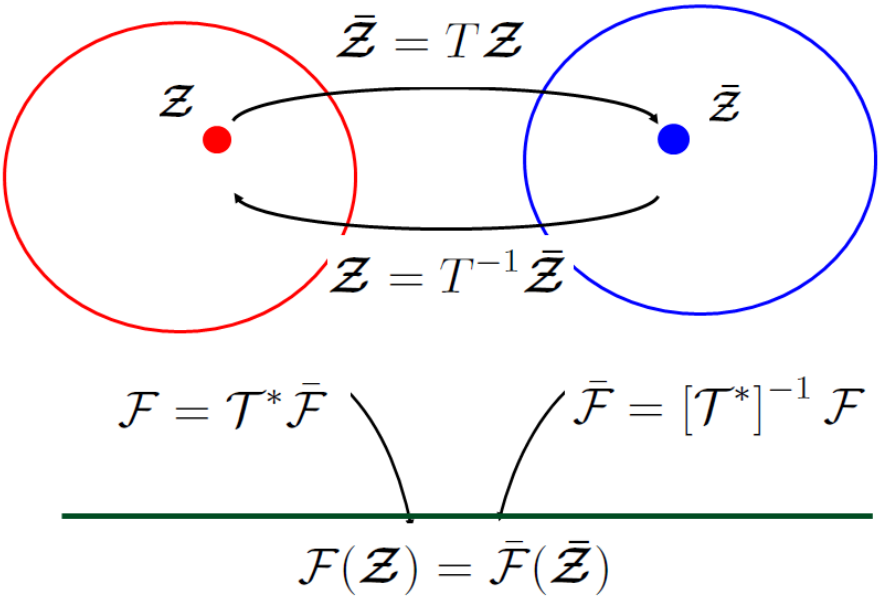


Figure 24: Change of variables $\bar{\boldsymbol{\mathcal{Z}}} = T\boldsymbol{\mathcal{Z}}$ and its inverse $\boldsymbol{\mathcal{Z}} = T^{-1}\bar{\boldsymbol{\mathcal{Z}}}$. The function $\mathcal{F}$ lives in the space of functions that depend on $\boldsymbol{\mathcal{Z}}$, and its counter part leaves in the space of functions that depend on $\bar{\boldsymbol{\mathcal{Z}}}$. They are related by the relationship $\bar{\mathcal{F}}(\bar{\boldsymbol{\mathcal{Z}}}) = \mathcal{F}(\boldsymbol{\mathcal{Z}})$. The function $\mathcal{F}$ is the image of $\bar{\mathcal{F}}$ via a pull-back operator $\mathcal{T}^*$, i.e. $\mathcal{F} = \mathcal{T}^*\bar{\mathcal{F}}$. Inversely $\bar{\mathcal{F}}$ is constructed from $\mathcal{F}$ via a push-forward operator that is the inverse of $\mathcal{T}$, i.e. $\bar{\mathcal{F}} = [\mathcal{T}^*]^{-1}\mathcal{F}$ - inspired from [2].

D.4 Pull-back of the gyroaverage operator

This section details the transformation of the functional

$$\begin{aligned} \mathcal{A}_{part} = & - \int d\Gamma^* \bar{\mathcal{F}}_1(\boldsymbol{\mathcal{Z}}^*) \mathcal{J}[\mathcal{H}_1](\boldsymbol{\mathcal{Z}}^*) \\ & - \frac{1}{2} \int d\Gamma^* \mathcal{F}'_{eq}(\boldsymbol{\mathcal{Z}}^*) \mathcal{J}[\mathcal{H}_1^2](\boldsymbol{\mathcal{Z}}^*) \\ & + \frac{1}{2} \int d\Gamma^* \mathcal{F}'_{eq}(\boldsymbol{\mathcal{Z}}^*) (\mathcal{J}[\mathcal{H}_1])^2(\boldsymbol{\mathcal{Z}}^*) \end{aligned} \tag{119}$$

in a functional that lives in the physical space. The functional is split in 3 contributions $\mathcal{A}_{part} = \mathcal{A}_1 + \mathcal{A}_2 + \mathcal{A}_3$, each term corresponds to one line of the r.h.s. in Eq.(119).

The first term is transformed as follows

$$\begin{aligned}
\mathcal{A}_1 &= -\int d^3\mathbf{X}dtdp_{\parallel}d\mu B^*_{||}(\boldsymbol{\mathcal{Z}}^*)\bar{\mathcal{F}}_1(\boldsymbol{\mathcal{Z}}^*)\mathcal{J}[\mathcal{H}_1](\boldsymbol{\mathcal{Z}}^*) \\
&- \int d^3\mathbf{X}dtdp_{\parallel}d\mu\int_0^{2\pi}\frac{d\varsigma}{2\pi}\left[B^*_{||}\bar{\mathcal{F}}_1\right](\mathbf{X},t,p_{\parallel},\mu)\mathcal{H}_1(\mathbf{X}+\boldsymbol{\rho}_0(\mathbf{X},\mu,\varsigma),t,p_{\parallel},\mu) \\
&= -\int d^3\mathbf{x}dtdp_{\parallel}d\mu\int_0^{2\pi}\frac{d\varsigma}{2\pi}\left[B^*_{||}\bar{\mathcal{F}}_1\right](\mathbf{x}-\boldsymbol{\rho}_0(\mathbf{x},\mu,\varsigma),t,p_{\parallel},\mu)\mathcal{H}_1(\mathbf{x},t,p_{\parallel},\mu) \\
&= -\int d^3\mathbf{x}dtd^3\mathbf{V}\frac{1}{B^*_{||}}\mathcal{J}^t\left\{B^*_{||}\bar{\mathcal{F}}_1\right\}(\mathbf{x},t,p_{\parallel},\mu)\,\mathcal{H}_1\left(\mathbf{x},t,p_{\parallel},\mu\right) \\
&= -\int d\gamma^*\frac{1}{B^*_{||}}\mathcal{J}^t\left\{B^*_{||}\bar{\mathcal{F}}_1\right\}(\boldsymbol{\mathfrak{z}}^*)\,\mathcal{H}_1\left(\boldsymbol{\mathfrak{z}}^*\right) \qquad (120)
\end{aligned}$$

where the operator $\mathcal{J}^t$ is defined as

$$\mathcal{J}^t[\Lambda]\left(\mathbf{x},t,p_{\parallel},\mu\right) = \int_0^{2\pi}\frac{d\varsigma}{2\pi}\Lambda\left(\mathbf{x}-\boldsymbol{\rho},t,p_{\parallel},\mu\right)$$

The second term is transformed a bit differently

$$\begin{aligned}
\mathcal{A}_2 &= -\frac{1}{2}\int d\Gamma^*\mathcal{F}'_{eq}(\boldsymbol{\mathcal{Z}}^*)\mathcal{J}[\mathcal{H}_1^2](\boldsymbol{\mathcal{Z}}^*) \\
&= -\frac{1}{2}\int d^3\mathbf{X}dtdp_{\parallel}d\mu\int_0^{2\pi}\frac{d\varsigma}{2\pi}\left[B^*_{||}\bar{\mathcal{F}}_1\right](\mathbf{X},t,p_{\parallel},\mu)\mathcal{H}_1^2(\mathbf{X}+\boldsymbol{\rho}_0(\mathbf{X},\mu,\varsigma),t,p_{\parallel},\mu) \\
&= -\frac{1}{2}\int d^3\mathbf{x}dtdp_{\parallel}d\mu\int_0^{2\pi}\frac{d\varsigma}{2\pi}\left[B^*_{||}\bar{\mathcal{F}}_1\right](\mathbf{x}-\boldsymbol{\rho}_0(\mathbf{x},\mu,\varsigma),t,p_{\parallel},\mu)\mathcal{H}_1^2(\mathbf{x},t,p_{\parallel},\mu) \\
&= -\frac{1}{2}\int d\gamma^*\frac{1}{B^*_{||}}\mathcal{J}^t\left[B^*_{||}\mathcal{F}'_{eq}\right](\boldsymbol{\mathfrak{z}}^*)\mathcal{H}_1^2(\boldsymbol{\mathfrak{z}}^*)
\end{aligned}$$

The last term follows the same path

$$\begin{aligned}
\mathcal{A}_3 &= \frac{1}{2}\int d\Gamma^*\mathcal{F}'_{eq}(\boldsymbol{\mathcal{Z}}^*)\mathcal{J}[\mathcal{H}_1](\boldsymbol{\mathcal{Z}}^*)\mathcal{J}[\mathcal{H}_1](\boldsymbol{\mathcal{Z}}^*) \\
&= \frac{1}{2}\int d^3\mathbf{X}dtdp_{\parallel}d\mu\int_0^{2\pi}\frac{d\varsigma}{2\pi}\left[B^*_{||}\mathcal{F}'_{eq}\mathcal{J}[\mathcal{H}_1]\right](\mathbf{X},t,p_{\parallel},\mu) \\
&\quad \mathcal{H}_1(\mathbf{X}+\boldsymbol{\rho}_0(\mathbf{X},\mu,\varsigma),t,p_{\parallel},\mu) \\
&= \frac{1}{2}\int d^3\mathbf{x}dtdp_{\parallel}d\mu\mathcal{H}_1(\mathbf{x},t,p_{\parallel},\mu) \\
&\quad \int_0^{2\pi}\frac{d\varsigma}{2\pi}\left[B^*_{||}\mathcal{F}'_{eq}\mathcal{J}[\mathcal{H}_1]\right](\mathbf{X}-\boldsymbol{\rho}_0(\mathbf{x},\mu,\varsigma),t,p_{\parallel},\mu) \\
&= \frac{1}{2}\int d\gamma^*\frac{1}{B^*_{||}(\boldsymbol{\mathfrak{z}}^*)}\mathcal{J}^t\left[B^*_{||}\mathcal{F}'_{eq}\mathcal{J}[\mathcal{H}_1]\right](\boldsymbol{\mathfrak{z}}^*)\mathcal{H}_1(\boldsymbol{\mathfrak{z}}^*)
\end{aligned}$$

D.5 Long wavelength expansion of the gyroaverage operator

This derivation does not raise special difficulties though somewhat cumbersome. Here are some details:

$$\begin{aligned}
\mathcal{J}^t\left[\mathcal{F}'_{eq}\mathcal{J}[\mathcal{H}_1]\right](\mathbf{X}) &= \int \frac{d\varsigma}{2\pi}\frac{d\varsigma'}{2\pi}\mathcal{F}'_{eq}(\mathbf{x}-\boldsymbol{\rho}')\mathcal{H}_1(\mathbf{x}+\boldsymbol{\rho}-\boldsymbol{\rho}') \\
&= \int \frac{d\varsigma}{2\pi}\frac{d\varsigma'}{2\pi}\Big[\mathcal{F}'_{eq}(\mathbf{x}) - \boldsymbol{\rho}'\cdot\nabla\mathcal{F}'_{eq}\big|_{\mathbf{x}} \\
&\quad +\frac{1}{2}\,\boldsymbol{\rho}'\cdot\nabla\left(\boldsymbol{\rho}'\cdot\nabla\mathcal{F}'_{eq}\right)\big|_{\mathbf{x}}\Big] \\
&\quad \big[\mathcal{H}_1(\mathbf{x}) + \left(\boldsymbol{\rho}-\boldsymbol{\rho}'\right)\cdot\nabla\mathcal{H}_1\big|_{\mathbf{x}} \\
&\quad +\frac{1}{2}\left(\boldsymbol{\rho}-\boldsymbol{\rho}'\right)\cdot\nabla\left[\left(\boldsymbol{\rho}-\boldsymbol{\rho}'\right)\cdot\nabla\mathcal{H}_1\big|_{\mathbf{x}}\right]\Big] \\
&= \left(\mathcal{F}'_{eq}(\mathbf{x}) + \frac{1}{4}\rho_c^2\nabla_\perp^2\mathcal{F}'_{eq}(\mathbf{x})\right)\mathcal{H}_1(\mathbf{x}) \\
&\quad +\frac{1}{2}\nabla\cdot\left[\rho_c^2\mathcal{F}'_{eq}(\mathbf{x})\nabla_\perp\mathcal{H}_1(\mathbf{x})\right]
\end{aligned}$$

where the vectors $\boldsymbol{\rho}$ and $\boldsymbol{\rho}'$ are the cyclotron displacements that correspond respectively to the cyclotron angles ς and ς'. The properties $\int \frac{d\varsigma}{2\pi}\rho_i\rho_j = \int \frac{d\varsigma'}{2\pi}\rho'_i\rho'_j = \frac{1}{2}\rho_c^2\delta_{ij}$, $\int \frac{d\varsigma}{2\pi}\rho_i = \int \frac{d\varsigma'}{2\pi}\rho'_j = 0$ have been used, where i,j are labels of directions perpendicular to the field, and $\rho_c^2 = \frac{2\mu}{e_a\Omega_c(\mathbf{x})}$. The square of the gyroradius ρ_c^2 has been permuted with the operator ∇ since its derivative is of order ϵ_B compared with the perturbed field derivatives. The last term is a polarisation term since for a Maxwellian distribution function, the following identity holds $\int d^3\mathbf{V}\frac{1}{2}\rho_c^2\mathcal{F}'_{eq} = \frac{N_{eq}(\mathbf{x})m_a}{e_a^2B^2(\mathbf{x})}$.

D.6 From guiding-centre to gyrocentre coordinates

The change of variables in the Hamiltonian approach is given by Eq.(14), i.e. is of the form $\bar{\boldsymbol{\mathcal{Z}}} = \boldsymbol{\mathcal{Z}} + \mathbf{G}_1$, with $\mathbf{G}_1 = \{S_1, \boldsymbol{\mathcal{Z}}\}_{\boldsymbol{\mathcal{Z}}}$. The components of $\mathbf{G}_1$ can be made explicit by using the expression of the Poisson bracket in guiding-centre coordinates

$$\begin{aligned}
\{F,G\}_{\mathbf{Z}} &= \frac{e_a}{m_a}\left(\frac{\partial F}{\partial\varsigma}\frac{\partial G}{\partial\mu} - \frac{\partial F}{\partial\mu}\frac{\partial G}{\partial\varsigma}\right) \\
&\quad +\frac{1}{m_a}\frac{\mathbf{B}^*}{B^*_{||}}\cdot\left(\frac{\partial F}{\partial\mathbf{X}}\frac{\partial G}{\partial p_{||}} - \frac{\partial F}{\partial p_{||}}\frac{\partial G}{\partial\mathbf{X}}\right) \\
&\quad -\frac{\mathbf{e}_{||}}{e_aB^*_{||}}\cdot\left(\frac{\partial F}{\partial\mathbf{X}}\times\frac{\partial G}{\partial\mathbf{X}}\right)
\end{aligned}$$

where

$$\mathbf{B}^* = \nabla\times\mathbf{A}^* = \mathbf{B}_{eq} + \frac{m_ap_{||}}{e_a}\nabla\times\mathbf{e}_{||}$$

and

$$B^*_{||} = \mathbf{e}_{||}\cdot\mathbf{B}^* = B_{eq} + \frac{m_ap_{||}}{e_a}\mathbf{e}_{||}\cdot\nabla\times\mathbf{e}_{||}$$

is the Jacobian of the guiding-centre transform. The following expressions are then found

$$\begin{aligned} G_1^{\mathbf{X}} &= \{S_1, \mathbf{X}\}_{\boldsymbol{\mathcal{Z}}} = -\frac{1}{m_a}\frac{\partial S_1}{\partial p_\parallel}\mathbf{e}_\parallel - \frac{\mathbf{e}_\parallel}{e_a B_\parallel^*} \times \frac{\partial S_1}{\partial \mathbf{X}} \\ G_1^{p_\parallel} &= \left\{S_1, p_\parallel\right\}_{\boldsymbol{\mathcal{Z}}} = \frac{1}{m_a}\frac{\mathbf{B}^*}{B_\parallel^*} \cdot \frac{\partial S_1}{\partial \mathbf{X}} \\ G_1^{\mu} &= \{S_1, \mu\}_{\boldsymbol{\mathcal{Z}}} = \frac{e_a}{m_a}\frac{\partial S_1}{\partial \varsigma} \\ G_1^{\varsigma} &= \{S_1, \varsigma\}_{\boldsymbol{\mathcal{Z}}} = -\frac{e_a}{m_a}\frac{\partial S_1}{\partial \mu} \end{aligned}$$

Since S_1 is a periodic function of ς with zero mean, these displacements are all periodic functions of ς. However, whenever one calculates the particle position vs the gyrocentre coordinates, expressed in guiding-centre coordinates,

$$\mathbf{x} = \mathbf{X} + \boldsymbol{\rho}\left(\mathbf{X}, \mu, \varsigma\right) = \bar{\mathbf{X}} - G_1^{\mathbf{X}}\left(\bar{\boldsymbol{\mathcal{Z}}}\right) - \frac{\partial \boldsymbol{\rho}}{\partial \mathbf{X}} \cdot G_1^{\mathbf{X}}\left(\bar{\boldsymbol{\mathcal{Z}}}\right) - \frac{\partial \boldsymbol{\rho}}{\partial \mu} G_1^{\mu}\left(\bar{\boldsymbol{\mathcal{Z}}}\right) - \frac{\partial \boldsymbol{\rho}}{\partial \varsigma} G_1^{\varsigma}\left(\bar{\boldsymbol{\mathcal{Z}}}\right)$$

terms with a finite gyroaverage mean (average over the gyroangle $\bar{\varsigma}$) appear due to the dependences of $\mathbf{G}_1$ on $\bar{\varsigma}$ (at given $(\bar{\mu}, \bar{\mathbf{X}}, \bar{p}_\parallel)$). The resulting mean displacement is responsible for the medium polarisation.

E Gyroaverage of the perturbed Hamiltonian

This section provides practical expressions for the gyroaverage of the perturbed Hamiltonian. From Eq.(11), the Hamiltonian associated with a perturbed electromagnetic field reads

$$\begin{aligned} \widetilde{\mathcal{H}}\left(\mathbf{X}, p_\parallel, \mu, \varsigma, t\right) &= e_a \tilde{\phi}(\mathbf{X} + \boldsymbol{\rho}_0, t) - e_a p_\parallel \tilde{A}_\parallel(\mathbf{X} + \boldsymbol{\rho}_0, t) \\ &- e_a \left(\dot{\mathbf{X}}_\perp + \dot{\boldsymbol{\rho}}_0\right) \cdot \tilde{\mathbf{A}}_\perp(\mathbf{X} + \boldsymbol{\rho}_0, t) + \frac{1}{2}\frac{e_a}{m_a}\tilde{\mathbf{A}}^2(\mathbf{X} + \boldsymbol{\rho}_0, t) \end{aligned}$$

The gyroaverage of $\widetilde{\mathcal{H}}(\boldsymbol{\mathcal{Z}}^*, \varsigma)$ is

$$\mathcal{J}[\widetilde{\mathcal{H}}](\boldsymbol{\mathcal{Z}}^*) = \int \frac{d\varsigma}{2\pi} \widetilde{\mathcal{H}}(\boldsymbol{\mathcal{Z}}^*, \varsigma)$$

where $\boldsymbol{\mathcal{Z}}^* = (\mathbf{X}, p_\parallel, \mu, t, E)$. The perturbed electric and vector potentials can be expanded in Fourier series

$$\tilde{\phi}\left(\mathbf{x}, t\right) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \tilde{\phi}\left(\boldsymbol{k}, t\right) e^{i\boldsymbol{k}\cdot\mathbf{x}}$$

and

$$\tilde{\mathbf{A}}\left(\mathbf{x}, t\right) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \tilde{\mathbf{A}}\left(\boldsymbol{k}, t\right) e^{i\boldsymbol{k}\cdot\mathbf{x}}$$

The cyclotron motion displacement and velocity read

$$\begin{aligned} \boldsymbol{\rho}_0(t) &= \rho_c\left(\cos\varsigma\, \mathbf{e}_1 - \sin\varsigma\, \mathbf{e}_2\right) \\ \dot{\boldsymbol{\rho}}_0(t) &= v_\perp\left(-\sin\varsigma\, \mathbf{e}_1 - \cos\varsigma\, \mathbf{e}_2\right) \end{aligned}$$

where $\mathbf{e}_1\left(\mathbf{X}\right), \mathbf{e}_2\left(\mathbf{X}\right)$ are two unit vectors locally orthogonal to the magnetic field $\mathbf{B}\left(\mathbf{X}, t\right)$. The wave vector $\boldsymbol{k}$ can be decomposed on the local basis $\mathbf{e}_\parallel, \mathbf{e}_1, \mathbf{e}_2$, i.e.

$$\boldsymbol{k} = k_\parallel \mathbf{e}_\parallel + k_1 \mathbf{e}_1 + k_2 \mathbf{e}_2$$

where all quantities, components and vectors, depend on the guiding-centre position $\mathbf{X}$. Using $\mathbf{x} = \mathbf{X} + \boldsymbol{\rho}_0$, it appears readily that

$$\mathcal{J}[\tilde{\phi}]\left(\boldsymbol{\mathcal{Z}}^*\right) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \tilde{\phi}\left(\boldsymbol{k}, t\right) e^{i\boldsymbol{k}\cdot\mathbf{X}} \int_0^{2\pi} \frac{d\varsigma}{2\pi} e^{i\boldsymbol{k}\cdot\boldsymbol{\rho}_0}$$

and

$$\mathcal{J}[\tilde{A}_\parallel]\left(\boldsymbol{\mathcal{Z}}^*\right) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \tilde{A}_\parallel\left(\boldsymbol{k}, t\right) e^{i\boldsymbol{k}\cdot\mathbf{X}} \int_0^{2\pi} \frac{d\varsigma}{2\pi} e^{i\boldsymbol{k}\cdot\boldsymbol{\rho}_0}$$

The components of the wave vector can be written

$$k_1 = k_\perp \sin(\varphi_k) \quad ; \quad k_2 = -k_\perp \cos(\varphi_k)$$

where $k_\perp = (k_1^2 + k_2^2)^{1/2}$ is usually called "perpendicular wave number", and φ_k is the wave number phase. It then appears that $\boldsymbol{k}\cdot\boldsymbol{\rho}_0 = k_\perp \rho_c \sin(\varsigma + \varphi_k)$, and

$$\mathcal{J}[\tilde{\phi}]\left(\boldsymbol{\mathcal{Z}}^*\right) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} J_0\left(k_\perp \rho_c\right) \tilde{\phi}\left(\boldsymbol{k}, t\right) e^{i\boldsymbol{k}\cdot\mathbf{X}}$$

and

$$\mathcal{J}[\tilde{A}_\parallel]\left(\boldsymbol{\mathcal{Z}}^*\right) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} J_0\left(k_\perp \rho_c\right) \tilde{A}_\parallel\left(\boldsymbol{k}, t\right) e^{i\boldsymbol{k}\cdot\mathbf{X}}$$

where J_0 is the Bessel function of index 0. Though approximate, this is a quite convenient explicit expression of the gyroaverage operator, since it consists in a simple multiplication by $J_0\left(k_\perp \rho_c\right)$ in the Fourier space. This is why the operator $\mathcal{J}$ is sometimes noted J_0. We move now to the piece that contains the perpendicular vector potential

$$\left\langle \left(\dot{\mathbf{X}}_\perp + \dot{\boldsymbol{\rho}}_0\right) \cdot \tilde{\mathbf{A}}_\perp(\mathbf{X} + \boldsymbol{\rho}_0, t) \right\rangle_\varsigma = \left\langle \dot{\mathbf{X}}_\perp \cdot \tilde{\mathbf{A}}_\perp(\mathbf{X} + \boldsymbol{\rho}_0, t) \right\rangle_\varsigma + \left\langle \dot{\boldsymbol{\rho}}_0 \cdot \tilde{\mathbf{A}}_\perp(\mathbf{X} + \boldsymbol{\rho}_0, t) \right\rangle_\varsigma$$

The first term in the r.h.s. is gyro-averaged like the electric and parallel vector potentials. The second bit is more delicate to handle. Let us detail it

$$\begin{aligned}
e_a \left\langle \dot{\boldsymbol{\rho}}_0 \cdot \tilde{\mathbf{A}}_\perp(\mathbf{X} + \boldsymbol{\rho}_0, t) \right\rangle_\varsigma &= -e_a v_\perp \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} e^{i\boldsymbol{k}\cdot\mathbf{X}} \int_0^{2\pi} \frac{d\varsigma}{2\pi} \underbrace{\left\{\sin\varsigma A_{1\mathbf{k}} + \cos\varsigma A_{2\mathbf{k}}\right\}}_{e^{i(k_1\rho_c\cos\varsigma - \sin\varsigma k_2\rho_c)}} \\
&= -e_a\Omega_c \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} e^{i\boldsymbol{k}\cdot\mathbf{X}} \left\{ iA_{1\mathbf{k}} \frac{\partial}{\partial k_2} - iA_{2\mathbf{k}} \frac{\partial}{\partial k_1} \right\} \underbrace{\int_0^{2\pi} \frac{d\varsigma}{2\pi}}_{e^{i(k_1\rho_c\cos\varsigma - \sin\varsigma k_2\rho_c)}} \\
&= -e_a\Omega_c \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} e^{i\boldsymbol{k}\cdot\mathbf{X}} \left\{ iA_{1\mathbf{k}} \frac{\partial}{\partial k_2} - iA_{2\mathbf{k}} \frac{\partial}{\partial k_1} \right\} J_0\left(k_\perp \rho_c\right) \\
&= -e_a\Omega_c\rho_c^2 \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} e^{i\boldsymbol{k}\cdot\mathbf{X}} \left\{ -ik_2 A_{1\mathbf{k}} + ik_1 A_{2\mathbf{k}} \right\} \frac{J_1\left(k_\perp \rho_c\right)}{k_\perp \rho_c} \\
&= -\mu \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} e^{i\boldsymbol{k}\cdot\mathbf{X}} \left(\mathbf{e}_\parallel \cdot \nabla \times \tilde{\mathbf{A}}_\perp\right)_{\mathbf{k}} \frac{2J_1\left(k_\perp \rho_c\right)}{k_\perp \rho_c} \\
&= -\mu\mathcal{B}\left[\tilde{B}_\parallel\right]
\end{aligned}$$

where $(A_{\mathbf{k}1}, A_{\mathbf{k}2})$ are the two components of the vector $A_{\mathbf{k}}$ in the vector basis $(\mathbf{e}_1, \mathbf{e}_2)$. The scalar $\mathbf{e}_\parallel \cdot \nabla \times \tilde{\mathbf{A}}_\perp$ is the parallel component of the perturbed magnetic field, noted $\tilde{B}_\parallel$, and $\mathcal{B}$ is the operator which consists in multiplying the Fourier component of the field by $2J_1\left(k_\perp \rho_c\right)/(k_\perp \rho_c)$. Let us note that $2J_1(x)/x$ approaches 1 when the argument x gets close to 0. Hence the operator $\mathcal{B}$ is the identity for perturbations with scale lengths much larger than the gyroradius, as does $\mathcal{J}$. Nevertheless its structure differs significantly from $\mathcal{J}$ applied on a scalar field. So finally the gyroaverage of the perturbed Hamiltonian reads

$$\mathcal{J}[\widetilde{\mathcal{H}}] = e_a\mathcal{J}[\tilde{\phi}] - e_a p_\parallel \mathcal{J}[\tilde{A}_\parallel] + \mu\mathcal{B}[\tilde{B}_\parallel] - e_a\dot{\mathbf{X}}_\perp \cdot \mathcal{J}[\tilde{\mathbf{A}}_\perp] + \frac{e_a^2}{2m_a}\mathcal{J}[\tilde{\mathbf{A}}^2]$$

The two last terms are of second order in the expansion parameter.
Finally, the same methodology can be used to estimate of the polarisation shift in the $(\tilde{\phi}, \tilde{A}_\parallel)$ representation

$$\langle \bar{\boldsymbol{\rho}}_1 \rangle_{\bar{\varsigma}} = -\frac{1}{B(\bar{\mathbf{X}})} \frac{\partial}{\partial \bar{\mu}} \int \frac{d\bar{\varsigma}}{2\pi} \boldsymbol{\rho}_0 \left\{ \mathcal{H}_1 - \mathcal{J}[\mathcal{H}_1] \right\}$$

Since $\mathcal{J}[\mathcal{H}_1]$ does not depend on the gyroangle, only $\mathcal{H}_1$ contributes. The Fourier expansion above gives

$$\mathcal{H}_1 \left(\bar{\boldsymbol{\mathcal{Z}}} \right) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \widetilde{\mathcal{H}}_1 \left(\boldsymbol{k}, \bar{p}_\parallel, t \right) e^{i\boldsymbol{k}\cdot(\bar{\mathbf{X}} + \boldsymbol{\rho}_0)}$$

where it has been used that $\widetilde{\mathcal{H}}_1$ does not depend on $\bar{\mu}$ in the $(\tilde{\phi}, \tilde{A}_\parallel)$ other than via the gyromotion. Hence

$$\begin{aligned}
\langle \bar{\boldsymbol{\rho}}_1 \rangle_{\bar{\varsigma}} &= -\frac{1}{B(\bar{\mathbf{X}})} \frac{\partial}{\partial \bar{\mu}} \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \mathcal{H}_1 \left(\boldsymbol{k}, \bar{p}_\parallel, t \right) \int \frac{d\bar{\varsigma}}{2\pi} \boldsymbol{\rho}_0 e^{i\boldsymbol{k}\cdot(\bar{\mathbf{X}} + \boldsymbol{\rho}_0)} \\
&= i \frac{1}{B(\bar{\mathbf{X}})} \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} e^{i\boldsymbol{k}\cdot\bar{\mathbf{X}}} \mathcal{H}_1 \left(\boldsymbol{k}, \bar{p}_\parallel, t \right) \frac{\partial}{\partial \bar{\mu}} \frac{\partial}{\partial \boldsymbol{k}} J_0(k_\perp \rho_c) \\
&= -i \frac{m_a}{e_a^2 B^2(\bar{\mathbf{X}})} \frac{\mathbf{k}_\perp}{k_\perp} \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} e^{i\boldsymbol{k}\cdot\bar{\mathbf{X}}} \mathcal{H}_1 \left(\boldsymbol{k}, \bar{p}_\parallel, t \right) \frac{1}{\rho_c} \frac{\partial}{\partial \rho_c} \left(\rho_c J_1(k_\perp \rho_c) \right) \\
&= -i \frac{m_a}{e_a^2 B^2(\bar{\mathbf{X}})} \mathbf{k}_\perp \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} e^{i\boldsymbol{k}\cdot\bar{\mathbf{X}}} \mathcal{H}_1 \left(\boldsymbol{k}, \bar{p}_\parallel, t \right) J_0(k_\perp \rho_c)
\end{aligned}$$

where it has been used that $\frac{\partial}{\partial \bar{\mu}} = \frac{m_a}{e_a^2 B_{eq}(\bar{\mathbf{X}})} \frac{\partial}{\rho_c \partial \rho_c}$, and also $J_0'(x) = -J_1(x)$, $\frac{d}{dx}\left(x J_1(x) \right) = x J_0(x)$. It appears finally that

$$\langle \boldsymbol{\rho}_1 \rangle_{\bar{\varsigma}} = -\frac{m_a}{e_a^2 B_{eq}^2(\bar{\mathbf{X}})} \frac{\partial \mathcal{J}[\mathcal{H}_1]}{\partial \bar{\mathbf{X}}_\perp}$$

F Charge and current densities in the $(\phi, A_\parallel, B_\parallel)$ representation

F.1 Calculation of the particle functional

The objective here is to write the functional

$$\mathcal{A}_{ad} = -\frac{1}{2} \int d\Gamma^* \mathcal{F}'_{eq}(\boldsymbol{\mathcal{Z}}^*) \mathcal{J}[\mathcal{H}_1^2](\boldsymbol{\mathcal{Z}}^*)$$

in the hybrid position space. It reads

$$\mathcal{A}_{ad} = -\frac{1}{2} \int d^3\mathbf{X} dt d^3\mathbf{V} \, \mathcal{F}'_{eq}(\mathbf{X}, p_\parallel, \mu) \mathcal{H}_1^2(\mathbf{X} + \boldsymbol{\rho}_0(\mathbf{X}, \mu, \varsigma), t, p_\parallel, \mu)$$

or equivalently

$$\mathcal{A}_{ad} = -\frac{1}{2} \int d^3\mathbf{x} dt d^3\mathbf{V} \frac{d\varsigma}{2\pi} \left[B_\parallel^* \mathcal{F}'_{eq} \right] (\mathbf{x} - \boldsymbol{\rho}_0(\mathbf{X}, \mu, \varsigma), p_\parallel, \mu) \mathcal{H}_1^2(\mathbf{x}, t, p_\parallel, \mu)$$

where

$$\mathcal{H}_1(\mathbf{x}, t, p_\parallel, \mu) = e_a \tilde{\phi}(\mathbf{x}, t) - e_a p_\parallel \tilde{A}_\parallel(\mathbf{x}, t) - e_a \dot{\boldsymbol{\rho}}_0(\mathbf{X}, \mu, \varsigma) \cdot \tilde{\mathbf{A}}_\perp(\mathbf{x}, t)$$

The unperturbed field $\left[B_\parallel^* \mathcal{F}'_{eq} \right] (\mathbf{x} - \boldsymbol{\rho}_0(\mathbf{X}, \mu, \varsigma), p_\parallel, \mu)$ can be replaced by $\left[B_\parallel^* \mathcal{F}'_{eq} \right] (\mathbf{x}, p_\parallel, \mu)$ assuming that $\mathcal{F}'_{eq}$ is a distribution function that evolves spatially on an equilibrium

scale L_p, so that a Taylor development in $\boldsymbol{\rho}_0/L_p$ leads to higher order corrections -same argument for the Jacobian $B^*_{||}$. Using the relations

$$\int d^3\mathbf{V}\mathcal{F}'_{eq}(\mathbf{x},p_\parallel,\mu)p_\parallel^2 = \frac{T_{eq}(\mathbf{x})}{m_a}$$
$$\int d^3\mathbf{V}\mathcal{F}'_{eq}(\mathbf{x},p_\parallel,\mu)\dot{\boldsymbol{\rho}}_0(\mathbf{X},\mu,\varsigma) = 0$$
$$\int d^3\mathbf{V}\mathcal{F}'_{eq}(\mathbf{x},p_\parallel,\mu)\dot{\boldsymbol{\rho}}_{0i}\dot{\boldsymbol{\rho}}_{0j}(\mathbf{X},\mu,\varsigma) = \frac{T_{eq}(\mathbf{x})}{m_a}\delta_{ij}$$

one gets

$$\begin{aligned}\mathcal{A}_{ad} &= \frac{1}{2}e_a^2\int d\gamma^*\frac{F_{eq}(\mathbf{x},p_\parallel,\mu)}{T_{eq}(\mathbf{x})}\left\{\tilde{\phi}^2(\mathbf{x},t) - 2p_\parallel\tilde{\phi}(\mathbf{x},t)\tilde{A}_\parallel(\mathbf{x},t)\right\}\\ &+ \frac{1}{2}\frac{e_a^2}{m_a}\int d^3\mathbf{x}dt\, N_{eq}(\mathbf{x})\tilde{\mathbf{A}}^2(\mathbf{x},t)\end{aligned}$$

F.2 Writing the action in the hybrid set of coordinates $(\mathbf{x}, t, p_\parallel, \mu)$

The action $\mathcal{A}$ Eq.(55) lives in the guiding-centre space, which means that a change of variables must be performed to write the various contributions in the hybrid set of coordinates $(\mathbf{x}, t, p_\parallel, \mu)$. The terms $\mathcal{J}[\tilde{\phi}]$ and $\mathcal{J}[\tilde{\mathbf{A}}]$ are treated in the same way as in the $(\phi, A_\parallel)$ representation. The operator $\mathcal{B}$ requires some care. Assuming that the perturbed field is null at infinity, it appears that[48]

$$-\int d\Gamma^*\bar{\mathcal{G}}_1\mu\mathcal{B}[\mathbf{e}_\parallel\cdot\nabla\times\tilde{\mathbf{A}}_\perp] = \int d^3\mathbf{x}dt\left(\nabla\times\tilde{\mathbf{M}}\right)\cdot\tilde{\mathbf{A}}_\perp$$

where $\bar{\mathcal{G}}_1$ is the "non-adiabatic" part of the perturbed distribution function[49]

$$\bar{\mathcal{G}}_1 = \bar{\mathcal{F}}_1 + \frac{F_{eq}}{T_{eq}}\mathcal{J}[\mathcal{H}_1]$$

$\tilde{M}\mathbf{e}_\parallel$ is a perturbed magnetisation field

$$\tilde{\mathbf{M}} = -\int d^3\mathbf{V}\mu\mathcal{B}^t[\bar{\mathcal{G}}_1]\mathbf{e}_\parallel$$

and the operator $\mathcal{B}^t$ satisfies the property

$$\int d\Gamma^* F\left(\boldsymbol{\mathcal{Z}}^*\right)\mathcal{B}[g\left(\boldsymbol{\mathfrak{z}}^*\right)] = \int d\gamma^*\mathcal{B}^t[F\left(\boldsymbol{\mathcal{Z}}^*\right)]g\left(\boldsymbol{\mathfrak{z}}^*\right) \tag{121}$$

So $\mathcal{B}^t$ is the adjoint of $\mathcal{B}$ in the sense of Eq.(121). It holds if the Jacobian of the guiding-centre transform exhibits large spatial scale only. A consequence of the Parseval theorem applied on Eq.(121) is that $\mathcal{B}^t$ operates in the Fourier space as $\mathcal{B}$, i.e. by multiplying the target field by $2J_1\left(k_\perp\rho_c\right)/k_\perp\rho_c$ and then moving back to the real space. In that sense $\mathcal{B}^t$ is the same as $\mathcal{B}$. However the notation $\mathcal{B}^t$ will be kept to mark the difference with $\mathcal{B}$

[48] This result is obtained the identity $\nabla\cdot(\mathbf{A}\times\mathbf{B}) = \mathbf{B}\cdot\nabla\times\mathbf{A} - \mathbf{A}\cdot\nabla\times\mathbf{B}$ for any couple of vectors $(\mathbf{A},\mathbf{B})$. The integration of the divergence operator over the volume introduces a surface term that cancels if one of the vectors (or both) $(\mathbf{A},\mathbf{B})$ vanishes on the surface. This assumption is retained here.

[49] The reader should be warned that there exists different definitions of an adiabatic response, in particular in the MHD context

since its mathematical meaning is different. An explicit expression of the magnetisation amplitude $\tilde{M} = \mathbf{e}_\parallel \cdot \tilde{\mathbf{M}}$ versus the perturbed distribution function $\mathcal{F}_1$ is

$$\begin{aligned}
\tilde{M} &= -\int d^3\mathbf{V} \mu \mathcal{B}^t[\bar{\mathcal{F}}_1] \\
&- \int d^3\mathbf{V} \frac{F_{eq} e}{T_{eq}} \mu \mathcal{B}^t \mathcal{J}[\tilde{\phi} - p_\parallel \tilde{A}_\parallel] \\
&- \int d^3\mathbf{V} \frac{F_{eq} e}{T_{eq}} \mu^2 \mathcal{B}^t \mathcal{B}[\mathbf{e}_\parallel \cdot \nabla \times \tilde{\mathbf{A}}_\perp]
\end{aligned}$$

F.3 Charge and current densities - general expressions

Once the appropriate change of variables has been done, the functional $\mathcal{A}$ Eq.(55) can be used to find the charge and current densities. It can be made explicit in he physical space as

$$\begin{aligned}
\mathcal{A} &= \frac{\epsilon_0}{2} \int d^3\mathbf{x} dt \, \tilde{\mathbf{E}}^2(\mathbf{x}, t) - \frac{1}{2\mu_0} \int d^3\mathbf{x} dt \, \tilde{\mathbf{B}}^2(\mathbf{x}, t) \\
&+ \frac{1}{2} \sum_{species} e_a^2 \int d^3\mathbf{x} dt d^3\mathbf{V} \frac{F_{eq}(\mathbf{x}, p_\parallel, \mu)}{T_{eq}(\mathbf{x})} \left\{ \tilde{\phi}^2(\mathbf{x}, t) - 2 p_\parallel \tilde{A}_\parallel(\mathbf{x}, t) \tilde{\phi}(\mathbf{x}, t) \right\} \\
&- \frac{1}{2} \sum_{species} \int d^3\mathbf{x} dt d^3\mathbf{V} \frac{F_{eq}(\mathbf{x})}{T_{eq}(\mathbf{x})} \left[e_a \mathcal{J}[\tilde{\phi}] \left(\mathfrak{z}^* \right) - e_a p_\parallel \mathcal{J}[\tilde{A}_\parallel] \left(\mathfrak{z}^* \right) \right. \\
&\quad \left. + 2\mu \mathcal{B}[\mathbf{e}_\parallel \cdot \nabla \times \tilde{\mathbf{A}}_\perp] \left(\mathfrak{z}^* \right) \right] \left[e_a \mathcal{J}[\tilde{\phi}] \left(\mathfrak{z}^* \right) - e_a p_\parallel \mathcal{J}[\tilde{A}_\parallel] \left(\mathfrak{z}^* \right) \right] \\
&- \sum_{species} \int d^3\mathbf{x} dt d^3\mathbf{V} \bar{\mathcal{F}}_1 \left(\mathfrak{z}^* \right) \left[e_a \mathcal{J}[\tilde{\phi}] \left(\mathfrak{z}^* \right) - e_a p_\parallel \mathcal{J}[\tilde{A}_\parallel] \left(\mathfrak{z}^* \right) \right] \\
&+ \sum_{species} \int d^3\mathbf{x} dt \left(\nabla \times \tilde{\mathbf{M}} \right) \cdot \tilde{\mathbf{A}}_\perp
\end{aligned}$$

The extremum in $\tilde{\phi}$ provides the charge density for each species

$$\begin{aligned}
\tilde{\varrho} &= e_a \int d^3\mathbf{V} \mathcal{J}^t[\bar{\mathcal{F}}_1] \\
&- e_a \int d^3\mathbf{V} \frac{F_{eq} e}{T_{eq}} \left\{ \tilde{\phi} - \mathcal{J}^t \mathcal{J}[\tilde{\phi}] \right\} \\
&+ e_a \int d^3\mathbf{V} \frac{F_{eq} e}{T_{eq}} p_\parallel \left(\tilde{A}_\parallel - \mathcal{J}^t \mathcal{J}[\tilde{A}_\parallel] \right) \\
&+ e_a \int d^3\mathbf{V} \frac{F_{eq}}{T_{eq}} \mu \mathcal{J}^t \mathcal{B}[\mathbf{e}_\parallel \cdot \nabla \times \tilde{\mathbf{A}}_\perp]
\end{aligned} \tag{122}$$

which holds as long as the Maxwellian equilibrium distribution function exhibits large spatial scales compared with fluctuations. The perturbed current for each species is given by the extremum of the action in $\tilde{\mathbf{A}}$ and reads

$$\begin{aligned}
\tilde{\mathbf{J}} &= e_a \int d^3\mathbf{V} \mathcal{J}^t[\bar{\mathcal{F}}_1] p_\parallel \mathbf{e}_\parallel \\
&- e_a \int d^3\mathbf{V} \frac{F_{eq} e}{T_{eq}} p_\parallel \left(\tilde{\phi} - \mathcal{J}^t \mathcal{J}[\tilde{\phi}] \right) \mathbf{e}_\parallel \\
&- e_a \int d^3\mathbf{V} \frac{F_{eq} e_a}{m_a} \mathcal{J}^t \mathcal{J}[\tilde{A}_\parallel] \mathbf{e}_\parallel \\
&+ e_a \int d^3\mathbf{V} \frac{F_{eq}}{T_{eq}} \mu p_\parallel \mathcal{J}^t \mathcal{B}[\mathbf{e}_\parallel \cdot \nabla \times \tilde{\mathbf{A}}_\perp] \mathbf{e}_\parallel \\
&+ \left(\nabla \times \tilde{\mathbf{M}} \right)_\perp
\end{aligned} \tag{123}$$

where it has been used that the average of $p_\parallel^2$ over a Maxwellian is T_{eq}/m. Eqs.(122,123) should be used to ensure the field coherence via the Maxwell equations, which become

$$\begin{aligned} \nabla \cdot \tilde{\mathbf{E}} &= \sum_{species} \frac{\tilde{\varrho}}{\epsilon_0} \\ \nabla \times \tilde{\mathbf{B}} &= \mu_0 \sum_{species} \tilde{\mathbf{J}} \end{aligned}$$

in the low frequency limit.

F.4 Some simplifications

The equations above are hardly tractable. Additional simplifications, consistent with the gyrokinetic ordering, are often used and discussed in this section. Since the vector potential is $\tilde{\mathbf{A}} = \tilde{A}_\parallel \mathbf{e}_\parallel + \tilde{\mathbf{A}}_\perp$, the perturbed magnetic field reads

$$\tilde{\mathbf{B}} = \nabla \tilde{A}_\parallel \times \mathbf{e}_\parallel + \tilde{A}_\parallel \nabla \times \mathbf{e}_\parallel + \nabla \times \tilde{\mathbf{A}}_\perp$$

The vector $\nabla \times \mathbf{e}_\parallel$ can be written as

$$\nabla \times \mathbf{e}_\parallel = \frac{\mu_0 J_{\parallel eq}}{B_{eq}} \mathbf{e}_\parallel + \boldsymbol{\tau}$$

where $\boldsymbol{\tau} = \mathbf{e}_\parallel \times \boldsymbol{\kappa}$, $\boldsymbol{\kappa}$ the field line curvature and $J_{\parallel eq}$ the total unperturbed current. As discussed before, its norm scales as the inverse of a macroscopic length, while the norms of $\nabla \tilde{A}_\parallel$ and $\nabla \times \tilde{\mathbf{A}}_\perp$ scale as $\tilde{A}_\parallel$ divided by a perturbation length scale. For turbulence, this scale length is typically an ion gyroradius, so that the term $\tilde{A}_\parallel \nabla \times \mathbf{e}_\parallel$ can be neglected since one order lower. It is thus recovered that $\tilde{B}_\parallel = \mathbf{e}_\parallel \cdot \nabla \times \tilde{\mathbf{A}}_\perp$. The perturbed parallel current reads[50]

$$\mu_0 \sum_{species} \tilde{J}_\parallel = \mathbf{e}_\parallel \cdot \nabla \times \tilde{\mathbf{B}} = -\nabla_\perp^2 \tilde{A}_\parallel + \nabla \cdot \left(\left(\nabla \times \tilde{\mathbf{A}}_\perp \right) \times \mathbf{e}_\parallel \right)$$

The last term of this identity can be ignored as well[51]. The Ampère equation then reduces to $\nabla_\perp^2 \tilde{A}_\parallel = -\mu_0 \tilde{J}_\parallel$. At this level of approximation, the perturbed magnetic field reads

$$\tilde{\mathbf{B}} = \nabla \tilde{A}_\parallel \times \mathbf{e}_\parallel + \tilde{B}_\parallel \mathbf{e}_\parallel$$

and the perpendicular current $\tilde{\mathbf{J}}_\perp = \nabla \times \tilde{B}_\parallel \mathbf{e}_\parallel$. After identification with the perpendicular component of the current in Eq.(123), the compact relation $\tilde{B}_\parallel = \mu_0 \sum_{species} \tilde{M}$ is

[50] We use here the identity $\mathbf{e}_\parallel \cdot \nabla \times \tilde{\mathbf{B}} = \nabla \cdot \left(\tilde{\mathbf{B}} \times \mathbf{e}_\parallel \right) + \tilde{\mathbf{B}} \cdot \nabla \times \mathbf{e}_\parallel$ - the last term can be ignored for small scale perturbations. The "perpendicular gradient" of a scalar $\tilde{A}_\parallel$ is defined as $\nabla_\perp \tilde{A}_\parallel = \nabla \tilde{A}_\parallel - \mathbf{e}_\parallel (\mathbf{e}_\parallel \cdot \nabla) \tilde{A}_\parallel = \mathbf{e}_\parallel \times (\nabla \tilde{A}_\parallel \times \mathbf{e}_\parallel)$, and consequently the "perpendicular" Laplacian is $\nabla_\perp^2 = \nabla \cdot \nabla_\perp$.

[51] Let us remind that the magnetic field can be written in Clebsh form $\mathbf{B}_{eq} = \nabla \alpha \times \nabla \psi$, where α and ψ are two suitable scalar fields. A system of coordinates (α, ψ, ℓ) is then chosen, where ℓ is the curvilinear abscissa along the unperturbed field lines. Hence the third contravariant vector is $\nabla \ell = \mathbf{e}_\parallel$, and the Jacobian $[(\nabla \alpha \times \nabla \psi) \cdot \nabla \ell]^{-1}$ is unity. The perpendicular potential vector is

$$\tilde{\mathbf{A}}_\perp = A_\psi \nabla \psi + A_\alpha \nabla \alpha$$

A bit of algebra shows that

$$\left(\nabla \times \tilde{\mathbf{A}}_\perp \right) \times \mathbf{e}_\parallel = \frac{\partial A_\psi}{\partial \ell} \nabla \psi + \frac{\partial A_\alpha}{\partial \ell} \nabla \alpha$$

Perturbations of the field tend to be aligned with the equilibrium field, so that parallel gradients (the operator ∂_ℓ) are ϵ_B smaller than transverse gradients (the operator $\nabla_\perp$). This allows neglecting this term against the perpendicular Laplacian of the perturbed parallel potential vector.

obtained. Finally, because the Debye length is smaller that the gyroradius in most magnetised plasmas of interest, the Gauss's law is reduced to a charge neutrality condition $\sum_{species} \tilde{\varrho} = 0$. So the final set of equations to be solved is

$$\begin{aligned}\sum_{species} \int d^3\mathbf{V} \frac{F_{eq} e_a^2}{T_{eq}} \left(1 - \mathcal{J}^t \mathcal{J}\right) [\tilde{\phi}] &= \sum_{species} \int d^3\mathbf{V} e_a \mathcal{J}^t [\bar{\mathcal{F}}_1] \\ &+ \sum_{species} \int d^3\mathbf{V} \frac{F_{eq} e_a^2}{T_{eq}} p_\parallel \left(1 - \mathcal{J}^t \mathcal{J}\right) [\tilde{A}_\parallel] \\ &+ \sum_{species} \int d^3\mathbf{V} \frac{F_{eq} e_a}{T_{eq}} \mu \mathcal{J}^t \mathcal{B} [\tilde{B}_\parallel] \end{aligned} \tag{124}$$

$$\begin{aligned} -\frac{1}{\mu_0} \nabla_\perp^2 \tilde{A}_\parallel + \sum_{species} \int d^3\mathbf{V} \frac{F_{eq} e_a^2}{m_a} \mathcal{J}^t \mathcal{J} [\tilde{A}_\parallel] &= \sum_{species} e_a \int d^3\mathbf{V} \mathcal{J}^t [\bar{\mathcal{F}}_1] p_\parallel \\ &- \sum_{species} \int d^3\mathbf{V} \frac{F_{eq} e_a^2}{T_{eq}} p_\parallel \left(\tilde{\phi} - \mathcal{J}^t \mathcal{J} [\tilde{\phi}]\right) \\ &+ \sum_{species} \int d^3\mathbf{V} \frac{F_{eq} e_a}{T_{eq}} \mu p_\parallel \mathcal{J}^t \mathcal{B} [\tilde{B}_\parallel] \end{aligned} \tag{125}$$

$$\begin{aligned} \frac{1}{\mu_0} \tilde{B}_\parallel + \sum_{species} \int d^3\mathbf{V} \frac{F_{eq} e_a}{T_{eq}} \mu^2 \mathcal{B}^t \mathcal{B} [\tilde{B}_\parallel] &= -\sum_{species} \int d^3\mathbf{V} \mu \mathcal{B}^t [\bar{\mathcal{F}}_1] \\ &- \sum_{species} \int d^3\mathbf{V} \frac{F_{eq} e_a}{T_{eq}} \mu \mathcal{B}^t \mathcal{J} [\tilde{\phi} - p_\parallel \tilde{A}_\parallel] \end{aligned} \tag{126}$$

These equations must be complemented by a solution of the gyrokinetic Vlasov equation. The equations of motion are the same as Eqs.(24,25), but with a new gyroaverage Hamiltonian[52]

$$\mathcal{J}[\widetilde{H}] = e_a \mathcal{J}[\tilde{\phi}] - e_a p_\parallel \mathcal{J}[\tilde{A}_\parallel] + \mu \mathcal{B}[\tilde{B}_\parallel] \tag{127}$$

F.5 A few more simplifications

The second term in the l.h.s. of Eq.(126) is of the order of the plasma β, and is therefore quite small. The first term of the r.h.s. is close to $\tilde{P}_\perp / B_{eq}$, where $\tilde{P}_\perp$ is the gyroaveraged perpendicular pressure defined as the average of μB_{eq} over the perturbed gyroaveraged distribution function $\mathcal{B}^t[\bar{\mathcal{F}}_1]$. Eq.(126) is then close to the usual MHD approximation $B_{eq}\tilde{B}_\parallel + \mu_0 \tilde{P}_\perp = 0$. Here the 2nd line of the r.h.s. of Eq.(57) is neglected on the basis of low β values. It turns out that when this relation holds, part of the effects due to $\tilde{B}_\parallel$ can be recovered by replacing the ∇B magnetic drift by a curvature drift[53]

$$\mathbf{v}_D = \left(m_a p_\parallel^2 + \mu B_{eq}\right) \frac{1}{e_a B_{||}^*} \mathbf{e}_\parallel \times \boldsymbol{\kappa}$$

[52] We again anticipate that 2nd order term should be kept to ensure some coherence with the symplectic approach - see below.

[53] This can be verified as follows. Still in the spirit of ideal MHD, let us look at the perturbed Hamiltonian in the low wavelength limit $\widetilde{H} = e_a\tilde{\phi} - e_a p_\parallel \tilde{A}_\parallel + \mu \tilde{B}_\parallel$ and assume that the ideal MHD constraint applies, i.e. null parallel electric field $-\partial_t \tilde{A}_\parallel - (\mathbf{e}_\parallel \cdot \nabla)\tilde{\phi} = 0$. The perturbed Hamiltonian can be recast as $\widetilde{H} = e_a \frac{d\tilde{\chi}}{dt} - e_a \mathbf{v}_D \cdot \nabla\tilde{\chi} + \mu \tilde{B}_\parallel$, where $\frac{d}{dt} = \frac{\partial}{\partial t} + p_\parallel \mathbf{e}_\parallel \cdot \nabla + \mathbf{v}_D \cdot \nabla$ is a Lagrangian (or material) derivative, and $\chi(\mathbf{X}, t) = \int^t dt' \phi(\mathbf{X}, t')$ is the primitive in time of the electric potential, so that $\tilde{A}_\parallel = -(\mathbf{e}_\parallel \cdot \nabla)\tilde{\chi}$. In first approximation, and in agreement with the MHD fluid pressure equation, the pressure perturbation is just given by the advection of the equilibrium pressure by the $E \times B$ electric drift, i.e. $\tilde{P}_\perp = -\frac{1}{B_{||}^*} \left(\mathbf{e}_\parallel \times \nabla\tilde{\chi}\right) \cdot \nabla P_{eq}$. Using

The relation $B_{eq}\tilde{B}_{\parallel}+\mu_0\tilde{P}_{\perp}=0$ is quite convenient, since it relates the compressional component of the magnetic field $\tilde{B}_{\parallel}$ to the perturbed pressure, which can be calculated easily from the perturbed distribution function. Besides, a further step consists in modifying the magnetic drift velocity and ignoring all contributions from the parallel compressibility of the field. Needless to say that this is a rather blunt approximation, which is in fact justified only for small scale modes (e.g. ballooning modes).

Some extra bold steps can be crossed. A number of terms in the charge and current densities are proportional to $p_{\parallel}(1-\mathcal{J}^t\mathcal{J})$. They vanish if the unperturbed distribution function is unshifted or if finite Larmor radius (FLR) effects are neglected. It would be unreasonable to ignore finite Larmor radius (FLR) effects on the ion species, since their gyroradius is commensurate with turbulence wavelengths. However it is quite licit to neglect the current carried by ion species in most practical cases (except fast ions though). Conversely the current density carried by electrons cannot be neglected, by electron finite Larmor radius (FLR) effects are negligible in practical situations since the electron gyroradius is much smaller than fluctuations wavelengths. These approximations lead to simplified, but tractable, Poisson and Ampère equations

$$\begin{aligned}\sum_{species}\int d^3\mathbf{V}\frac{F_{eq}e_a^2}{T_{eq}}\left(1-\mathcal{J}^t\mathcal{J}\right)[\tilde{\phi}] &= \sum_{species}e_a\int d^3\mathbf{V}\mathcal{J}^t[\bar{\mathcal{F}}_1]\\ &+ \sum_{species}e_a\int d^3\mathbf{V}\frac{F_{eq}}{T_{eq}}\mu\mathcal{J}^t\mathcal{B}[\tilde{B}_{\parallel}]\end{aligned}$$

$$-\frac{1}{\mu_0}\nabla_{\perp}^2\tilde{A}_{\parallel}+\sum_{species}\int d^3\mathbf{V}\frac{F_{eq}e_a^2}{m_a}\mathcal{J}^t\mathcal{J}[\tilde{A}_{\parallel}]=\sum_{species}e_a\int d^3\mathbf{V}p_{\parallel}\mathcal{J}^t[\bar{\mathcal{F}}_1]$$

to be completed with an equation over the compressional component of field

$$\begin{aligned}\frac{1}{\mu_0}\tilde{B}_{\parallel}+\sum_{species}\int d^3\mathbf{V}\frac{F_{eq}e_a}{T_{eq}}\mu^2\mathcal{B}^t\mathcal{B}[\tilde{B}_{\parallel}] &= -\sum_{species}\int d^3\mathbf{V}\mu\mathcal{B}^t[\bar{\mathcal{F}}_1]\\ &- \sum_{species}\int d^3\mathbf{V}\frac{F_{eq}e_a}{T_{eq}}\mu\mathcal{B}^t\mathcal{J}[\tilde{\phi}]\end{aligned}\qquad(129)$$

F.6 Explicit expressions of gyroaverage operators

The choice of Maxwellian distribution functions allows performing analytic integration over the velocity space. It was shown in Appendix E that the gyroaverage operator $\mathcal{J}$ consists in multiplying the Fourier transform of the field by $J_0(k_\perp\rho_c)$ and the operator $\mathcal{B}$ does a similar operation, but with a multiplication by $\frac{2J_1(k_\perp\rho_c)}{k_\perp\rho_c}$. It then appears that the average over a Maxwellian of the operators $\mathcal{J}^t\mathcal{J}$, $\mathcal{J}^t\mathcal{B}$ and $\mathcal{B}^t\mathcal{B}$ are explicit, thanks to the relationships

$$2\int_0^{+\infty}dvve^{-v^2}J_0^2(av)=\Gamma_0(b)$$

$$2\int_0^{+\infty}dvv^3e^{-v^2}J_0^2(av)=\Gamma_0(b)-b\left(\Gamma_0(b)-\Gamma_1(b)\right)$$

$\tilde{B}_{\parallel}=-\mu_0\tilde{P}_{\perp}/B_{eq}$ the perturbed Hamitonian becomes

$$\widetilde{H}=e_a\frac{d\tilde{\chi}}{dt}-e_a\mathbf{v}_D\cdot\nabla\tilde{\chi}-\mu\frac{\mu_0}{B_{eq}B_{||}^*}\left(\mathbf{e}_{\parallel}\times\nabla P_{eq}\right)\cdot\nabla\tilde{\chi}\qquad(128)$$

The first piece $\frac{d\tilde{\chi}}{dt}$ can be incorporated in the adiabatic response of the perturbed distribution and will not be discussed further. Using the expression of the field line curvature $\boldsymbol{\kappa}=\frac{\nabla_\perp B_{eq}}{B_{eq}}+\frac{\mu_0\nabla_\perp P_{eq}}{B_{eq}^2}$, it appears that the two last term of the Hamiltonian can be regrouped as the advection of $\tilde{\chi}$ by a modified magnetic drift velocity where the ∇B drift piece is replaced by the curvature.

$$2\int_0^{+\infty} dv v^3 e^{-v^2} J_1^2(av) = b\left(\Gamma_0(b) - \Gamma_1(b)\right)$$

where $a = \sqrt{2}k_\perp \rho_T$, $b = \frac{a^2}{2} = k_\perp^2 \rho_T^2$, $\rho_T = \frac{\sqrt{m_a T_{eq}}}{eB}$ is the thermal gyroradius and $\Gamma_\nu(b) = I_\nu(b)e^{-b}$, I_ν the modified Bessel function of index ν. The following identities can then be derived

$$\frac{1}{N_{eq}}\int d^3\mathbf{V} F_{eq} \mathcal{J}^t \mathcal{J} = \Gamma_0(b)$$

$$\frac{1}{N_{eq}}\int d^3\mathbf{V} F_{eq} \left(\frac{\mu B_{eq}}{T_{eq}}\right)^2 \mathcal{B}^t \mathcal{B} = 2\left(\Gamma_0(b) - \Gamma_1(b)\right)$$

$$\frac{1}{N_{eq}}\int d^3\mathbf{V} F_{eq} \left(\frac{\mu B_{eq}}{T_{eq}}\right) \mathcal{B}^t \mathcal{J} = \Gamma_0(b) - \Gamma_1(b)$$

These relations have to be understood as operations in the Fourier space. For electrons, $b \ll 1$, and $\Gamma_0 \sim 1$, $\Gamma_1 \sim 0$. The Maxwell equations then become

$$\begin{aligned}\sum_{species} \frac{N_{eq} e_a^2}{T_{eq}} \left(1 - \Gamma_0\right)[\tilde{\phi}] &= \sum_{species} e_a \int d^3\mathbf{V} \mathcal{J}^t [\bar{\mathcal{F}}_1] \\ &+ \sum_{species} \frac{N_{eq} e}{B_{eq}} \left(\Gamma_0 - \Gamma_1\right)[\tilde{B}_\parallel]\end{aligned}$$

$$-\frac{1}{\mu_0}\nabla_\perp^2 \tilde{A}_\parallel + \sum_{species} \frac{N_{eq} e_a^2}{m_a} \Gamma_0[\tilde{A}_\parallel] = \sum_{species} e_a \int d^3\mathbf{V} p_\parallel \mathcal{J}^t [\bar{\mathcal{F}}_1]$$

$$\begin{aligned}\frac{\tilde{B}_\parallel}{B_{eq}} &= -\sum_{species} \int d^3\mathbf{V} \frac{2\mu_0 N_{eq} T_{eq}}{B_{eq}^2} \frac{1}{B_{eq}} \left(\Gamma_0 - \Gamma_1\right)[\tilde{B}_\parallel] \\ &- \frac{\mu_0}{B_{eq}^2} \sum_{species} \int d^3\mathbf{V} \mu B_{eq} \mathcal{B}^t [\bar{\mathcal{F}}_1] \\ &- \sum_{species} \frac{\mu_0 N_{eq} T_{eq}}{B_{eq}^2} \frac{e_a}{T_{eq}} \left(\Gamma_0 - \Gamma_1\right)[\tilde{\phi}]\end{aligned}$$

where the last equation has been normalised to show explicitly the dependence on the parameter $\frac{\mu_0 N_{eq} T_{eq}}{B_{eq}^2}$ that scales like the plasma β.

G Near identity canonical change of variables

We start first with the second order theory in the Hamiltonian case. Near identity canonical changes of coordinates at all orders were studied extensively in the context of celestial mechanics. A powerful approach is based on Deprit series, which can be summarised as follows. We call here "canonical change of variables" a relationship between old and new variables of the form which bears the same structure as the one for canonical variables (see Appendix D.1)

$$\boldsymbol{\mathcal{Z}} = \bar{\boldsymbol{\mathcal{Z}}} - \left\{S, \bar{\boldsymbol{\mathcal{Z}}}\right\}_{\bar{\boldsymbol{\mathcal{Z}}}}$$

where the generating function S can be expanded as a series in ϵ_δ

$$S(\bar{\mathcal{Z}}) = \sum_{n=0}^{+\infty} S_n(\bar{\mathcal{Z}}) \epsilon_\delta^n$$

All calculations are run in gyrocentre variables. As in the first order theory, the old Hamiltonian $\mathcal{H}(\boldsymbol{\mathcal{Z}})$ is related to the new one $\bar{\mathcal{H}}(\bar{\boldsymbol{\mathcal{Z}}})$ via the relation $\mathcal{H}(\boldsymbol{\mathcal{Z}}) = \bar{\mathcal{H}}(\bar{\boldsymbol{\mathcal{Z}}})$. The Hamiltonians can also be expanded in ϵ_δ, e.g.

$$\begin{aligned}\mathcal{H}(\boldsymbol{\mathcal{Z}}) &= \sum_{n=0}^{+\infty} \mathcal{H}_n(\boldsymbol{\mathcal{Z}})\epsilon_\delta^n \\ \bar{\mathcal{H}}(\bar{\boldsymbol{\mathcal{Z}}}) &= \sum_{n=0}^{+\infty} \bar{\mathcal{H}}(\bar{\boldsymbol{\mathcal{Z}}})\epsilon_\delta^n\end{aligned}$$

Equating $\mathcal{H}(\boldsymbol{\mathcal{Z}})$ with $\bar{\mathcal{H}}(\bar{\boldsymbol{\mathcal{Z}}})$ order by order, one finds

$$\begin{aligned}\bar{\mathcal{H}}_0(\bar{\boldsymbol{\mathcal{Z}}}) &= \mathcal{H}_0(\bar{\boldsymbol{\mathcal{Z}}}) \\ \bar{\mathcal{H}}_1(\bar{\boldsymbol{\mathcal{Z}}}) &= \mathcal{H}_1(\bar{\boldsymbol{\mathcal{Z}}}) - \{S_1, \mathcal{H}_0\}_{\bar{\boldsymbol{\mathcal{Z}}}} \\ \bar{\mathcal{H}}_2(\bar{\boldsymbol{\mathcal{Z}}}) &= \mathcal{H}_2(\bar{\boldsymbol{\mathcal{Z}}}) - \{S_1, \mathcal{H}_1\}_{\bar{\boldsymbol{\mathcal{Z}}}} + \frac{1}{2}\{S_1, \{S_1, \mathcal{H}_0\}\}_{\bar{\boldsymbol{\mathcal{Z}}}} - \{S_2, \mathcal{H}_0\}_{\bar{\boldsymbol{\mathcal{Z}}}}\end{aligned}$$

The order 0 is the same as before. Requesting that the new Hamiltonian of order 1 does not depend on the gyroangle implies that $\bar{\mathcal{H}}_1 = \langle \mathcal{H}_1 \rangle$[54]. As shown in section 3.2, this operation is a gyroaverage over a cyclotron circular motion, so that

$$\bar{\mathcal{H}}_1(\bar{\boldsymbol{\mathcal{Z}}}^*) = \mathcal{J}[\mathcal{H}_1](\bar{\boldsymbol{\mathcal{Z}}}^*)$$

Hence the generating function S_1 is given by the equation

$$\{S_1, \mathcal{H}_0\}_{\bar{\boldsymbol{\mathcal{Z}}}} = \widetilde{\mathcal{H}}_1(\bar{\boldsymbol{\mathcal{Z}}}) \tag{130}$$

where

$$\widetilde{\mathcal{H}}_1(\bar{\boldsymbol{\mathcal{Z}}}) = \mathcal{H}_1(\bar{\boldsymbol{\mathcal{Z}}}^*, \bar{\varsigma}) - \mathcal{J}[\mathcal{H}_1](\bar{\boldsymbol{\mathcal{Z}}}^*)$$

This is the same as Eq.(16) derived in section 3.2. The expression Eq.(13) of the Poisson bracket leads to the Lagrangian derivative Eq.(17), which can also be written

$$\{S_1, \mathcal{H}_0\}_{\bar{\boldsymbol{\mathcal{Z}}}} = \{S_1, \mathcal{H}_0\}_c + \{S_1, \mathcal{H}_0\}_{\bar{\boldsymbol{\mathcal{Z}}}^*}$$

where

$$\{S_1, \mathcal{H}_0\}_c = \Omega_c \frac{\partial S_1}{\partial \bar{\varsigma}}$$

is the Poisson bracket that corresponds to the cyclotron motion and

$$\{S_1, \mathcal{H}_0\}_{\bar{\boldsymbol{\mathcal{Z}}}^*} = \frac{\partial S_1}{\partial t} + \bar{p}_\parallel (\mathbf{e}_\parallel \cdot \nabla) S_1 + (\dot{\bar{\mathbf{X}}}_\perp \cdot \nabla) S_1$$

is the Lagrangian derivative along unperturbed gyrocentre trajectories. This "slow" Lagrangian derivative is of order ϵ_ω compared with the first one. Hence the solution Eq.(17) is recovered from Eq.(130) at lowest order

$$\frac{\partial S_1}{\partial \bar{\varsigma}} = \frac{1}{\Omega_c}\widetilde{\mathcal{H}} + o(\epsilon_\omega) \tag{131}$$

In the traditional "gyroBohm" ordering, one has $\epsilon_\omega \sim \epsilon_\delta$. However this level of accuracy is insufficient for a second order calculation is one chooses an ordering[55]. One solution, advocated by Tronko and Chandre [4], consists in choosing the ordering, namely$\epsilon_\omega \sim \epsilon_\delta^2$. Another solution is to estimate the error that is made when restricting the Poisson bracket

[54]The average is on the gyrocentre gyroangle $\bar{\varsigma}$. Since all calculations are run in gyrocentre coordinates, the subscript $\bar{\varsigma}$ is omitted for clarity in the averages.

[55]N.B.:The solution of Eq.(130) is explicit when using angle/action variables. If all functions are expanded as Fourier series in $\boldsymbol{\alpha}$, e.g. $\widetilde{\mathcal{H}}_1 = \sum_{\mathbf{n}} \widetilde{\mathcal{H}}_{1,\mathbf{n}} e^{i\mathbf{n}\cdot\boldsymbol{\alpha}}$, then $S_{1,\mathbf{n}} = \frac{\widetilde{\mathcal{H}}_{1,\mathbf{n}}}{i\mathbf{n}\cdot\boldsymbol{\Omega}}$

to the cyclotron motion. Hence, we do not use for now a specific solution S_1 until the ordering is clarified, i.e. we keep Eq.(130) in its full generality. The equation that rules S_2 is

$$\{S_2, \mathcal{H}_0\}_{\bar{\boldsymbol{\mathcal{Z}}}} = \widetilde{\mathcal{H}}_2 - \{S_1, \mathcal{H}_1\}_{\bar{\boldsymbol{\mathcal{Z}}}} + \frac{1}{2}\{S_1, \{S_1, \mathcal{H}_0\}_{\bar{\boldsymbol{\mathcal{Z}}}}\}_{\bar{\boldsymbol{\mathcal{Z}}}}$$

where $\widetilde{\mathcal{H}}_2 = \mathcal{H}_2 - \mathcal{J}[\mathcal{H}_2]$ and $\{S_2, \mathcal{H}_0\}_{\bar{\boldsymbol{\mathcal{Z}}}} = \Omega_c \frac{\partial S_2}{\partial \bar{\varsigma}} + o(\epsilon_\omega)$, and all Poisson brackets are calculated with the new gyrocentre variables $\bar{\boldsymbol{\mathcal{Z}}}$. A solubility constraint is obtained by averaging this equation over $\bar{\varsigma}$. It provides the expression of the second order Hamiltonian

$$\bar{\mathcal{H}}_2 = \langle \mathcal{H}_2 \rangle - \langle \{S_1, \mathcal{H}_1\}_{\bar{\boldsymbol{\mathcal{Z}}}} \rangle + \frac{1}{2}\left\langle \{S_1, \{S_1, \mathcal{H}_0\}_{\bar{\boldsymbol{\mathcal{Z}}}}\}_{\bar{\boldsymbol{\mathcal{Z}}}} \right\rangle$$

where $\langle \mathcal{H} \rangle = \int \frac{d\bar{\varsigma}}{2\pi} \mathcal{H}$ for any function $\mathcal{H}(\bar{\boldsymbol{\mathcal{Z}}})$[56]. Using

$$\left\langle \{S_1, \{S_1, \mathcal{H}_0\}_{\bar{\boldsymbol{\mathcal{Z}}}}\}_{\bar{\boldsymbol{\mathcal{Z}}}} \right\rangle = \left\langle \left\{S_1, \mathcal{H}_1 - \bar{\mathcal{H}}_1\right\}_{\bar{\boldsymbol{\mathcal{Z}}}} \right\rangle$$

leads to the following expression

$$\bar{\mathcal{H}}_2 = \langle \mathcal{H}_2 \rangle - \frac{1}{2}\left\langle \left\{S_1, \mathcal{H}_1 + \bar{\mathcal{H}}_1\right\}_{\bar{\boldsymbol{\mathcal{Z}}}} \right\rangle$$

Let us now assess the order of magnitude of the various terms. The second order Hamiltonians $\mathcal{H}_2$ and $\bar{\mathcal{H}}_2$ scale as $\epsilon_\delta^2 T_{eq}$. Poisson brackets that appear in the second term of the r.h.s. of Eq.(132) are of the form

$$\{S_1, \mathcal{H}_1\}_{\bar{\mathbf{Z}}} = \{S_1, \mathcal{H}_1\}_c + \{S_1, \mathcal{H}_1\}_{\bar{\mathbf{Z}}^*}$$

where the cyclotron part is

$$\{S_1, \mathcal{H}_1\}_c = \frac{e_a}{m_a}\left(\frac{\partial S_1}{\partial \bar{\varsigma}}\frac{\partial \mathcal{H}_1}{\partial \bar{\mu}} - \frac{\partial S_1}{\partial \bar{\mu}}\frac{\partial \mathcal{H}_1}{\partial \bar{\varsigma}}\right)$$

and the gyrocentre contribution is

$$\begin{aligned} \{S_1, \mathcal{H}_1\}_{\bar{\mathbf{Z}}^*} &= \frac{1}{m_a}\frac{\mathbf{B}^*}{B_{||}^*} \cdot \left(\frac{\partial S_1}{\partial \bar{\mathbf{X}}}\frac{\partial \mathcal{H}_1}{\partial \bar{p}_{||}} - \frac{\partial S_1}{\partial \bar{p}_{||}}\frac{\partial \mathcal{H}_1}{\partial \bar{\mathbf{X}}}\right) \\ &\quad - \frac{\mathbf{e}_{||}}{e_a B_{||}^*} \cdot \left(\frac{\partial S_1}{\partial \bar{\mathbf{X}}} \times \frac{\partial \mathcal{H}_1}{\partial \bar{\mathbf{X}}}\right) \end{aligned} \tag{132}$$

The Poisson bracket $\left\{S_1, \bar{\mathcal{H}}_1\right\}$ is decomposed in the same way. After an integration by part, the average of the cyclotron part of the Poisson bracket is recast as

$$\left\langle \left\{S_1, \mathcal{H}_1 + \bar{\mathcal{H}}_1\right\}_c \right\rangle = \frac{e_a}{m_a}\frac{\partial}{\partial \bar{\mu}}\left\langle \frac{\partial S_1}{\partial \bar{\varsigma}}\left(\mathcal{H}_1 + \bar{\mathcal{H}}_1\right) \right\rangle$$

It appears that the lowest order solution S_1 Eq.(131) is sufficient to find the $\epsilon_\delta^2 T_{eq}$ term, i.e.

$$\left\langle \left\{S_1, \mathcal{H}_1 + \bar{\mathcal{H}}_1\right\}_c \right\rangle = \frac{1}{B_{eq}}\frac{\partial}{\partial \bar{\mu}}\left\langle \widetilde{\mathcal{H}}_1^2 \right\rangle$$

where the property $\left\langle \widetilde{\mathcal{H}}_1^2 \right\rangle = \left\langle \left(\mathcal{H}_1^2 - \bar{\mathcal{H}}_1^2\right) \right\rangle$ has been used. The bracket $\left\{S_1, \mathcal{H}_1 + \bar{\mathcal{H}}_1\right\}_{\bar{\mathbf{Z}}^*}$, given by Eq.(132), does not raise special difficulty. The first term, which involves parallel gradients of fluctuations provides a term of order $\epsilon_\delta^3 T_{eq}$, which does not contribute. The last term has the right scaling $\epsilon_\delta^2 T_{eq}$

$$\left\{S_1, \mathcal{H}_1 + \bar{\mathcal{H}}_1\right\}_{\bar{\mathbf{Z}}^*} = -\frac{1}{2B_{eq}}\left\langle \frac{\mathbf{e}_{||}}{e_a B_{||}^*} \cdot \left(\frac{\partial S_1}{\partial \bar{\mathbf{X}}} \times \frac{\partial\left(\mathcal{H}_1 + \bar{\mathcal{H}}_1\right)}{\partial \bar{\mathbf{X}}}\right) \right\rangle + o(\epsilon_\omega)$$

[56]Let us recall that $\langle \mathcal{H} \rangle_{\bar{\varsigma}}$ is also the gyroaverage of $\mathcal{H}(\bar{\boldsymbol{\mathcal{Z}}})$ noted $\mathcal{J}[\mathcal{H}]$. Both notations are used.

Now because it is the lowest order expression of S_1 that should be used in this expression, which is a periodic function of $\bar{\varsigma}$, the contribution of the term $\bar{\mathcal{H}}_1$ vanishes because $\bar{\mathcal{H}}_1$ does not depend on $\bar{\varsigma}$ and S_1 is periodic in $\bar{\varsigma}$ with zero mean. For the same reason, $\mathcal{H}_1$ can be replaced by $\widetilde{\mathcal{H}}_1$. Hence the final expression of the second order Hamiltonian is

$$\bar{\mathcal{H}}_2 = \left\langle \widetilde{\mathcal{H}}_2 \right\rangle - \frac{1}{2B_{eq}} \frac{\partial}{\partial \bar{\mu}} \left\langle \widetilde{\mathcal{H}}_1^2 \right\rangle - \frac{1}{2} \left\langle \left\{ S_1, \widetilde{\mathcal{H}}_1 \right\}_{\bar{\mathbf{Z}}^*} \right\rangle \tag{133}$$

where

$$\left\{ S_1, \widetilde{\mathcal{H}}_1 \right\}_{\bar{\mathbf{Z}}^*} = - \left\langle \frac{\mathbf{e}_\parallel}{e_a B_\parallel^*} \cdot \left(\frac{\partial S_1}{\partial \bar{\mathbf{X}}} \times \frac{\partial \widetilde{\mathcal{H}}_1}{\partial \bar{\mathbf{X}}} \right) \right\rangle$$

and S_1 is given by Eq.(131). The last term of Eq.(133) cannot be neglected in principle. However it is in practice a small term. This can be understood as follows. Indeed following the recipe of Appendix E, $\widetilde{\mathcal{H}}_1$ can be developed in the Fourier space

$$\widetilde{\mathcal{H}}_1 = \sum_{n_1=-\infty}^{+\infty} \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} J_{n_1}\left(k_\perp \rho_c\right) \widetilde{\mathcal{H}}_{1\boldsymbol{k}}\left(\bar{\mu}, \bar{p}_\parallel, t\right) e^{i\boldsymbol{k}\cdot\bar{\mathbf{X}} + in_1(\bar{\varsigma}+\varphi_{\boldsymbol{k}})}$$

Using the relation $\frac{\partial S_1}{\partial \bar{\varsigma}} = \frac{1}{\Omega_c}\widetilde{\mathcal{H}}$ provides a similar expansion for the generating function

$$S_1 = \sum_{n_1=-\infty}^{+\infty} \frac{1}{in_1\Omega_c} \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} J_{n_1}\left(k_\perp \rho_c\right) \widetilde{\mathcal{H}}_{1\boldsymbol{k}}\left(\bar{\mu}, \bar{p}_\parallel, t\right) e^{i\boldsymbol{k}\cdot\bar{\mathbf{X}} + in_1(\bar{\varsigma}+\varphi_{\boldsymbol{k}})}$$

This gives the following identity

$$\begin{aligned} \left\{ S_1, \widetilde{\mathcal{H}}_1 \right\}_{\bar{\mathbf{Z}}^*} \;=\; & - \sum_{n_1=-\infty, n_1\neq 0}^{+\infty} \frac{1}{in_1\Omega_c} \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \frac{d^3\boldsymbol{k}'}{(2\pi)^3} J_{n_1}\left(k_\perp \rho_c\right) J_{n_1}\left(k'_\perp \rho_c\right) \\ & \frac{\mathbf{e}_\parallel}{e_a B_\parallel^*} \cdot \left(\boldsymbol{k} \times \boldsymbol{k}'\right) \widetilde{\mathcal{H}}_{1\boldsymbol{k}} \widetilde{\mathcal{H}}^*_{1\boldsymbol{k}'} e^{i\left[(\boldsymbol{k}-\boldsymbol{k}')\cdot\bar{\mathbf{X}} + n_1(\varphi_{\boldsymbol{k}}-\varphi_{\boldsymbol{k}'})\right]} \end{aligned} \tag{134}$$

This sum is non null, and of order 2 in ϵ_δ^2. It is sometimes called "ponderomotive" force. However, it is expected to be small in practical cases. Indeed for small value of its argument $k_\perp\rho_c$, the Bessel function of index n varies as $(k_\perp\rho_c)^n$ and is small because the sum involves only $n_1 \neq 0$ terms. Same for $J_{n_1}\left(k'_\perp\rho_c\right)$. In the limit of large arguments, $J_{n_1}\left(k_\perp\rho_c\right)$ and $J_{n_1}\left(k'_\perp\rho_c\right)$ are oscillating functions of their argument - the integral over the wave numbers $\boldsymbol{k}$ and $\boldsymbol{k}'$ concentrate most of the contribution to the line $\boldsymbol{k} = \boldsymbol{k}'$ for which the cross-product is small.

H Non canonical gyrokinetic transform

H.1 Action principle for gyrocentre equations of motion

We start first with a change of non canonical coordinates at first order in the perturbed field, before moving to a generalisation at all orders. Let us remember that the equations of motion of a particle guiding-centre satisfy a principle of minimum action $\mathcal{A}_{part} = \int L dt$. This formulation can be recast as an action principle that uses the 1-form $\gamma = (m_a\mathbf{v} + e_a\mathbf{A}) \cdot d\mathbf{x} - H dt$. This form can be written in guiding-centre coordinates. It is noted Γ, in accordance with the conventions adopted throughout the paper, and depends on $(\mathbf{X}, u_\parallel, \mu, \varsigma, t, E)$[57]. The objective is to derive a new gyrocentre form $\bar{\Gamma}$ such that the new magnetic moment $\bar{\mu}$ is a invariant of motion. The form Γ contains unperturbed and

[57]We choose here the non canonical variables t and E since the whole section is dedicated to non canonical variables

perturbed parts, i.e. $\Gamma = \Gamma_0 + \epsilon_\delta \Gamma_1$. The unperturbed part is the one derived by Littlejohn [7]

$$\Gamma_0 = e_a \mathbf{A}^*_{eq} \cdot d\mathbf{X} + \frac{m_a}{e_a} \mu d\varsigma - H_0 dt$$

where[58]

$$H_0 = \frac{1}{2} m_a u_\parallel^2 + \mu B_{eq}(\mathbf{X}) + e_a \phi_{eq}(\mathbf{X})$$

and the symplectic vector potential is

$$\mathbf{A}^*_{eq}(\mathbf{X}) = \mathbf{A}_{eq}(\mathbf{X}) + \frac{m_a u_\parallel}{e_a} \mathbf{e}_\parallel(\mathbf{X})$$

The perturbed part is

$$\Gamma_1 = e_a \tilde{\mathbf{A}}(\mathbf{X} + \boldsymbol{\rho}, t) \cdot d(\mathbf{X} + \boldsymbol{\rho}) - e_a \tilde{\phi}(\mathbf{X} + \boldsymbol{\rho}, t) dt$$

Terms can be regrouped to yield the total form

$$\Gamma = e_a \left(\mathbf{A}^*_{eq} + \tilde{\mathbf{A}} \right) \cdot d\mathbf{X} + e_a \tilde{\mathbf{A}} \cdot \frac{\partial \boldsymbol{\rho}}{\partial \mu} d\mu + \left(\frac{m_a}{e_a} \mu + e_a \tilde{\mathbf{A}} \cdot \frac{\partial \boldsymbol{\rho}}{\partial \varsigma} \right) d\varsigma - H dt \tag{135}$$

where $H = H_0 + H_1$ and $H_1 = e_a \tilde{\phi}(\mathbf{X} + \boldsymbol{\rho}, t)$. Hence the covariant components of the form Γ_1 are[59]

$$\Gamma_{1,\alpha} = e_a \tilde{\mathbf{A}}_{gc,\alpha} \quad ; \quad \Gamma_{1,\mu} = e_a \tilde{\mathbf{A}}_{gc} \cdot \frac{\partial \boldsymbol{\rho}}{\partial \mu} \quad ; \quad \Gamma_{1,\varsigma} = e_a \tilde{\mathbf{A}}_{gc} \cdot \frac{\partial \boldsymbol{\rho}}{\partial \varsigma} \quad ; \quad \Gamma_{1,t} = -e_a \tilde{\phi}_{gc}$$

where $\tilde{\mathbf{A}}_{gc}$ and $\tilde{\phi}_{gc}$ are the vector and electric potentials written in guiding-centre coordinates, i.e.

$$\begin{aligned} \tilde{\mathbf{A}}_{gc}(\mathbf{X}, \mu, \varsigma) &= \tilde{\mathbf{A}}(\mathbf{X} + \boldsymbol{\rho}) \\ \tilde{\phi}_{gc}(\mathbf{X}, \mu, \varsigma) &= \tilde{\phi}_a(\mathbf{X} + \boldsymbol{\rho}) \end{aligned}$$

Eq.(135) clearly shows that μ is no longer an invariant of motion in presence of a perturbed electromagnetic field. The objective remains to build a new invariant of motion. Formally the form Γ can be written as $\Gamma = \Gamma_i d\mathcal{Z}^i$ in the extended phase space, where Γ is now a 1-form that lives in the 8D extended phase space. The aim of the gyrocentre transform is to find a set of variables $\bar{\mathcal{Z}}^i = (\bar{\mathbf{X}}, \bar{p}_\parallel, \bar{\mu}, \bar{\varsigma})$ such that the new Hamiltonian does not depend on the gyroangle $\bar{\varsigma}$. The new form is $\bar{\Gamma} = \bar{\Gamma}_i d\bar{\mathcal{Z}}^i$, with some flexibility on the choice of $\bar{\Gamma}_i$. The two forms Γ and $\bar{\Gamma}$ may differ by the differential of a scalar function S, i.e. $\bar{\Gamma}_i d\bar{\mathcal{Z}}^i = \Gamma_i d\mathcal{Z}^i - dS$, or formally $\Gamma = \bar{\Gamma} + dS$. This is sometimes called a "gauge transformation" - instead of generating function - we will use indifferently both names. The forms Γ and $\bar{\Gamma}$ and the function S are expanded in ϵ_δ

$$\begin{aligned} \Gamma(\mathcal{Z}) &= \sum_{n=0}^{+\infty} \Gamma_n(\mathcal{Z}) \epsilon_\delta^n \\ \bar{\Gamma}(\bar{\mathcal{Z}}) &= \sum_{n=0}^{+\infty} \bar{\Gamma}_n(\bar{\mathcal{Z}}) \epsilon_\delta^n \\ S(\mathcal{Z}) &= \sum_{n=0}^{+\infty} S_n(\mathcal{Z}) \epsilon_\delta^n \end{aligned}$$

From the expression above, it appears that $\Gamma_n = 0$, for $n \geq 2$. As mentioned in a previous section for the distribution functions $\mathcal{F}$ and $\bar{\mathcal{F}}$, the two forms Γ and $\bar{\Gamma}$ cannot be compared

[58] To simplify we assume that the equilibrium quantities do not depend on time. In fact, slow variations are allowed.

[59] Greek indices indicate spatial coordinates.

directly, since one is a function of $\mathcal{Z}$, while the other one is a function $\bar{\mathcal{Z}}$. Two compare Γ and $\bar{\Gamma}$, one has to apply rules for change of variables. The coordinate transform from $\mathcal{Z}$ to $\bar{\mathcal{Z}}$ is a push-forward transform, that will be noted T, i.e.

$$\bar{\mathcal{Z}} = T\mathcal{Z}$$

The operator that transform a function $\bar{\Gamma}(\bar{\mathcal{Z}})$ to $\Gamma(\mathcal{Z})$ is called a pull-back operator $\mathcal{T}^*$, i.e.

$$\bar{\Gamma} = \mathcal{T}^*\Gamma$$

For a scalar function, e.g. the Hamiltonian (or the distribution function), one has $\mathcal{H}(\mathcal{Z}) = \bar{\mathcal{H}}(\bar{\mathcal{Z}})$, and therefore

$$\bar{\mathcal{H}}(T\mathcal{Z}) = \mathcal{H}(\mathcal{Z}) = \left[\mathcal{T}^*\bar{\mathcal{H}}\right](\bar{\mathcal{Z}})$$

H.2 First order theory

Let us now consider a first order change of variables

$$\bar{\mathcal{Z}}^i = \mathcal{Z}^i + \epsilon_\delta G_1^i(\mathcal{Z})$$

Using

$$\begin{aligned}
\Gamma_i(\mathcal{Z}) &= \Gamma_i(\bar{\mathcal{Z}}) - \epsilon_\delta G_1^j \frac{\partial \Gamma_i}{\partial \mathcal{Z}^j}(\bar{\mathcal{Z}}) \\
d\mathcal{Z}^i &= d\bar{\mathcal{Z}}^i - \epsilon_\delta \frac{\partial G_1^i}{\partial \mathcal{Z}^j}(\bar{\mathcal{Z}}^j) d\bar{\mathcal{Z}}^j
\end{aligned}$$

Equating the various orders, one finds at lowest order

$$\bar{\Gamma}_0 = \Gamma_0 + dS_0$$

The next order yields

$$\bar{\Gamma}_{1i} = \Gamma_{1i} - G_1^j \frac{\partial \Gamma_{0i}}{\partial \mathcal{Z}^j} - \Gamma_{0j} \frac{\partial G_1^j}{\partial \mathcal{Z}^i} + \frac{\partial S_1}{\partial \mathcal{Z}^i}$$

The term $-\Gamma_{0j}\frac{\partial G_1^j}{\partial \mathcal{Z}^i}$ can be written as $-\frac{\partial}{\partial \mathcal{Z}^i}\left(\Gamma_{0j}G_1^j\right) + G_1^j\frac{\partial \Gamma_{0j}}{\partial \mathcal{Z}^i}$. The first term can be incorporated in the gauge function S (the name remains unchanged to avoid a multiplication of notations), so that finally

$$\bar{\Gamma}_1 = \Gamma_1 - L_1\Gamma_0 + dS_1$$

where $L_1\Gamma_0$ is called a Lie derivative [26, 27, 28, 2] of the form Γ_0 and is defined as

$$\left(L_1\Gamma_0\right)_i = G_1^j\left\{\frac{\partial \Gamma_{0i}}{\partial \mathcal{Z}^j} - \frac{\partial \Gamma_{0j}}{\partial \mathcal{Z}^i}\right\} = G_1^j\omega_{0,ji} = \mathbf{G}_1 \cdot \boldsymbol{\omega}$$

where $\omega_{0,ji}$ is the Lagrange brackets associated with the form Γ_0. Hence one arrives to the equation

$$\mathbf{G}_1 \cdot \boldsymbol{\omega} = \Gamma_1 - \bar{\Gamma}_1 + dS_1$$

Let us introduce the Poisson brackets $J_0^{ij} = \left\{\mathcal{Z}^i, \mathcal{Z}^j\right\}_0$. The elements J_0^{ij} are Poisson brackets and are the components of a tensor. Its inverse is also a tensor, the elements of which are Lagrange brackets $\omega_{0,jk}$. Hence Poisson and Lagrange brackets satisfy the relationship $J_0^{ij}\omega_{0,jk} = \delta_k^i$. This provides an explicit expression of the displacement $\mathbf{G}_1$

$$G_1^i = \left\{S_1, \mathcal{Z}^i\right\}_0 + \left(\Gamma_{1j} - \bar{\Gamma}_{1j}\right) J_0^{ji}$$

Since $\bar{\mathcal{Z}}^i = \mathcal{Z}^i + G_1^i$, and $\bar{\mathcal{H}}(\bar{\mathcal{Z}}) = \mathcal{H}(\mathcal{Z})$, one has $\bar{\mathcal{H}}_1(\mathcal{Z}) = \mathcal{H}_1(\mathcal{Z}) - G_1^i\frac{\partial \mathcal{H}_0}{\partial Z^i}$ or equivalently

$$\bar{\mathcal{K}}_1(\mathcal{Z}) = \mathcal{K}_1(\mathcal{Z}) - \{S_1, \mathcal{H}_0\}_0$$

where $\mathcal{K}_1 = \mathcal{H}_1 - \dot{\mathcal{Z}}_0^i \Gamma_{1i}$ is an effective Hamiltonian, and $\bar{\mathcal{K}}_1 = \bar{\mathcal{H}}_1 - \dot{\bar{\mathcal{Z}}}_0^i \bar{\Gamma}_{1i}$ its gyrocentre counterpart[60]. We now impose that the new effective Hamiltonian is independent of the gyroangle. In the gyrokinetic ordering

$$\{S_1, \mathcal{H}_0\}_0 = \Omega_c \frac{\partial S_1}{\partial \varsigma} + o(\epsilon_\delta)$$

This imposes a constraint on $\bar{\mathcal{K}}_1$

$$\bar{\mathcal{K}}_1(\mathcal{Z}) = \mathcal{J}[\mathcal{K}_1(\mathcal{Z})]$$

and the expression of the generating function S_1

$$S_1 = \frac{1}{\Omega_c} \int^{\varsigma} d\varsigma' \left(\mathcal{K}_1 - \mathcal{J}[\mathcal{K}_1(\mathcal{Z})] \right)$$

H.3 "Hamiltonian" and "symplectic" approaches

We are now in position to derive the various expressions found in the literature. From the form Eq.(135) it appears a suitable choice of the gyrocentre form is[61]

$$\bar{\Gamma} = e_a \left(\mathbf{A}^*_{eq} + \epsilon_s \mathcal{J}[\tilde{\mathbf{A}}] \right) \cdot d\bar{\mathbf{X}} + \frac{m_a}{e_a} \bar{\mu} d\bar{\varsigma} - \bar{H} dt$$

where $\bar{H} = H_0 + \bar{H}_1$, and $\bar{H}_1$ has yet to be determined, and $\epsilon_s = 0$ or 1 is a switch, of which the meaning is given below. Using Eq.(136), the effective Hamiltonian $\mathcal{K}_1 = \mathcal{H}_1 - \dot{\mathcal{Z}}_0^i \Gamma_{1i}$ can be written[62]

$$\mathcal{K}_1 = e_a \tilde{\phi}_{gc} - e_a \left(\dot{\mathbf{X}} + \dot{\boldsymbol{\rho}} \right) \cdot \tilde{\mathbf{A}}_{gc}$$

The perturbed gyrocentre form is $\bar{\Gamma}_1 = e_a \epsilon_s \mathcal{J}[\tilde{\mathbf{A}}] \cdot d\bar{\mathbf{X}}$, so that the corresponding vector field $\bar{\Gamma}_{1\alpha}$ is defined by

$$\bar{\Gamma}_{1\alpha} = e_a \epsilon_s \mathcal{J}[\tilde{\mathbf{A}}]_\alpha$$

The case $\epsilon_s = 0$ corresponds to $\bar{\Gamma}_{1\alpha} = 0$. It is called "Hamiltonian" since the symplectic part of the form vanishes. In this case $\bar{\mathcal{K}}_1 = \bar{\mathcal{H}}_1$. Another choice consists in setting $\epsilon_s = 1$, such that $\bar{\Gamma}_1 = e_a \mathcal{J}[\tilde{\mathbf{A}}] \cdot d\bar{\mathbf{X}}$. In this case, $\bar{\Gamma}_{1\alpha} = e_a \mathcal{J}[\tilde{\mathbf{A}}]_\alpha$, and it is called "symplectic" because the symplectic part of the perturbed form, i.e. $\mathcal{J}[\tilde{\mathbf{A}}] \cdot d\bar{\mathbf{X}}$ is finite. The effective Hamiltonian reads

$$\bar{\mathcal{K}}_1 = \bar{\mathcal{H}}_1 + e_a \dot{\mathbf{X}} \cdot \mathcal{J}[\tilde{\mathbf{A}}_{gc}]$$

Since $\bar{\mathcal{K}}_1 = \mathcal{J}[\mathcal{K}_1]$, one arrives to an explicit expression of the gyrocentre Hamiltonian

$$\bar{\mathcal{H}}_1 = e_a \mathcal{J}[\tilde{\phi}_{gc}] - e_a \mathcal{J}\left[\left(\dot{\mathbf{X}} + \dot{\boldsymbol{\rho}} \right) \cdot \tilde{\mathbf{A}}_{gc} \right] + \epsilon_s e_a \dot{\mathbf{X}} \cdot \mathcal{J}[\tilde{\mathbf{A}}_{gc}]$$

The displacement $\mathbf{G}_1$ is given by the relation

$$G_1^i = \left\{ S_1, \mathcal{Z}^i \right\}_0 + e_a \tilde{\mathbf{A}}_{gc} \cdot \left\{ \mathbf{X} + \boldsymbol{\rho}, \mathcal{Z}^i \right\}_0 - \epsilon_s e_a \mathcal{J}[\tilde{\mathbf{A}}_{gc}] \cdot \left\{ \mathbf{X}, \mathcal{Z}^i \right\}_0 \tag{136}$$

[60] The velocity $\dot{\mathcal{Z}}_0^i$ must be understood here as $\dot{\mathcal{Z}}_0^i = -\left\{ \mathcal{H}_0, \mathcal{Z}^i \right\}$, i.e. the velocity of the unperturbed motion, hence the index 0 - in the same way $\dot{\bar{\mathcal{Z}}}_0^i = -\left\{ \mathcal{H}_0, \bar{\mathcal{Z}}^i \right\}$

[61] The full vector potential is implemented here in the symplectic part when $\epsilon_S = 1$. An alternative is to only take a part of it, .e.g. $\tilde{A}_\parallel$, and put the rest in the Hamiltonian.

[62] The reader should be warned that notations have been changed compared with the rest of the document. The perturbed Hamiltonian is $H_1 = e_a \tilde{\phi}$, i.e. the electric potential only, whereas the perturbed Hamiltonian at order one was until now $e_a \tilde{\phi} - e_a \left(\dot{\mathbf{X}} + \dot{\boldsymbol{\rho}} \right) \cdot \tilde{\mathbf{A}}(\mathbf{X} + \boldsymbol{\rho})$, i.e. closer to the field $\mathcal{K}_1$. Also in the Hamiltonian approach, a second order Hamiltonian $\frac{e_a^2}{2m_a} \tilde{\mathbf{A}}^2$ pops in, whereas in the present approach there is no perturbed Hamiltonian $\mathcal{H}_2$ nor form Γ_2.

The structure of the Poisson bracket in the guiding centre extended phase space is reminded

$$\begin{aligned}\{\mathcal{F},\mathcal{G}\}_{gc} &= \frac{e_a}{m_a}\left(\frac{\partial\mathcal{F}}{\partial\varsigma}\frac{\partial\mathcal{G}}{\partial\mu}-\frac{\partial\mathcal{F}}{\partial\mu}\frac{\partial\mathcal{G}}{\partial\varsigma}\right)\\ &+\frac{1}{m_a}\frac{\mathbf{B}^*}{B^*_{||}}\cdot\left(\frac{\partial\mathcal{F}}{\partial\mathbf{X}}\frac{\partial\mathcal{G}}{\partial u_{||}}-\frac{\partial\mathcal{F}}{\partial u_{||}}\frac{\partial\mathcal{G}}{\partial\mathbf{X}}\right)\\ &-\frac{\mathbf{e}_{||}}{e_aB^*_{||}}\cdot\left(\frac{\partial\mathcal{F}}{\partial\mathbf{X}}\times\frac{\partial\mathcal{G}}{\partial\mathbf{X}}\right)\\ &+\left(\frac{\partial\mathcal{F}}{\partial E}\frac{\partial\mathcal{G}}{\partial t}-\frac{\partial\mathcal{F}}{\partial t}\frac{\partial\mathcal{G}}{\partial E}\right)\end{aligned} \tag{137}$$

The spatial and parallel velocity components of the displacement G_1 play an important role in the theory. Their expressions can be obtained by combining Eq.(136) and Eq.(137) [27]

$$\begin{aligned}G_1^{\mathbf{X}} &= -\frac{1}{m_a}\frac{\partial S_1}{\partial u_{||}}\frac{\mathbf{B}^*}{B^*_{||}}-\frac{\mathbf{e}_{||}}{e_aB^*_{||}}\times\nabla S_1\\ &\quad -\frac{\mathbf{e}_{||}}{e_aB^*_{||}}\times\left(\tilde{\mathbf{A}}_{gc}-\epsilon_s\mathcal{J}[\tilde{\mathbf{A}}_{gc}]\right)\\ G_1^{p_{||}} &= \left(\mathbf{e}_{||}\cdot\nabla\right)S_1+\frac{e_a}{m_a}\left(\tilde{A}_{gc,||}-\epsilon_s\mathcal{J}[\tilde{A}_{gc,||}]\right)\end{aligned}$$

The other components yield the modifications of the magnetic moment, and gyroangle. It is reminded here that the parallel gradient of fluctuations is smaller and introduces an extra ϵ_δ so that the relationship between the new gyrocentre parallel "velocity" and the guiding-centre velocity is given by the relationship

$$\bar{p}_{||}=u_{||}+\frac{e_a}{m_a}\left(\tilde{A}_{gc,||}-\epsilon_s\mathcal{J}[\tilde{A}_{gc,||}]\right)$$

It appears that the relation between $\bar{p}_{||}$ and $u_{||}$ is $\bar{p}_{||}=u_{||}+\frac{e_a}{m_a}\tilde{A}_{gc,||}$ in the Hamiltonian case. Hence the gyrocentre velocity $\bar{p}_{||}$ coincides with the "Hamiltonian velocity" $\bar{p}_{||}$ used in this note. This approach provides a rigorous derivation of the gyrocentre parallel velocity. In the symplectic approach, the gyrocentre parallel velocity is different, related to $p_{||}$ via the relation

$$\bar{p}_{||}=u_{||}+\frac{e_a}{m_a}\left(\tilde{A}_{gc,||}-\mathcal{J}[\tilde{A}_{gc,||}]\right)$$

H.4 Arbitrary order

This scheme is extended to arbitrary order by writing the push-forward operator T, defined as $\bar{\mathcal{Z}}=T\mathcal{Z}$, as a product of successive transforms of increasing order, i.e.

$$T=\prod_{n=0}^{\infty}T_n=...T_nT_{n-1}...T_2T_1T_0$$

where $T_n=\exp\left(\epsilon_\delta^nG_n\right)$. It then appears that the pullback operator T^*, that transform the function $\bar{\Gamma}(\bar{\mathcal{Z}})$ into a function $\Gamma(\mathcal{Z})$ is

$$\mathcal{T}^*=\mathcal{T}_1^*\mathcal{T}_2^*...\mathcal{T}_{n-1}^*\mathcal{T}_n^*...$$

where $\mathcal{T}_n^*=\exp(\epsilon_\delta^nL_n)$, and L_n is the Lie derivative associated with G_n. Note that the succession of $\mathcal{T}_n^*$ moves in opposite order compared with the T_n. Since we need to express the new form $\bar{\Gamma}$ as a function of the old one Γ, and therefore the inverse pullback operator $\mathcal{T}^{*-1}$, which reads

$$\mathcal{T}^{*-1}=...\mathcal{T}_n^{*-1}\mathcal{T}_{n-1}^{*-1}...\mathcal{T}_2^{*-1}\mathcal{T}_1^{*-1}$$

where $\mathcal{T}_n^{*-1} = \exp\{-\epsilon_\delta^n L_n\}$. When applied to the form $\Gamma = \mathcal{T}^*\bar{\Gamma}$, this yields the following relations [26, 27, 28, 2]

$$\begin{aligned}
\bar{\Gamma}_0 &= \Gamma_0 + dS_0 \\
\bar{\Gamma}_1 &= \Gamma_1 - L_1\Gamma_0 + dS_1 \\
\bar{\Gamma}_2 &= \Gamma_2 - L_1\Gamma_1 + \left(\frac{1}{2}L_1^2 - L_2\right)\Gamma_0 + dS_2
\end{aligned}$$

where $L_n\Gamma_p = \mathbf{G}_n \cdot \boldsymbol{\omega}_p$. The pullback operator acts on a scalar in a way similar to its action on a form, but with the Lie derivative defined as $L_n = G_n^i \frac{\partial}{\partial \mathcal{Z}^i}$ instead of a scalar product with Lagrange brackets. Hence the new Hamiltonian $\bar{\mathcal{H}} = \mathcal{T}^{*-1}\mathcal{H}$ verifies the following sequence

$$\begin{aligned}
\bar{\mathcal{H}}_0 &= \mathcal{H}_0 \\
\bar{\mathcal{H}}_1 &= \mathcal{H}_1 - L_1\mathcal{H}_0 \\
\bar{\mathcal{H}}_2 &= \mathcal{H}_2 - L_1\mathcal{H}_1 + \left(\frac{1}{2}L_1^2 - L_2\right)\mathcal{H}_0
\end{aligned}$$

These relations allow to solve the problem at all orders, at least formally.

H.5 Equations of motion

It must be realised that the guiding-centre Poisson bracket must be modified to account for the modification of the symplectic form by the perturbed vector potential $\tilde{\mathbf{A}}$ in the symplectic case. The new effective magnetic field and Jacobian of the transform become[63]

$$\mathbf{B}^* = \nabla \times \mathbf{A}^* = \mathbf{B}_{eq} + \frac{m_a \bar{p}_\parallel}{e_a}\nabla \times \mathbf{e}_\parallel + \epsilon_s \nabla \times \mathcal{J}[\tilde{\mathbf{A}}_{gc}]$$

and

$$B_\parallel^* = \mathbf{e}_\parallel \cdot \mathbf{B}^* = B_{eq} + \frac{m_a \bar{p}_\parallel}{e_a}\mathbf{e}_\parallel \cdot \nabla \times \mathbf{e}_\parallel + \epsilon_s \mathbf{e}_\parallel \cdot \nabla \times \mathcal{J}[\tilde{\mathbf{A}}_{gc}]$$

where the index "eq" is used for unperturbed quantities, and $\mathbf{e}_\parallel$ remains equal to $\mathbf{B}_{eq}/B_{eq}$. The gyrocentre Poisson bracket in the phase space takes the form[64]

$$\begin{aligned}
\{\mathcal{F}, \mathcal{G}\}_{gc} &= \frac{e_a}{m_a}\left(\frac{\partial \mathcal{F}}{\partial \bar{\varsigma}}\frac{\partial \mathcal{G}}{\partial \bar{\mu}^*} - \frac{\partial \mathcal{F}}{\partial \bar{\mu}^*}\frac{\partial \mathcal{G}}{\partial \bar{\varsigma}}\right) \\
&\quad + \frac{1}{m_a}\frac{\mathbf{B}^*}{B_\parallel^*} \cdot \left(\bar{\nabla}^*\mathcal{F}\frac{\partial \mathcal{G}}{\partial \bar{p}_\parallel} - \frac{\partial \mathcal{F}}{\partial \bar{p}_\parallel}\bar{\nabla}^*\mathcal{G}\right) \\
&\quad - \frac{\mathbf{e}_\parallel}{e_a B_\parallel^*} \cdot \left(\bar{\nabla}^*\mathcal{F} \times \bar{\nabla}^*\mathcal{G}\right) \\
&\quad + \left(\frac{\partial \mathcal{F}}{\partial E}\frac{\partial \mathcal{G}}{\partial t} - \frac{\partial \mathcal{F}}{\partial t}\frac{\partial \mathcal{G}}{\partial E}\right)
\end{aligned} \tag{138}$$

where

$$\bar{\nabla}^*\mathcal{F} = \bar{\nabla}\mathcal{F} - \epsilon_s e_a \left(\frac{\partial \mathcal{J}\left[\tilde{\mathbf{A}}_{gc}\right]}{\partial t}\frac{\partial \mathcal{F}}{\partial E} - \frac{e_a}{m_a}\frac{\partial \mathcal{J}\left[\tilde{\mathbf{A}}_{gc}\right]}{\partial \bar{\mu}}\frac{\partial \mathcal{F}}{\partial \bar{\varsigma}}\right)$$

$$\frac{\partial \mathcal{F}}{\partial \bar{\mu}^*} = \frac{\partial \mathcal{F}}{\partial \bar{\mu}} - \epsilon_s e_a \mathbf{e}_\parallel \cdot \left(\frac{\partial \mathcal{J}\left[\tilde{\mathbf{A}}_{gc}\right]}{\partial t} \times \frac{\partial \mathcal{J}\left[\tilde{\mathbf{A}}_{gc}\right]}{\partial \mu}\right)\frac{\partial \mathcal{F}}{\partial E}$$

[63] Note that $\bar{p}_\parallel$ replaces $u_\parallel$ in $\mathbf{B}^*$ and the Jacobian $B_\parallel^*$ since $\bar{\Gamma}_0(\bar{\mathcal{Z}}) = \Gamma_0(\bar{\mathcal{Z}})$. Same remark for the Hamiltonian $\bar{\mathcal{H}}_0$, equal to $\mathcal{H}_0$ with $u_\parallel$ replaced by $\bar{p}_\parallel$.

[64] The Poisson bracket is the inverse of the tensor $\omega_0 + \omega_1$, where $\omega_1 = e_a \mathcal{J}\left[\tilde{\mathbf{A}}_{gc}\right]$. At first order, the inverse of this matrix is $J_0 - J_0\omega_1 J_0$, where $J_0 = \omega_0^{-1}$ is the zeroth order Poisson bracket $\left\{\mathcal{Z}^i, \mathcal{Z}^j\right\}_0$.

and $\bar{\nabla}\mathcal{F} = \frac{\partial \mathcal{F}}{\partial \bar{\mathbf{X}}}$. The trajectory equations are given by the equations

$$\begin{aligned}
\frac{d\bar{\mathbf{X}}}{dt} &= -\left\{\bar{H}, \bar{\mathbf{X}}\right\} = \frac{1}{m_a}\frac{d\bar{H}}{d\bar{p}_\parallel}\frac{\mathbf{B}^*}{B^*_\parallel} + \frac{1}{e_a B^*_\parallel}\mathbf{e}_\parallel \times \left(\nabla \bar{H} + \epsilon_s \frac{\partial \mathcal{J}[\tilde{\mathbf{A}}_{gc}]}{\partial t}\right) \\
m_a \frac{d\bar{p}_\parallel}{dt} &= -\left\{\bar{H}, m_a \bar{p}_\parallel\right\} = -\frac{\mathbf{B}^*}{B^*_\parallel} \cdot \left(\nabla \bar{H} + \epsilon_s \frac{\partial \mathcal{J}[\tilde{\mathbf{A}}_{gc}]}{\partial t}\right)
\end{aligned}$$

where $\bar{H} = \bar{H}_0 + \epsilon_\delta \bar{H}_1 + \epsilon_\delta^2 \bar{H}_2$, with

$$\begin{aligned}
\bar{H}_0 &= \frac{1}{2} m_a \bar{p}_\parallel^2 + \bar{\mu} B_{eq}(\bar{\mathbf{X}}) + e_a \phi_{eq}(\bar{\mathbf{X}}) \\
\bar{H}_1 &= e_a \mathcal{J}[\tilde{\phi}_{gc}] - (1-\epsilon_s)\, e_a \bar{p}_\parallel \mathcal{J}[\tilde{A}_{gc,\parallel}] - e_a \mathcal{J}\left[\left(\dot{\mathbf{X}}_\perp + \dot{\boldsymbol{\rho}}\right) \cdot \tilde{\mathbf{A}}_{gc,\perp}\right]
\end{aligned}$$

The calculation of the second order Hamiltonian $\bar{H}_2$ is quite cumbersome, and can be found in the overview by Brizard and Hahm [2][65]. We give here the expression in the $(\tilde{\phi}, \tilde{A}_\parallel)$ representation

$$\begin{aligned}
\bar{H}_2 &= \frac{e_a^2}{2m_a}\mathcal{J}[\tilde{A}^2_{gc,\parallel}] - \epsilon_s \frac{e_a^2}{2m_a}\left(\mathcal{J}[\tilde{A}_{gc,\parallel}]\right)^2 \\
&- \frac{e_a^2}{2\Omega_c}\oint \frac{d\varsigma}{2\pi}\left\{\int^\varsigma d\varsigma' \tilde{\psi}, \tilde{\psi}\right\}
\end{aligned} \tag{139}$$

where

$$\tilde{\psi} = \tilde{\phi}_{gc} - (1-\epsilon_s)\,\bar{p}_\parallel \tilde{A}_{gc,\parallel}$$

The expression Eq.(59) is recovered in the "Hamiltonian" case $\epsilon_s = 0$ - more precisely the first line of Eq.(139) corresponds to $\left\langle \widetilde{\mathcal{H}}_2 \right\rangle$ in Eq.(59), while the second line is identical to the sum of the two last terms. These equations of motion coincide with Eqs.(24,25) when $\epsilon_s = 0$.

[65] The derivation of the general expression of the Hamiltonian $\bar{H}_2$ is given in Brizard and Hahm page 446 and reproduced here in extenso

$$\begin{aligned}
\bar{H}_2 &= -\frac{1}{2}\left\langle \{S_1, \{S_1, \mathcal{H}_0\}\}\right\rangle \\
&- \left\langle \Delta\Gamma_{1i}\left\{Z^i, \mathcal{K}_1\right\}\right\rangle \\
&- \frac{1}{2}\left\langle \Delta\Gamma_{1i} J_0^{ij}\left\{\Gamma_{1j} + \bar{\Gamma}_{1j}, \mathcal{H}_0\right\}\right\rangle \\
&- \frac{1}{2}\left\langle \Delta\Gamma_{1i}\left\{Z^i, \dot{\bar{\mathcal{Z}}}_0^i\right\}\left(\Gamma_{1j} + \bar{\Gamma}_{1j}\right)\right\rangle
\end{aligned}$$

where $\Delta\Gamma_{1i} = \Gamma_{1i} - \bar{\Gamma}_{1i}$, brackets are average over the gyroangle. The first line can easily be shown to be identical to the "ponderomotive" force with the same calculation done for the Hamiltonian theory at second order. Obtaining the term in $\tilde{A}^2_{gc,\parallel}$ is a bit trickier. The second line is readily shown to be equal to

$$\begin{aligned}
-\left\langle \Delta\Gamma_{1i}\left\{Z^i, \mathcal{K}_1\right\}\right\rangle &= -\frac{e_a^2}{m_a}\left\langle \left(\tilde{A}_{gc,\parallel} - \epsilon_s \mathcal{J}\left[\tilde{A}_{gc,\parallel}\right]\right)\left\{\ell_G, m_a p_\parallel\right\}\frac{\partial \tilde{\psi}}{\partial p_\parallel}\right\rangle \\
&= \frac{e_a^2}{m_a}\left(\mathcal{J}\left[\tilde{A}^2_{gc,\parallel}\right] - \epsilon_s\left(\mathcal{J}\left[\tilde{A}_{gc,\parallel}\right]\right)^2\right)
\end{aligned}$$

where ℓ_G is the position of the guiding-centre along the field line. The third line is easily shown to be equal to zero, while the last term reads

$$\begin{aligned}
-\frac{1}{2}\left\langle \Delta\Gamma_{1i}\left\{Z^i, \dot{\bar{\mathcal{Z}}}_0^i\right\}\left(\Gamma_{1j} + \bar{\Gamma}_{1j}\right)\right\rangle &= -\frac{e_a^2}{2m_a}\left\langle \left(\tilde{A}_{gc,\parallel} - \epsilon_s \mathcal{J}\left[\tilde{A}_{gc,\parallel}\right]\right)\left\{\ell_G, m_a p_\parallel\right\}\left(\tilde{A}_{gc,\parallel} + \epsilon_s \mathcal{J}\left[\tilde{A}_{gc,\parallel}\right]\right)\right\rangle \\
&= -\frac{e_a^2}{2m_a}\left(\mathcal{J}\left[\tilde{A}^2_{gc,\parallel}\right] - \epsilon_s\left(\mathcal{J}\left[\tilde{A}_{gc,\parallel}\right]\right)^2\right)
\end{aligned}$$

Combining the two results yield the result that was looked for.

I Useful relations for the derivation of energy conservation

I.1 Fast derivation of a global law of energy conservation

In this derivation, time plays a special role. Hence time dependences are left explicit to get useful conservation equations. We leave for a while the extended phase space. Let us multiply the gyrokinetic Vlasov equation Eq.(23) by $\bar{\mathcal{H}}$, integrate over the phase space, and sum over all species. The resulting equation is

$$\sum_{species} \int d\bar{\Gamma} \frac{\partial}{\partial t} \bar{\mathcal{F}}(\bar{\mathbf{Z}}, t) \bar{\mathcal{H}}(\bar{\mathbf{Z}}, t) = 0$$

where $d\bar{\Gamma} = \frac{B_{||}^*}{m_a} d^3\bar{\mathbf{X}}\, d\bar{p}_{||}\, d\bar{\mu}\, d\varsigma$ is the element of volume in the gyrocentre phase space[66]. The gyrocentre variables $\bar{\mathbf{Z}}$ are dummy variables in this integral. We can therefore replace them safely by guiding-centre variables $\mathbf{Z} = (\mathbf{X}, p_{||}, \mu)$. The Hamiltonian is split in unperturbed and first order perturbed components $\bar{\mathcal{H}}(\mathbf{Z}^*, t) = \bar{\mathcal{H}}_{eq}(\mathbf{Z}^*, t) + \bar{\mathcal{H}}_1(\mathbf{Z}^*, t)$. Defining a "particle energy" as[67]

$$\mathcal{E}_{part} = \sum_{species} \int d\Gamma^* \bar{\mathcal{F}}(\mathbf{Z}^*, t) \bar{\mathcal{H}}_{eq}(\mathbf{Z}^*)$$

and using the relationship $\bar{\mathcal{H}}_1(\mathbf{Z}^*, t) = \mathcal{J}[H_1](\mathbf{Z}^*, t)$, this equation is reformulated as

$$\frac{\partial}{\partial t} \sum_{species} \int d\Gamma^* \bar{\mathcal{F}}(\mathbf{Z}^*, t) \bar{\mathcal{H}}_{eq}(\mathbf{Z}^*) + \sum_{species} \int d\Gamma^* \frac{\partial}{\partial t} \bar{\mathcal{F}}_1(\mathbf{Z}^*, t) \mathcal{J}[\mathcal{H}_1](\mathbf{Z}^*, t) = 0 \qquad (140)$$

where $d\Gamma^* = \frac{B_{||}^*}{m_a} d^3\mathbf{X}\, 2\pi\, dp_{||}\, d\mu$. The Maxwell equations in the gyrokinetic framework are derived by having the action $\mathcal{A} = \mathcal{A}_{field} + \mathcal{A}_{part} + \mathcal{A}_{skin}$, given by Eqs.(53,54,54), extremum with respect to $(\tilde{\phi}, \tilde{\mathbf{A}})$. Hence the electromagnetic potentials $(\tilde{\phi}, \tilde{\mathbf{A}})$ are given by linear operators applied to charge and current densities. This is just the consequence of the linearity of the Maxwell equations that relate fields to sources, i.e. charge and current densities. A consequence is the following property, where one moves from the guiding-centre coordinates $\mathbf{Z}^* = (\mathbf{X}, p_{||}, \mu)$ to the "hybrid" set $\mathbf{z}^* = (\mathbf{x}, p_{||}, \mu)$

$$\begin{aligned} \int d\Gamma^* \frac{\partial}{\partial t} \bar{\mathcal{F}}_1(\mathbf{Z}^*, t) \mathcal{J}[\mathcal{H}_1](\mathbf{Z}^*, t) &= \int d\gamma^* \frac{\partial}{\partial t} \mathcal{J}^t[\bar{\mathcal{F}}_1](\mathbf{z}^*, t) \mathcal{H}_1(\mathbf{z}^*, t) \\ &= \frac{1}{2} \frac{\partial}{\partial t} \int d\gamma^* \mathcal{J}^t[\bar{\mathcal{F}}_1](\mathbf{z}^*, t) \mathcal{H}_1(\mathbf{z}^*, t) \\ &= \frac{1}{2} \frac{\partial}{\partial t} \int d\Gamma^* \bar{\mathcal{F}}_1(\mathbf{Z}^*, t) \mathcal{J}[\mathcal{H}_1(\mathbf{Z}^*, t)] \qquad (141) \end{aligned}$$

Combining yields Eqs.(140, 141, 66) yields the conservation equation Eq.(67).

[66] This identity is readily derived by noting that $\int d\Gamma \{F, G\}_{\mathbf{Z}} = 0$ for any functions F and G of $\bar{Z}$ and also using $\bar{\mathcal{H}} \{\bar{\mathcal{H}}, \bar{\mathcal{F}}\}_{\mathbf{Z}} = \frac{1}{2} \{\bar{\mathcal{H}}^2, \bar{\mathcal{F}}\}$.

[67] We will see that it differs significantly from the usual kinetic energy - hence this weird name.

I.2 Calculation of field-particle actions

The action that corresponds to the potential-charge density interaction $\int dt\, d^3\mathbf{x}\varrho\phi$ can be reworked as follows

$$\begin{aligned}
\int dt\, d^3\mathbf{x}\, \varrho(\mathbf{x},t)\phi(\mathbf{x},t) &= \sum_{species} e_a \int dt\, d^3\mathbf{x}\, d^3\mathbf{p}\, f(\mathbf{x},\mathbf{p},t)\phi(\mathbf{x},t) \\
&= \sum_{species} e_a \int d\bar{\mathcal{Z}}\, \bar{\mathcal{F}}(\bar{\mathcal{Z}})\phi(\bar{\mathbf{X}} + \bar{\boldsymbol{\rho}}(\bar{\mathbf{X}},\bar{\mu},\bar{\varsigma}),t) \\
&= \sum_{species} e_a \int d\mathcal{Z}\, \bar{\mathcal{F}}(\mathcal{Z})\phi(\mathbf{X} + \boldsymbol{\rho}_0(\mathbf{X},\mu,\varsigma),t) \\
&= \sum_{species} e_a \int dt\, d^3\mathbf{x}\, d^3\mathbf{V}\, \bar{\mathcal{F}}(\mathbf{x},\mu,p_\parallel,t)\mathcal{J}[\phi](\mathbf{x},\mu,t)
\end{aligned}$$

where $\bar{\boldsymbol{\rho}}$ has been replaced by $\boldsymbol{\rho}_0$ since the polarisation shift introduces higher corrections in the action. Similarly, the magnetic part of the field-particle action $\int d^3\mathbf{x}\, \tilde{\mathbf{J}}\cdot\tilde{\mathbf{A}}$ can be reworked as follows

$$\begin{aligned}
\int d^3\mathbf{x}\, \tilde{\mathbf{J}}(\mathbf{x},t)\cdot\tilde{\mathbf{A}}(\mathbf{x},t) &= \sum_{species} e_a \int dt\, d^3\mathbf{x}\, d^3\mathbf{p}\, f(\mathbf{x},\mathbf{p},t)\mathbf{v}(\mathbf{x},\mathbf{p},t)\cdot\tilde{\mathbf{A}}(\mathbf{x},t) \\
&= \sum_{species} e_a \int d\bar{\mathcal{Z}}\, \bar{\mathcal{F}}(\bar{\mathbf{X}},\bar{\mu},\bar{p}_\parallel) \\
&\quad \left(\dot{\bar{\mathbf{X}}}_\perp(\bar{\mathbf{X}},\bar{\mu},\bar{p}_\parallel) + \dot{\bar{\boldsymbol{\rho}}}(\bar{\mathbf{X}},\bar{\mu},\bar{\varsigma}) + p_\parallel \mathbf{e}_\parallel\right)\cdot\tilde{\mathbf{A}}(\bar{\mathbf{X}} + \bar{\boldsymbol{\rho}}(\bar{\mathbf{X}},\bar{\mu},\bar{\varsigma}),t) \\
&\quad - \sum_{species} \int d\bar{\mathcal{Z}}\, \bar{\mathcal{F}}_{eq}(\bar{\mathbf{X}},\bar{\mu},\bar{p}_\parallel)\frac{e_a^2}{m_a}\tilde{\mathbf{A}}^2(\bar{\mathbf{X}} + \bar{\boldsymbol{\rho}}(\bar{\mathbf{X}},\bar{\mu},\bar{\varsigma}),t) \\
&= \sum_{species} e_a \int d\mathcal{Z}\, \bar{\mathcal{F}}(\mathbf{X},\mu,p_\parallel) \\
&\quad \left(\mathcal{J}\left[\dot{\boldsymbol{\rho}}_0(\mathbf{X},\mu,\varsigma)\cdot\tilde{\mathbf{A}}_\perp(\mathbf{X} + \boldsymbol{\rho}_0(\mathbf{X},\mu,\varsigma),t)\right] + p_\parallel\mathcal{J}[\tilde{A}_\parallel](\mathbf{X},\mu)\right) \\
&\quad - \sum_{species} \int d\mathcal{Z}\, \bar{\mathcal{F}}_{eq}(\mathbf{X},\mu,p_\parallel)\frac{e_a^2}{m_a}\mathcal{J}[\tilde{\mathbf{A}}^2](\mathbf{X},\mu,t) \\
&= \sum_{species} e_a \int dt\, d^3\mathbf{x}\, d^3\mathbf{V}\, \bar{\mathcal{F}}\left(p_\parallel\mathcal{J}[\tilde{A}_\parallel] - \mu\mathcal{B}[\tilde{B}_\parallel]\right) \\
&\quad - \sum_{species} \int d^3\mathbf{x}\, d^3\mathbf{V}\, \bar{\mathcal{F}}_{eq}\frac{e_a^2}{m_a}\mathcal{J}[\tilde{\mathbf{A}}^2]
\end{aligned}$$

The term $\dot{\bar{\mathbf{X}}}_\perp\cdot\tilde{\mathbf{A}}$ is second order in ϵ_B, hence is multiplied by $\bar{\mathcal{F}}_{eq}$. This terms disappears through the average over time, since it is linear in $\tilde{\mathbf{A}}$.

References

[1] E. A. Frieman and Liu Chen. Nonlinear gyrokinetic equations for low frequency electromagnetic waves in general plasma equilibria. *The Physics of Fluids*, 25(3):502–508, 1982.

[2] A. J. Brizard and T. S. Hahm. Foundations of nonlinear gyrokinetic theory. *Rev. Mod. Phys.*, 79:421–468, Apr 2007.

[3] N. Tronko, A. Bottino, C. Chandre, and E. Sonnendruecker. Hierarchy of second order gyrokinetic hamiltonian models for particle-in-cell codes. *Plasma Physics and Controlled Fusion*, 59(6):064008, may 2017.

[4] N. Tronko and C. Chandre. Second-order nonlinear gyrokinetic theory: from the particle to the gyrocentre. *Journal of Plasma Physics*, 84(3):925840301, 2018.

[5] X. Garbet, Y. Idomura, L. Villard, and T.H. Watanabe. Gyrokinetic simulations of turbulent transport. *Nuclear Fusion*, 50(4):043002, mar 2010.

[6] V.I. Arnold. *Mathematical Methods of Classical Mechanics*. Springer, 1978.

[7] R. G. Littlejohn. Variational principles of guiding centre motion. *Journal of Plasma Physics*, 29(1):111–125, 1983.

[8] H. Goldstein. *Classical Mechanics*. Addison-Wesley, 1980.

[9] A. J. Lichtenberg and M. A. Lieberman. *Regular and Stochastic Motion*. Springer, 1983.

[10] J. R. Cary. Time dependent canonical transformations and the symmetry equals invariant theorem. *Journal of Mathematical Physics*, 18(12):2432–2435, 1977.

[11] H. Sugama. Gyrokinetic field theory. *Physics of Plasmas*, 7(2):466–480, 2000.

[12] Alain J. Brizard. New variational principle for the vlasov maxwell equations. *Phys. Rev. Lett.*, 84:5768–5771, Jun 2000.

[13] Alain J. Brizard. Variational principle for nonlinear gyrokinetic vlasov maxwell equations. *Physics of Plasmas*, 7(12):4816–4822, 2000.

[14] J.D. Jackson. *Classical Electrodynamics*. Wiley, 1975.

[15] L.D. Landau and E.M. Lifshitz. *The Classical Theory of Fields*. Pergamon Press, 1971.

[16] J P Graves, D Zullino, D Brunetti, S Lanthaler, and C Wahlberg. Reduced models for parallel magnetic field fluctuations and their impact on pressure gradient driven MHD instabilities in axisymmetric toroidal plasmas. *Plasma Physics and Controlled Fusion*, 61(10):104003, aug 2019.

[17] Ge Dong, Jian Bao, Amitava Bhattacharjee, Alain Brizard, Zhihong Lin, and Peter Porazik. Gyrokinetic particle simulations of the effects of compressional magnetic perturbations on drift-alfvenic instabilities in tokamaks. *Physics of Plasmas*, 24(8):081205, 2017.

[18] A. Deprit. Canonical transformations depending on a small parameter. *Celestial mechanics*, 1(1):12–30, 1969.

[19] Alain J. Brizard and Natalia Tronko. Exact momentum conservation laws for the gyrokinetic vlasov-poisson equations. *Physics of Plasmas*, 18(8):082307, 2011.

[20] E. Lanti, N. Ohana, N. Tronko, T. Hayward-Schneider, A. Bottino, B.F. McMillan, A. Mishchenko, A. Scheinberg, A. Biancalani, P. Angelino, S. Brunner, J. Dominski, P. Donnel, C. Gheller, R. Hatzky, A. Jocksch, S. Jolliet, Z.X. Lu, J.P. Martin Collar, I. Novikau, E. SonnendrÃ$\frac{1}{4}$cker, T. Vernay, and L. Villard. Orb5: A global electromagnetic gyrokinetic code using the pic approach in toroidal geometry. *Computer Physics Communications*, 251:107072, 2020.

[21] N. Tronko, A. Bottino, and E. Sonnendruecker. Second order gyrokinetic theory for particle-in-cell codes. *Physics of Plasmas*, 23(8):082505, 2016.

[22] L.D. Landau and E.M. Lifshitz. *Electrodynamics of continuous media*. Pergamon Press, 1971.

[23] R. Kleiber, R. Hatzky, A. Koenies, A. Mishchenko, and E. Sonnendruecker. An explicit large time step particle-in-cell scheme for nonlinear gyrokinetic simulations in the electromagnetic regime. *Physics of Plasmas*, 23(3):032501, 2016.

[24] A. Mishchenko, A. Bottino, R. Hatzky, E. Sonnendruecker, R. Kleiber, and A. Koenies. Mitigation of the cancellation problem in the gyrokinetic particle-in-cell simulations of global electromagnetic modes. *Physics of Plasmas*, 24(8):081206, 2017.

[25] John R Cary and Alain J Brizard. Hamiltonian theory of guiding-center motion. *Reviews of modern physics*, 81(2):693, 2009.

[26] R. G. Littlejohn. Hamiltonian perturbation theory in noncanonical coordinates. *Journal of Mathematical Physics*, 23(5):742–747, 1982.

[27] A. Brizard. Nonlinear gyrokinetic maxwell vlasov equations using magnetic coordinates. *Journal of Plasma Physics*, 41(3):541–559, 1989.

[28] T. S. Hahm. Nonlinear gyrokinetic equations for tokamak microturbulence. *The Physics of Fluids*, 31(9):2670–2673, 1988.