Current coupling scheme for reduced kinetic/hydrodynamic models in strongly-magnetized, low-beta plasmas with large currents

Stefan Possanner[1, a)]

Max Planck Institute for Plasma Physics, Boltzmannstraße 2, 85748 Garching, Germany[b)]

(Dated: 10 January 2019)

We develop model equations for the study of energetic-particle effects on ideal MHD stability in the context of the current-coupling approach. A two-species, electromagnetic fluid is considered as the bulk plasma, coupled to one Vlasov equation for the energetic particles. The targeted physical scenario relates to strongly-magnetized, low-beta plasmas in the quasi-neutral, low-Mach regime. We first derive an extended MHD system for the bulk fluid starting from the two-species Euler equations with small ion and electron inertia. Our generalized Ohm's law features pressure and inertia terms. Moreover, both pressure equations are kept, thus allowing for large bulk current densities (Hall MHD). On the kinetic side, and under the same ordering assumptions, we perform a drift-kinetic reduction using Littlejohn's technique of variational averaging. The derived drift-kinetic-MHD hybrid model satisfies an exact energy theorem and is suitable for the study of energetic particle effects on MHD stability in the described physical scenario. We finally state the dispersion relation for the new model and discuss its relevance with respect to existing hybrid approaches.

[a)]stefan.possanner@ma.tum.de; http://www-m16.ma.tum.de/Allgemeines/StefanPossanner

[b)]Technical University of Munich, Department of Mathematics, Boltzmannstraße 3, 85748 Garching, Germany

CONTENTS

I. INTRODUCTION

sss

II. MODEL DERIVATION

A. Multi-species plasma with energetic particles

In this section we derive a reduced kinetic/hydrodynamic model starting from the Vlasov-Euler-Maxwell equations for plasma composed of bulk electrons and ions (Euler equations) and of energetic ions (Vlasov equation). Maxwell's equations for the electromagnetic fields

$\mathbf{E}$ and $\mathbf{B}$ in SI units read

$$\nabla \times \mathbf{B} = \mu_0 \left(\varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \mathbf{j} \right) \qquad \text{Ampère's law}\,, \tag{1a}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \qquad \text{Faraday's law}\,, \tag{1b}$$

$$\nabla \cdot \mathbf{E} = \frac{\varrho}{\varepsilon_0} \qquad \text{Poisson equation}\,, \tag{1c}$$

$$\nabla \cdot \mathbf{B} = 0 \qquad \text{no magnetic monopoles}\,, \tag{1d}$$

where ε_0 and μ_0 stand for the vacuum permittivity and permeability, respectively, and $\varepsilon_0 \mu_0 = c^{-2}$ with $c \approx 3\ 10^8$ m/s denoting the speed of light in vacuum. From time to time it may be useful to write the fields $\mathbf{E}$ and $\mathbf{B}$ in terms of the electromagnetic potentials ϕ and $\mathbf{A}$, defined via (in SI units)

$$\mathbf{E} = -\nabla \phi - \frac{\partial \mathbf{A}}{\partial t}\,, \qquad \mathbf{B} = \nabla \times \mathbf{A}\,. \tag{2}$$

Maxwell's equations are coupled to the plasma via the charge density ϱ and the current density $\mathbf{j}$,

$$\varrho := \sum_{\mathrm{s}} q_{\mathrm{s}} n_{\mathrm{s}} + q_{\mathrm{h}} n_{\mathrm{h}}\,, \qquad \mathbf{j} := \sum_{\mathrm{s}} q_{\mathrm{s}} n_{\mathrm{s}} \mathbf{u}_{\mathrm{s}} + \mathbf{j}_{\mathrm{h}}\,, \tag{3}$$

where $s = i, e$ denotes the fluid species (ions or electrons), q_{s} the charge, n_{s} the number density, $\mathbf{u}_{\mathrm{s}}$ the mean flow and n_{h} and $\mathbf{j}_{\mathrm{h}}$ stand for the number density and current density of the energetic (hot) particles, respectively. The bulk of the plasma is assumed to be in local thermal equilibrium, satisfying the Euler equations under the Lorentz force:

$$\frac{\partial n_{\mathrm{s}}}{\partial t} + \nabla \cdot (n_{\mathrm{s}} \mathbf{u}_{\mathrm{s}}) = 0\,, \tag{4a}$$

$$\frac{\partial}{\partial t}(n_{\mathrm{s}} \mathbf{u}_{\mathrm{s}}) + \nabla \cdot (n_{\mathrm{s}} \mathbf{u}_{\mathrm{s}} \otimes \mathbf{u}_{\mathrm{s}}) + \frac{1}{m_{\mathrm{s}}} \nabla p_{\mathrm{s}} = \frac{q_{\mathrm{s}}}{m_{\mathrm{s}}} (n_{\mathrm{s}} \mathbf{E} + n_{\mathrm{s}} \mathbf{u}_{\mathrm{s}} \times \mathbf{B})\,, \tag{4b}$$

$$\frac{\partial p_{\mathrm{s}}}{\partial t} + \mathbf{u}_{\mathrm{s}} \cdot \nabla p_{\mathrm{s}} + \gamma\, p_{\mathrm{s}} \nabla \cdot \mathbf{u}_{\mathrm{s}} = 0\,. \tag{4c}$$

Here, m_{s} denotes the mass, p_{s} the pressure and $\gamma = 5/3$ is the heat capacity ratio. The energetic particles are assumed to satisfy the Vlasov equation:

$$\frac{\partial f_{\mathrm{h}}}{\partial t} + \mathbf{v} \cdot \nabla f_{\mathrm{h}} + \frac{q_{\mathrm{h}}}{m_{\mathrm{h}}} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \frac{\partial f_{\mathrm{h}}}{\partial \mathbf{v}} = 0\,. \tag{5}$$

The corresponding number density and current are defined by

$$n_{\mathrm{h}} := \int f_{\mathrm{h}}\, \mathrm{d}\mathbf{v}\,, \qquad \mathbf{j}_{\mathrm{h}} := q_{\mathrm{h}} \int \mathbf{v} f_{\mathrm{h}}\, \mathrm{d}\mathbf{v}\,. \tag{6}$$

The system (1)-(6) is closed and provides a rather sophisticated description of a multi-species plasma. Our first aim is the reduction of degrees of freedom of this system, thereby enabling a more efficient numerical solution of some relevant physics properties, while filtering the high-frequency phenomena.

B. Ordering assumptions

The reduction of the system (1)-(6) starts with its normalization and via the implementation of a suitable ordering scheme that reflects the physics under consideration. In what follows we indicate the characteristic size of a variable by a hat, e.g. $\widehat{\omega}$ stands for the characteristic frequency (inverse time scale) and $\widehat{u}$ stands for the characteristic mean flow of the system. We shall work with a low-frequency MHD ordering, characterized by the following assumptions:

(a) Non-relativistic: $v_{\mathrm{th,i}} \ll c$. The ion thermal velocity $v_{\mathrm{th,i}} = (k_{\mathrm{B}}T_{\mathrm{i}}/m_{\mathrm{i}})^{1/2}$, where k_{B} denotes the Boltzmann constant, is much smaller than the speed of light.

(b) Quasi-neutral: $\lambda_{\mathrm{D}} \ll \widehat{x}$. The characteristic length scale $\widehat{x}$ is much larger than the Debye length of the ions, $\lambda_{\mathrm{D}} = (\varepsilon_0 k_{\mathrm{B}} T_{\mathrm{i}}/(e^2\widehat{n}))^{1/2}$, where $\widehat{n}$ stands for the characteristic density of the bulk plasma.

(c) Low-frequency: $\widehat{\omega} \ll \Omega_{\mathrm{ci}}$. The considered frequencies are well below the ion cyclotron frequency $\Omega_{\mathrm{ci}} = e\widehat{B}/m_{\mathrm{i}}$, where e is the elementary charge.

(d) MHD ordering: $\widehat{E}/\widehat{B} \sim \widehat{u}$. The characteristic mean flow is on the scale of the $E\times B$-drift velocity.

(e) Low Mach: $\widehat{u} \ll v_{\mathrm{th,i}}$. The mean flow is much smaller than the ion thermal velocity.

(f) Rarefied energetic ions: $\widehat{n}_{\mathrm{h}} \ll \widehat{n}$. The energetic particles are ions with low density.

(g) Large currents: $\widehat{j}/e \sim \widehat{n}\widehat{u}$. The currents of bulk plasma and energetic particles are of the order of the bulk mean flow (center of mass velocity).

(h) Low plasma beta: $\beta = \widehat{n} k_{\mathrm{B}} T_{\mathrm{i}}/(2\widehat{B}^2\mu_0) \ll 1$.

We remark in particular the assumption (g) on the currents, which deviates from the usual assumption of small currents in conventional derivations of MHD models[2]. Indeed, large currents will lead to an "extended MHD" model in which both the ion and the electron pressure are featured, leading to a quasi-one-fluid description where charge density and current are determined from quasi-neutrality and Ampère's law, respectively. This is known as Hall-MHD.

Let us now quantify the above ordering assumptions. Important parameters in a magnetized plasma are the transit frequency ω_i and the Larmor radius ρ_i of the ions,

$$\omega_\mathrm{i} = \frac{v_\mathrm{th,i}}{\widehat{x}}\,, \qquad \rho_\mathrm{i} = \frac{v_\mathrm{th,i}}{\Omega_\mathrm{ci}}\,. \tag{7}$$

Hence, $\omega_\mathrm{i}/\Omega_\mathrm{ci} = \rho_\mathrm{i}/\widehat{x}$. We suppose $\widehat{\omega}^{-1}$ and $\widehat{x}$ to be macroscopic time and space scales, in particular

$$\varepsilon := \frac{\rho_\mathrm{i}}{\widehat{x}} \ll 1\,, \qquad \widehat{\omega} = \frac{\omega_\mathrm{i}^2}{\widehat{\omega}_\mathrm{c}} = \varepsilon\,\omega_\mathrm{i}\,, \tag{8}$$

where ε is a small ordering parameter. It represents the "degree of magnetization" of the plasma. The time scale $\widehat{\omega}^{-1}$ is known as the Bohm time scale, characterized also by $\widehat{\omega}/\Omega_\mathrm{ci} = \varepsilon^2$. We will be interested in modes with variations on these particular scales, ie.

$$\frac{\partial}{\partial t} = \widehat{\omega}\,\frac{\partial}{\partial t'}\,, \qquad \frac{\partial}{\partial x} = \frac{1}{\widehat{x}}\frac{\partial}{\partial x'}\,, \tag{9}$$

where $t' := t\,\widehat{\omega}$ and $x' := x/\widehat{x}$ are dimensionless variables of order one. The scale of the mean flow $\widehat{u}$ is set by these time and space scales and the scale for the electrostatic energy is given by the thermal energy ($T_\mathrm{i} = T_\mathrm{e}$):

$$\widehat{u} = \widehat{x}\,\widehat{\omega}\,, \qquad e\,\widehat{\phi} = k_\mathrm{B}T_\mathrm{e} = k_\mathrm{B}T_\mathrm{i}\,. \tag{10}$$

From the first of these relations we obtain $\widehat{u}/v_\mathrm{th,i} = \varepsilon$ while the second relation yields

$$\begin{aligned}
& e\,\widehat{\phi} = k_\mathrm{B}T_\mathrm{i} \\
\Leftrightarrow\quad & \frac{e\,\widehat{\phi}}{\widehat{x}} = \frac{k_\mathrm{B}T_\mathrm{i}}{\widehat{x}} \\
\Leftrightarrow\quad & \frac{\widehat{E}}{\widehat{B}} = \frac{k_\mathrm{B}T_\mathrm{i}}{e\,\widehat{B}\,\widehat{x}} = \frac{m_\mathrm{i}\,v_\mathrm{th,i}^2}{e\,\widehat{B}\,\widehat{x}} = \frac{\omega_\mathrm{i}}{\Omega_\mathrm{ci}}\,v_\mathrm{th,i} = \varepsilon\,v_\mathrm{th,i} \\
\Leftrightarrow\quad & \frac{\widehat{E}}{\widehat{u}\widehat{B}} = 1\,,
\end{aligned} \tag{11}$$

which is consistent with the MHD ordering $\widehat{E}/\widehat{B} \sim \widehat{u}$, assumption (d). The thermal speed of the energetic particles, described by the Vlasov equation (5), is assumed to be on the scale of the ion thermal speed, $\widehat{v} \sim v_{\mathrm{th,i}}$. To complete the ordering we quantify the non-relativistic assumption (a), the quasi-neutrality (b) and the rarefied ions assumption (f) as

$$\frac{v_{\mathrm{th,i}}}{c} \sim \varepsilon\,, \qquad \frac{\lambda_{\mathrm{D}}}{\widehat{x}} \sim \varepsilon^{3/2}\,, \qquad \frac{\widehat{n}_{\mathrm{h}}}{\widehat{n}} \sim \varepsilon\,. \tag{12}$$

Note in particular the second of these relations, which guarantees a low-beta plasma:

$$\beta = \frac{\widehat{n} k_{\mathrm{B}} T_i}{2\widehat{B}^2 \mu_0} = \left(\frac{\rho_i}{\lambda_{\mathrm{D}}}\right)^2 \left(\frac{v_{\mathrm{th,i}}}{c}\right)^2 \sim \varepsilon\,. \tag{13}$$

Hence, writing $\mathbf{E}(t,x) = \widehat{E}\,\mathbf{E}'(t',x')$ for the dependent variables and inserting the ordering assumptions (8)-(12) into Maxwell's equations (1), then omitting the primes for a clearer notation, yields

$$\nabla \times \mathbf{B} = \mathbf{j} + \varepsilon^3 \frac{\partial \mathbf{E}}{\partial t}\,, \tag{14a}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}\,, \tag{14b}$$

$$\varepsilon^3\, \nabla \cdot \mathbf{E} = \varrho\,, \tag{14c}$$

$$\nabla \cdot \mathbf{B} = 0\,. \tag{14d}$$

From (3) we find that the normalized current reads

$$\mathbf{j}' = \frac{\mathbf{j}}{\widehat{j}} = \sum_{\mathrm{s}} Z_{\mathrm{s}} n'_{\mathrm{s}} \mathbf{u}'_{\mathrm{s}} + \frac{\widehat{v}}{\widehat{u}}\frac{\widehat{n}_{\mathrm{h}}}{\widehat{n}}\,\mathbf{j}'_{\mathrm{h}}\,, \qquad \mathbf{j}'_{\mathrm{h}} = Z_{\mathrm{h}} \int \mathbf{v}' f'_{\mathrm{h}}\, \mathrm{d}^3\mathbf{v}'\,, \tag{15}$$

where $Z_{\mathrm{s}} = q_{\mathrm{s}}/e$ denotes the charge number. Since $\widehat{v}/\widehat{u} = 1/\varepsilon$ and $\widehat{n}_{\mathrm{h}}/\widehat{n} = \varepsilon$, the factor in front of the normalized hot current density $\mathbf{j}'_{\mathrm{h}}$ is one. Had we demanded that $\widehat{v} \sim v_{\mathrm{th,i}}/\varepsilon$, which is reasonable for energetic particles, we would have to assume $\widehat{n}_{\mathrm{h}}/\widehat{n} \sim \varepsilon^2$ in order to have a hot current density on the scale of the bulk mean velocity. The theory presented below allows for such a scaling and might be even better adapted to it than to the case $\widehat{n}_{\mathrm{h}}/\widehat{n} = \varepsilon$, since terms of order n_{h} will be neglected in the final model for the sake of energy conservation. For the charge density the ordering $\widehat{n}_{\mathrm{h}}/\widehat{n} = \varepsilon$ means

$$\rho' = \frac{\rho}{e\widehat{n}} = \sum_{\mathrm{s}} Z_{\mathrm{s}} n'_{\mathrm{s}} + \varepsilon\, Z_{\mathrm{h}} n'_{\mathrm{h}}\,, \qquad n'_{\mathrm{h}} = \int f'_{\mathrm{h}}\, \mathrm{d}^3\mathbf{v}'\,. \tag{16}$$

The normalized fluid equations can be obtained from (4) by inserting (8)-(12):

$$\frac{\partial n_\mathrm{s}}{\partial t} + \nabla\cdot(n_\mathrm{s}\mathbf{u}_\mathrm{s}) = 0\,, \tag{17a}$$

$$\varepsilon^2\,\frac{m_\mathrm{s}}{m_\mathrm{i}}\left[\frac{\partial}{\partial t}(n_\mathrm{s}\mathbf{u}_\mathrm{s}) + \nabla\cdot(n_\mathrm{s}\mathbf{u}_\mathrm{s}\otimes\mathbf{u}_\mathrm{s})\right] + \nabla p_\mathrm{s} = Z_\mathrm{s}\left(n_\mathrm{s}\mathbf{E} + n_\mathrm{s}\mathbf{u}_\mathrm{s}\times\mathbf{B}\right), \tag{17b}$$

$$\frac{\partial p_\mathrm{s}}{\partial t} + \mathbf{u}_\mathrm{s}\cdot\nabla p_\mathrm{s} + \gamma\,p_\mathrm{s}\nabla\cdot\mathbf{u}_\mathrm{s} = 0\,. \tag{17c}$$

We will complete the ordering by assuming the mass ration to be

$$\frac{m_\mathrm{e}}{m_\mathrm{i}} \sim \varepsilon\,. \tag{18}$$

Finally, the normalized version of the kinetic equation (5) reads

$$\varepsilon\,\frac{\partial f_\mathrm{h}}{\partial t} + \mathbf{v}\cdot\nabla f_\mathrm{h} + C_\mathrm{h}\left(\mathbf{E} + \frac{\mathbf{v}\times\mathbf{B}}{\varepsilon}\right)\cdot\frac{\partial f_\mathrm{h}}{\partial\mathbf{v}} = 0\,, \tag{19}$$

where $C_\mathrm{h} = Z_\mathrm{h}/M_\mathrm{h}$ with $M_\mathrm{h} = m_\mathrm{h}/m_\mathrm{i}$ is the charge to mass ratio normalized to m_i/e.

C. MHD variables and truncation

Our aim is to derive a quasi-one-fluid description similar to ideal MHD from the normalized set of equations (14) and (17), assuming for the time being that the distribution of energetic particles is known. Let us introduce as MHD variables the bulk mass density τ and the bulk momentum density $\tau\mathbf{U}$:

$$\tau := m_\mathrm{i}\,n_\mathrm{i} + m_\mathrm{e}\,n_\mathrm{e}\,, \tag{20a}$$

$$\tau\mathbf{U} := m_\mathrm{i}\,n_\mathrm{i}\,\mathbf{u}_\mathrm{i} + m_\mathrm{e}\,n_\mathrm{e}\,\mathbf{u}_\mathrm{e}\,. \tag{20b}$$

Normalized to the ion mass and to the ion momentum they read

$$\tau = n_\mathrm{i} + \varepsilon\,n_\mathrm{e}\,, \qquad \tau\mathbf{U} = n_\mathrm{i}\mathbf{u}_\mathrm{i} + \varepsilon\,n_\mathrm{e}\mathbf{u}_\mathrm{e}\,. \tag{21}$$

In terms of these variables, the sum of the mass conservation laws (17a) yields

$$\frac{\partial\tau}{\partial t} + \nabla\cdot(\tau\mathbf{U}) = 0\,. \tag{22}$$

The sum of the momentum conservation laws (17b) can be approximated with an error of order $O(\varepsilon^3)$ in terms of the MHD variables:

$$\varepsilon^2\left[\frac{\partial}{\partial t}(\tau\mathbf{U}) + \nabla\cdot(\tau\mathbf{U}\otimes\mathbf{U})\right] + \nabla(p_\mathrm{i} + p_\mathrm{e}) = \sum_\mathrm{s} Z_\mathrm{s}n_\mathrm{s}\mathbf{E} + \sum_\mathrm{s} Z_\mathrm{s}n_\mathrm{s}\mathbf{u}_\mathrm{s}\times\mathbf{B} + O(\varepsilon^3)\,. \tag{23}$$

The quasi-one-fluid description is further enabled by inserting Poisson's equation (14c) and Ampère's law (14a) into the definition (3) of the charge and the current density, respectively,

$$
\begin{aligned}
\sum_{\mathrm{s}} Z_{\mathrm{s}} n_{\mathrm{s}} &= \rho - \varepsilon\, Z_{\mathrm{h}} n_{\mathrm{h}} = -\varepsilon\, Z_{\mathrm{h}} n_{\mathrm{h}} + O(\varepsilon^3)\,, \\
\sum_{\mathrm{s}} Z_{\mathrm{s}} n_{\mathrm{s}} \mathbf{u}_{\mathrm{s}} &= \mathbf{j} - \mathbf{j}_{\mathrm{h}} = \nabla \times \mathbf{B} - \mathbf{j}_{\mathrm{h}} + O(\varepsilon^3)\,.
\end{aligned}
\tag{24}
$$

Inserting this into (23) and dividing by τ yields

$$
\varepsilon^2 \left(\frac{\partial \mathbf{U}}{\partial t} + \mathbf{U} \cdot \nabla \mathbf{U} \right) + \frac{\nabla (p_{\mathrm{i}} + p_{\mathrm{e}})}{\tau} = -\varepsilon\, \frac{Z_{\mathrm{h}} n_{\mathrm{h}}}{\tau}\, \mathbf{E} + \frac{(\nabla \times \mathbf{B} - \mathbf{j}_{\mathrm{h}})}{\tau} \times \mathbf{B} + O(\varepsilon^3)\,. \tag{25}
$$

The difference of the momentum conservation laws (17b) leads to an approximate expression for the electric field (generalized Ohm's law). Dividing the ion equation by Z_{i}, further multiplying the electron equation by ε and taking the difference allows us to write

$$
\begin{aligned}
\tau \left(\frac{1}{Z_{\mathrm{i}}} - \varepsilon \right) \varepsilon^2 \left(\frac{\partial \mathbf{U}}{\partial t} + \mathbf{U} \cdot \nabla \mathbf{U} \right) + \nabla \left(\frac{p_{\mathrm{i}}}{Z_{\mathrm{i}}} - \varepsilon\, p_{\mathrm{e}} \right) &= (n_{\mathrm{i}} + \varepsilon\, n_{\mathrm{e}}) \mathbf{E} \\
&\quad + (n_{\mathrm{i}} \mathbf{u}_{\mathrm{i}} + \varepsilon\, n_{\mathrm{e}} \mathbf{u}_{\mathrm{e}}) \times \mathbf{B} + O(\varepsilon^3)\,.
\end{aligned}
\tag{26}
$$

Note that the error in this equation is $O(\varepsilon^3)$ and therefore we are allowed subtract ε in the factor $(1/Z_{\mathrm{i}} - \varepsilon)$ in front of the inertia term. This subtraction is necessary to obtain a consistent energy theorem, as will become clear later on. Equation (26) leads to the following approximation for the electric field:

$$
\mathbf{E} = -\mathbf{U} \times \mathbf{B} + \frac{1}{\tau} \nabla \left(\frac{p_{\mathrm{i}}}{Z_{\mathrm{i}}} - \varepsilon\, p_{\mathrm{e}} \right) + \left(\frac{1}{Z_{\mathrm{i}}} - \varepsilon \right) \varepsilon^2 \left(\frac{\partial \mathbf{U}}{\partial t} + \mathbf{U} \cdot \nabla \mathbf{U} \right) + O(\varepsilon^3)\,. \tag{27}
$$

If one solves this equation for $\mathbf{U}_\perp = \mathbf{b} \times \mathbf{U} \times \mathbf{b}$, then the left-hand side yields the $E \times B$-drift and the two terms on the right-hand side yield the diamagnetic drift and the polarization drift, respectively. Note that for energy conservation it is crucial to keep the polarization drift, as demonstrated below. The ion and electron pressures are computed from the original equations,

$$
\begin{aligned}
\frac{\partial p_{\mathrm{i}}}{\partial t} + \mathbf{u}_{\mathrm{i}} \cdot \nabla p_{\mathrm{i}} + \gamma\, p_{\mathrm{i}} \nabla \cdot \mathbf{u}_{\mathrm{i}} &= 0\,, \\
\frac{\partial p_{\mathrm{e}}}{\partial t} + \mathbf{u}_{\mathrm{e}} \cdot \nabla p_{\mathrm{e}} + \gamma\, p_{\mathrm{e}} \nabla \cdot \mathbf{u}_{\mathrm{e}} &= 0\,.
\end{aligned}
\tag{28}
$$

Here, the flows $\mathbf{u}_{\mathrm{i}}$ and $\mathbf{u}_{\mathrm{e}}$ have to be expressed in terms of $\mathbf{U}$ and $\mathbf{j}$. In what follows we consider ions and electrons with charge numbers $Z_{\mathrm{i}} \geq 1$ and $Z_{\mathrm{e}} = -1$. From

$$
\begin{aligned}
\tau \mathbf{U} &= n_{\mathrm{i}} \mathbf{u}_{\mathrm{i}} + \varepsilon\, n_{\mathrm{e}} \mathbf{u}_{\mathrm{e}}\,, \\
\mathbf{j} &= Z_{\mathrm{i}} n_{\mathrm{i}} \mathbf{u}_{\mathrm{i}} - n_{\mathrm{e}} \mathbf{u}_{\mathrm{e}} + \mathbf{j}_{\mathrm{h}}\,,
\end{aligned}
\tag{29}
$$

we obtain

$$
\begin{aligned}
n_\mathrm{i}\mathbf{u}_\mathrm{i} &= \frac{1}{1+\varepsilon Z_\mathrm{i}}\Big[\tau\mathbf{U} + \varepsilon\,(\mathbf{j}-\mathbf{j}_\mathrm{h})\Big]\,,\\
n_\mathrm{e}\mathbf{u}_\mathrm{e} &= \frac{1}{1+\varepsilon Z_\mathrm{i}}\Big[Z_\mathrm{i}\,\tau\mathbf{U} - (\mathbf{j}-\mathbf{j}_\mathrm{h})\Big]\,.
\end{aligned}
\tag{30}
$$

Similarly, for the number densities, from

$$
\begin{aligned}
\tau &= n_\mathrm{i} + \varepsilon\, n_\mathrm{e}\,,\\
\rho &= Z_\mathrm{i} n_\mathrm{i} - n_\mathrm{e} + \varepsilon\, Z_\mathrm{h} n_\mathrm{h}\,,
\end{aligned}
\tag{31}
$$

we compute

$$
\begin{aligned}
n_\mathrm{i} &= \frac{1}{1+\varepsilon Z_\mathrm{i}}\Big[\tau + \varepsilon\,(\rho - \varepsilon\, Z_\mathrm{h} n_\mathrm{h})\Big]\\
n_\mathrm{e} &= \frac{1}{1+\varepsilon Z_\mathrm{i}}\Big[Z_\mathrm{i}\,\tau - (\rho - \varepsilon\, Z_\mathrm{h} n_\mathrm{h})\Big]\,.
\end{aligned}
\tag{32}
$$

Using also $\rho = O(\varepsilon^4)$ from Poisson's equation and $\mathbf{j} = \nabla\times\mathbf{B} + O(\varepsilon^4)$ from Ampère's law we finally obtain

$$
\begin{aligned}
\mathbf{u}_\mathrm{i} &= \frac{\tau\mathbf{U} + \varepsilon\,(\nabla\times\mathbf{B} - \mathbf{j}_\mathrm{h})}{\tau - \varepsilon^2\, Z_\mathrm{h} n_\mathrm{h}} + O(\varepsilon^5)\,,\\
\mathbf{u}_\mathrm{e} &= \frac{Z_\mathrm{i}\,\tau\mathbf{U} - (\nabla\times\mathbf{B} - \mathbf{j}_\mathrm{h})}{Z_\mathrm{i}\,\tau + \varepsilon\, Z_\mathrm{h} n_\mathrm{h}} + O(\varepsilon^4)\,.
\end{aligned}
\tag{33}
$$

Let us summarize the results for the reduced hybrid description obtained so far: the fluid variables are the mass density τ, the mass flux $\tau\mathbf{U}$, the ion and electron pressure p_i and p_e, coupled to evolution equations for the magnetic field $\mathbf{B}$ and for the energetic particle distribution f_h. After truncation, the obtained hybrid model reads

$$
\left\{
\begin{aligned}
&\frac{\partial\tau}{\partial t} + \nabla\cdot(\tau\mathbf{U}) = 0\,, && \text{(34a)}\\
&\varepsilon^2\left(\frac{\partial\mathbf{U}}{\partial t} + \mathbf{U}\cdot\nabla\mathbf{U}\right) + \frac{\nabla(p_\mathrm{i}+p_\mathrm{e})}{\tau} = -\varepsilon\,\frac{Z_\mathrm{h} n_\mathrm{h}}{\tau}\,\mathbf{E} + \frac{(\nabla\times\mathbf{B} - \mathbf{j}_\mathrm{h})}{\tau}\times\mathbf{B} && \text{(34b)}\\
&\frac{\partial p_\mathrm{i}}{\partial t} + \left[\frac{\tau\mathbf{U} + \varepsilon\,(\nabla\times\mathbf{B} - \mathbf{j}_\mathrm{h})}{\tau - \varepsilon^2\, Z_\mathrm{h} n_\mathrm{h}}\right]\cdot\nabla p_\mathrm{i} + \gamma\, p_\mathrm{i}\nabla\cdot\left[\frac{\tau\mathbf{U} + \varepsilon\,(\nabla\times\mathbf{B} - \mathbf{j}_\mathrm{h})}{\tau - \varepsilon^2\, Z_\mathrm{h} n_\mathrm{h}}\right] = 0\,, && \text{(34c)}\\
&\frac{\partial p_\mathrm{e}}{\partial t} + \left[\frac{Z_\mathrm{i}\,\tau\mathbf{U} - (\nabla\times\mathbf{B} - \mathbf{j}_\mathrm{h})}{Z_\mathrm{i}\,\tau + \varepsilon\, Z_\mathrm{h} n_\mathrm{h}}\right]\cdot\nabla p_\mathrm{e} + \gamma\, p_\mathrm{e}\nabla\cdot\left[\frac{Z_\mathrm{i}\,\tau\mathbf{U} - (\nabla\times\mathbf{B} - \mathbf{j}_\mathrm{h})}{Z_\mathrm{i}\,\tau + \varepsilon\, Z_\mathrm{h} n_\mathrm{h}}\right] = 0\,, && \text{(34d)}\\
&\frac{\partial\mathbf{B}}{\partial t} = -\nabla\times\mathbf{E}\,, && \text{(34e)}\\
&\varepsilon\,\frac{\partial f_\mathrm{h}}{\partial t} + \mathbf{v}\cdot\nabla f_\mathrm{h} + C_\mathrm{h}\left(\mathbf{E} + \frac{\mathbf{v}\times\mathbf{B}}{\varepsilon}\right)\cdot\frac{\partial f_\mathrm{h}}{\partial\mathbf{v}} = 0\,, && \text{(34f)}\\
&\quad\mathbf{E} = -\mathbf{U}\times\mathbf{B} + \frac{1}{\tau}\nabla\left(\frac{p_\mathrm{i}}{Z_\mathrm{i}} - \varepsilon\, p_\mathrm{e}\right) + \left(\frac{1}{Z_\mathrm{i}} - \varepsilon\right)\varepsilon^2\left(\frac{\partial\mathbf{U}}{\partial t} + \mathbf{U}\cdot\nabla\mathbf{U}\right)\,. && \text{(34g)}
\end{aligned}
\right.
$$

The system (34) is a quasi-one-fluid model derived by truncation from the original set of two-fluid equations. However, this model is not satisfactory from the standpoint of energy conservation and has to be simplified in order to be energetically consistent. We shall discuss the different levels of complexity for energetically consistent models derived from (34) in the following sections.

D. Energy theorem without energetic particles

At first we study the energy conservation law for the system (34) in case that energetic particles are absent. For this we suppose that $n_\mathrm{h} = j_\mathrm{h} = 0$ and we neglect the Vlasov equation for the energetic particles. The momentum conservation law (34b) then reads

$$\varepsilon^2 \left(\frac{\partial \mathbf{U}}{\partial t} + \mathbf{U} \cdot \nabla \mathbf{U} \right) + \frac{\nabla (p_\mathrm{i} + p_\mathrm{e})}{\tau} = \frac{(\nabla \times \mathbf{B})}{\tau} \times \mathbf{B} \,. \tag{35}$$

We thus obtain the extended MHD system

$$\text{extd. MHD} \begin{cases} \dfrac{\partial \tau}{\partial t} + \nabla \cdot (\tau \mathbf{U}) = 0 \,, & \text{(36a)} \\[2ex] \varepsilon^2 \left(\dfrac{\partial \mathbf{U}}{\partial t} + \mathbf{U} \cdot \nabla \mathbf{U} \right) + \dfrac{\nabla (p_\mathrm{i} + p_\mathrm{e})}{\tau} = \dfrac{(\nabla \times \mathbf{B})}{\tau} \times \mathbf{B} \,, & \text{(36b)} \\[2ex] \dfrac{\partial p_\mathrm{i}}{\partial t} + \left(\mathbf{U} + \varepsilon \dfrac{\nabla \times \mathbf{B}}{\tau} \right) \cdot \nabla p_\mathrm{i} + \gamma \, p_\mathrm{i} \nabla \cdot \left(\mathbf{U} + \varepsilon \dfrac{\nabla \times \mathbf{B}}{\tau} \right) = 0 \,, & \text{(36c)} \\[2ex] \dfrac{\partial p_\mathrm{e}}{\partial t} + \left(\mathbf{U} - \dfrac{\nabla \times \mathbf{B}}{Z_\mathrm{i} \, \tau} \right) \cdot \nabla p_\mathrm{e} + \gamma \, p_\mathrm{e} \nabla \cdot \left(\mathbf{U} - \dfrac{\nabla \times \mathbf{B}}{Z_\mathrm{i} \, \tau} \right) = 0 \,, & \text{(36d)} \\[2ex] \dfrac{\partial \mathbf{B}}{\partial t} = - \nabla \times \mathbf{E} \,, & \text{(36e)} \\[2ex] \mathbf{E} = - \mathbf{U} \times \mathbf{B} + \dfrac{1}{\tau} \nabla \left(\varepsilon \, p_\mathrm{i} - \dfrac{p_\mathrm{e}}{Z_\mathrm{i}} \right) + \left(\dfrac{1}{Z_\mathrm{i}} - \varepsilon \right) \dfrac{(\nabla \times \mathbf{B})}{\tau} \times \mathbf{B} \,. & \text{(36f)} \end{cases}$$

We now show that the system (36) conserves the following (normalized) energy exactly:

$$\mathcal{E}_\mathrm{MHD} := \frac{3}{2} \int (p_\mathrm{i} + p_\mathrm{e}) \, \mathrm{d}^3\mathbf{x} + \varepsilon^2 \int \tau \, \frac{|\mathbf{U}|^2}{2} \, \mathrm{d}^3\mathbf{x} + \int \frac{|\mathbf{B}|^2}{2} \, \mathrm{d}^3\mathbf{x} \,, \tag{37}$$

Let us compute the time derivative of this expression term by term. We assume suitable boundary conditions such that all occurring boundary integrals vanish. Recalling that

$\gamma = 5/3$, for the ion internal energy we have

$$\frac{3}{2}\int \frac{\partial p_\mathrm{i}}{\partial t}\,\mathrm{d}^3\mathbf{x} = -\frac{3}{2}\int \nabla\cdot\left[\left(\mathbf{U}+\varepsilon\,\frac{\nabla\times\mathbf{B}}{\tau}\right)p_\mathrm{i}\right]\mathrm{d}^3\mathbf{x} - \int p_\mathrm{i}\nabla\cdot\left(\mathbf{U}+\varepsilon\,\frac{\nabla\times\mathbf{B}}{\tau}\right)\mathrm{d}^3\mathbf{x}$$

$$= \int \nabla p_\mathrm{i}\cdot\left(\mathbf{U}+\varepsilon\,\frac{\nabla\times\mathbf{B}}{\tau}\right)\mathrm{d}^3\mathbf{x} =: \mathcal{P}_\mathrm{i}^0\,. \tag{38}$$

Similarly, for the electron internal energy we compute

$$\frac{3}{2}\int \frac{\partial p_\mathrm{e}}{\partial t}\,\mathrm{d}^3\mathbf{x} = \int \nabla p_\mathrm{e}\cdot\left(\mathbf{U}-\frac{\nabla\times\mathbf{B}}{Z_\mathrm{i}\,\tau}\right)\mathrm{d}^3\mathbf{x} =: \mathcal{P}_\mathrm{e}^0\,. \tag{39}$$

The change in the kinetic plasma energy consists of two parts, namely

$$\varepsilon^2\int \frac{\partial\tau}{\partial t}\,\frac{|\mathbf{U}|^2}{2}\,\mathrm{d}^3\mathbf{x} = -\varepsilon^2\int \nabla\cdot(\tau\mathbf{U})\,\frac{|\mathbf{U}|^2}{2}\,\mathrm{d}^3\mathbf{x} = \varepsilon^2\int \tau\mathbf{U}\cdot\nabla\frac{|\mathbf{U}|^2}{2}\,\mathrm{d}^3\mathbf{x}\,, \tag{40}$$

and

$$\varepsilon^2\int \tau\,\frac{\partial}{\partial t}\frac{|\mathbf{U}|^2}{2}\,\mathrm{d}^3\mathbf{x} = \varepsilon^2\int \tau\,\frac{\partial\mathbf{U}}{\partial t}\cdot\mathbf{U}\,\mathrm{d}^3\mathbf{x} \tag{41}$$

$$= \varepsilon^2\int\left[-\tau\mathbf{U}\cdot\nabla\mathbf{U} - \frac{1}{\varepsilon^2}\,\nabla(p_\mathrm{i}+p_\mathrm{e}) + \frac{1}{\varepsilon^2}\,(\nabla\times\mathbf{B})\times\mathbf{B}\right]\cdot\mathbf{U}\,\mathrm{d}^3\mathbf{x}$$

$$= -\varepsilon^2\int \tau\mathbf{U}\cdot\nabla\frac{|\mathbf{U}|^2}{2}\,\mathrm{d}^3\mathbf{x} - \int \nabla(p_\mathrm{i}+p_\mathrm{e})\cdot\mathbf{U}\,\mathrm{d}^3\mathbf{x} - \int \mathbf{B}\cdot\nabla\times(\mathbf{U}\times\mathbf{B})\,\mathrm{d}^3\mathbf{x}\,.$$

Therefore,

$$\frac{\partial}{\partial t}\,\varepsilon^2\int \tau\,\frac{|\mathbf{U}|^2}{2}\,\mathrm{d}^3\mathbf{x} = -\int \nabla(p_\mathrm{i}+p_\mathrm{e})\cdot\mathbf{U}\,\mathrm{d}^3\mathbf{x} - \int \mathbf{B}\cdot\nabla\times(\mathbf{U}\times\mathbf{B})\,\mathrm{d}^3\mathbf{x} =: \mathcal{K}^0\,. \tag{42}$$

Finally, for the magnetic field energy we have

$$\int \frac{\partial}{\partial t}\frac{|\mathbf{B}|^2}{2}\,\mathrm{d}^3\mathbf{x} = \int \frac{\partial\mathbf{B}}{\partial t}\cdot\mathbf{B}\,\mathrm{d}^3\mathbf{x} \tag{43}$$

$$= \int \mathbf{B}\cdot\nabla\times(\mathbf{U}\times\mathbf{B})\,\mathrm{d}^3\mathbf{x} - \int \nabla\times\mathbf{B}\cdot\frac{1}{\tau}\nabla\left(\varepsilon\,p_\mathrm{i} - \frac{p_\mathrm{e}}{Z_\mathrm{i}}\right)\mathrm{d}^3\mathbf{x} =: \mathcal{B}^0\,.$$

Summing up these pieces yields

$$\frac{\mathrm{d}\mathcal{E}_\mathrm{MHD}}{\mathrm{d}t} = \mathcal{P}_\mathrm{i}^0 + \mathcal{P}_\mathrm{e}^0 + \mathcal{K}^0 + \mathcal{B}^0 = 0\,. \tag{44}$$

E. First current coupling scheme

The simplest current coupling scheme can be deduced from (34) by setting $n_\mathrm{h} = 0$ in the MHD part. The momentum conservation law (34b) then reads

$$\varepsilon^2\left(\frac{\partial\mathbf{U}}{\partial t} + \mathbf{U}\cdot\nabla\mathbf{U}\right) + \frac{\nabla(p_\mathrm{i}+p_\mathrm{e})}{\tau} = \frac{(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{\tau}\times\mathbf{B}\,. \tag{45}$$

We thus obtain the first current coupling scheme:

$$
\mathrm{CC1}\begin{cases}
\dfrac{\partial\tau}{\partial t}+\nabla\cdot(\tau\mathbf{U})=0\,, & \text{(46a)}\\[2ex]
\varepsilon^2\left(\dfrac{\partial\mathbf{U}}{\partial t}+\mathbf{U}\cdot\nabla\mathbf{U}\right)+\dfrac{\nabla(p_\mathrm{i}+p_\mathrm{e})}{\tau}=\dfrac{(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{\tau}\times\mathbf{B}\,, & \text{(46b)}\\[2ex]
\dfrac{\partial p_\mathrm{i}}{\partial t}+\left[\mathbf{U}+\varepsilon\,\dfrac{(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{\tau}\right]\cdot\nabla p_\mathrm{i}+\gamma\,p_\mathrm{i}\nabla\cdot\left[\mathbf{U}+\varepsilon\,\dfrac{(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{\tau}\right]=0\,, & \text{(46c)}\\[2ex]
\dfrac{\partial p_\mathrm{e}}{\partial t}+\left[\mathbf{U}-\dfrac{(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{Z_\mathrm{i}\,\tau}\right]\cdot\nabla p_\mathrm{e}+\gamma\,p_\mathrm{e}\nabla\cdot\left[\mathbf{U}-\dfrac{(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{Z_\mathrm{i}\,\tau}\right]=0\,, & \text{(46d)}\\[2ex]
\dfrac{\partial\mathbf{B}}{\partial t}=-\nabla\times\mathbf{E}\,, & \text{(46e)}\\[2ex]
\mathbf{E}=-\mathbf{U}\times\mathbf{B}+\dfrac{1}{\tau}\nabla\left(\varepsilon\,p_\mathrm{i}-\dfrac{p_\mathrm{e}}{Z_\mathrm{i}}\right)+\left(\dfrac{1}{Z_\mathrm{i}}-\varepsilon\right)\dfrac{(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{\tau}\times\mathbf{B}\,, & \text{(46f)}\\[2ex]
\varepsilon\,\dfrac{\partial f_\mathrm{h}}{\partial t}+\mathbf{v}\cdot\nabla f_\mathrm{h}+C_\mathrm{h}\left(\mathbf{E}+\dfrac{\mathbf{v}\times\mathbf{B}}{\varepsilon}\right)\cdot\dfrac{\partial f_\mathrm{h}}{\partial\mathbf{v}}=0\,, & \text{(46g)}
\end{cases}
$$

where the current density of the energetic particles reads

$$
\mathbf{j}_\mathrm{h}=Z_\mathrm{h}\int\mathbf{v}\,f_\mathrm{h}\,\mathrm{d}^3\mathbf{v}\,. \tag{47}
$$

The MHD part of this current-coupling scheme can also be obtained from (36) by replacing $\nabla\times\mathbf{B}\to(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})$. We now prove that the following energy is conserved exactly by the system (46):

$$
\mathcal{E}_\mathrm{tot}:=\mathcal{E}_\mathrm{MHD}+\varepsilon\,\mathcal{E}_\mathrm{h}\,, \tag{48}
$$

where the contribution of the hot particles reads

$$
\mathcal{E}_\mathrm{h}:=\int w_\mathrm{h}\,\mathrm{d}^3\mathbf{x}=\frac{M_\mathrm{h}}{2}\int\int|\mathbf{v}|^2 f_\mathrm{h}\,\mathrm{d}^3\mathbf{x}\,\mathrm{d}^3\mathbf{v}\,. \tag{49}
$$

As evident from (48), our ordering implies that the contribution $\mathcal{E}_\mathrm{h}$ to the total energy is small compared to the contribution of the bulk part $\mathcal{E}_\mathrm{MHD}$. This is because we assumed low density, $n_\mathrm{h}=O(\varepsilon)$. From (46g) we obtain

$$
\begin{aligned}
\varepsilon\,\frac{\mathrm{d}\mathcal{E}_\mathrm{h}}{\mathrm{d}t}&=\varepsilon\int\frac{\partial w_\mathrm{h}}{\partial t}\,\mathrm{d}^3\mathbf{x} \qquad\qquad (50)\\
&=\int\mathbf{j}_\mathrm{h}\cdot\mathbf{E}\,\mathrm{d}^3\mathbf{x}\\
&=\int\mathbf{j}_\mathrm{h}\cdot\left[-\mathbf{U}\times\mathbf{B}+\frac{1}{\tau}\nabla\left(\varepsilon\,p_\mathrm{i}-\frac{p_\mathrm{e}}{Z_\mathrm{i}}\right)+\left(\frac{1}{Z_\mathrm{i}}-\varepsilon\right)\frac{(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{\tau}\times\mathbf{B}\right]\mathrm{d}^3\mathbf{x}=:\mathcal{H}^1\,.
\end{aligned}
$$

Let us now compute the time derivative of the MHD energy defined in (37), by using the previous results (38)-(43). From (46b) we have

$$\frac{\partial}{\partial t}\,\varepsilon^2\int\tau\,\frac{|\mathbf{U}|^2}{2}\,\mathrm{d}^3\mathbf{x}=\mathcal{K}^0-\int\mathbf{j}_\mathrm{h}\times\mathbf{B}\cdot\mathbf{U}\,\mathrm{d}^3\mathbf{x}=:\mathcal{K}^1\,, \tag{51}$$

which cancels the first term on the right-hand-side of (50). From the new terms in the pressure equations we obtain

$$\frac{3}{2}\int\frac{\partial p_\mathrm{i}}{\partial t}\,\mathrm{d}^3\mathbf{x}=\mathcal{P}_\mathrm{i}^0-\varepsilon\int\nabla p_\mathrm{i}\cdot\frac{\mathbf{j}_\mathrm{h}}{\tau}\,\mathrm{d}^3\mathbf{x}=:\mathcal{P}_\mathrm{i}^1\,, \tag{52}$$

$$\frac{3}{2}\int\frac{\partial p_\mathrm{e}}{\partial t}\,\mathrm{d}^3\mathbf{x}=\mathcal{P}_\mathrm{e}^0+\frac{1}{Z_\mathrm{i}}\int\nabla p_\mathrm{e}\cdot\frac{\mathbf{j}_\mathrm{h}}{\tau}\,\mathrm{d}^3\mathbf{x}=:\mathcal{P}_2^1\,. \tag{53}$$

Finally, in the magnetic energy we have

$$\int\frac{\partial}{\partial t}\frac{|\mathbf{B}|^2}{2}\,\mathrm{d}^3\mathbf{x}=\mathcal{B}-\left(\frac{1}{Z_\mathrm{i}}-\varepsilon\right)\int(\nabla\times\mathbf{B})\cdot\frac{(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{\tau}\times\mathbf{B}=:\mathcal{B}^1\,. \tag{54}$$

Summing up all contributions and using (44) from the previous section we obtain

$$\begin{aligned}\frac{\mathrm{d}\mathcal{E}_\mathrm{tot}}{\mathrm{d}t}&=\mathcal{P}_\mathrm{i}^1+\mathcal{P}_\mathrm{e}^1+\mathcal{K}^1+\mathcal{B}^1+\mathcal{H}^1\\&=-\left(\frac{1}{Z_\mathrm{i}}-\varepsilon\right)\int(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})\cdot\frac{(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{\tau}\times\mathbf{B}\\&=0\,.\end{aligned} \tag{55}$$

F. Second current coupling scheme

We now remove the assumption $n_\mathrm{h}=0$ and allow for non-vanishing energetic particle density in the system (34). However, this needs to be done with care in order not to break energy conservation. The momentum conservation law (34b) is now given by

$$\varepsilon^2\left(\frac{\partial\mathbf{U}}{\partial t}+\mathbf{U}\cdot\nabla\mathbf{U}\right)+\frac{\nabla(p_\mathrm{i}+p_\mathrm{e})}{\tau}=-\varepsilon\,\frac{Z_\mathrm{h}n_\mathrm{h}}{\tau}\,\mathbf{E}+\frac{(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{\tau}\times\mathbf{B}\,, \tag{56}$$

which, inserted into (34g), leads to the tentative electric field $\mathbf{E}'$ of the form

$$\mathbf{E}'=\frac{1}{1+\varepsilon\,\eta}\left[-\mathbf{U}\times\mathbf{B}+\frac{1}{\tau}\nabla\left(\varepsilon\,p_\mathrm{i}-\frac{p_\mathrm{e}}{Z_\mathrm{i}}\right)+\left(\frac{1}{Z_\mathrm{i}}-\varepsilon\right)\frac{(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{\tau}\times\mathbf{B}\right]\,, \tag{57}$$

where η is given by

$$\eta=\frac{Z_\mathrm{h}n_\mathrm{h}}{\tau}\left(\frac{1}{Z_\mathrm{i}}-\varepsilon\right)\,. \tag{58}$$

It turns out that inserting the tentative electric field $\mathbf{E}'$ into the system (34) breaks energy conservation. Nevertheless, we can still find an energy theorem by truncating $\mathbf{E}'$ at the appropriate places. Instead of (57), we will use

$$\mathbf{E} = -\mathbf{U}\times\mathbf{B}\left(1-\varepsilon\,\eta\right) + \frac{1}{\tau}\nabla\left(\varepsilon\,p_\mathrm{i} - \frac{p_\mathrm{e}}{Z_\mathrm{i}}\right) + \left(\frac{1}{Z_\mathrm{i}} - \varepsilon\right)\frac{(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{\tau}\times\mathbf{B}\,, \tag{59}$$

where in comparison with (46f) from the first scheme only the term $-\mathbf{U}\times\mathbf{B}$ is altered by a factor obtained by expanding the pre-factor in (57) in ε. Moreover, in the pressure equations we need to approximate as follows:

$$\frac{\tau\mathbf{U} + \varepsilon\,(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{\tau - \varepsilon^2\,Z_\mathrm{h}n_\mathrm{h}} = \mathbf{U} + \varepsilon\,\frac{(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{\tau} + \varepsilon^2\,\frac{Z_\mathrm{h}n_\mathrm{h}}{\tau}\,\mathbf{U} + O(\varepsilon^3)\,, \tag{60}$$

$$\frac{Z_\mathrm{i}\,\tau\mathbf{U} - (\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{Z_\mathrm{i}\,\tau + \varepsilon\,Z_\mathrm{h}n_\mathrm{h}} = \mathbf{U} - \frac{(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{Z_\mathrm{i}\,\tau} - \varepsilon\,\frac{Z_\mathrm{h}n_\mathrm{h}}{Z_\mathrm{i}\,\tau}\,\mathbf{U} + O(\varepsilon)\,. \tag{61}$$

While the first line is consistent with the ordering, in the second line we truncate one term featuring the current that is of order ε. This truncation will turn out necessary for energy conservation. The second current coupling scheme thus reads:

$$\text{CC2}\begin{cases}
\dfrac{\partial\tau}{\partial t} + \nabla\cdot(\tau\mathbf{U}) = 0\,, & \text{(62a)}\\[2ex]
\varepsilon^2\left(\dfrac{\partial\mathbf{U}}{\partial t} + \mathbf{U}\cdot\nabla\mathbf{U}\right) + \dfrac{\nabla(p_\mathrm{i}+p_\mathrm{e})}{\tau} = -\varepsilon\,\dfrac{Z_\mathrm{h}n_\mathrm{h}}{\tau}\,\mathbf{E} + \dfrac{(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{\tau}\times\mathbf{B}\,, & \text{(62b)}\\[2ex]
\dfrac{\partial p_\mathrm{i}}{\partial t} + \left[\mathbf{U}\left(1+\varepsilon^2\,\dfrac{Z_\mathrm{h}n_\mathrm{h}}{\tau}\right) + \varepsilon\,\dfrac{(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{\tau}\right]\cdot\nabla p_\mathrm{i} & \text{(62c)}\\[2ex]
\qquad +\gamma\,p_\mathrm{i}\nabla\cdot\left[\mathbf{U}\left(1+\varepsilon^2\,\dfrac{Z_\mathrm{h}n_\mathrm{h}}{\tau}\right) + \varepsilon\,\dfrac{(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{\tau}\right] = 0\,, & \\[2ex]
\dfrac{\partial p_\mathrm{e}}{\partial t} + \left[\mathbf{U}\left(1-\varepsilon\,\dfrac{Z_\mathrm{h}n_\mathrm{h}}{Z_\mathrm{i}\,\tau}\right) - \dfrac{(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{Z_\mathrm{i}\,\tau}\right]\cdot\nabla p_\mathrm{e} & \text{(62d)}\\[2ex]
\qquad +\gamma\,p_\mathrm{e}\nabla\cdot\left[\mathbf{U}\left(1-\varepsilon\,\dfrac{Z_\mathrm{h}n_\mathrm{h}}{Z_\mathrm{i}\,\tau}\right) - \dfrac{(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{Z_\mathrm{i}\,\tau}\right] = 0\,, & \\[2ex]
\dfrac{\partial\mathbf{B}}{\partial t} = -\nabla\times\mathbf{E}\,, & \text{(62e)}\\[2ex]
\mathbf{E} = -\mathbf{U}\times\mathbf{B}\left(1-\varepsilon\,\eta\right) + \dfrac{1}{\tau}\nabla\left(\varepsilon\,p_\mathrm{i} - \dfrac{p_\mathrm{e}}{Z_\mathrm{i}}\right) + \left(\dfrac{1}{Z_\mathrm{i}} - \varepsilon\right)\dfrac{(\nabla\times\mathbf{B}-\mathbf{j}_\mathrm{h})}{\tau}\times\mathbf{B}\,, & \text{(62f)}\\[2ex]
\varepsilon\,\dfrac{\partial f_\mathrm{h}}{\partial t} + \mathbf{v}\cdot\nabla f_\mathrm{h} + C_\mathrm{h}\left(\mathbf{E} + \dfrac{\mathbf{v}\times\mathbf{B}}{\varepsilon}\right)\cdot\dfrac{\partial f_\mathrm{h}}{\partial\mathbf{v}} = 0\,. & \text{(62g)}
\end{cases}$$

Let us now prove that the system (62) conserves the energy $\mathcal{E}_{\mathrm{tot}}$ defined in (48) exactly. To keep a concise notation we will use the previous results (50)-(54). The contribution of the energetic particles is similar to (50) and reads

$$\varepsilon\,\frac{\mathrm{d}\mathcal{E}_{\mathrm{h}}}{\mathrm{d}t} = \mathcal{H}^1 + \varepsilon \int \eta\,\mathbf{j}_{\mathrm{h}} \cdot \mathbf{U} \times \mathbf{B}\,. \tag{63}$$

Moreover,

$$\begin{aligned} \frac{\partial}{\partial t}\,\varepsilon^2 \int \tau\,\frac{|\mathbf{U}|^2}{2}\,\mathrm{d}^3\mathbf{x} &= \mathcal{K}^1 - \varepsilon \int Z_{\mathrm{h}} n_{\mathrm{h}}\,\mathbf{E} \cdot \mathbf{U}\,\mathrm{d}^3\mathbf{x} \qquad (64)\\ &= \mathcal{K}^1 - \varepsilon \int Z_{\mathrm{h}} n_{\mathrm{h}} \left[\frac{1}{\tau}\nabla\left(\varepsilon\, p_{\mathrm{i}} - \frac{p_{\mathrm{e}}}{Z_{\mathrm{i}}}\right) + \left(\frac{1}{Z_{\mathrm{i}}} - \varepsilon\right)\frac{(\nabla \times \mathbf{B} - \mathbf{j}_{\mathrm{h}})}{\tau} \times \mathbf{B}\right] \cdot \mathbf{U}\,\mathrm{d}^3\mathbf{x}\,, \end{aligned}$$

$$\frac{3}{2}\int \frac{\partial p_{\mathrm{i}}}{\partial t}\,\mathrm{d}^3\mathbf{x} = \mathcal{P}_{\mathrm{i}}^1 + \varepsilon \int \frac{Z_{\mathrm{h}} n_{\mathrm{h}}}{\tau}\nabla(\varepsilon\, p_{\mathrm{i}}) \cdot \mathbf{U}\,\mathrm{d}^3\mathbf{x}\,, \tag{65}$$

$$\frac{3}{2}\int \frac{\partial p_{\mathrm{e}}}{\partial t}\,\mathrm{d}^3\mathbf{x} = \mathcal{P}_{\mathrm{e}}^1 - \varepsilon \int \frac{Z_{\mathrm{h}} n_{\mathrm{h}}}{\tau}\nabla\left(\frac{p_{\mathrm{e}}}{Z_{\mathrm{i}}}\right) \cdot \mathbf{U}\,\mathrm{d}^3\mathbf{x}\,, \tag{66}$$

$$\int \frac{\partial}{\partial t}\frac{|\mathbf{B}|^2}{2}\,\mathrm{d}^3\mathbf{x} = \mathcal{B}^1 - \varepsilon \int \eta\,\nabla \times \mathbf{B} \cdot (\mathbf{U} \times \mathbf{B})\,\mathrm{d}^3\mathbf{x}\,. \tag{67}$$

Summing up all contributions and using (55) from the previous section we obtain

$$\begin{aligned} \frac{\mathrm{d}\mathcal{E}_{\mathrm{tot}}}{\mathrm{d}t} &= -\varepsilon \int \eta\,(\nabla \times \mathbf{B} - \mathbf{j}_{\mathrm{h}}) \cdot (\mathbf{U} \times \mathbf{B})\,\mathrm{d}^3\mathbf{x} \\ &\quad - \varepsilon \int Z_{\mathrm{h}} n_{\mathrm{h}} \left(\frac{1}{Z_{\mathrm{i}}} - \varepsilon\right)\frac{(\nabla \times \mathbf{B} - \mathbf{j}_{\mathrm{h}})}{\tau} \cdot (\mathbf{B} \times \mathbf{U})\,\mathrm{d}^3\mathbf{x} \qquad (68)\\ &= 0\,. \end{aligned}$$

G. Drift-kinetic-MHD current coupling

The next step is to approximate the Vlasov equation (62g) by a drift-kinetic (DK) model which describes the energetic particles as an ensemble of gyro-centers subject to magnetic drifts, where the fast gyration of particles around magnetic field lines is averaged out. In appendix A we derive this DK equation on the level of the Lagrangian function of single particle motion via the method of variational averaging[3], adapted to the particular ordering of the Vlasov equation (62g). Therefore, our reduction procedure is consistent with the physical assumptions stated in II B on the fluid and on the kinetic level. Moreover, we will prove an energy theorem for the reduced DK-MHD hybrid model.

The DK equation derived in appendix A is formulated in the gyro-center (GY) phase space $\widehat{\Omega}$ with coordinates $\widehat{\mathbf{q}} = (\mathbf{r}, q_\parallel, \widehat{\mu}, \alpha) \in \widehat{\Omega} \subset \mathbb{R}^6$. Here, $\mathbf{r} \in \mathbb{R}^3$ stands for the GY-position, $q_\parallel$ for its parallel velocity, $\widehat{\mu}$ denotes the generalized magnetic moment, a constant of the GY-motion, and α is the gyro-angle. The distribution of GY-centers, denoted by $f_{\mathrm{h}}^{\mathrm{D}} : \mathbb{R} \times \widehat{\Omega} \to \mathbb{R}_+$, is assumed independent of α and a solution of the following DK equation:

$$\varepsilon \frac{\partial f_{\mathrm{h}}^{\mathrm{D}}}{\partial t} + \mathbf{u}_{\mathrm{g}} \cdot \frac{\partial f_{\mathrm{h}}^{\mathrm{D}}}{\partial \mathbf{r}} + E_{\mathrm{g}} \frac{\partial f_{\mathrm{h}}^{\mathrm{D}}}{\partial q_\parallel} = 0\,, \tag{69}$$

where the GY-velocity $\mathbf{u}_{\mathrm{g}}$ and the GY-acceleration E_{g} are defined by

$$\mathbf{u}_{\mathrm{g}} := q_\parallel \frac{\mathbf{B}^*}{B_\parallel^*} + \varepsilon \frac{\mathbf{E}^* \times \mathbf{b}}{B_\parallel^*}\,, \qquad E_{\mathrm{g}} := C_{\mathrm{h}} \frac{\mathbf{E}^* \cdot \mathbf{B}^*}{B_\parallel^*}\,. \tag{70}$$

Here,

$$\mathbf{B}^* := \mathbf{B} + \varepsilon \frac{q_\parallel}{C_{\mathrm{h}}} \nabla \times \mathbf{b}\,, \qquad \mathbf{E}^* := \mathbf{E} - \nabla(\widehat{\mu} B) - \varepsilon \frac{q_\parallel}{C_{\mathrm{h}}} \frac{\partial \mathbf{b}}{\partial t}\,, \tag{71}$$

and $B_\parallel^* = \mathbf{B}^* \cdot \mathbf{b}$ is the Jacobian determinant of the GY-transformation leading to (69). Therefore, introducing $F_{\mathrm{h}} := f_{\mathrm{h}}^{\mathrm{D}} B_\parallel^*$, from the Liouville property (A63) we have

$$\frac{\partial F_{\mathrm{h}}}{\partial t} + \nabla \cdot (\mathbf{u}_{\mathrm{g}} F_{\mathrm{h}}) + \frac{\partial}{\partial q_\parallel}(E_{\mathrm{g}} F_{\mathrm{h}}) = 0\,. \tag{72}$$

We remark that the generalized fields $\mathbf{B}^*$ and $\mathbf{E}^*$ contain terms due to magnetic curvature and due to the time dependence of $\mathbf{B}$. The full GY-drifts, including the ∇B-drift and the geodesic curvature drift, can be made explicit by expanding $B_\parallel^*$ in ε.

Our goal is to maintain the energy conservation (68) also for the reduced DK-MHD model, where (69) is solved instead of (62g) to obtain the contribution of the energetic particles. For this, we need to identify at first the DK energy $\mathcal{E}_{\mathrm{h}}^{\mathrm{D}}$ and then define the DK current density $\mathbf{j}_{\mathrm{h}}^{\mathrm{D}}$ properly in order to achieve energy conservation. From (50) and (63) we deduce that a sufficient condition for energy conservation is

$$\varepsilon \frac{\mathrm{d}}{\mathrm{d}t} \mathcal{E}_{\mathrm{h}}^{\mathrm{D}} = \int \mathbf{j}_{\mathrm{h}}^{\mathrm{D}} \cdot \mathbf{E}\, \mathrm{d}^3\mathbf{x}\,, \tag{73}$$

where $\mathbf{E}$ is approximated by (59). Remark that integration is here over particle position $\mathbf{x}$ and not over GY-position $\mathbf{r}$. To compute the energy of hot particles, let us insert the GY-transform $\widehat{\mathcal{T}}_{\mathrm{gy}}^{\varepsilon} : \widehat{\Omega} \to \Omega$, $\widehat{\mathbf{q}} \mapsto \mathbf{q} = (\mathbf{x}, \mathbf{v})$, determined by the generating functions (A70), under the energy integral in (49):

$$
\begin{aligned}
\mathcal{E}_\mathrm{h}(t) &= \int w_\mathrm{h}(\mathbf{x},t)\,\mathrm{d}^3\mathbf{x} \\
&= \frac{M_\mathrm{h}}{2}\int |\mathbf{v}|^2\, f_\mathrm{h}(\mathbf{q},t)\,\mathrm{d}^6\mathbf{q} \\
&= \frac{M_\mathrm{h}}{2}\int |\mathbf{v}(\widehat{\mathbf{q}},t)|^2\, f_\mathrm{h}(\widehat{\tau}^\varepsilon_\mathrm{gy}(\widehat{\mathbf{q}}),t)\, B^*_\parallel\,\mathrm{d}^6\widehat{\mathbf{q}} \\
&\approx \frac{M_\mathrm{h}}{2\pi}\int \left[\frac{v_\parallel^2(\widehat{\mathbf{q}},t)}{2} + \mu(\widehat{\mathbf{q}},t)B(\widehat{\mathbf{q}},t)\right] F_\mathrm{h}(\widehat{\mathbf{q}},t)\,\mathrm{d}^6\widehat{\mathbf{q}}\,.
\end{aligned}
\tag{74}
$$

Here we used the definition (A68) of the magnetic moment. Note in particular the approximation in the last line, which speaks to the fact that the solution f_h^D of the DK equation is merely an approximation of the true solution, $f_\mathrm{h}^\mathrm{D}/2\pi \approx f_\mathrm{h}\circ\widehat{\tau}^\varepsilon_\mathrm{gy}$. This is because because the characteristics of the DK equation stem form a truncated Lagrangian, see[3] for details. In order to proceed let us introduce some notation:

$$
\int_{\widehat{\mu}} = \int\int_0^\infty \mathrm{d}\widehat{\mu}\,, \qquad \langle\cdot\rangle = \frac{1}{2\pi}\int_0^{2\pi}(\cdot)\,\mathrm{d}\alpha\,. \tag{75}
$$

If we now use that $F_\mathrm{h} = F_\mathrm{h}(\mathbf{r},q_\parallel,\widehat{\mu})$ is independent of the gyro-angle, the energy (74) can be written as

$$
\begin{aligned}
\mathcal{E}_\mathrm{h} &\approx \frac{M_\mathrm{h}}{2\pi}\int\left[\frac{v_\parallel^2(\widehat{\mathbf{q}})}{2} + \mu(\widehat{\mathbf{q}})B(\widehat{\mathbf{q}})\right] F_\mathrm{h}\,\mathrm{d}^6\widehat{\mathbf{q}}\,. \\
&= M_\mathrm{h}\int_{\widehat{\mu}}\left\langle \frac{(q_\parallel + \varepsilon\, G_1^\parallel)^2}{2} + (\widehat{\mu} + \varepsilon\, G_1^\mu)(B + \varepsilon\,\boldsymbol{\rho}_1\cdot\nabla B)\right\rangle F_\mathrm{h}\,\mathrm{d}^3\mathbf{r}\,\mathrm{d}q_\parallel + O(\varepsilon^2) \\
&= M_\mathrm{h}\int_{\widehat{\mu}}\left(\frac{q_\parallel^2}{2} + \widehat{\mu}B\right) F_\mathrm{h}\,\mathrm{d}^3\mathbf{r}\,\mathrm{d}q_\parallel + O(\varepsilon^2)\,,
\end{aligned}
\tag{76}
$$

where we used the generators (A70) to compute

$$
\left\langle q_\parallel G_1^\parallel + G_1^\mu B + \widehat{\mu}\,\boldsymbol{\rho}_1\cdot\nabla B\right\rangle = 0\,. \tag{77}
$$

Now let us define the DK energy as

$$
\mathcal{E}_\mathrm{h}^\mathrm{D} := M_\mathrm{h}\int_{\widehat{\mu}}\left(\frac{q_\parallel^2}{2} + \widehat{\mu}B\right) F_\mathrm{h}\,\mathrm{d}^3\mathbf{r}\,\mathrm{d}q_\parallel\,, \tag{78}
$$

then $\mathcal{E}_\mathrm{h}^\mathrm{D}$ approximates the true energy as

$$
\mathcal{E}_\mathrm{h}(t) \approx \mathcal{E}_\mathrm{h}^\mathrm{D}(t) + O(\varepsilon^2)\,. \tag{79}
$$

Therefore, we want to choose the DK current density $\mathbf{j}_\mathrm{h}^\mathrm{D}$ such that the relation (73) holds exactly, thereby leading to an exact energy conservation law. A first hint towards $\mathbf{j}_\mathrm{h}^\mathrm{D}$ comes from the charge conservation law imposed by (72): since the total amount of energetic particles in a volume V is

$$
\begin{aligned}
\mathcal{Q}_{\mathrm{h},V}(t) &= \int_V f_\mathrm{h}(\mathbf{q},t)\,\mathrm{d}^6\mathbf{q} \\
&= \int_{\widehat{V}} f_\mathrm{h}(\widehat{\tau}^\varepsilon_\mathrm{gy}(\widehat{\mathbf{q}}),t)\,B_\parallel^*\,\mathrm{d}^6\widehat{\mathbf{q}} \\
&\approx \frac{1}{2\pi}\int_{\widehat{V}} F_\mathrm{h}(\widehat{\mathbf{q}},t)\,\mathrm{d}^6\widehat{\mathbf{q}}\,,
\end{aligned}
\tag{80}
$$

where the approximation in the last line is of the same nature as in (74), it makes sense to define the density of gyro-centers as

$$
n_\mathrm{h}^\mathrm{D}(\mathbf{r},t) := \int_{\widehat{\mu}} F_\mathrm{h}(\mathbf{r},q_\parallel,\widehat{\mu})\,\mathrm{d}q_\parallel\,. \tag{81}
$$

The DK equation (72) then yields the charge conservation law

$$
\frac{\partial(Z_\mathrm{h} n_\mathrm{h}^\mathrm{D})}{\partial t} + \nabla\cdot\mathbf{j}_{\mathrm{h},0}^\mathrm{D} = 0\,, \tag{82}
$$

where the zeroth-order DK current density reads

$$
\begin{aligned}
\mathbf{j}_{\mathrm{h},0}^\mathrm{D}(\mathbf{r},t) &:= Z_\mathrm{h}\int_{\widehat{\mu}} \mathbf{u}_\mathrm{g}\,F_\mathrm{h}(\mathbf{r},q_\parallel,\widehat{\mu})\,\mathrm{d}q_\parallel \\
&= Z_\mathrm{h}\int_{\widehat{\mu}} \left(q_\parallel\,\frac{\mathbf{B}^*}{B_\parallel^*} + \varepsilon\,\frac{\mathbf{E}^*\times\mathbf{b}}{B_\parallel^*}\right) F_\mathrm{h}\,\mathrm{d}q_\parallel\,.
\end{aligned}
\tag{83}
$$

Our aim is now to prove the formula (73) for a suitable current density. As a first step for determining $\mathbf{j}_\mathrm{h}^\mathrm{D}$ we compute

$$
\begin{aligned}
\int \mathbf{j}_{\mathrm{h},0}^\mathrm{D}\cdot\mathbf{E}\,\mathrm{d}^3\mathbf{r} &= Z_\mathrm{h}\int_{\widehat{\mu}} \left(q_\parallel\,\frac{\mathbf{B}^*\cdot\mathbf{E}}{B_\parallel^*} + \varepsilon\,\frac{\mathbf{E}^*\times\mathbf{b}\cdot\mathbf{E}}{B_\parallel^*}\right) F_\mathrm{h}\,\mathrm{d}^3\mathbf{r}\,\mathrm{d}q_\parallel \\
&= Z_\mathrm{h}\int_{\widehat{\mu}} \left(q_\parallel\,\frac{\mathbf{B}^*\cdot\mathbf{E}^*}{B_\parallel^*} + \frac{q_\parallel}{C_\mathrm{h}}\,\frac{\mathbf{B}^*\cdot\nabla(\widehat{\mu}B)}{B_\parallel^*} + \varepsilon\,\frac{q_\parallel^2}{C_\mathrm{h}}\,\frac{\mathbf{B}^*\cdot\partial_t\mathbf{b}}{B_\parallel^*}\right) F_\mathrm{h}\,\mathrm{d}^3\mathbf{r}\,\mathrm{d}q_\parallel \\
&\quad + Z_\mathrm{h}\int_{\widehat{\mu}} \left(\varepsilon\,\frac{1}{C_\mathrm{h}}\,\frac{\mathbf{E}^*\times\mathbf{b}\cdot\nabla(\widehat{\mu}B)}{B_\parallel^*} + \varepsilon^2\,\frac{q_\parallel}{C_\mathrm{h}}\,\frac{\mathbf{E}^*\times\mathbf{b}\cdot\partial_t\mathbf{b}}{B_\parallel^*}\right) F_\mathrm{h}\,\mathrm{d}^3\mathbf{r}\,\mathrm{d}q_\parallel \\
&= M_\mathrm{h}\int_{\widehat{\mu}} q_\parallel\,E_\mathrm{g} F_\mathrm{h}\,\mathrm{d}^3\mathbf{r}\,\mathrm{d}q_\parallel + M_\mathrm{h}\int_{\widehat{\mu}} \mathbf{u}_\mathrm{g}\cdot\left[\nabla(\widehat{\mu}B) + \varepsilon\,q_\parallel\,\frac{\partial\mathbf{b}}{\partial t}\right] F_\mathrm{h}\,\mathrm{d}^3\mathbf{r}\,\mathrm{d}q_\parallel\,.
\end{aligned}
\tag{84}
$$

On the other hand, the time evolution of the DK energy (100) consists of three pieces:

$$
\begin{aligned}
\varepsilon \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\mathrm{h}}^{\mathrm{D}} = \varepsilon\, M_{\mathrm{h}} \int_{\widehat{\mu}} \frac{q_\parallel^2}{2}\, \frac{\partial F_{\mathrm{h}}}{\partial t}\, \mathrm{d}^3\mathbf{r}\, \mathrm{d}q_\parallel \\
+ \varepsilon\, M_{\mathrm{h}} \int_{\widehat{\mu}} \widehat{\mu} B\, \frac{\partial F_{\mathrm{h}}}{\partial t}\, \mathrm{d}^3\mathbf{r}\, \mathrm{d}q_\parallel \\
+ \varepsilon\, M_{\mathrm{h}} \int_{\widehat{\mu}} \widehat{\mu}\, \frac{\partial B}{\partial t}\, F_{\mathrm{h}}\, \mathrm{d}^3\mathbf{r}\, \mathrm{d}q_\parallel\,.
\end{aligned}
\tag{85}
$$

The first two pieces can be computed from the DK equation (72):

$$
\begin{aligned}
\varepsilon\, M_{\mathrm{h}} \int_{\widehat{\mu}} \frac{q_\parallel^2}{2}\, \frac{\partial F_{\mathrm{h}}}{\partial t}\, \mathrm{d}^3\mathbf{r}\, \mathrm{d}q_\parallel &= -M_{\mathrm{h}} \int_{\widehat{\mu}} \frac{q_\parallel^2}{2}\, \frac{\partial (E_{\mathrm{g}} F_{\mathrm{h}})}{\partial q_\parallel}\, \mathrm{d}^3\mathbf{r}\, \mathrm{d}q_\parallel \\
&= M_{\mathrm{h}} \int_{\widehat{\mu}} q_\parallel\, E_{\mathrm{g}} F_{\mathrm{h}}\, \mathrm{d}^3\mathbf{r}\, \mathrm{d}q_\parallel\,,
\end{aligned}
\tag{86}
$$

and

$$
\begin{aligned}
\varepsilon\, M_{\mathrm{h}} \int_{\widehat{\mu}} \widehat{\mu} B\, \frac{\partial F_{\mathrm{h}}}{\partial t}\, \mathrm{d}^3\mathbf{r}\, \mathrm{d}q_\parallel &= -M_{\mathrm{h}} \int_{\widehat{\mu}} \widehat{\mu} B\, \nabla\cdot(\mathbf{u}_{\mathrm{g}} F_{\mathrm{h}})\, \mathrm{d}^3\mathbf{r}\, \mathrm{d}q_\parallel \\
&= M_{\mathrm{h}} \int_{\widehat{\mu}} \nabla(\widehat{\mu} B)\cdot \mathbf{u}_{\mathrm{g}} F_{\mathrm{h}}\, \mathrm{d}^3\mathbf{r}\, \mathrm{d}q_\parallel\,.
\end{aligned}
\tag{87}
$$

By comparison of (84) with (85)-(87) we obtain

$$
\varepsilon \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\mathrm{h}}^{\mathrm{D}} - \varepsilon\, M_{\mathrm{h}} \int_{\widehat{\mu}} \widehat{\mu}\, \frac{\partial B}{\partial t}\, F_{\mathrm{h}}\, \mathrm{d}^3\mathbf{r}\, \mathrm{d}q_\parallel = \int \mathbf{j}_{\mathrm{h},0}^{\mathrm{D}}\cdot \mathbf{E}\, \mathrm{d}^3\mathbf{r} - \varepsilon\, M_{\mathrm{h}} \int_{\widehat{\mu}} q_\parallel\, \mathbf{u}_{\mathrm{g}}\cdot \frac{\partial \mathbf{b}}{\partial t}\, F_{\mathrm{h}}\, \mathrm{d}^3\mathbf{r}\, \mathrm{d}q_\parallel \tag{88}
$$

Moreover, from Faraday's law $\partial_t \mathbf{B} = -\nabla\times\mathbf{E}$ we compute

$$
\frac{\partial B}{\partial t} = -\nabla\times\mathbf{E}\cdot\mathbf{b}\,, \qquad \frac{\partial \mathbf{b}}{\partial t} = -\frac{(\nabla\times\mathbf{E})_\perp}{B}\,, \tag{89}
$$

which, after inserting into (88) and integrating by parts, yields

$$
\varepsilon \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\mathrm{h}}^{\mathrm{D}} = \int \left[\mathbf{j}_{\mathrm{h},0}^{\mathrm{D}} - \varepsilon\, \nabla\times \int_{\widehat{\mu}} M_{\mathrm{h}} \left(\widehat{\mu}\, \mathbf{b} - \frac{q_\parallel}{B}\, \mathbf{u}_{\mathrm{g},\perp}\right) F_{\mathrm{h}}\, \mathrm{d}q_\parallel\right]\cdot \mathbf{E}\, \mathrm{d}^3\mathbf{r} \tag{90}
$$

Assuming that the domain of integration in $\mathbf{r}$ is the whole $\mathbb{R}^3$ we can relabel $\mathbf{r}\to\mathbf{x}$ to obtain the desired result

$$
\varepsilon \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\mathrm{h}}^{\mathrm{D}} = \int \left[\mathbf{j}_{\mathrm{h},0}^{\mathrm{D}} + \varepsilon\, \nabla\times\mathbf{M}\right]\cdot \mathbf{E}\, \mathrm{d}^3\mathbf{x}\,, \tag{91}
$$

with the magnetization

$$
\mathbf{M} := -M_{\mathrm{h}} \int_{\widehat{\mu}} \left(\widehat{\mu}\, \mathbf{b} - \frac{q_\parallel}{B}\, \mathbf{u}_{\mathrm{g},\perp}\right) F_{\mathrm{h}}\, \mathrm{d}q_\parallel\,. \tag{92}
$$

The second term in the magnetization $\mathbf{M}$ is known as the moving-dipole correction[1]. Therefore, if in (62) we replace $n_\mathrm{h} \to n_\mathrm{h}^\mathrm{D}$ and $\mathbf{j}_\mathrm{h} \to \mathbf{j}_\mathrm{h}^\mathrm{D}$, where

$$\mathbf{j}_\mathrm{h}^\mathrm{D} := \mathbf{j}_{\mathrm{h},0}^\mathrm{D} + \varepsilon\, \nabla \times \mathbf{M}\,, \tag{93}$$

we obtain the following energy theorem for the reduced DK-MHD system:

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{E}_\mathrm{MHD} + \varepsilon\, \mathcal{E}_\mathrm{h}^\mathrm{D}) = 0\,. \tag{94}$$

H. The DK-MHD model in SI units

The reduced DK-MHD model in the current coupling regime with finite density of energetic particles, derived in the previous section, is stated here in SI units. Denoting by $a_\mathrm{s} = q_\mathrm{s}/m_\mathrm{s}$ and $a_\mathrm{h} = q_\mathrm{h}/m_\mathrm{h}$ the charge to mass ratio of the respective species, we have

$$\frac{\partial \tau}{\partial t} + \nabla \cdot (\tau \mathbf{U}) = 0\,, \tag{95a}$$

$$\frac{\partial \mathbf{U}}{\partial t} + \mathbf{U}\cdot\nabla\mathbf{U} + \frac{\nabla(p_\mathrm{i} + p_\mathrm{e})}{\tau} = -\frac{q_\mathrm{h} n_\mathrm{h}^\mathrm{D}}{\tau}\,\mathbf{E} + \frac{1}{\mu_0}\,\frac{(\nabla\times\mathbf{B} - \mu_0\,\mathbf{j}_\mathrm{h}^\mathrm{D})}{\tau} \times \mathbf{B}\,, \tag{95b}$$

$$\begin{aligned}\frac{\partial p_\mathrm{i}}{\partial t} &+ \left[\mathbf{U}\left(1 + a_\mathrm{e}\,\frac{q_\mathrm{h} n_\mathrm{h}^\mathrm{D}}{\tau}\right) + \frac{a_\mathrm{e}}{\mu_0}\,\frac{(\nabla\times\mathbf{B} - \mu_0\,\mathbf{j}_\mathrm{h}^\mathrm{D})}{\tau}\right]\cdot\nabla p_\mathrm{i} \\ &\quad + \gamma\, p_\mathrm{i} \nabla\cdot\left[\mathbf{U}\left(1 + a_\mathrm{e}\,\frac{q_\mathrm{h} n_\mathrm{h}^\mathrm{D}}{\tau}\right) + \frac{a_\mathrm{e}}{\mu_0}\,\frac{(\nabla\times\mathbf{B} - \mu_0\,\mathbf{j}_\mathrm{h}^\mathrm{D})}{\tau}\right] = 0\,,\end{aligned} \tag{95c}$$

$$\begin{aligned}\frac{\partial p_\mathrm{e}}{\partial t} &+ \left[\mathbf{U}\left(1 - a_\mathrm{i}\,\frac{q_\mathrm{h} n_\mathrm{h}^\mathrm{D}}{\tau}\right) - \frac{a_\mathrm{i}}{\mu_0}\,\frac{(\nabla\times\mathbf{B} - \mu_0\,\mathbf{j}_\mathrm{h}^\mathrm{D})}{\tau}\right]\cdot\nabla p_\mathrm{e} \\ &\quad + \gamma\, p_\mathrm{e} \nabla\cdot\left[\mathbf{U}\left(1 - a_\mathrm{i}\,\frac{q_\mathrm{h} n_\mathrm{h}^\mathrm{D}}{\tau}\right) - \frac{a_\mathrm{i}}{\mu_0}\,\frac{(\nabla\times\mathbf{B} - \mu_0\,\mathbf{j}_\mathrm{h}^\mathrm{D})}{\tau}\right] = 0\,,\end{aligned} \tag{95d}$$

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla\times\mathbf{E}\,, \tag{95e}$$

$$\begin{aligned}\mathbf{E} = -\mathbf{U}\times\mathbf{B}&\left[1 - \frac{q_\mathrm{h} n_\mathrm{h}^\mathrm{D}}{\tau}\,(a_\mathrm{i} - a_\mathrm{e})\right] \\ &+ \frac{1}{\tau}\nabla\,(a_\mathrm{e}\, p_\mathrm{i} - a_\mathrm{i}\, p_\mathrm{e}) + (a_\mathrm{i} - a_\mathrm{e})\,\frac{1}{\mu_0}\,\frac{(\nabla\times\mathbf{B} - \mu_0\,\mathbf{j}_\mathrm{h}^\mathrm{D})}{\tau}\times\mathbf{B}\,,\end{aligned} \tag{95f}$$

$$\frac{\partial f_\mathrm{h}^\mathrm{D}}{\partial t} + \mathbf{u}_\mathrm{g}\cdot\frac{\partial f_\mathrm{h}^\mathrm{D}}{\partial \mathbf{r}} + E_\mathrm{g}\,\frac{\partial f_\mathrm{h}^\mathrm{D}}{\partial q_\parallel} = 0\,. \tag{95g}$$

Here, the GY-velocity and the GY-acceleration are given by

$$\mathbf{u}_\mathrm{g} = q_\| \frac{\mathbf{B}^*}{B_\|^*} + \frac{\mathbf{E}^* \times \mathbf{b}}{B_\|^*}\,, \qquad E_\mathrm{g} = a_\mathrm{h} \frac{\mathbf{E}^* \cdot \mathbf{B}^*}{B_\|^*}\,, \tag{96}$$

with the fields $\mathbf{B}^*$ and $\mathbf{E}^*$ defined as

$$\begin{aligned} \mathbf{B}^* &= \mathbf{B} + \frac{q_\|}{a_\mathrm{h}} \nabla \times \mathbf{b}\,, \\ \mathbf{E}^* &= \mathbf{E} - \frac{\widehat{\mu}}{a_\mathrm{h}} \nabla B - \frac{q_\|}{a_\mathrm{h}} \frac{\partial \mathbf{b}}{\partial t}\,. \end{aligned} \tag{97}$$

The density and the current of the energetic particles read

$$n_\mathrm{h}^\mathrm{D} = \int_{\widehat{\mu}} f_\mathrm{h}^\mathrm{D}\, B_\|^*\, \mathrm{d}q_\|\,, \qquad \mathbf{j}_\mathrm{h}^\mathrm{D} = q_\mathrm{h} \int_{\widehat{\mu}} \mathbf{u}_\mathrm{g}\, f_\mathrm{h}^\mathrm{D}\, B_\|^*\, \mathrm{d}q_\| + \nabla \times \mathbf{M}\,, \tag{98}$$

where the magnetization $\mathbf{M}$ is defined by

$$\mathbf{M} := -m_\mathrm{h} \int_{\widehat{\mu}} \left[\widehat{\mu}\, \mathbf{b} - \frac{q_\|}{B}\, \mathbf{u}_{\mathrm{g},\perp} \right] f_\mathrm{h}^\mathrm{D}\, B_\|^*\, \mathrm{d}q_\|\,, \tag{99}$$

The following energy is conserved exactly by the system (95):

$$\begin{aligned} \mathcal{E}_\mathrm{tot} = \frac{3}{2} \int (p_\mathrm{i} + p_\mathrm{e})\, \mathrm{d}^3\mathbf{x} + \int \tau\, \frac{|\mathbf{U}|^2}{2}\, \mathrm{d}^3\mathbf{x} + \frac{1}{\mu_0} \int \frac{|\mathbf{B}|^2}{2}\, \mathrm{d}^3\mathbf{x} \\ + m_\mathrm{h} \int_{\widehat{\mu}} \left(\frac{q_\|^2}{2} + \widehat{\mu} B \right) f_\mathrm{h}^\mathrm{D}\, B_\|^*\, \mathrm{d}^3\mathbf{x}\, \mathrm{d}q_\|\,. \end{aligned} \tag{100}$$

III. SUB-MODELS AND DISPERSION RELATIONS

In this section we state some sub-models that are contained in the hybrid model (95). We start from the standard MHD equations and keep adding terms in an energetically consistent way so to eventually arrive at the full model (95) again.

A. Standard MHD

1. Equations

The simplest sub-model of (62) is standard MHD. In standard MHD the energetic particles are neglected, $n_{\mathrm{h}}^{\mathrm{D}} = \mathbf{j}_{\mathrm{h}}^{\mathrm{D}} = 0$, and the electric field takes the simple form

$$\mathbf{E} = -\mathbf{U} \times \mathbf{B}\,. \tag{101}$$

Writing $p = p_{\mathrm{i}} + p_{\mathrm{e}}$, the standard MHD equations read

$$\frac{\partial \tau}{\partial t} + \nabla \cdot (\tau \mathbf{U}) = 0\,, \tag{102a}$$

$$\frac{\partial \mathbf{U}}{\partial t} + \mathbf{U} \cdot \nabla \mathbf{U} + \frac{\nabla p}{\tau} = \frac{1}{\mu_0} \frac{\nabla \times \mathbf{B}}{\tau} \times \mathbf{B}\,, \tag{102b}$$

$$\frac{\partial p}{\partial t} + \mathbf{U} \cdot \nabla p + \gamma\, p\, \nabla \cdot \mathbf{U} = 0\,, \tag{102c}$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{U} \times \mathbf{B})\,. \tag{102d}$$

2. Dispersion relation for homogeneous plasma

Linearized MHD equations are obtained from (102) by keeping only the first order terms in an expansion of the solution around the homogeneous equilibrium

$$\tau_0 = const.\,, \qquad \mathbf{U}_0 = 0\,, \qquad p_0 = const.\,, \qquad \mathbf{B}_0 = B_0\, \mathbf{e}_z\,. \tag{103}$$

Writing $\tau = \tau_0 + \tau_1$ and relabeling $\tau_1 \to \tau$ after linearization yields the system

$$\frac{\partial \tau}{\partial t} + \tau_0 \nabla \cdot \mathbf{U} = 0\,, \tag{104a}$$

$$\tau_0 \frac{\partial \mathbf{U}}{\partial t} + \nabla p = \frac{B_0}{\mu_0} (\nabla \times \mathbf{B}) \times \mathbf{e}_z\,, \tag{104b}$$

$$\frac{\partial p}{\partial t} + \gamma\, p_0\, \nabla \cdot \mathbf{U} = 0\,, \tag{104c}$$

$$\frac{\partial \mathbf{B}}{\partial t} = B_0\, \nabla \times (\mathbf{U} \times \mathbf{e}_z)\,. \tag{104d}$$

We are interested in waves propagating in the direction of the magnetic field, hence with wave vector $\mathbf{k} = k_\parallel \mathbf{e}_z$. Writing the unknowns as $\mathbf{w} = (\tau, \mathbf{U}, p, \mathbf{B})^\top \in \mathbb{R}^8$, we insert into (104) the ansatz $\mathbf{w}(z,t) = \widehat{\mathbf{w}}\, e^{i(k_\parallel z - \omega t)}$ with $\widehat{\mathbf{w}} \in \mathbb{R}^8$ constant to obtain the eigenvalue problem

$$-\,\omega\, \mathbf{w} + \mathbf{A} \cdot \mathbf{w} = 0\,, \tag{105}$$

where the matrix $\mathbf{A} \in \mathbb{R}^{8\times 8}$ can be deduced from

$$-\,\omega\, \tau + \tau_0\, k_\parallel U_\parallel = 0\,, \tag{106a}$$

$$-\,\omega\, \mathbf{U} + \frac{1}{\tau_0} k_\parallel\, p\, \mathbf{e}_z - \frac{B_0}{\tau_0 \mu_0}\, k_\parallel \mathbf{B}_\perp = 0\,, \tag{106b}$$

$$-\,\omega\, p + \gamma\, p_0\, k_\parallel U_\parallel = 0\,, \tag{106c}$$

$$-\,\omega\, \mathbf{B} - B_0\, k_\parallel \mathbf{U}_\perp = 0\,, \tag{106d}$$

and we used the notation $U_\parallel = \mathbf{U} \cdot \mathbf{e}_z$ and $\mathbf{U}_\perp = \mathbf{e}_z \times \mathbf{U} \times \mathbf{e}_z = (U_x, U_y, 0)^\top$. Solving (105) is straightforward: from (106) we obtain

$$\tau = \frac{\tau_0}{\omega}\, k_\parallel U_\parallel\,, \qquad p = \frac{\gamma\, p_0}{\omega}\, k_\parallel U_\parallel\,, \qquad \mathbf{B} = \mathbf{B}_\perp = -\frac{B_0}{\omega}\, k_\parallel \mathbf{U}_\perp\,, \tag{107}$$

which inserted into the momentum equation leads to

$$\omega^2\, \mathbf{U} - \frac{\gamma\, p_0}{\tau_0}\, k_\parallel^2\, U_\parallel\, \mathbf{e}_z - \frac{B_0^2}{\tau_0 \mu_0}\, k_\parallel^2\, \mathbf{U}_\perp = 0\,. \tag{108}$$

The two relevant velocities are the sound speed v_s and the Alfvén velocity v_A,

$$v_\mathrm{s} = \sqrt{\frac{\gamma\, p_0}{\tau_0}}\,, \qquad v_\mathrm{A} = \sqrt{\frac{B_0^2}{\tau_0 \mu_0}}\,, \tag{109}$$

such that (108) reads

$$\begin{pmatrix} \omega^2 - v_{\mathrm{A}}^2\, k_\parallel^2 & 0 & 0 \\ 0 & \omega^2 - v_{\mathrm{A}}^2\, k_\parallel^2 & 0 \\ 0 & 0 & \omega^2 - v_{\mathrm{s}}^2\, k_\parallel^2 \end{pmatrix} \begin{pmatrix} U_x \\ U_y \\ U_\parallel \end{pmatrix} = 0\,. \tag{110}$$

Hence, there are two kinds of eigenmodes:

1. Sound waves: $\omega = \pm c_{\mathrm{s}} k_\parallel$. In this case $\mathbf{U}_\perp = 0$ and thus there is no perturbation of the magnetic field due to (107).

2. Shear Alfvén waves: $\omega = \pm v_{\mathrm{A}} k_\parallel$. In this case $U_\parallel = 0$, the magnetic wave is transversal and there is no perturbation of the density and the pressure due to (107).

3. Dispersion relation with radial profiles

B. Standard MHD with Vlasov current coupling

1. Equations

In standard MHD, the electric field is given by (101) and we can set the total pressure $p = p_\mathrm{i} + p_\mathrm{e}$ a an unknown. The energetic particle contribution is computed from the Vlasov equation (5). The simplest current-coupling scheme is thus

$$\frac{\partial \tau}{\partial t} + \nabla \cdot (\tau \mathbf{U}) = 0\,, \tag{111a}$$

$$\frac{\partial \mathbf{U}}{\partial t} + \mathbf{U} \cdot \nabla \mathbf{U} + \frac{\nabla p}{\tau} = \frac{1}{\mu_0} \frac{\nabla \times \mathbf{B}}{\tau} \times \mathbf{B} + \frac{(q_\mathrm{h} n_\mathrm{h} \mathbf{U} - \mathbf{j}_\mathrm{h})}{\tau} \times \mathbf{B}\,, \tag{111b}$$

$$\frac{\partial p}{\partial t} + \mathbf{U} \cdot \nabla p + \gamma\, p\, \nabla \cdot \mathbf{U} = 0\,, \tag{111c}$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{U} \times \mathbf{B})\,, \tag{111d}$$

$$\frac{\partial f_\mathrm{h}}{\partial t} + \mathbf{v} \cdot \nabla f_\mathrm{h} + \frac{q_\mathrm{h}}{m_\mathrm{h}} (\mathbf{v} - \mathbf{U}) \times \mathbf{B} \cdot \frac{\partial f_\mathrm{h}}{\partial \mathbf{v}} = 0\,, \tag{111e}$$

where

$$n_\mathrm{h} := \int f_\mathrm{h}\, \mathrm{d}\mathbf{v}\,, \qquad \mathbf{j}_\mathrm{h} := q_\mathrm{h} \int \mathbf{v} f_\mathrm{h}\, \mathrm{d}\mathbf{v}\,. \tag{112}$$

2. Dispersion relation for homogeneous plasma

We now compute the linear wave-particle interaction, where the word "wave" refers to a plane wave solution for the MHD part as in (105) and the word "particle" refers to the hot particles evolved by the Vlasov equation. The Vlasov equation is thus linearized around the homogeneous equilibrium (103):

$$\frac{\partial f_\mathrm{h}}{\partial t} + \mathbf{v} \cdot \nabla f_\mathrm{h} + \Omega_\mathrm{ch} \mathbf{v} \times \mathbf{e}_z \cdot \frac{\partial f_\mathrm{h}}{\partial \mathbf{v}} = -a_\mathrm{h} \left(\mathbf{E} + \mathbf{v} \times \mathbf{B} \right) \cdot \frac{\partial f_0}{\partial \mathbf{v}}\,, \tag{113}$$

where $\Omega_\mathrm{ch} = q_\mathrm{h} B_0 / m_\mathrm{h}$ denotes the hot particles' cyclotron frequency, $f_0 = f_{\mathrm{h},0}(\mathbf{v})$ is a homogeneous equilibrium distribution and we decided to use a general electric field perturbation $\mathbf{E}$ instead of $\mathbf{E} = -\mathbf{U} \times \mathbf{B}_0$. This will be convenient later when more terms are added to the electric field (extended MHD). We can solve equation (113) exactly as follows: the left hand side describes advection along the so-called unperturbed characteristics $\mathbf{X} = \mathbf{X}(s; t, \mathbf{x}, \mathbf{v})$

and $\mathbf{V} = \mathbf{V}(s; t, \mathbf{x}, \mathbf{v})$, solutions of

$$\begin{aligned} \frac{\mathrm{d}\mathbf{X}}{\mathrm{d}s} &= \mathbf{V}\,, & \mathbf{X}(t; t, \mathbf{x}, \mathbf{v}) &= \mathbf{x}\,, \\ \frac{\mathrm{d}\mathbf{V}}{\mathrm{d}s} &= \Omega_{\mathrm{ch}} \mathbf{V} \times \mathbf{e}_z\,, & \mathbf{V}(t; t, \mathbf{x}, \mathbf{v}) &= \mathbf{v}\,. \end{aligned} \tag{114}$$

We observe that the parallel part $V_\parallel = \mathbf{V} \cdot \mathbf{e}_z$ of the velocity is constant, hence $V_\parallel = v_\parallel$, and that the perpendicular part reads

$$\mathbf{V}_\perp = e^{\Omega_{\mathrm{ch}}(s-t)\mathbf{R}}\, \mathbf{v}_\perp\,, \qquad \mathbf{R} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}\,. \tag{115}$$

The series representation of the matrix exponential leads to

$$e^{\Omega_{\mathrm{ch}}(s-t)\mathbf{R}} = \cos[\Omega_{\mathrm{ch}}(s-t)]\, \mathbf{I}_2 + \sin[\Omega_{\mathrm{ch}}(s-t)]\, \mathbf{R} = \begin{pmatrix} \cos[\Omega_{\mathrm{ch}}(s-t)] & \sin[\Omega_{\mathrm{ch}}(s-t)] \\ -\sin[\Omega_{\mathrm{ch}}(s-t)] & \cos[\Omega_{\mathrm{ch}}(s-t)] \end{pmatrix}\,,$$

and, therefore, the velocity part of the characteristic is

$$\begin{pmatrix} V_x \\ V_y \\ V_\parallel \end{pmatrix} = \begin{pmatrix} \cos[\Omega_{\mathrm{ch}}(s-t)]\, v_x + \sin[\Omega_{\mathrm{ch}}(s-t)]\, v_y \\ -\sin[\Omega_{\mathrm{ch}}(s-t)]\, v_x + \cos[\Omega_{\mathrm{ch}}(s-t)]\, v_y \\ v_\parallel \end{pmatrix}\,. \tag{116}$$

The positional part $\mathbf{X} = (X, Y, Z)^\top$ is then obtained by integration and reads

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \mathbf{x} + \begin{pmatrix} \frac{1}{\Omega_{\mathrm{ch}}} \sin[\Omega_{\mathrm{ch}}(s-t)]\, v_x - \frac{1}{\Omega_{\mathrm{ch}}} \cos[\Omega_{\mathrm{ch}}(s-t)]\, v_y + \frac{1}{\Omega_{\mathrm{ch}}}\, v_y \\ \frac{1}{\Omega_{\mathrm{ch}}} \cos[\Omega_{\mathrm{ch}}(s-t)]\, v_x + \frac{1}{\Omega_{\mathrm{ch}}} \sin[\Omega_{\mathrm{ch}}(s-t)\, v_y - \frac{1}{\Omega_{\mathrm{ch}}}\, v_x \\ (s-t)\, v_\parallel \end{pmatrix}\,. \tag{117}$$

According to the linear Vlasov equation (113), the rate of change of f_{h} when moving along a characteristic is given by the right-hand-side of (113), evaluated at the characteristic:

$$\frac{\mathrm{d}}{\mathrm{d}s} f_{\mathrm{h}}(\mathbf{X}, \mathbf{V}, s) = -a_{\mathrm{h}}\, [\mathbf{E}(\mathbf{X}, s) + \mathbf{V} \times \mathbf{B}(\mathbf{X}, s)] \cdot \frac{\partial f_0}{\partial \mathbf{v}}(\mathbf{V})\,. \tag{118}$$

We can integrate this equation in time from $s = -\infty$ to $s = t$ and assume $f_{\mathrm{h}}(\mathbf{X}, \mathbf{V}, -\infty) = 0$ to obtain

$$f_{\mathrm{h}}(\mathbf{x}, \mathbf{v}, t) = -a_{\mathrm{h}} \int_{-\infty}^{t} [\mathbf{E}(\mathbf{X}, s) + \mathbf{V} \times \mathbf{B}(\mathbf{X}, s)] \cdot \frac{\partial f_0}{\partial \mathbf{v}}(\mathbf{V})\, \mathrm{d}s\,. \tag{119}$$

Assuming an isotropic equilibrium, $f_0 = f_0(v)$, where $v = |\mathbf{v}|$, we obtain

$$\frac{\partial f_0}{\partial \mathbf{v}} = \frac{\partial f_0}{\partial v} \frac{\partial v}{\partial \mathbf{v}} = \frac{\partial f_0}{\partial v} \frac{\mathbf{v}}{v}\,, \tag{120}$$

and thus

$$\mathbf{V}\times\mathbf{B}(\mathbf{X},s)\cdot\frac{\partial f_0}{\partial\mathbf{v}}(\mathbf{V})=\mathbf{V}\times\mathbf{B}(\mathbf{X},s)\cdot\mathbf{V}\,\frac{\partial f_0}{\partial v}\frac{1}{v}(\mathbf{V})=0\,. \tag{121}$$

Moreover, the equations (114) tell us that $|\mathbf{V}|$ is a constant of the motion, $|\mathbf{V}|=v$, thus

$$\frac{\partial f_0}{\partial v}\frac{1}{v}(\mathbf{V})=\frac{1}{v}\frac{\partial f_0}{\partial v}(v)\,. \tag{122}$$

Expressing the electric field as a plane wave moving along the magnetic field,

$$\mathbf{E}(\mathbf{X},s)=\widehat{\mathbf{E}}\,e^{i(k_\parallel Z-\omega s)}\,, \tag{123}$$

the solution f_h from (119) can be written as

$$f_\mathrm{h}(\mathbf{x},\mathbf{v},t)=-a_\mathrm{h}\,\frac{1}{v}\frac{\partial f_0}{\partial v}\,\widehat{\mathbf{E}}\cdot\int_{-\infty}^{t}\mathbf{V}\,e^{i(k_\parallel Z-\omega s)}\,\mathrm{d}s\,. \tag{124}$$

Performing the change of variables $s'=s-t$ and inserting the characteristics (116) and Z from (117), respectively, leads to

$$\begin{aligned}
f_\mathrm{h}(\mathbf{x},\mathbf{v},t)&=-a_\mathrm{h}\,\frac{1}{v}\frac{\partial f_0}{\partial v}\,\widehat{\mathbf{E}}\,e^{i(k_\parallel z-\omega t)}\cdot\int_{-\infty}^{0}\mathbf{V}\,e^{i(k_\parallel v_\parallel-\omega)s'}\,\mathrm{d}s'\\
&=-a_\mathrm{h}\,\frac{1}{v}\frac{\partial f_0}{\partial v}\,\mathbf{E}\cdot\int_{-\infty}^{0}\begin{pmatrix}\cos(\Omega_\mathrm{ch}s')\,v_x+\sin(\Omega_\mathrm{ch}s')\,v_y\\-\sin(\Omega_\mathrm{ch}s')\,v_x+\cos(\Omega_\mathrm{ch}s')\,v_y\\v_\parallel\end{pmatrix}e^{i(k_\parallel v_\parallel-\omega)s'}\,\mathrm{d}s'\,.
\end{aligned} \tag{125}$$

Here, assuming that ω has a positive imaginary part, $Im(\omega)>0$ (growing modes), we can solve the following three integrals:

$$\int_{-\infty}^{0}e^{i(k_\parallel v_\parallel-\omega)s'}\,\mathrm{d}s'=-i\,\frac{e^{i(k_\parallel v_\parallel-\omega)s'}}{k_\parallel v_\parallel-\omega}\Bigg|_{-\infty}^{0}=-i\,\frac{1}{k_\parallel v_\parallel-\omega}=:-i\,P_0(v_\parallel)\,, \tag{126}$$

$$\begin{aligned}
\int_{-\infty}^{0}\cos(\Omega_\mathrm{ch}s')\,e^{i(k_\parallel v_\parallel-\omega)s'}\,\mathrm{d}s'&=\int_{-\infty}^{0}\frac{1}{2}\Big(e^{i\Omega_\mathrm{ch}s'}+e^{-i\Omega_\mathrm{ch}s'}\Big)\,e^{i(k_\parallel v_\parallel-\omega)s'}\,\mathrm{d}s'\\
&=\frac{1}{2}\Bigg[-i\,\frac{e^{i(k_\parallel v_\parallel+\Omega_\mathrm{ch}-\omega)s'}}{k_\parallel v_\parallel+\Omega_\mathrm{ch}-\omega}-i\,\frac{e^{i(k_\parallel v_\parallel-\Omega_\mathrm{ch}-\omega)s'}}{k_\parallel v_\parallel-\Omega_\mathrm{ch}-\omega}\Bigg]_{-\infty}^{0}\\
&=\frac{-i}{2}\Bigg(\frac{1}{k_\parallel v_\parallel+\Omega_\mathrm{ch}-\omega}+\frac{1}{k_\parallel v_\parallel-\Omega_\mathrm{ch}-\omega}\Bigg)=:-\frac{i}{2}\,P_1(v_\parallel)\,,
\end{aligned} \tag{127}$$

$$\begin{aligned}
\int_{-\infty}^{0}\sin(\Omega_\mathrm{ch}s')\,e^{i(k_\parallel v_\parallel-\omega)s'}\,\mathrm{d}s'&=\int_{-\infty}^{0}\frac{1}{2i}\Big(e^{i\Omega_\mathrm{ch}s'}-e^{-i\Omega_\mathrm{ch}s'}\Big)\,e^{i(k_\parallel v_\parallel-\omega)s'}\,\mathrm{d}s'\\
&=\frac{1}{2}\Bigg[-\frac{e^{i(k_\parallel v_\parallel+\Omega_\mathrm{ch}-\omega)s'}}{k_\parallel v_\parallel+\Omega_\mathrm{ch}-\omega}+\frac{e^{i(k_\parallel v_\parallel-\Omega_\mathrm{ch}-\omega)s'}}{k_\parallel v_\parallel-\Omega_\mathrm{ch}-\omega}\Bigg]_{-\infty}^{0}\\
&=-\frac{1}{2}\Bigg(\frac{1}{k_\parallel v_\parallel+\Omega_\mathrm{ch}-\omega}-\frac{1}{k_\parallel v_\parallel-\Omega_\mathrm{ch}-\omega}\Bigg)=:-\frac{1}{2}\,P_2(v_\parallel)\,.
\end{aligned} \tag{128}$$

The solution (125) can now be written in the compact form

$$f_{\mathrm{h}}(\mathbf{x},\mathbf{v},t) = -a_{\mathrm{h}}\,\frac{1}{v}\frac{\partial f_0}{\partial v}\,\mathbf{E}\cdot\begin{pmatrix} -\frac{i}{2}\,P_1(v_\parallel)\,v_x - \frac{1}{2}\,P_2(v_\parallel)\,v_y \\ \frac{1}{2}\,P_2(v_\parallel)\,v_x - \frac{i}{2}\,P_1(v_\parallel)\,v_y \\ -i\,P_0(v_\parallel)\,v_\parallel \end{pmatrix}. \tag{129}$$

The linearized MHD part is similar to (104), now with a modified momentum equation of the form

$$\tau_0\,\frac{\partial \mathbf{U}}{\partial t} + \nabla p = \frac{B_0}{\mu_0}\,(\nabla\times\mathbf{B})\times\mathbf{e}_z + B_0\,(q_{\mathrm{h}} n_{\mathrm{h},0}\mathbf{U} - \mathbf{j}_{\mathrm{h}})\times\mathbf{e}_z\,, \tag{130}$$

where $n_{\mathrm{h},0}$ denotes a constant equilibrium density the hot particles and $\mathbf{j}_{\mathrm{h}}$ is computed from f_{h} via (112). Inserting the plane wave solution yields

$$-\,\omega\,\mathbf{U} + \frac{1}{\tau_0}k_\parallel\,p\,\mathbf{e}_z - \frac{B_0}{\tau_0\mu_0}\,k_\parallel\mathbf{B}_\perp + i\,\frac{B_0}{\tau_0}\,(q_{\mathrm{h}} n_{\mathrm{h},0}\mathbf{U} - \mathbf{j}_{\mathrm{h}})\times\mathbf{e}_z = 0\,, \tag{131}$$

which, after inserting (107) and multiplying by $-\omega$ becomes

$$\omega^2\,\mathbf{U} - v_{\mathrm{s}}^2\,k_\parallel^2\,U_\parallel\,\mathbf{e}_z - v_{\mathrm{A}}^2\,k_\parallel^2\,\mathbf{U}_\perp - i\,\omega\,\frac{B_0}{\tau_0}\,(q_{\mathrm{h}} n_{\mathrm{h},0}\mathbf{U} - \mathbf{j}_{\mathrm{h}})\times\mathbf{e}_z = 0\,. \tag{132}$$

It remains to express the current density $\mathbf{j}_{\mathrm{h}}$ in terms of the flow velocity $\mathbf{U}$. For this, let us assume that f_0 is Maxwellian,

$$f_0 = f_0(v) = \frac{n_{\mathrm{h},0}}{(2\pi T_{\mathrm{h},0})^{3/2}}\,e^{-v^2/(2T_{\mathrm{h},0})}\,,\qquad \frac{\partial f_0}{\partial v} = -f_0\,\frac{v}{T_{\mathrm{h},0}}\,, \tag{133}$$

where $T_{\mathrm{h},0}$ stands for a constant equilibrium temperature of the hot particles. We shall be aware of the following integrals:

$$\int_{-\infty}^{\infty}\begin{pmatrix}1\\ u\\ u^2\end{pmatrix}\frac{1}{(2\pi T_{\mathrm{h},0})^{1/2}}\,e^{-u^2/(2T_{\mathrm{h},0})}\,\mathrm{d}u = \begin{pmatrix}1\\ 0\\ T_{\mathrm{h},0}\end{pmatrix}. \tag{134}$$

Hence, from (112) and (129) we obtain

$$\begin{aligned}\mathbf{j}_{\mathrm{h}} &= \frac{q_{\mathrm{h}} a_{\mathrm{h}}}{T_{\mathrm{h},0}}\,\mathbf{E}\cdot\int_{\mathbb{R}^3}\begin{pmatrix} -\frac{i}{2}\,P_1(v_\parallel)\,v_x - \frac{1}{2}\,P_2(v_\parallel)\,v_y \\ \frac{1}{2}\,P_2(v_\parallel)\,v_x - \frac{i}{2}\,P_1(v_\parallel)\,v_y \\ -i\,P_0(v_\parallel)\,v_\parallel \end{pmatrix}\mathbf{v}\,\frac{n_{\mathrm{h},0}}{(2\pi T_{\mathrm{h},0})^{3/2}}\,e^{-(v_x^2+v_y^2+v_\parallel^2)/(2T_{\mathrm{h},0})}\,\mathrm{d}^3\mathbf{v} \\ &=: \widehat{\boldsymbol{\sigma}}_{\mathrm{h}}(\omega,k_\parallel)\cdot\mathbf{E}\,, \end{aligned} \tag{135}$$

where $\widehat{\boldsymbol{\sigma}}_{\mathrm{h}} \in \mathbb{R}^{3\times 3}$ denotes the conductivity tensor, given by

$$\begin{aligned}
\widehat{\boldsymbol{\sigma}}_{\mathrm{h}}(\omega, k_\parallel) &= q_{\mathrm{h}} a_{\mathrm{h}} \frac{n_{\mathrm{h},0}}{T_{\mathrm{h},0}} \int_{\mathbb{R}^3} \begin{pmatrix} v_x \\ v_y \\ v_\parallel \end{pmatrix} \begin{pmatrix} -\frac{i}{2} P_1(v_\parallel)\, v_x - \frac{1}{2} P_2(v_\parallel)\, v_y \\ \frac{1}{2} P_2(v_\parallel)\, v_x - \frac{i}{2} P_1(v_\parallel)\, v_y \\ -i\, P_0(v_\parallel)\, v_\parallel \end{pmatrix}^{\top} \frac{e^{-(v_x^2+v_y^2+v_\parallel^2)/(2T_{\mathrm{h},0})}}{(2\pi T_{\mathrm{h},0})^{3/2}} \,\mathrm{d}^3\mathbf{v} \\
&= q_{\mathrm{h}} a_{\mathrm{h}}\, n_{\mathrm{h},0} \int_{-\infty}^{\infty} \begin{pmatrix} -\frac{i}{2} P_1(v_\parallel) & \frac{1}{2} P_2(v_\parallel) & 0 \\ -\frac{1}{2} P_2(v_\parallel) & -\frac{i}{2} P_1(v_\parallel) & 0 \\ 0 & 0 & -i\, P_0(v_\parallel)\, \frac{v_\parallel^2}{T_{\mathrm{h},0}} \end{pmatrix} \frac{e^{-v_\parallel^2/(2T_{\mathrm{h},0})}}{(2\pi T_{\mathrm{h},0})^{1/2}} \,\mathrm{d}v_\parallel \,. \qquad (136)
\end{aligned}$$

Here, the dependence on $(\omega, k_\parallel)$ is hidden in the terms P_0, P_1 and P_2 defined in (126)-(128). It is clear from (132) that only $\mathbf{j}_{\mathrm{h},\perp}$ is needed for the linear theory. Moreover, we shall now make use of the fact that in standard MHD the electric field is given by $\mathbf{E} = -\mathbf{U} \times \mathbf{B}_0$. Denoting the components of the conductivity tensor by $\widehat{\boldsymbol{\sigma}}_{\mathrm{h}} = (\sigma_{ij})_{1\leq i,j\leq 3}$, we obtain

$$j_{\mathrm{h},x} = -\sigma_{11}(\omega, k_\parallel)\, B_0\, U_y + \sigma_{12}(\omega, k_\parallel)\, B_0\, U_x \,, \qquad (137)$$

$$j_{\mathrm{h},y} = -\sigma_{21}(\omega, k_\parallel)\, B_0\, U_y + \sigma_{22}(\omega, k_\parallel)\, B_0\, U_x \,. \qquad (138)$$

Equation (132) is then written in matrix form as

$$\begin{pmatrix} \omega^2 - v_{\mathrm{A}}^2\, k_\parallel^2 + i\, \mu_0 v_{\mathrm{A}}^2\, \sigma_{22}(\omega, k_\parallel)\, \omega & -i\, \mu_0 v_{\mathrm{A}}^2 \left[\frac{q_{\mathrm{h}} n_{\mathrm{h},0}}{B_0} + \sigma_{21}(\omega, k_\parallel)\right] \omega & 0 \\ i\, \mu_0 v_{\mathrm{A}}^2 \left[\frac{q_{\mathrm{h}} n_{\mathrm{h},0}}{B_0} - \sigma_{12}(\omega, k_\parallel)\right] \omega & \omega^2 - v_{\mathrm{A}}^2\, k_\parallel^2 + i\, \mu_0 v_{\mathrm{A}}^2\, \sigma_{11}(\omega, k_\parallel)\, \omega & 0 \\ 0 & 0 & \omega^2 - v_{\mathrm{s}}^2\, k_\parallel^2 \end{pmatrix} \begin{pmatrix} U_x \\ U_y \\ U_\parallel \end{pmatrix} = 0 \,. \qquad (139)$$

Clearly, the sound waves are not affected by the wave-particle interaction. However, the shear-Alfvén waves change: the condition for $\mathbf{U}_\perp \neq 0$ reads

$$D(\omega, k_\parallel) := \left[\omega^2 - v_{\mathrm{A}}^2\, k_\parallel^2 + i\, \mu_0 v_{\mathrm{A}}^2\, \sigma_{11}(\omega, k_\parallel)\, \omega\right]^2 - \mu_0^2\, v_{\mathrm{A}}^4 \left[\frac{q_{\mathrm{h}} n_{\mathrm{h},0}}{B_0} + \sigma_{12}(\omega, k_\parallel)\right]^2 \omega^2 = 0 \,. \qquad (140)$$

Here, we used that the conductivity tensor satisfies $\sigma_{11} = \sigma_{22}$ and $\sigma_{21} = -\sigma_{12}$. The dispersion relation $D(\omega, k_\parallel) = 0$ can be solved numerically for a given value of $k_\parallel$.

3. Dispersion relation with radial profiles

For the hot particle density, from (129) we obtain

$$
\begin{aligned}
n_\mathrm{h} &= -\frac{a_\mathrm{h}}{T_{\mathrm{h},0}}\,\mathbf{E}\cdot\int_{\mathbb{R}^3}\begin{pmatrix}-\frac{i}{2}\,P_1(v_\parallel)\,v_x-\frac{1}{2}\,P_2(v_\parallel)\,v_y\\ \frac{1}{2}\,P_2(v_\parallel)\,v_x-\frac{i}{2}\,P_1(v_\parallel)\,v_y\\ -i\,P_0(v_\parallel)\,v_\parallel\end{pmatrix}\,\frac{n_{\mathrm{h},0}}{(2\pi T_{\mathrm{h},0})^{3/2}}\,e^{-(v_x^2+v_y^2+v_\parallel^2)/(2T_{\mathrm{h},0})}\,\mathrm{d}^3\mathbf{v}\\
&= i\,a_\mathrm{h}\frac{n_{\mathrm{h},0}}{T_{\mathrm{h},0}}\,E_\parallel\int_{-\infty}^{\infty}P_0(v_\parallel)\,v_\parallel\,\frac{1}{(2\pi T_{\mathrm{h},0})^{1/2}}\,e^{-v_\parallel^2/(2T_{\mathrm{h},0})}\,\mathrm{d}v_\parallel\\
&= i\,a_\mathrm{h}\frac{n_{\mathrm{h},0}}{T_{\mathrm{h},0}}\,E_\parallel\,G_0(\omega,k_\parallel)\,, \qquad (141)
\end{aligned}
$$

where we defined

$$
G_0(\omega,k_\parallel) := \frac{1}{(2\pi T_{\mathrm{h},0})^{1/2}}\int_{-\infty}^{\infty}\frac{v_\parallel}{k_\parallel v_\parallel-\omega}\,e^{-v_\parallel^2/(2T_{\mathrm{h},0})}\,\mathrm{d}v_\parallel\,. \qquad (142)
$$

C. Standard MHD with DK current coupling

1. Equations

Replacing in (111) the Vlasov equation with the drift-kinetic equation stated in (62) leads to the following current-coupling scheme:

$$\frac{\partial \tau}{\partial t} + \nabla \cdot (\tau \mathbf{U}) = 0\,, \tag{143a}$$

$$\frac{\partial \mathbf{U}}{\partial t} + \mathbf{U} \cdot \nabla \mathbf{U} + \frac{\nabla p}{\tau} = \frac{1}{\mu_0} \frac{\nabla \times \mathbf{B}}{\tau} \times \mathbf{B} + \frac{(q_\mathrm{h} n_\mathrm{h}^\mathrm{D}\, \mathbf{U} - \mathbf{j}_\mathrm{h}^\mathrm{D})}{\tau} \times \mathbf{B}\,, \tag{143b}$$

$$\frac{\partial p}{\partial t} + \mathbf{U} \cdot \nabla p + \gamma\, p\, \nabla \cdot \mathbf{U} = 0\,, \tag{143c}$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{U} \times \mathbf{B})\,, \tag{143d}$$

$$\frac{\partial f_\mathrm{h}^\mathrm{D}}{\partial t} + \mathbf{u}_\mathrm{g} \cdot \nabla f_\mathrm{h}^\mathrm{D} + E_\mathrm{g}\, \frac{\partial f_\mathrm{h}^\mathrm{D}}{\partial q_\parallel} = 0\,, \tag{143e}$$

where

$$\mathbf{u}_\mathrm{g} = q_\parallel \frac{\mathbf{B}^*}{B_\parallel^*} + \frac{\mathbf{E}^* \times \mathbf{b}}{B_\parallel^*}\,, \qquad E_\mathrm{g} = a_\mathrm{h} \frac{\mathbf{E}^* \cdot \mathbf{B}^*}{B_\parallel^*}\,, \tag{144}$$

and

$$\mathbf{B}^* = \mathbf{B} + \frac{q_\parallel}{a_\mathrm{h}}\, \nabla \times \mathbf{b}\,, \tag{145}$$

$$\mathbf{E}^* = -\mathbf{U} \times \mathbf{B} - \frac{\widehat{\mu}}{a_\mathrm{h}}\, \nabla B - \frac{q_\parallel}{a_\mathrm{h}} \frac{\partial \mathbf{b}}{\partial t}\,, \tag{146}$$

$$n_\mathrm{h}^\mathrm{D} = \int_{\widehat{\mu}} f_\mathrm{h}^\mathrm{D}\, B_\parallel^*\, \mathrm{d}q_\parallel\,, \tag{147}$$

$$\mathbf{j}_\mathrm{h}^\mathrm{D} = q_\mathrm{h} \int_{\widehat{\mu}} \mathbf{u}_\mathrm{g}\, f_\mathrm{h}^\mathrm{D}\, B_\parallel^*\, \mathrm{d}q_\parallel + \nabla \times \mathbf{M}\,, \tag{148}$$

$$\mathbf{M} = -m_\mathrm{h} \int_{\widehat{\mu}} \left[\widehat{\mu}\, \mathbf{b} - \frac{q_\parallel}{B}\, \mathbf{u}_{\mathrm{g},\perp} \right] f_\mathrm{h}^\mathrm{D}\, B_\parallel^*\, \mathrm{d}q_\parallel\,, \tag{149}$$

with $B_\parallel^* = \mathbf{B}^* \cdot \mathbf{b}$.

2. Dispersion relation for homogeneous plasma

The only thing that changes with respect to the Vlasov-MHD case is the current density, $\mathbf{j}_\mathrm{h} \to \mathbf{j}_\mathrm{h}^\mathrm{D}$. We thus need to compute the solution f_h^D of the linearized DK equation:

$$\frac{\partial f_\mathrm{h}^\mathrm{D}}{\partial t} + q_\parallel \, \mathbf{e}_z \cdot \nabla f_\mathrm{h}^\mathrm{D} = -E_\mathrm{g} \, \frac{\partial f_0^\mathrm{D}}{\partial q_\parallel} \,, \tag{150}$$

where $f_0^\mathrm{D} = f_{\mathrm{h},0}^\mathrm{D}(q_\parallel, \widehat{\mu})$ is a homogeneous DK equilibrium distribution and E_g stands for the linearized gyro-center acceleration:

$$E_\mathrm{g} = a_\mathrm{h} \, \mathbf{E}^* \cdot \mathbf{e}_z = -a_\mathrm{h} \, B_0 \, \mathbf{U} \times \mathbf{e}_z \cdot \mathbf{e}_z - \widehat{\mu} \, \mathbf{e}_z \cdot \nabla B - q_\parallel \, \mathbf{e}_z \cdot \frac{\partial \mathbf{b}}{\partial t} = 0 \,, \tag{151}$$

where we used that $B = \sqrt{2 B_0 \, \mathbf{e}_z \cdot \mathbf{B}}$ and $\mathbf{e}_z \cdot \mathbf{b} = \mathbf{e}_z \cdot \mathbf{B}/B_0$, which are both zero because of (107), ie. $\mathbf{B} = \mathbf{B}_\perp$. We thus obtain $f_\mathrm{h}^\mathrm{D} = 0$ but the current has also contributions from the linearized magnetization $\mathbf{M}$ and gyro-center velocity $\mathbf{u}_\mathrm{g}$, respectively. Let us assume a Gaussian equilibrium,

$$f_0^\mathrm{D} = \frac{n_{\mathrm{h},0}}{(2\pi T_{\mathrm{h},0})^{3/2}} \, e^{-(q_\parallel^2/2 + \widehat{\mu} B_0)/T_{\mathrm{h},0}} \,, \tag{152}$$

such that

$$\int_{\widehat{\mu}} \begin{pmatrix} 1 \\ q_\parallel \\ q_\parallel^2 \end{pmatrix} f_0^\mathrm{D} \, B_0 \, \mathrm{d}q_\parallel = \begin{pmatrix} n_{\mathrm{h},0} \\ 0 \\ n_{\mathrm{h},0} \, T_{\mathrm{h},0} \end{pmatrix} , \qquad \int_{\widehat{\mu}} \widehat{\mu} \, f_0^\mathrm{D} \, B_0 \, \mathrm{d}q_\parallel = \frac{n_{\mathrm{h},0} \, T_{\mathrm{h},0}}{B_0} \,. \tag{153}$$

With this we compute the linear expressions corresponding to (144)-(149):

$$\mathbf{B}^* = \mathbf{B} + \frac{q_\parallel}{a_\mathrm{h}} \, \nabla \times \mathbf{b} \,, \tag{154}$$

$$B_\parallel^* = \frac{q_\parallel}{a_\mathrm{h}} \, \nabla \times \mathbf{b} \cdot \mathbf{e}_z \,, \tag{155}$$

$$\mathbf{E}^* = -B_0 \, \mathbf{U} \times \mathbf{e}_z - \frac{q_\parallel}{a_\mathrm{h}} \frac{\partial \mathbf{b}}{\partial t} \,, \tag{156}$$

$$\mathbf{u}_\mathrm{g} = q_\parallel \left(\frac{\mathbf{B}^*}{B_0} - \frac{B_\parallel^*}{B_0} \, \mathbf{e}_z \right) + \frac{\mathbf{E}^* \times \mathbf{e}_z}{B_0} \,, \qquad E_\mathrm{g} = 0 \,, \tag{157}$$

$$n_\mathrm{h}^\mathrm{D} = 0 \,, \tag{158}$$

$$\mathbf{j}_\mathrm{h}^\mathrm{D} = q_\mathrm{h} \int_{\widehat{\mu}} \mathbf{u}_\mathrm{g} \, f_0^\mathrm{D} \, B_0 \, \mathrm{d}q_\parallel + \nabla \times \mathbf{M} \,, \tag{159}$$

$$\mathbf{M} = -m_\mathrm{h} \int_{\widehat{\mu}} \left[\widehat{\mu} \, \mathbf{b} - \frac{q_\parallel}{B_0} \, \mathbf{u}_{\mathrm{g},\perp} \right] f_0^\mathrm{D} \, B_0 \, \mathrm{d}q_\parallel \,. \tag{160}$$

In particular,

$$
\begin{aligned}
\int_{\widehat{\mu}} \mathbf{u}_{\rm g}\, f_0^{\rm D}\, B_0\, {\rm d}q_\parallel &= \int_{\widehat{\mu}} \left[q_\parallel \left(\frac{\mathbf{B}^*}{B_0} - \frac{B_\parallel^*}{B_0}\, \mathbf{e}_z \right) + \frac{\mathbf{E}^* \times \mathbf{e}_z}{B_0} \right] f_0^{\rm D}\, B_0\, {\rm d}q_\parallel \\
&= \int_{\widehat{\mu}} \left[q_\parallel\, \frac{\mathbf{B}}{B_0} + \frac{q_\parallel^2}{a_{\rm h} B_0}\, (\nabla \times \mathbf{b})_\perp + \mathbf{U}_\perp - \frac{q_\parallel}{a_{\rm h} B_0}\, \frac{\partial \mathbf{b}}{\partial t} \times \mathbf{e}_z \right] f_0^{\rm D}\, B_0\, {\rm d}q_\parallel \\
&= \frac{n_{\rm h,0}\, T_{\rm h,0}}{a_{\rm h} B_0}\, (\nabla \times \mathbf{b})_\perp + n_{\rm h,0}\, \mathbf{U}_\perp\,,
\end{aligned}
\tag{161}
$$

and

$$
\begin{aligned}
\mathbf{M} &= -m_{\rm h} \int_{\widehat{\mu}} \left[\widehat{\mu}\, \mathbf{b} - \frac{q_\parallel}{B_0} \left(q_\parallel\, \frac{\mathbf{B}_\perp^*}{B_0} + \frac{\mathbf{E}^* \times \mathbf{e}_z}{B_0} \right) \right] f_0^{\rm D}\, B_0\, {\rm d}q_\parallel \\
&= -m_{\rm h} \int_{\widehat{\mu}} \left[\widehat{\mu}\, \mathbf{b} - \frac{q_\parallel}{B_0} \left(q_\parallel \frac{\mathbf{B}_\perp}{B_0} + \frac{q_\parallel^2}{a_{\rm h} B_0}\, (\nabla \times \mathbf{b})_\perp + \mathbf{U}_\perp - \frac{q_\parallel}{a_{\rm h} B_0}\, \frac{\partial \mathbf{b}}{\partial t} \times \mathbf{e}_z \right) \right] f_0^{\rm D}\, B_0\, {\rm d}q_\parallel \\
&= -m_{\rm h} \left[\frac{n_{\rm h,0}\, T_{\rm h,0}}{B_0}\, \mathbf{b} - \frac{n_{\rm h,0}\, T_{\rm h,0}}{B_0^2}\, \mathbf{B}_\perp - \frac{n_{\rm h,0}\, T_{\rm h,0}}{a_{\rm h} B_0^2}\, \frac{\partial \mathbf{b}}{\partial t} \times \mathbf{e}_z \right].
\end{aligned}
\tag{162}
$$

This yields the current density

$$
\begin{aligned}
\mathbf{j}_{\rm h}^{\rm D} &= m_{\rm h} \frac{n_{\rm h,0}\, T_{\rm h,0}}{B_0}\, (\nabla \times \mathbf{b})_\perp + q_{\rm h} n_{\rm h,0}\, \mathbf{U}_\perp \\
&\quad - m_{\rm h} \left[\frac{n_{\rm h,0}\, T_{\rm h,0}}{B_0}\, \nabla \times \mathbf{b} - \frac{n_{\rm h,0}\, T_{\rm h,0}}{B_0^2}\, \nabla \times \mathbf{B}_\perp - \frac{n_{\rm h,0}\, T_{\rm h,0}}{a_{\rm h} B_0^2}\, \nabla \times \frac{\partial (\mathbf{b} \times \mathbf{e}_z)}{\partial t} \right].
\end{aligned}
\tag{163}
$$

3. Dispersion relation with radial profiles

D. Extended MHD

1. Equations

The "extended MHD system" features additional terms in the electric field:

$$\mathbf{E} = -\mathbf{U} \times \mathbf{B} + \frac{1}{\tau}\nabla\left(a_\mathrm{e}\, p_\mathrm{i} - a_\mathrm{i}\, p_\mathrm{e}\right) + \left(a_\mathrm{i} - a_\mathrm{e}\right)\frac{1}{\mu_0}\frac{\nabla \times \mathbf{B}}{\tau} \times \mathbf{B}\,. \tag{164}$$

The additional terms are related to the diamagnetic drift and to the polarization drift derived from generalized Ohm's law; their inclusion is consistent with the ordering assumptions from section II B. The addition of these terms makes it necessary to consider two pressure equations, one for ions and one for electrons. The current density proportional to $\nabla \times \mathbf{B}$ has to appear in the pressure fluxes to achieve energy conservation. The extended MHD system reads

$$\frac{\partial \tau}{\partial t} + \nabla \cdot (\tau \mathbf{U}) = 0\,, \tag{165a}$$

$$\frac{\partial \mathbf{U}}{\partial t} + \mathbf{U} \cdot \nabla \mathbf{U} + \frac{\nabla(p_\mathrm{i} + p_\mathrm{e})}{\tau} = \frac{1}{\mu_0}\frac{(\nabla \times \mathbf{B})}{\tau} \times \mathbf{B}\,, \tag{165b}$$

$$\frac{\partial p_\mathrm{i}}{\partial t} + \left[\mathbf{U} + \frac{a_\mathrm{e}}{\mu_0}\frac{\nabla \times \mathbf{B}}{\tau}\right] \cdot \nabla p_\mathrm{i} + \gamma p_\mathrm{i} \nabla \cdot \left[\mathbf{U} + \frac{a_\mathrm{e}}{\mu_0}\frac{\nabla \times \mathbf{B}}{\tau}\right] = 0\,, \tag{165c}$$

$$\frac{\partial p_\mathrm{e}}{\partial t} + \left[\mathbf{U} - \frac{a_\mathrm{i}}{\mu_0}\frac{\nabla \times \mathbf{B}}{\tau}\right] \cdot \nabla p_\mathrm{e} + \gamma p_\mathrm{e} \nabla \cdot \left[\mathbf{U} - \frac{a_\mathrm{i}}{\mu_0}\frac{\nabla \times \mathbf{B}}{\tau}\right] = 0\,, \tag{165d}$$

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E}\,, \tag{165e}$$

where $\mathbf{E}$ is given by (164).

2. Dispersion relation for homogeneous plasma

3. Dispersion relation with radial profiles

E. Extended MHD: DK current coupling with $n_\mathrm{h} \to 0$

1. Equations

The extended MHD system (165) can be coupled in a simple way with the DK model for the energetic particles, namely under the assumption that $n_\mathrm{h} \to 0$ and only $\mathbf{j}_\mathrm{h}$ is finite. In this case the electric field is given by

$$\mathbf{E} = -\mathbf{U}\times\mathbf{B} + \frac{1}{\tau}\nabla\left(a_\mathrm{e}\,p_\mathrm{i} - a_\mathrm{i}\,p_\mathrm{e}\right) + (a_\mathrm{i} - a_\mathrm{e})\,\frac{1}{\mu_0}\,\frac{(\nabla\times\mathbf{B} - \mu_0\,\mathbf{j}_\mathrm{h}^\mathrm{D})}{\tau}\times\mathbf{B}\,, \tag{166}$$

and the energy-conserving current coupling scheme reads

$$\frac{\partial\tau}{\partial t} + \nabla\cdot(\tau\mathbf{U}) = 0\,, \tag{167a}$$

$$\frac{\partial\mathbf{U}}{\partial t} + \mathbf{U}\cdot\nabla\mathbf{U} + \frac{\nabla(p_\mathrm{i}+p_\mathrm{e})}{\tau} = \frac{1}{\mu_0}\,\frac{(\nabla\times\mathbf{B} - \mu_0\,\mathbf{j}_\mathrm{h}^\mathrm{D})}{\tau}\times\mathbf{B}\,, \tag{167b}$$

$$\frac{\partial p_\mathrm{i}}{\partial t} + \left[\mathbf{U} + \frac{a_\mathrm{e}}{\mu_0}\frac{(\nabla\times\mathbf{B} - \mu_0\,\mathbf{j}_\mathrm{h}^\mathrm{D})}{\tau}\right]\cdot\nabla p_\mathrm{i} + \gamma p_\mathrm{i}\nabla\cdot\left[\mathbf{U} + \frac{a_\mathrm{e}}{\mu_0}\frac{(\nabla\times\mathbf{B} - \mu_0\,\mathbf{j}_\mathrm{h}^\mathrm{D})}{\tau}\right] = 0\,, \tag{167c}$$

$$\frac{\partial p_\mathrm{e}}{\partial t} + \left[\mathbf{U} - \frac{a_\mathrm{i}}{\mu_0}\frac{(\nabla\times\mathbf{B} - \mu_0\,\mathbf{j}_\mathrm{h}^\mathrm{D})}{\tau}\right]\cdot\nabla p_\mathrm{e} + \gamma p_\mathrm{e}\nabla\cdot\left[\mathbf{U} - \frac{a_\mathrm{i}}{\mu_0}\frac{(\nabla\times\mathbf{B} - \mu_0\,\mathbf{j}_\mathrm{h}^\mathrm{D})}{\tau}\right] = 0\,, \tag{167d}$$

$$\frac{\partial\mathbf{B}}{\partial t} = -\nabla\times\mathbf{E}\,, \tag{167e}$$

$$\frac{\partial f_\mathrm{h}^\mathrm{D}}{\partial t} + \mathbf{u}_\mathrm{g}\cdot\frac{\partial f_\mathrm{h}^\mathrm{D}}{\partial\mathbf{r}} + E_\mathrm{g}\,\frac{\partial f_\mathrm{h}^\mathrm{D}}{\partial q_\parallel} = 0\,, \tag{167f}$$

where the relevant terms for the DK equation are stated in (**??**)-(**??**).

2. Dispersion relation

The unperturbed characteristics $\mathbf{X} = \mathbf{X}(s;t,\mathbf{x},q_\parallel)$ and $Q_\parallel = Q_\parallel(s;t,\mathbf{x},q_\parallel)$ satisfy

$$\begin{aligned} \frac{\mathrm{d}\mathbf{X}}{\mathrm{d}s} &= Q_\parallel\,\mathbf{e}_z\,, \qquad & \mathbf{X}(t;t,\mathbf{x},q_\parallel) &= \mathbf{x}\,,\\ \frac{\mathrm{d}Q_\parallel}{\mathrm{d}s} &= 0\,, \qquad & Q_\parallel(t;t,\mathbf{x},q_\parallel) &= q_\parallel\,, \end{aligned} \tag{168}$$

and are particularly simple in this case:

$$Q_\parallel = q_\parallel\,, \qquad \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \mathbf{x} + \begin{pmatrix} x \\ y \\ (s-t)\,q_\parallel \end{pmatrix}. \tag{169}$$

Equation (150) then yields

$$\frac{\mathrm{d}}{\mathrm{d}s} f_\mathrm{h}^\mathrm{D}(\mathbf{X}, Q_\parallel, \widehat{\mu}, s) = -E_\mathrm{g}(\mathbf{X}, Q_\parallel, \widehat{\mu}, s)\,\frac{\partial f_0^\mathrm{D}}{\partial q_\parallel}(Q_\parallel, \widehat{\mu})\,, \tag{170}$$

which can be integrated from $s=-\infty$ to $s=t$ to obtain

$$f_{\mathrm{h}}^{\mathrm{D}}(\mathbf{x}, q_\parallel, \widehat{\mu}, t) = -\int_{-\infty}^{t} E_{\mathrm{g}}(\mathbf{X}, Q_\parallel, \widehat{\mu}, s)\, \frac{\partial f_0^{\mathrm{D}}}{\partial q_\parallel}(Q_\parallel, \widehat{\mu})\, \mathrm{d}s\,, \tag{171}$$

where $f_{\mathrm{h}}^{\mathrm{D}}(\mathbf{X}, Q_\parallel, \widehat{\mu}, -\infty) = 0$ was assumed. The next step is to insert into E_{g} the plane wave solution $\mathbf{B}(\mathbf{X}, s) = \widehat{\mathbf{B}}\, e^{i(k_\parallel Z - \omega s)}$, which leads to

F. Extended MHD: DK current coupling

1. Equations

The most sophisticated DK-MHD current coupling scheme we derived is the system (95). We now compute the corresponding dispersion relations.

2. Dispersion relation

IV. DISCRETIZATION

A. Finite element exterior calculus

B. Spline basis functions

C. Standard MHD

1. Two-dimensional case

2. Three-dimensional case

D. Standard MHD with DK current coupling

Appendix A: Derivation of the drift-kinetic equation

We give a step-by-step derivation of the drift-kinetic model used in this work.

1. Characteristics and variational principle

Our aim is to approximate the normalized Vlasov equation (19),

$$\varepsilon \frac{\partial f_{\mathrm{h}}}{\partial t} + \mathbf{v} \cdot \nabla f_{\mathrm{h}} + C_{\mathrm{h}} \left(\mathbf{E} + \frac{\mathbf{v} \times \mathbf{B}}{\varepsilon} \right) \cdot \frac{\partial f_{\mathrm{h}}}{\partial \mathbf{v}} = 0 \,, \tag{A1}$$

by a gyro-center (GY) model. The gyro-period in this equation is of order $O(\varepsilon^2)$ such that we expect accurate approximations from rather simple GY-models, where magnetic curvature effects are rather small. We follow the methodology of variational averaging (VA) reviewed in[3], which is briefly outlined in what follows. The basic principle is to approximate the characteristic equations of (A1), which are given by

$$\begin{aligned} \varepsilon \frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} &= \mathbf{v}(t) \,, \\ \varepsilon \frac{\mathrm{d}\mathbf{v}(t)}{\mathrm{d}t} &= C_{\mathrm{h}} \left[\mathbf{E}\big(\mathbf{x}(t), t\big) + \frac{\mathbf{v}(t) \times \mathbf{B}\big(\mathbf{x}(t), t\big)}{\varepsilon} \right] . \end{aligned} \tag{A2}$$

A solution (or characteristic) is a curve in the phase space $\Omega \subset \mathbb{R}^6$, which is assumed a six-dimensional differentiable manifold. The kinetic solution f_{h} is constant along these characteristics. This is known as the transport theorem: functions that are constant along solutions of single-particle trajectories are called kinetic solutions; the satisfy a kinetic PDE (without collision operator). The equations (A2) can be obtained from the variational principle $\delta \mathcal{L}[\mathbf{x}(t), \mathbf{v}(t)] = 0$, where the functional $\mathcal{L}$ is defined by

$$\mathcal{L}[\mathbf{x}(t), \mathbf{v}(t)] := \int_{t_0}^{t_1} L \left(\mathbf{x}(t), \mathbf{v}(t), \frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t}, \frac{\mathrm{d}\mathbf{v}(t)}{\mathrm{d}t} \right) \mathrm{d}t \,, \tag{A3}$$

and the Lagrangian L is in principle a function of twelve coordinates (and of time, which we view as a parameter in the following),

$$L = L(\mathbf{x}, \mathbf{v}, \dot{\mathbf{x}}, \dot{\mathbf{v}}) := \big[\varepsilon\, \mathbf{v} + C_{\mathrm{h}}\, \mathbf{A}(\mathbf{x}, t) \big] \cdot \dot{\mathbf{x}} - \frac{|\mathbf{v}|^2}{2} - C_{\mathrm{h}}\, \phi(\mathbf{x}, t) \,, \tag{A4}$$

where $\dot{\mathbf{x}} \in \mathbb{R}^3$ and $\dot{\mathbf{v}} \in \mathbb{R}^3$ are coordinates of the tangent space at $(\mathbf{x}, \mathbf{v}) \in \Omega$, and the Lagrangian (A4) does not depend on $\dot{\mathbf{v}}$ in the given formulation. We urge the reader to pay attention to the difference between $\mathbf{x}(t)$, $\mathbf{v}(t)$, $\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t}$, $\frac{\mathrm{d}\mathbf{v}(t)}{\mathrm{d}t}$ on the one hand and $\mathbf{x}$, $\mathbf{v}$, $\dot{\mathbf{x}}$, $\dot{\mathbf{v}}$ on the other hand: while the former denote trajectories and their derivatives determined by the actual dynamics (A2), such that the derivatives are of order $1/\varepsilon$ and $1/\varepsilon^2$, respectively, the latter are coordinates on the tangent bundle $T\Omega$ and are of order one. Moreover, in (A4) we introduced the electromagnetic potentials $\mathbf{A}$ and ϕ which satisfy

$$\mathbf{B} = \nabla \times \mathbf{A}\,, \qquad \mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}\,. \tag{A5}$$

It is straightforward to verify that the Euler-Lagrange equations following from the variational principle with the Lagrangian L are indeed the characteristic equations (A2).

According to the transport theorem, we can formulate a different (but equivalent) kinetic equation if we merely change the coordinates on the phase space and write down the characteristic equations (A2) in the new coordinates. VA is based on applying such a change of coordinates in the Lagrangian L rather than the equations of motion. Denoting the coordinate change by $\widehat{\tau}^\varepsilon_{\text{gy}} : \widehat{\Omega} \to \Omega$, where $\widehat{\mathbf{q}} \in \widehat{\Omega} \subset \mathbb{R}^6$ are called the GY-coordinates and we denote the standard coordinates by $(\mathbf{x}, \mathbf{v}) = \mathbf{q} \in \Omega$, such that $(\mathbf{x}, \mathbf{v}) = \mathbf{q} = \widehat{\tau}^\varepsilon_{\text{gy}}(\widehat{\mathbf{q}})$, leads to an expression of the Lagrangian (A4) in the new coordinates on the tangent bundle, $L(\mathbf{q}, \dot{\mathbf{q}}) = \widehat{L}^\varepsilon(\widehat{\mathbf{q}}, \dot{\widehat{\mathbf{q}}})$. The new Lagrangian $\widehat{L}^\varepsilon$ follows from the "tangent map" between the tangent bundles of the coordinate spaces[3],

$$T\widehat{\tau}^\varepsilon_{\text{gy}} : TU \to T\Omega\,, \qquad (\mathbf{q}, \dot{\mathbf{q}}) = T\widehat{\tau}^\varepsilon_{\text{gy}}(\widehat{\mathbf{q}}, \dot{\widehat{\mathbf{q}}}) = \left(\widehat{\tau}^\varepsilon_{\text{gy}}(\widehat{\mathbf{q}})\,,\, D\widehat{\tau}^\varepsilon_{\text{gy}}(\widehat{\mathbf{q}}) \cdot \dot{\widehat{\mathbf{q}}} + \frac{\partial \widehat{\tau}^\varepsilon_{\text{gy}}(\widehat{\mathbf{q}})}{\partial t}\right)\,, \tag{A6}$$

where $D\widehat{\tau}^\varepsilon_{\text{gy}}$ denotes the Jacobian. Then, simply define

$$\widehat{L}^\varepsilon(\widehat{\mathbf{q}}, \dot{\widehat{\mathbf{q}}}) := L\left(T\widehat{\tau}^\varepsilon_{\text{gy}}(\widehat{\mathbf{q}}, \dot{\widehat{\mathbf{q}}})\right)\,. \tag{A7}$$

Choosing and appropriate GY-transform $\widehat{\tau}^\varepsilon_{\text{gy}}$, and using then the equivalence of Lagrangians under the addition of differentials ("total time derivatives"), one can show that the new $\widehat{L}^\varepsilon$ is equivalent to a simpler Lagrangian in the sense that one of the coordinates, called the gyro-angle α, is quasi-cyclic,

$$\frac{\partial \widehat{L}^\varepsilon}{\partial \alpha} = O(\varepsilon^{N+1})\,, \qquad N \geq 1\,. \tag{A8}$$

This means that $\widehat{\mu} := \frac{\partial L'}{\partial \dot{\alpha}}$ is an adiabatic invariant, known as the generalized magnetic moment. Dimensional reduction is then achieved a) by adopting $\widehat{\mu}$ as one of the coordinates and b) by considering only the gyro-averaged part of the kinetic equation corresponding to the new characteristics. This leads to the drift-kinetic equation on a four-dimensional phase space. We refer to[3] for more details regarding the methodology.

In the approach used in this work, the GY-transformation will be a composition of three mappings, $\widehat{\tau}^\varepsilon_{\text{gy}} = \tau' \circ \tau^\varepsilon \circ \widehat{\tau}$, where

1. τ' is a preliminary transformation to new velocity coordinates, $(v_\parallel,\ v_\perp,\theta) \mapsto \mathbf{v}$, in order to single out the fast variable θ of the system,

2. τ^ε is the core of the GY-transformation in the form of a finite power series in ε, containing the so-called generating functions as coefficients,

3. $\widehat{\tau}$ makes then generalized magnetic moment $\widehat{\mu}$ one of the coordinates.

The preliminary transformation depends on the choice of a local orthonormal basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{b})$, where $\mathbf{b} = \mathbf{B}/|\mathbf{B}| = \mathbf{B}/B$ is the unit vector pointing in the direction of the magnetic field at a given point and the other two unit vectors satisfy $\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{b}$, thus they span the plane perpendicular to $\mathbf{b}$. The particular directions of $\mathbf{e}_1$ is arbitrary and is called the "gyro-gauge". The drift-kinetic model we derive will be independent of the gyro-gauge. The preliminary transformation $\tau' : \Omega' \to \Omega$, where $\mathbf{q}' = (\mathbf{x}, \mathbf{v}') = (\mathbf{x}, v_\parallel, v_\perp, \theta) \in \Omega' \subset \mathbb{R}^6$ are the new coordinates, is defined by

$$\mathbf{x} = \mathbf{x}\,, \qquad \mathbf{v} = v_\parallel \mathbf{b} + v_\perp \mathbf{c}\,, \tag{A9}$$

where $\mathbf{c} = \mathbf{e}_1 \cos(\theta) - \mathbf{e}_2 \sin(\theta)$. Hence,

$$v_\parallel = \mathbf{v}\cdot\mathbf{b}\,, \qquad v_\perp = |\mathbf{v} - (\mathbf{v}\cdot\mathbf{b})\mathbf{b}|\,, \qquad \theta = -\arctan\left(\frac{\mathbf{v}\cdot\mathbf{e}_2}{\mathbf{v}\cdot\mathbf{e}_1}\right)\,, \tag{A10a}$$

and $\det(D\tau') = -v_\perp$. We will also need the vector $\mathbf{a} = \mathbf{e}_1 \sin(\theta) + \mathbf{e}_2 \cos(\theta)$ such that $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ form an orthonormal basis and $\partial_\theta \mathbf{c} = -\mathbf{a}$. In the new coordinates the Lagrangian (A4) reads

$$L'(\mathbf{q}', \dot{\mathbf{q}}') = \big[\varepsilon\, v_\parallel \mathbf{b}(\mathbf{x}, t) + \varepsilon\, v_\perp \mathbf{c}(\mathbf{x}, \theta, t) + C_\mathrm{h}\, \mathbf{A}(\mathbf{x}, t)\big]\cdot\dot{\mathbf{x}} - \frac{v_\parallel^2}{2} - \frac{v_\perp^2}{2} - C_\mathrm{h}\, \phi(\mathbf{x}, t)\,. \tag{A11}$$

2. Generating functions

The core of the GY-transformation is the mapping $\tau^\varepsilon : \Omega_\mathrm{gy} \to \Omega'$, where the new coordinates are denoted by $\mathbf{q}_\mathrm{gy} \in \Omega_\mathrm{gy} \subset \mathbb{R}^6$. The mapping is defined as a finite power series in ε,

$$\mathbf{q}' = \tau^\varepsilon(\mathbf{q}_\mathrm{gy}) = \mathbf{q}_\mathrm{gy} + \sum_{n=1}^{N+1} \varepsilon^n\, \mathbf{G}_n(\mathbf{q}_\mathrm{gy}, t)\,, \tag{A12}$$

where the $\mathbf{G}_n$ are called the generating functions of the transformation and N denotes the order of the transformation. Component-wise,

$$\mathbf{q}' = \begin{pmatrix} \mathbf{x} \\ v_\parallel \\ v_\perp \\ \theta \end{pmatrix}, \qquad \mathbf{q}_{\mathrm{gy}} = \begin{pmatrix} \mathbf{r} \\ q_\parallel \\ q_\perp \\ \alpha \end{pmatrix}, \qquad \mathbf{G}_n = \begin{pmatrix} \boldsymbol{\rho}_n \\ G_n^\parallel \\ G_n^\perp \\ G_n^\theta \end{pmatrix}, \tag{A13}$$

where $\mathbf{x}, \mathbf{r}, \boldsymbol{\rho}_n \in \mathbb{R}^3$ denote the particle position, the GY-position and the position generators, respectively. Moreover, the tangent map yields

$$\dot{\mathbf{x}} = \dot{\mathbf{r}} + \sum_{n=1}^{N+1} \varepsilon^n \, \dot{\boldsymbol{\rho}}_n \,, \qquad \dot{\boldsymbol{\rho}}_n = \left(\frac{\partial \boldsymbol{\rho}_n}{\partial \mathbf{q}_{\mathrm{gy}}} \cdot \dot{\mathbf{q}}_{\mathrm{gy}} + \frac{\partial \boldsymbol{\rho}_n}{\partial t} \right). \tag{A14}$$

Let us now insert the transformation (A12)-(A14) into the Lagrangian L' given in (A11) and expand in powers of ε to obtain a new Lagrangian

$$L^\varepsilon(\mathbf{q}_{\mathrm{gy}}, \dot{\mathbf{q}}_{\mathrm{gy}}) = L_0 + \varepsilon \, L_1 + \varepsilon^2 \, L_2 + \dots \,. \tag{A15}$$

In the present theory we go up to second order, which means we compute L_2 but not L_3, and we assume the electromagnetic fields to be sufficiently regular for Taylor expansion up to the desired order, for instance

$$\mathbf{A}(\mathbf{x}, t) = \mathbf{A}(\mathbf{r}, t) + \varepsilon \, \boldsymbol{\rho}_1 \cdot \nabla \mathbf{A}(\mathbf{r}, t) + \varepsilon^2 \left[\frac{1}{2} (\boldsymbol{\rho}_1 \cdot \nabla)^2 \mathbf{A}(\mathbf{r}, t) + \boldsymbol{\rho}_2 \cdot \nabla \mathbf{A}(\mathbf{r}, t) \right] + \dots \,, \tag{A16}$$

From (A11) we compute

$$L_0 = C_{\mathrm{h}} \, \mathbf{A}(\mathbf{r}, t) \cdot \dot{\mathbf{r}} - H_0 \,, \tag{A17a}$$

$$L_1 = \left[q_\parallel \mathbf{b}(\mathbf{r}, t) + q_\perp \mathbf{c}(\mathbf{r}, \alpha, t) + C_{\mathrm{h}} \, \boldsymbol{\rho}_1 \cdot \nabla \mathbf{A}(\mathbf{r}, t) \right] \cdot \dot{\mathbf{r}} + C_{\mathrm{h}} \, \mathbf{A}(\mathbf{r}, t) \cdot \dot{\boldsymbol{\rho}}_1 - H_1 \,, \tag{A17b}$$

$$\begin{aligned} L_2 = & \left[G_1^\parallel \, \mathbf{b}(\mathbf{r}, t) + G_1^\perp \, \mathbf{c}(\mathbf{r}, \alpha, t) + q_\parallel \boldsymbol{\rho}_1 \cdot \nabla \mathbf{b}(\mathbf{r}, t) + q_\perp \boldsymbol{\rho}_1 \cdot \nabla \mathbf{c}(\mathbf{r}, \alpha, t) - q_\perp G_1^\theta \, \mathbf{a}(\mathbf{r}, t) \right] \cdot \dot{\mathbf{r}} \\ & + C_{\mathrm{h}} \left[\frac{1}{2} (\boldsymbol{\rho}_1 \cdot \nabla)^2 \mathbf{A}(\mathbf{r}, t) + \boldsymbol{\rho}_2 \cdot \nabla \mathbf{A}(\mathbf{r}, t) \right] \cdot \dot{\mathbf{r}} \qquad \text{(A17c)} \\ & + \left[q_\parallel \mathbf{b}(\mathbf{r}, t) + q_\perp \mathbf{c}(\mathbf{r}, \alpha, t) + C_{\mathrm{h}} \, \boldsymbol{\rho}_1 \cdot \nabla \mathbf{A}(\mathbf{r}, t) \right] \cdot \dot{\boldsymbol{\rho}}_1 + C_{\mathrm{h}} \, \mathbf{A}(\mathbf{r}, t) \cdot \dot{\boldsymbol{\rho}}_2 - H_2 \,, \end{aligned}$$

where the Hamiltonians are given by

$$H_0 = \frac{q_\parallel^2}{2} + \frac{q_\perp^2}{2} + C_\mathrm{h}\,\phi(\mathbf{r},t)\,, \tag{A17d}$$

$$H_1 = q_\parallel\, G_1^\parallel + q_\perp G_1^\perp + C_\mathrm{h}\,\boldsymbol{\rho}_1\cdot\nabla\phi(\mathbf{r},t)\,, \tag{A17e}$$

$$H_2 = q_\parallel\, G_2^\parallel + q_\perp G_2^\perp + \frac{(G_1^\parallel)^2}{2} + \frac{(G_1^\perp)^2}{2} + C_\mathrm{h}\,\left[\frac{1}{2}(\boldsymbol{\rho}_1\cdot\nabla)^2\phi(\mathbf{r},t) + \boldsymbol{\rho}_2\cdot\nabla\phi(\mathbf{r},t)\right]. \tag{A17f}$$

It has been shown in[3] that it possible to make α the quasi-cyclic variable satisfying (A8) by choosing the generating functions such that α is eliminated from the Lagrangian L^ε order by order, up to arbitrary order N. Indeed, this can be achieved from the expansion (A17) in an algebraic manner, without having to solve any differential equations for the generators. The gyro-angle α is already cyclic in L_0. In L_1 we have a gyro-angle dependence in $\mathbf{c}(\mathbf{r},\alpha,t)$ and possibly in the Hamiltonian H_1. Using the product rule on the tangent space (ie. elements of the tangent space behave like "derivatives"), as well as the equivalence "$\sim$" of Lagrangians under the addition of total derivatives, we obtain

$$\begin{aligned}\mathbf{A}(\mathbf{r},t)\cdot\dot{\boldsymbol{\rho}}_1 &= \frac{\mathrm{d}}{\mathrm{d}t}\big[\mathbf{A}(\mathbf{r},t)\cdot\boldsymbol{\rho}_1\big] - \dot{\mathbf{A}}(\mathbf{r},t)\cdot\boldsymbol{\rho}_1\\ &\sim -\dot{\mathbf{r}}\cdot\nabla\mathbf{A}(\mathbf{r},t)\cdot\boldsymbol{\rho}_1 - \frac{\partial\mathbf{A}(\mathbf{r},t)}{\partial t}\cdot\boldsymbol{\rho}_1\,.\end{aligned} \tag{A18}$$

Here, the term with $\partial_t\mathbf{A}$ will be added to the Hamiltonian H_1. Then, using in L_1 of (A17b) the identity

$$\nabla\mathbf{A}(\mathbf{r},t)\cdot\boldsymbol{\rho}_1 - \boldsymbol{\rho}_1\cdot\nabla\mathbf{A}(\mathbf{r},t) = \boldsymbol{\rho}_1\times[\nabla\times\mathbf{A}(\mathbf{r},t)] = \boldsymbol{\rho}_1\times\mathbf{B}(\mathbf{r},t)\,, \tag{A19}$$

we can choose $\boldsymbol{\rho}_1$ such that $q_\perp\mathbf{c}(\mathbf{r},\alpha,t) - C_\mathrm{h}\,\boldsymbol{\rho}_1\times\mathbf{B}(\mathbf{r},t) = 0$, that is

$$\boldsymbol{\rho}_1 = \frac{1}{C_\mathrm{h}}\frac{q_\perp}{B(\mathbf{r},t)}\,\mathbf{a}(\mathbf{r},\alpha,t)\,. \tag{A20}$$

In order to remove the α-dependence from H_1 we choose the generators such that $H_1 = 0$, that is

$$G_1^\perp = -\frac{q_\parallel}{q_\perp}\,G_1^\parallel + \frac{C_\mathrm{h}}{q_\perp}\,\boldsymbol{\rho}_1\cdot\mathbf{E}(\mathbf{r},t)\,. \tag{A21}$$

This leads to a first order Lagrangian of the form

$$L_1 \sim q_\parallel\mathbf{b}(\mathbf{r},t)\cdot\dot{\mathbf{r}}\,. \tag{A22}$$

At second order, in L_2 we can use (A18)-(A19) with $\boldsymbol{\rho}_2$ in place of $\boldsymbol{\rho}_1$. Moreover, omitting from now on the arguments of fields and basis vectors,

$$\begin{aligned}\boldsymbol{\rho}_1 \cdot \nabla\mathbf{A} \cdot \dot{\boldsymbol{\rho}}_1 &= \frac{1}{2}\,\boldsymbol{\rho}_1 \cdot \nabla\mathbf{A} \cdot \dot{\boldsymbol{\rho}}_1 + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{\rho}_1 \cdot \nabla\mathbf{A} \cdot \boldsymbol{\rho}_1\right) \\ &\quad - \frac{1}{2}\,\dot{\boldsymbol{\rho}}_1 \cdot \nabla\mathbf{A} \cdot \boldsymbol{\rho}_1 - \frac{1}{2}\,\boldsymbol{\rho}_1 \cdot \nabla\left[\left(\dot{\mathbf{r}} \cdot \nabla + \frac{\partial}{\partial t}\right)\mathbf{A}\right] \cdot \boldsymbol{\rho}_1 \\ &\sim -\frac{1}{2}\,\dot{\boldsymbol{\rho}}_1 \cdot \left[\boldsymbol{\rho}_1 \times \mathbf{B}\right] - \frac{1}{2}\,\boldsymbol{\rho}_1 \cdot \nabla\left[\left(\dot{\mathbf{r}} \cdot \nabla + \frac{\partial}{\partial t}\right)\mathbf{A}\right] \cdot \boldsymbol{\rho}_1\end{aligned} \tag{A23}$$

Then, in (A17c) we have the equivalence

$$\begin{aligned}\frac{1}{2}(\boldsymbol{\rho}_1 \cdot \nabla)^2\mathbf{A} \cdot \dot{\mathbf{r}} + \boldsymbol{\rho}_1 \cdot \nabla\mathbf{A} \cdot \dot{\boldsymbol{\rho}}_1 &\sim -\frac{1}{2}\,\dot{\boldsymbol{\rho}}_1 \cdot \left[\boldsymbol{\rho}_1 \times \mathbf{B}\right] - \frac{1}{2}\,\boldsymbol{\rho}_1 \cdot \nabla\frac{\partial\mathbf{A}}{\partial t} \cdot \boldsymbol{\rho}_1 \\ &\quad - \frac{1}{2}\,\dot{\mathbf{r}} \cdot \left[\boldsymbol{\rho}_1 \times (\boldsymbol{\rho}_1 \cdot \nabla\mathbf{B})\right],\end{aligned} \tag{A24}$$

where from the result (A20) for $\boldsymbol{\rho}_1$ we compute

$$-\frac{1}{2}\,\dot{\boldsymbol{\rho}}_1 \cdot \left[\boldsymbol{\rho}_1 \times \mathbf{B}\right] = -\frac{1}{C_{\mathrm{h}}^2}\frac{q_\perp^2}{2B}\left(\dot{\mathbf{r}} \cdot \nabla\mathbf{a} \cdot \mathbf{c} + \frac{\partial\mathbf{a}}{\partial t} \cdot \mathbf{c} + \dot{\alpha}\right). \tag{A25}$$

Furthermore, in the last line of (A17c) we have the equivalence

$$\begin{aligned}(q_\parallel\mathbf{b} + q_\perp\mathbf{c}) \cdot \dot{\boldsymbol{\rho}}_1 &= \frac{\mathrm{d}}{\mathrm{d}t}\left[(q_\parallel\mathbf{b} + q_\perp\mathbf{c}) \cdot \boldsymbol{\rho}_1\right] - q_\parallel\dot{\mathbf{b}} \cdot \boldsymbol{\rho}_1 - q_\perp\dot{\mathbf{c}} \cdot \boldsymbol{\rho}_1 \\ &\sim -\frac{1}{C_{\mathrm{h}}}\frac{q_\parallel q_\perp}{B}\left[\dot{\mathbf{r}} \cdot \nabla\mathbf{b} + \frac{\partial\mathbf{b}}{\partial t}\right] \cdot \mathbf{a} \\ &\quad - \frac{1}{C_{\mathrm{h}}}\frac{q_\perp^2}{B}\left[\dot{\mathbf{r}} \cdot \nabla\mathbf{c} + \frac{\partial\mathbf{c}}{\partial t} - \dot{\alpha}\,\mathbf{a}\right] \cdot \mathbf{a}.\end{aligned} \tag{A26}$$

Hence, as an intermediate result, the Lagrangian L_2 is equivalent to

$$\begin{aligned}L_2 \sim{}& \left[G_1^\parallel\,\mathbf{b} + G_1^\perp\,\mathbf{c} - q_\perp G_1^\theta\,\mathbf{a} - \frac{1}{C_{\mathrm{h}}}\frac{q_\parallel q_\perp}{B}\,\mathbf{a} \times (\nabla \times \mathbf{b}) - \frac{1}{C_{\mathrm{h}}}\frac{q_\perp^2}{B}\,\mathbf{a} \times (\nabla \times \mathbf{c})\right] \cdot \dot{\mathbf{r}} \\ &+ C_{\mathrm{h}}\left[-\boldsymbol{\rho}_2 \times \mathbf{B} - \frac{1}{2}\,\boldsymbol{\rho}_1 \times (\boldsymbol{\rho}_1 \cdot \nabla\mathbf{B}) - \frac{1}{C_{\mathrm{h}}^2}\frac{q_\perp^2}{2B}\,\nabla\mathbf{a} \cdot \mathbf{c}\right] \cdot \dot{\mathbf{r}} \\ &- \frac{1}{C_{\mathrm{h}}}\frac{q_\parallel q_\perp}{B}\frac{\partial\mathbf{b}}{\partial t} \cdot \mathbf{a} - \frac{1}{C_{\mathrm{h}}}\frac{q_\perp^2}{2B}\frac{\partial\mathbf{c}}{\partial t} \cdot \mathbf{a} + \frac{1}{C_{\mathrm{h}}}\frac{q_\perp^2}{2B}\,\dot{\alpha} - \frac{1}{2}\,C_{\mathrm{h}}\,\boldsymbol{\rho}_1 \cdot \nabla\frac{\partial\mathbf{A}}{\partial t} \cdot \boldsymbol{\rho}_1 - H_2.\end{aligned} \tag{A27}$$

We remark that the terms $\nabla\mathbf{a} \cdot \mathbf{c}$ and $\partial_t\mathbf{c} \cdot \mathbf{a}$ are gyro-gauge terms which make the dynamics independent of the chosen basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{b})$. The parallel components of the terms multiplying

$\dot{\mathbf{r}}$ can be absorbed into $G_1^{\parallel}$ by setting

$$
\begin{aligned}
G_1^{\parallel} &= \frac{1}{C_{\rm h}}\frac{q_\parallel q_\perp}{B}\,\mathbf{b}\cdot\mathbf{a}\times(\nabla\times\mathbf{b}) + \frac{1}{C_{\rm h}}\frac{q_\perp^2}{B}\,\mathbf{b}\cdot\mathbf{a}\times(\nabla\times\mathbf{c}) + \frac{1}{2}\,C_{\rm h}\,\mathbf{b}\cdot\boldsymbol{\rho}_1\times(\boldsymbol{\rho}_1\cdot\nabla\mathbf{B}) \\
&\qquad + \frac{1}{C_{\rm h}}\frac{q_\perp^2}{2B}\,\mathbf{b}\cdot\nabla\mathbf{a}\cdot\mathbf{c} \qquad\qquad \text{(A28)}\\
&= -\frac{1}{C_{\rm h}}\frac{q_\parallel q_\perp}{B}\,\mathbf{c}\cdot\nabla\times\mathbf{b} - \frac{1}{C_{\rm h}}\frac{q_\perp^2}{B}\,\mathbf{c}\cdot\nabla\times\mathbf{c} - \frac{1}{C_{\rm h}}\frac{q_\perp^2}{B}\,\mathbf{a}\cdot\nabla\mathbf{b}\cdot\mathbf{c} + \frac{1}{C_{\rm h}}\frac{q_\perp^2}{2B}\,\mathbf{b}\cdot\nabla\mathbf{a}\cdot\mathbf{c} \\
&= -\frac{1}{C_{\rm h}}\frac{q_\parallel q_\perp}{B}\,\mathbf{c}\cdot\nabla\times\mathbf{b} - \frac{1}{C_{\rm h}}\frac{q_\perp^2}{2B}\,\mathbf{b}\cdot\nabla\mathbf{a}\cdot\mathbf{c}\,,
\end{aligned}
$$

where to get to the last line we used

$$
\mathbf{c}\cdot\nabla\times\mathbf{c} = \mathbf{c}\cdot\nabla\times(\mathbf{a}\times\mathbf{b}) = \mathbf{b}\cdot\nabla\mathbf{a}\cdot\mathbf{c} - \mathbf{a}\cdot\nabla\mathbf{b}\cdot\mathbf{c}\,. \tag{A29}
$$

The components of the terms multiplying $\dot{\mathbf{r}}$ which are perpendicular to $\mathbf{b}$ can be absorbed into $\boldsymbol{\rho}_2$, by multiplying $\mathbf{b}\times$ and setting

$$
\begin{aligned}
C_{\rm h}B\,\boldsymbol{\rho}_{2,\perp} &= C_{\rm h}\,\mathbf{b}\times\boldsymbol{\rho}_2\times\mathbf{B} \qquad\qquad \text{(A30)}\\
&= G_1^{\perp}\,\mathbf{b}\times\mathbf{c} - q_\perp G_1^{\theta}\,\mathbf{b}\times\mathbf{a} - \frac{1}{C_{\rm h}}\frac{q_\parallel q_\perp}{B}\mathbf{b}\times\big[\mathbf{a}\times(\nabla\times\mathbf{b})\big] - \frac{1}{C_{\rm h}}\frac{q_\perp^2}{B}\mathbf{b}\times\big[\mathbf{a}\times(\nabla\times\mathbf{c})\big] \\
&\quad - \frac{1}{C_{\rm h}}\frac{q_\perp^2}{2B^2}\,\mathbf{b}\times\big[\mathbf{a}\times(\mathbf{a}\cdot\nabla\mathbf{B})\big] - \frac{1}{C_{\rm h}}\frac{q_\perp^2}{2B}\,\mathbf{b}\times\nabla\mathbf{a}\cdot\mathbf{c} \\
&= G_1^{\perp}\,\mathbf{a} + q_\perp G_1^{\theta}\,\mathbf{c} - \frac{1}{C_{\rm h}}\frac{q_\parallel q_\perp}{B}\big(\mathbf{b}\cdot\nabla\times\mathbf{b}\big)\,\mathbf{a} - \frac{1}{C_{\rm h}}\frac{q_\perp^2}{B}\big(\mathbf{b}\cdot\nabla\times\mathbf{c}\big)\,\mathbf{a} \\
&\quad - \frac{1}{C_{\rm h}}\frac{q_\perp^2}{2B^2}\,(\mathbf{a}\cdot\nabla B)\,\mathbf{a} - \frac{1}{C_{\rm h}}\frac{q_\perp^2}{2B}\,\mathbf{b}\times\nabla\mathbf{a}\cdot\mathbf{c}\,.
\end{aligned}
$$

The Hamiltonian part of (A27) is absorbed into $G_2^{\perp}$ from H_2 by setting

$$
\begin{aligned}
q_\perp G_2^{\perp} &= -q_\parallel\,G_2^{\parallel} - \frac{(G_1^{\parallel})^2}{2} - \frac{(G_1^{\perp})^2}{2} + C_{\rm h}\,\left[\frac{1}{2}\boldsymbol{\rho}_1\cdot\nabla\mathbf{E}\cdot\boldsymbol{\rho}_1 + \boldsymbol{\rho}_2\cdot\nabla\phi(\mathbf{r},t)\right] \\
&\qquad - \frac{1}{C_{\rm h}}\frac{q_\parallel q_\perp}{B}\,\frac{\partial\mathbf{b}}{\partial t}\cdot\mathbf{a} - \frac{1}{C_{\rm h}}\frac{q_\perp^2}{2B}\,\frac{\partial\mathbf{c}}{\partial t}\cdot\mathbf{a}\,.
\end{aligned} \tag{A31}
$$

Inserting now into L_2 from (A27) the generators $G_1^{\parallel}$, $\boldsymbol{\rho}_{2,\perp}$ and $G_2^{\perp}$ from (A28), (A30) and (A31), respectively, yields

$$
L_2 \sim \frac{1}{C_{\rm h}}\frac{q_\perp^2}{2B}\,\dot{\alpha}\,. \tag{A32}
$$

From L_0 in (A17a), L_1 in (A22) and L_2 in (A32) we obtain the Lagrangian L^ε to second order as

$$
L^\varepsilon \sim \big[\varepsilon\,q_\parallel\mathbf{b}(\mathbf{r},t) + C_{\rm h}\,\mathbf{A}(\mathbf{r},t)\big]\cdot\dot{\mathbf{r}} + \varepsilon^2\frac{1}{C_{\rm h}}\frac{q_\perp^2}{2B(\mathbf{r},t)}\,\dot{\alpha} - \frac{q_\parallel^2}{2} - \frac{q_\perp^2}{2} - C_{\rm h}\,\phi(\mathbf{r},t) + O(\varepsilon^3)\,. \tag{A33}
$$

3. Equations of motion

The drift-kinetic equations of motion are obtained from the Lagrangian (A33) by adopting the coefficient of $\dot{\alpha}$ as one of the coordinates, called the "generalized magnetic moment":

$$\widehat{\mu} := \frac{q_\perp^2}{2B(\mathbf{r},t)}\,. \tag{A34}$$

The third piece $\widehat{\tau}$ of the GY-transformation is thus defined by

$$\widehat{\tau}: \quad \widehat{\mu} \mapsto q_\perp\,, \qquad q_\perp = \sqrt{2\,\widehat{\mu}\,B}\,. \tag{A35}$$

The GY-coordinates are thus denoted by $\widehat{\mathbf{q}} = (\mathbf{r}, q_\parallel, \widehat{\mu}, \alpha) \in \widehat{\Omega}_{\text{gy}} \subset \mathbb{R}^6$. This yields the second-order drift-kinetic Lagrangian

$$L_{\text{D}} = \big[\varepsilon\, q_\parallel \mathbf{b}(\mathbf{r},t) + C_{\text{h}}\, \mathbf{A}(\mathbf{r},t)\big] \cdot \dot{\mathbf{r}} + \varepsilon^2\, \frac{\widehat{\mu}}{C_{\text{h}}}\, \dot{\alpha} - \frac{q_\parallel^2}{2} - \widehat{\mu} B(\mathbf{r},t) - C_{\text{h}}\, \phi(\mathbf{r},t)\,. \tag{A36}$$

Defining

$$\mathbf{A}^* := \mathbf{A} + \varepsilon\, \frac{q_\parallel}{C_{\text{h}}} \mathbf{b}\,, \qquad \mathbf{B}^* := \nabla \times \mathbf{A}^*\,, \qquad \mathbf{E}^* := \mathbf{E} - \frac{\widehat{\mu}}{C_{\text{h}}} \nabla B - \varepsilon\, \frac{q_\parallel}{C_{\text{h}}}\, \frac{\partial \mathbf{b}}{\partial t}\,, \tag{A37}$$

The Euler-Lagrange equations of L_{D} for the variable $\mathbf{r}$ are obtained from

$$\left.\frac{\partial L_{\text{D}}}{\partial \mathbf{r}}\right|_{\widehat{\mathbf{q}}(t)} = C_{\text{h}}\, \nabla \mathbf{A}^* \cdot \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} - \widehat{\mu}\, \nabla B - C_{\text{h}}\, \nabla \phi\,, \tag{A38}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} \left.\frac{\partial L_{\text{D}}}{\partial \dot{\mathbf{r}}}\right|_{\widehat{\mathbf{q}}(t)} = C_{\text{h}}\, \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} \cdot \nabla \mathbf{A}^* + C_{\text{h}}\, \frac{\partial \mathbf{A}}{\partial t} + \varepsilon\, q_\parallel\, \frac{\partial \mathbf{b}}{\partial t} + \varepsilon\, \frac{\mathrm{d}q_\parallel}{\mathrm{d}t}\, \mathbf{b}\,, \tag{A39}$$

and read

$$0 = \left.\frac{\partial L_{\text{D}}}{\partial \mathbf{r}}\right|_{\widehat{\mathbf{q}}(t)} - \frac{\mathrm{d}}{\mathrm{d}t} \left.\frac{\partial L_{\text{D}}}{\partial \dot{\mathbf{r}}}\right|_{\widehat{\mathbf{q}}(t)} = C_{\text{h}}\, \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} \times \mathbf{B}^* + C_{\text{h}}\, \mathbf{E}^* - \varepsilon\, \frac{\mathrm{d}q_\parallel}{\mathrm{d}t}\, \mathbf{b}\,. \tag{A40}$$

Multiplying $\mathbf{b}\times$ from the left and defining $B_\parallel^* := \mathbf{B}^* \cdot \mathbf{b}$ we obtain

$$0 = B_\parallel^*\, \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} - \left(\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} \cdot \mathbf{b}\right) \mathbf{B}^* + \mathbf{b} \times \mathbf{E}^*\,. \tag{A41}$$

Projecting (A40) on $\mathbf{B}^*$ yields

$$0 = C_{\text{h}}\, \mathbf{E}^* \cdot \mathbf{B}^* - \varepsilon\, \frac{\mathrm{d}q_\parallel}{\mathrm{d}t}\, B_\parallel^*\,. \tag{A42}$$

Moreover, for the velocity degrees of freedom we obtain

$$0 = \left.\frac{\partial L_{\text{D}}}{\partial q_\parallel}\right|_{\widehat{\mathbf{q}}(t)} - \frac{\mathrm{d}}{\mathrm{d}t} \left.\frac{\partial L_{\text{D}}}{\partial \dot{q}_\parallel}\right|_{\widehat{\mathbf{q}}(t)} = \varepsilon\, \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} \cdot \mathbf{b} - q_\parallel\,, \tag{A43}$$

$$0 = \left.\frac{\partial L_{\text{D}}}{\partial \widehat{\mu}}\right|_{\widehat{\mathbf{q}}(t)} - \frac{\mathrm{d}}{\mathrm{d}t} \left.\frac{\partial L_{\text{D}}}{\partial \dot{\widehat{\mu}}}\right|_{\widehat{\mathbf{q}}(t)} = \varepsilon^2\, \frac{1}{C_{\text{h}}}\, \frac{\mathrm{d}\alpha}{\mathrm{d}t} - B\,, \tag{A44}$$

$$0 = \left.\frac{\partial L_{\text{D}}}{\partial \alpha}\right|_{\widehat{\mathbf{q}}(t)} - \frac{\mathrm{d}}{\mathrm{d}t} \left.\frac{\partial L_{\text{D}}}{\partial \dot{\alpha}}\right|_{\widehat{\mathbf{q}}(t)} = -\varepsilon^2\, \frac{1}{C_{\text{h}}}\, \frac{\mathrm{d}\widehat{\mu}}{\mathrm{d}t}\,, \tag{A45}$$

so that the drift-kinetic equations of motions read

$$\varepsilon\,\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \mathbf{u}_\mathrm{g}(\widehat{\mathbf{q}},t)\,, \qquad \mathbf{u}_\mathrm{g}(\widehat{\mathbf{q}},t) := q_\parallel\,\frac{\mathbf{B}^*}{B_\parallel^*} + \varepsilon\,\frac{\mathbf{E}^*\times\mathbf{b}}{B_\parallel^*}\,, \tag{A46a}$$

$$\varepsilon\,\frac{\mathrm{d}q_\parallel}{\mathrm{d}t} = E_\mathrm{g}(\widehat{\mathbf{q}},t)\,, \qquad E_\mathrm{g}(\widehat{\mathbf{q}},t) := C_\mathrm{h}\,\frac{\mathbf{E}^*\cdot\mathbf{B}^*}{B_\parallel^*}\,, \tag{A46b}$$

$$\varepsilon\,\frac{\mathrm{d}\widehat{\mu}}{\mathrm{d}t} = 0\,, \tag{A46c}$$

$$\varepsilon\,\frac{\mathrm{d}\alpha}{\mathrm{d}t} = \frac{C_\mathrm{h}\,B}{\varepsilon}\,. \tag{A46d}$$

Here, we call $\mathbf{u}_\mathrm{g}$ the GY-velocity and E_g the GY-acceleration. Solutions of this system approximate the true characteristics obtained from (A2) with an error of order $O(\varepsilon)$.

Let us elaborate on the geometric structure of the drift-kinetic equations of motion (A46). This is useful for a better understanding of the ensuing drift-kinetic equation and for the corresponding conservation laws of mass, momentum and energy. Due to its origin from the Lagrangian L_D in (A36), the system (A46) is in fact a non-canonical symplectic system. The symplectic structure is manifested by a Poisson bracket, defined via a skew-symmetric Poisson matrix. This has important consequences, for instance the conservation of phase space volume under the flow of the system. We will now clarify these notions in more detail.

Let us write the Lagrangian (A36) in the form

$$L_\mathrm{D} = \boldsymbol{\gamma}\cdot\dot{\widehat{\mathbf{q}}} - H\,, \tag{A47}$$

where $\boldsymbol{\gamma}$ denotes the symplectic form and H stands for the Hamiltonian, given by

$$\boldsymbol{\gamma} = (C_\mathrm{h}\,\mathbf{A}^*, 0, 0, \varepsilon^2\,\widehat{\mu}) \in \mathbb{R}^6\,, \qquad H = \frac{q_\parallel^2}{2} + C_\mathrm{h}\,(\phi + \widehat{\mu}B)\,. \tag{A48}$$

The Euler-Lagrange equations read

$$0 = \left.\frac{\partial L_\mathrm{D}}{\partial\widehat{\mathbf{q}}}\right|_{\widehat{\mathbf{q}}(t)} - \frac{\mathrm{d}}{\mathrm{d}t}\left.\frac{\partial L_\mathrm{D}}{\partial\dot{\widehat{\mathbf{q}}}}\right|_{\widehat{\mathbf{q}}(t)} = \left(\frac{\partial\boldsymbol{\gamma}}{\partial\widehat{\mathbf{q}}}^\top - \frac{\partial\boldsymbol{\gamma}}{\partial\widehat{\mathbf{q}}}\right)\cdot\frac{\mathrm{d}\widehat{\mathbf{q}}}{\mathrm{d}t} - \frac{\partial H}{\partial\widehat{\mathbf{q}}} - \frac{\partial\boldsymbol{\gamma}}{\partial t}\,. \tag{A49}$$

Here, $\partial\boldsymbol{\gamma}/\partial\widehat{\mathbf{q}}$ stands for the Jacobian. The Lagrange matrix is defined as

$$\boldsymbol{\omega} := \frac{\partial\boldsymbol{\gamma}}{\partial\widehat{\mathbf{q}}}^\top - \frac{\partial\boldsymbol{\gamma}}{\partial\widehat{\mathbf{q}}} = \left(\begin{array}{cc|c} C_\mathrm{h}\,\mathcal{B}^* & -\varepsilon\,\mathbf{b} & \multirow{2}{*}{$\mathbf{0}$} \\ \varepsilon\,\mathbf{b}^\top & 0 & \\ \hline \multicolumn{2}{c|}{\multirow{2}{*}{$\mathbf{0}$}} & \begin{matrix} 0 & \varepsilon^2 \end{matrix} \\ \multicolumn{2}{c|}{} & \begin{matrix} -\varepsilon^2 & 0 \end{matrix} \end{array}\right) \in \mathbb{R}^{6\times 6}\,, \tag{A50}$$

where $\mathcal{B}^* \in \mathbb{R}^{3\times 3}$ is the skew-symmetric matrix $\nabla \mathbf{A}^* - (\nabla \mathbf{A}^*)^\top$, such that $\mathcal{B}^* \cdot \mathbf{v} = \mathbf{v} \times \mathbf{B}^*$, that is

$$\mathcal{B}^* = \begin{pmatrix} 0 & B_3^* & -B_2^* \\ -B_3^* & 0 & B_1^* \\ B_2^* & -B_1^* & 0 \end{pmatrix} \in \mathbb{R}^{3\times 3} . \tag{A51}$$

The Euler-Lagrange equations (A49) can thus be written as

$$\frac{\boldsymbol{\omega}}{\varepsilon} \cdot \varepsilon \frac{\mathrm{d}\widehat{\mathbf{q}}}{\mathrm{d}t} = \frac{\partial H}{\partial \widehat{\mathbf{q}}} + \frac{\partial \boldsymbol{\gamma}}{\partial t} . \tag{A52}$$

We factorized an ε in front of the time derivative in order to stay consistent with the time scale from (A46). Now the first interesting question is whether $\boldsymbol{\omega}/\varepsilon$ is invertible. For this we check the determinant:

$$\det\left(\frac{\boldsymbol{\omega}}{\varepsilon}\right) = \varepsilon^2 \det \begin{pmatrix} C_\mathrm{h}\, \mathcal{B}^*/\varepsilon & -\mathbf{b} \\ \mathbf{b}^\top & 0 \end{pmatrix} = (B_\parallel^*)^2 . \tag{A53}$$

For $B_\parallel^* \neq 0$ we define the Poisson matrix $\mathbf{J} := (\boldsymbol{\omega}/\varepsilon)^{-1}$, which is also skew symmetric, and write the dynamics as

$$\varepsilon \frac{\mathrm{d}\widehat{\mathbf{q}}}{\mathrm{d}t} = \mathbf{C}(\widehat{\mathbf{q}}) := \mathbf{J} \cdot \left(\frac{\partial H}{\partial \widehat{\mathbf{q}}} + \frac{\partial \boldsymbol{\gamma}}{\partial t} \right) . \tag{A54}$$

This is a symplectic system of which the direction field $\mathbf{C}(\widehat{\mathbf{q}})$ has been computed explicitly in (A46). However, due to the time dependence of the electromagnetic fields, its geometric properties can only be derived in the "extended phase space", where time and energy are added as additional coordinates. Hence, let $\widehat{\mathbf{q}}_e = (\widehat{\mathbf{q}}, t, w) \in \widehat{\Omega}_e \subset \mathbb{R}^8$ denote the extended coordinates and consider the extended Lagrangian

$$L_{\mathrm{D},e} = \boldsymbol{\gamma} \cdot \dot{\widehat{\mathbf{q}}} - \varepsilon\, w\, \dot{t} - H + \varepsilon\, w\, , \tag{A55}$$

where we can define the extended symplectic form $\boldsymbol{\gamma}_e$ and extended Hamiltonian H_e as

$$\boldsymbol{\gamma}_e = (\boldsymbol{\gamma}, -\varepsilon\, w, 0) \in \mathbb{R}^8\, , \qquad H_e = H - \varepsilon\, w\, . \tag{A56}$$

We remark that neither $\boldsymbol{\gamma}_e$ nor H_e carry an explicit dependence on the dynamical parameter (time-like), as time is now one of the coordinates of the manifold. The extended Lagrange matrix and its determinant read

$$\boldsymbol{\omega}_e = \left(\begin{array}{c|c} \boldsymbol{\omega} & \mathbf{0} \\ \hline \mathbf{0} & \begin{matrix} 0 & \varepsilon \\ -\varepsilon & 0 \end{matrix} \end{array}\right) , \qquad \det\left(\frac{\boldsymbol{\omega}_e}{\varepsilon}\right) = \det\left(\frac{\boldsymbol{\omega}}{\varepsilon}\right) = (B_\parallel^*)^2 . \tag{A57}$$

Hence, for $B_\parallel^*$ we define the extended Poisson matrix $\mathbf{J}_e := (\boldsymbol{\omega}_e/\varepsilon)^{-1}$ and write the extended dynamics as

$$\varepsilon\,\frac{\mathrm{d}\widehat{\mathbf{q}}_e}{\mathrm{d}t} = \mathbf{C}_\mathrm{e}(\widehat{\mathbf{q}}_\mathrm{e}) := \mathbf{J}_e \cdot \frac{\partial H_e}{\partial \widehat{\mathbf{q}}_e}\,. \tag{A58}$$

Explicitly, the two additional Euler-Lagrange equations are

$$0 = \frac{\partial L_{\mathrm{D},e}}{\partial t}\bigg|_{\widehat{\mathbf{q}}_e(s)} - \frac{\mathrm{d}}{\mathrm{d}s}\frac{\partial L_{\mathrm{D},e}}{\partial \dot{t}}\bigg|_{\widehat{\mathbf{q}}_e(s)} = \frac{\partial \boldsymbol{\gamma}}{\partial t}\cdot\frac{\mathrm{d}\widehat{\mathbf{q}}}{\mathrm{d}s} - \frac{\partial H}{\partial t} + \varepsilon\,\frac{\mathrm{d}w}{\mathrm{d}s}\,. \tag{A59}$$

$$0 = \frac{\partial L_{\mathrm{D},e}}{\partial w}\bigg|_{\widehat{\mathbf{q}}_e(s)} - \frac{\mathrm{d}}{\mathrm{d}s}\frac{\partial L_{\mathrm{D},e}}{\partial \dot{w}}\bigg|_{\widehat{\mathbf{q}}_e(s)} = -\varepsilon\,\frac{\mathrm{d}t}{\mathrm{d}s} + \varepsilon\,. \tag{A60}$$

The "Poisson bracket" for two functions $F, G : \widehat{\Omega}_e \to \mathbb{R}$ in the extended phase space is defined as

$$\{F, G\}_e := \frac{\partial F}{\partial \widehat{\mathbf{q}}_e}\cdot \mathbf{J}_e \cdot \frac{\partial G}{\partial \widehat{\mathbf{q}}_e}\,. \tag{A61}$$

From (A58) it then follows that all functions in the kernel of $\{\cdot, H_e\}_e$ are constants of the motion. The trivial one is the energy H_e itself and all functions thereof. However, there are also less obvious ones like ...

Another important property of non-canonical symplectic systems is the conservation of phase space volume. This is equivalent to satisfying the "Liouville property": let $\mathbf{q} \in \Omega \subset \mathbb{R}^{2n}$ stand for some coordinates on an even-dimensional phase space, further let $\mathbf{J} : \Omega \to \mathbb{R}^{2\times 2n}$ denote a skew-symmetric invertible Poisson matrix and let $H : \Omega \to \mathbb{R}$ be the Hamiltonian. Then, assuming non-canonical symplectic dynamics, we have

$$\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t} = \mathbf{C}(\mathbf{q}) := \mathbf{J}\cdot\frac{\partial H}{\partial \mathbf{q}} \qquad \Longrightarrow \qquad \mathrm{div}_\mathbf{q}\left(\sqrt{\det \mathbf{J}^{-1}}\,\mathbf{C}(\mathbf{q})\right) = 0\,. \tag{A62}$$

The proof of this property relies on Darboux's theorem and is left as an exercise to the reader. Applying the Liouville property to the extended dynamics (A58) means

$$\begin{aligned} 0 &= \mathrm{div}_{\widehat{\mathbf{q}}_e}\big(B_\parallel^*\,\mathbf{C}_\mathrm{e}(\widehat{\mathbf{q}}_e)\big) \\ &= \mathrm{div}_{\widehat{\mathbf{q}}}\big(B_\parallel^*\,\mathbf{C}(\widehat{\mathbf{q}})\big) + \varepsilon\,\frac{\partial}{\partial t}B_\parallel^* \\ &= \frac{\partial}{\partial \mathbf{r}}\big(B_\parallel^*\,\mathbf{u}_\mathrm{g}\big) + \frac{\partial}{\partial q_\parallel}\big(B_\parallel^*\,E_\mathrm{g}\big) + \varepsilon\,\frac{\partial}{\partial t}B_\parallel^*\,, \end{aligned} \tag{A63}$$

where $\mathbf{C}(\widehat{\mathbf{q}})$ denotes the direction field (ie. the right-hand side) of (A46) and we also inserted

(A59) and (A60). Let us verify (A63) also by direct computation:

$$
\begin{aligned}
\varepsilon \frac{\partial}{\partial t} B_\parallel^* &= \varepsilon\, \mathbf{b} \cdot \frac{\partial \mathbf{B}^*}{\partial t} + \varepsilon\, \mathbf{B}^* \cdot \frac{\partial \mathbf{b}}{\partial t}\,,\\
&= \varepsilon\, \mathbf{b} \cdot \frac{\partial \mathbf{B}}{\partial t} + \varepsilon^2 \frac{q_\parallel}{C_\mathrm{h}}\, \mathbf{b} \cdot \nabla \times \frac{\partial \mathbf{b}}{\partial t} + \varepsilon\, \mathbf{B}^* \cdot \frac{\partial \mathbf{b}}{\partial t}\,,\\
\nabla \cdot (B_\parallel^*\, \mathbf{u}_\mathrm{g}) &= \nabla \cdot \big(q_\parallel\, \mathbf{B}^* + \varepsilon\, \mathbf{E}^* \times \mathbf{b}\big)\\
&= q_\parallel\, \nabla \cdot \mathbf{B}^* + \varepsilon\, \mathbf{b} \cdot \nabla \times \mathbf{E}^* - \varepsilon\, \mathbf{E}^* \cdot \nabla \times \mathbf{b}\\
&= \varepsilon\, \mathbf{b} \cdot \nabla \times \mathbf{E} - \varepsilon^2 \frac{q_\parallel}{C_\mathrm{h}}\, \mathbf{b} \cdot \nabla \times \frac{\partial \mathbf{b}}{\partial t} + \varepsilon\, \mathbf{E}^* \cdot \nabla \times \mathbf{b}\,,\\
\frac{\partial}{\partial q_\parallel}(E_\mathrm{g}\, B_\parallel^*) &= C_\mathrm{h}\, \mathbf{B}^* \cdot \frac{\partial \mathbf{E}^*}{\partial q_\parallel} + C_\mathrm{h}\, \mathbf{E}^* \cdot \frac{\partial \mathbf{B}^*}{\partial q_\parallel}\\
&= -\varepsilon\, \mathbf{B}^* \cdot \frac{\partial \mathbf{b}}{\partial t} + \varepsilon\, \mathbf{E}^* \cdot \nabla \times \mathbf{b}\,.
\end{aligned}
$$

Therefore, summing up these terms, we obtain that the Liouville property holds if

$$
\varepsilon\, \mathbf{b} \cdot \left(\frac{\partial \mathbf{B}}{\partial t} + \nabla \times \mathbf{E} \right) = 0\,, \tag{A64}
$$

which is satisfied because of Faraday's law.

4. Drift-kinetic equation

Suppose that $F_\mathrm{h} = F_\mathrm{h}(t, \widehat{\mathbf{q}}) = F_\mathrm{h}(t, \mathbf{r}, q_\parallel, \widehat{\mu}, \alpha)$ with $F_\mathrm{h} = f_\mathrm{h}^\mathrm{D} B_\parallel^*$ denotes a distribution of energetic particles in the GY-phase-space. Demanding conservation of mass (transport theorem) then yields

$$
\frac{\partial F_\mathrm{h}}{\partial t} + \mathrm{div}_{\widehat{\mathbf{q}}} \big(\mathbf{C}(\widehat{\mathbf{q}})\, F_\mathrm{h} \big) = 0\,, \tag{A65}
$$

where $\mathbf{C}(\widehat{\mathbf{q}})$ denotes the right-hand-side of the drift-kinetic characteristic equations (A46). Due to the Liouville property (A63), this equation is equivalent to

$$
\frac{\partial f_\mathrm{h}^\mathrm{D}}{\partial t} + \mathbf{C}(\widehat{\mathbf{q}}) \cdot \frac{\partial f_\mathrm{h}^\mathrm{D}}{\partial \widehat{\mathbf{q}}} = 0\,. \tag{A66}
$$

Supposing that initially f_h^D is independent of the gyro-angle α yields the drift-kinetic equation

$$
\frac{\partial f_\mathrm{h}^\mathrm{D}}{\partial t} + \mathbf{u}_\mathrm{g} \cdot \frac{\partial f_\mathrm{h}^\mathrm{D}}{\partial \mathbf{r}} + E_\mathrm{g} \frac{\partial f_\mathrm{h}^\mathrm{D}}{\partial q_\parallel} = 0\,, \tag{A67}
$$

where $\mathbf{u}_\mathrm{g}$ and E_g have been defined in (A46a) and (A46b), respectively.

5. First order generators: summary

In practice, instead of starting from the coordinates $\mathbf{q}' = (\mathbf{x}, v_\parallel, v_\perp, \theta)$, it will be convenient to do the preliminary transform to a set of coordinates $\mathbf{q}'' = (\mathbf{x}, v_\parallel, \mu, \theta)$, which features the magnetic moment μ, defined by

$$\mu := \frac{v_\perp^2}{2B(\mathbf{x},t)}\,. \tag{A68}$$

This is in fact the leading order of the generalized magnetic moment $\widehat{\mu}$ defined in (A34). The corresponding Gy-transform reads

$$\begin{aligned}\mu &= \frac{v_\perp^2}{2B(\mathbf{x},t)}\\ &= \frac{(q_\perp + \varepsilon\, G_1^\perp)^2}{2B(\mathbf{r}+\varepsilon\,\boldsymbol{\rho}_1,t)} + O(\varepsilon^2)\\ &= \widehat{\mu} + \varepsilon\,\left[\frac{q_\perp\, G_1^\perp}{B(\mathbf{r},t)} - \frac{q_\perp^2}{2B^2(\mathbf{r},t)}\,\boldsymbol{\rho}_1\cdot\nabla B(\mathbf{r},t)\right] + O(\varepsilon^2)\,,\end{aligned} \tag{A69}$$

where the generating functions are to be evaluated at the coordinates $\widehat{\mathbf{q}}$. In summary, the first order inverse GY-transform is determined by the following generating functions, derived in (A20), (A21) and in (A28):

$$\boldsymbol{\rho}_1 = \left(\frac{2\widehat{\mu}}{B}\right)^{1/2}\mathbf{a}\,, \tag{A70a}$$

$$G_1^\parallel = -\frac{q_\parallel}{C_\mathrm{h}}\left(\frac{2\widehat{\mu}}{B}\right)^{1/2}\mathbf{c}\cdot\nabla\times\mathbf{b} - \frac{\widehat{\mu}}{C_\mathrm{h}}\,\mathbf{b}\cdot\nabla\mathbf{a}\cdot\mathbf{c} \tag{A70b}$$

$$\begin{aligned}G_1^\mu = \frac{q_\parallel^2}{B}\left(\frac{2\widehat{\mu}}{B}\right)^{1/2}\mathbf{c}\cdot\nabla\times\mathbf{b} + \frac{q_\parallel\,\widehat{\mu}}{C_\mathrm{h}\,B}\,\mathbf{b}\cdot\nabla\mathbf{a}\cdot\mathbf{c}\\ + \frac{1}{B}\left(\frac{2\widehat{\mu}}{B}\right)^{1/2}\mathbf{a}\cdot(\mathbf{E} - \widehat{\mu}\,\nabla B)\,,\end{aligned} \tag{A70c}$$

where G_1^μ has been defined from (A69) via

$$G_1^\mu := \left(\frac{2\widehat{\mu}}{B}\right)^{1/2} G_1^\perp - \frac{\widehat{\mu}}{B}\left(\frac{2\widehat{\mu}}{B}\right)^{1/2}\mathbf{a}\cdot\nabla B\,,$$

with

$$\begin{aligned}G_1^\perp &= -\frac{q_\parallel}{(2\widehat{\mu}\,B)^{1/2}}\,G_1^\parallel + \frac{1}{B}\,\mathbf{a}\cdot\mathbf{E}\\ &= \frac{q_\parallel^2}{B}\,\mathbf{c}\cdot\nabla\times\mathbf{b} + \frac{q_\parallel}{C_\mathrm{h}}\left(\frac{\widehat{\mu}}{2B}\right)^{1/2}\mathbf{b}\cdot\nabla\mathbf{a}\cdot\mathbf{c} + \frac{1}{B}\,\mathbf{a}\cdot\mathbf{E}\,.\end{aligned} \tag{A71}$$

REFERENCES

[1]Joshua W Burby and Cesare Tronci. Variational approach to low-frequency kinetic-MHD in the current coupling scheme. *Plasma Physics and Controlled Fusion*, 59(4):045013, 2017.

[2]Johan Peter Goedbloed and Stefaan Poedts. *Principles of magnetohydrodynamics: with applications to laboratory and astrophysical plasmas*. Cambridge university press, 2004.

[3]Stefan Possanner. Gyrokinetics from variational averaging: existence and error bounds. *arXiv preprint arXiv:1711.09620*, November 2017.